

Albrecht Schmidt, Paul Holleis
Matthias Kranz, Andreas Butz (Hrsg.)

Kollaboration, Nachrichtenaustausch und Communities

Hauptseminar Medieninformatik SS 2006

Technical Report
LMU-MI-2006-2, Aug. 2006
ISSN 1862-5207



University of Munich
Department of Computer Science
Media Informatics Group

Albrecht Schmidt, Paul Holleis, Matthias Kranz, Andreas Butz (Herausgeber)

Kollaboration, Nachrichtenaustausch und Communities

Ein Überblick über Forschungsarbeiten im Bereich der
Mensch-Maschine-Interaktion

Vorwort

In den letzten 10 Jahren haben sich massive Veränderungen im Bereich der Benutzungsschnittstellen vollzogen. Mit diesem Bericht über *Kollaboration, Nachrichtenaustausch und Communities* wollen wir kompakt vergangene und aktuelle Entwicklungen und Forschungstrends im Bereich Mensch-Maschine-Interaktion einem interessierten Fachpublikum zugänglich machen. Dazu analysieren Studenten und Mitarbeiter des Lehrstuhl Medieninformatik an der Ludwig-Maximilians-Universität München Beiträge und Veröffentlichungen aus zahlreichen Workshops, Konferenzen und wissenschaftlichen Zeitschriften. Diese Reihe an Texten wurde im Zusammenhang mit einem Hauptseminar im Sommersemester 2006 von Studenten erstellt.

Jedes der einzelnen Kapitel greift ein Thema auf und stellt wesentliche Forschungstrends in diesem Bereich in kurzen Artikeln in deutscher Sprache vor. Im Rahmen des Seminars wurde von den Studenten zu den jeweiligen Themen Vorträge gehalten, die die wesentlichen Inhalte noch einmal zusammenfassten. Die Folien zu diesen Vorträgen und weitere Informationen über die Veranstaltung können auf der Webseite <http://www.hcilab.org/events/collaboration/> eingesehen werden.

Dieser Bericht richtet sich in erster Linie an Informatiker, Medieninformatiker, User Interface Designer und Studenten der Informatik und Medieninformatik. Speziell der Aspekt der Unterstützung von Gemeinschaften mit Hilfe verschiedenen Technologien sowohl im privaten wie im Arbeitsumfeld findet in aktuellster Forschung große Beachtung. Im weiteren Sinne richtet sich der Bericht an alle die sich für neue Formen der Interaktion zwischen Mensch und Maschine interessieren.

Veranstalter und Herausgeber sind die beiden von der DFG geförderten Nachwuchsforschungsgruppen „Embedded Interaction“ (<http://www.hcilab.org>) und „Fluidum“ (<http://www.fluidum.org>) an der Ludwig-Maximilians-Universität München.

München, August 2006

Die Herausgeber

Albrecht Schmidt

Paul Holleis

Matthias Kranz

Andreas Butz

Inhaltsverzeichnis

<i>Maximilian Lindner</i> Kontextsensitive Unterbrechungstechniken von Benachrichtigungssystemen	9
<i>Tobias Emrich</i> Privatsphäre in der Multimediakommunikation am Beispiel Instant Messaging	35
<i>Andreas Attenberger</i> Entwicklungsansätze für CSCW-Anwendungen und verteilte Benutzerschnittstellen.....	54
<i>Daniel Hilkert</i> Collaborative Filtering.....	77
<i>Johannes Jüngst</i> Kollaboration mit Hilfe von physikalischen Objekten	95
<i>Thorsten Fritsche</i> Mehrbenutzer-Kollaboration mit großen Displays	115
<i>Hendrik Richter</i> Kollaboration mittels direkter Interaktion mit großen Displays.....	135
<i>Stefan Seitz</i> Kollaboration mittels entfernter Interaktion mit großen Displays.....	153
<i>Lucie Drasch</i> Die Geschichte der E-Mail und des Instant Messaging	175
<i>Ana Monika Ticaloiu</i> Informationsvisualisierung in E-Mail Anwendungen	191
<i>Claudia Ruch</i> Instant Messaging und Anwendungen.....	211
<i>Ina Zumbruch</i> „Generation Blogging“: Weblogs erobern das Internet.....	229

<i>Thomas Lederer</i>	
Web-basierte Kollaborations-Systeme	249
<i>Gert Kauntz</i>	
Verteilte Musik-Sammlungen und Empfehlungssysteme	267
<i>Benjamin Rabinowitsch</i>	
Mobile Communities und Freundschafts-Suchdienste	285
<i>Christoph Pahre</i>	
Mobile Multimedia Nachrichten-Übertragung	303
Autorenverzeichnis	321

Aufmerksamkeit gewinnen auf dem richtigen Level

Kontextsensitive Unterbrechungstechniken von Benachrichtigungssystemen

Maximilian Lindner

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
maximilian.lindner@gmx.net

Zusammenfassung Benachrichtigungssysteme finden im heutigen mobilen Computerzeitalter immer größere Bedeutung. Unangemessene Benachrichtigungen und daraus resultierende Unterbrechungen des Benutzers sind immer unerwünschter. Um diesem Problem zu begegnen, sollen neue Systeme darauf ausgerichtet sein, in Abhängigkeit der aktuellen Benutzersituation eine passende Unterbrechungsstrategie bereitzustellen. Diese kann neben der klassischen visuellen und auditiven Darstellung auch andere Kanäle der menschlichen Sinneswahrnehmung miteinbeziehen. Weitere wichtige Prinzipien von Benachrichtigungssystemen sind die Integration von Datenabstraktion, Benachrichtigungslevels und Transitionen.

1 Einleitung

Der Mensch des 21. Jahrhunderts lebt im Zeitalter des *Ubiquitous Computing*, der ständigen Allgegenwärtigkeit von intelligenten, sprich datenverarbeitenden, Computersystemen aller Art. Viele dieser Systeme haben die Aufgabe, Informationen zu übermitteln oder nur zu bestimmten Zeitpunkten, abhängig von eingehenden Ereignissen, auf sich aufmerksam zu machen und somit den Benutzer bei seiner aktuellen Beschäftigung zu unterbrechen. Derartige Anwendungen sollen im folgenden Benachrichtigungssysteme genannt werden.

Beispiele hierfür sind jedem geläufig: Der abgespielte Sound des EMail-Client bei einer eingehenden Nachricht, ebenso wie der Alarm eines Weckers oder das Klingeln eines Telefons. Die Entwicklung und Verbreitung solcher Systeme nimmt weiterhin zu: Heutzutage tragen wir das empfangsbereite Handy und den virtuellen Terminkalender im Organizer mit uns herum, sind über Skype und andere Instant Messaging Applications auf dem Notebook mit UMTS-Verbindung erreichbar, sehen auf der Straße dynamische Displays zu den Verspätungen der Straßenbahnen, und werden von Beep-Tönen im Auto daran erinnert, dass wir uns noch nicht angeschnallt haben. Irgendwann werden wir intelligente Kleidung tragen, die uns über den schwachen Akkustand des integrierten MP3-Players informieren will.

Da sich der Mensch nicht auf alles gleichzeitig konzentrieren kann und nicht alle Informationen zu jedem Zeitpunkt relevant sind, operieren viele dieser Systeme oft in der Peripherie unserer Wahrnehmung. Dennoch sind die Leute es heutzutage ge-

wohnt, neben ihrer Haupttätigkeit direkten Zugang zu anderen, davon unabhängigen Ereignissen zu haben. Dafür werden dem Benutzer auf sogenannten peripheren Displays Informationen angezeigt, während er auf andere wichtigere Dinge konzentriert ist. Beim Begriff Display muss man sich von der Vorstellung lösen, dass es sich ausschließlich um einen Computerbildschirm handelt. Vielmehr bezeichnet der Begriff in diesem Artikel sämtliche Objekte, die Zustände und Zustandsänderungen vermitteln können.

Bisher gingen die meisten Benachrichtigungssysteme von der Prämisse aus, dass sie die einzigen Geräte sind, die von einer Person benutzt werden. Wie bereits dargelegt, ist dies immer weniger der Fall, und unkontrollierte, störende Benachrichtigungen werden immer problematischer. Wer will schon gerne in die peinliche Situation kommen, wenn sich während eines Referats das Handy in der Hosentasche mit einem lauten Klingeln zu Wort meldet? In 99 Prozent der Fälle wäre eine solche Unterbrechung mehr als unerwünscht, es sei denn, bei dem Anruf ginge es um den tragischen Unfall eines Familienmitglieds. Anhand dieses Beispiels sieht man, dass je nach Benutzerkontext und Dringlichkeit eines Ereignisses ein anderes Benachrichtigungslevel nötig ist: Unterbrechungen in angemessener Form, abhängig von der aktuellen Situation des Benutzers und der Wichtigkeit der Nachricht, sind daher eine Herausforderung für statische und mobile Computersysteme.

Zur Erklärung und Illustration einiger Aspekte dieses Themas soll der *Motion Monitor* (in Abbildung 1 dargestellt) als ein einfaches Benachrichtigungssystem herangezogen werden: Dabei geht es um die Übermittlung der Aktivität in einem anderen Raum, um eine Art Verbundenheit mit einem entfernten Freund oder Kollegen aufrecht zu erhalten. Eine Kamera ermittelt mit der Technik des Framedifferencing Bewegungen und kann so den Aktivitätsgrad in dem Raum bestimmen. Dieser wird durch unterschiedlich schnelle Farbänderungen eines peripheren Displays in Form einer Kugel für die andere Person visualisiert. Das System basiert auf dem kommerziell vertriebenen *Ambient Orb*, einer Kugel, die Übergänge zwischen vielen tausend verschiedenen Farben darstellen kann, um Veränderungen von Informationen zu vermitteln.



Abbildung 1. Der Motion Monitor (aus [2])

Der Motion Monitor ist ein Repräsentant der *ambient displays*. Ambient displays, zu deutsch Umgebungsanzeigen, können als eine Unterklasse von peripheren Displays angesehen werden: Sie sind ästhetisch ansprechende Anzeigegeräte, die für die Darstellung von nicht-kritischen Informationen ausgelegt sind und dafür konzipiert sind, Hintergrund- oder Kontextinformationen, die der Benutzer nicht ständig aufnehmen möchte, zu präsentieren. Ihr Ziel ist es, Datenverläufe darzustellen, ohne dabei unterbrechend oder ablenkend zu wirken. Sie erfordern nur ein Minimum an Aufmerksamkeit, zeigen Veränderungen einfach, kontinuierlich, in langsamen Bewegungen und meist ohne Interaktionsmöglichkeit an. Die Abbildung der Daten auf das Anzeigegerät ist möglichst natürlich und intuitiv [1]. Umgebungsanzeigen werden im Laufe dieses Artikels öfters angesprochen werden.

2 Eingabe: Kontext und Unterbrechbarkeit des Benutzers

Jede Nachricht hat für die empfangende Person zwei Seiten: Zum einen hat sie einen gewissen Wert, zum anderen stört sie den Benutzer bei seiner aktuellen Beschäftigung. Jeder weiß, dass Unterbrechungen die Arbeitsleistung merklich mindern. Im alltäglichen Leben hat ein Benachrichtigungssystem im Hintergrund meist keinen Bezug zur aktuell ausgeführten Aufgabe, daher kann eine Unterbrechung auch sehr ärgerlich oder gar gefährlich sein [3].

Es erscheint deshalb sehr zweckmäßig, die Art einer Benachrichtigung mit der aktuellen Situation des Benutzers, und natürlich auch mit der Wichtigkeit des Ereignisses, zu gewichten. Die Wichtigkeit der Nachricht ist ein sehr relevanter Punkt; da jedoch deren automatische Erkennung ein sehr komplexes Thema ist und sie auch stark von der individuellen Anwendung abhängt, wird die Dringlichkeit eines Ereignisses in dieser Abhandlung als gegebene Variable angesehen.

Für die Bestimmung des sogenannten user contexts, also der aktuellen Situation einer Person, gibt es zwei Möglichkeiten: Zum einen kann der Benutzer sein gewünschtes Unterbrechungslevel selbst einstellen, was jedoch sehr umständlich und unkomfortabel für den Benutzer wäre und zu eigenverschuldeten Fehlern führen könnte, zum Beispiel beim Vergessen, das Handy während der Vorlesung auszuschalten. Zum anderen besteht die Möglichkeit, bestimmte Sensoren zu verwenden, die dem System Informationen für die Errechnung des wahrscheinlichen Benutzerkontextes liefern.

Verschiedene Kombinationen von Variablen sind denkbar, die zum einen messbar sind und auch eine relevante Aussage über den user context machen. Zwei Ansätze sollen im folgenden vorgestellt werden.

2.1 Ermittlung des situativen Benutzerkontexts durch Sensoren

Der Grundgedanke des Experiments von Nicky Kern und Bernt Schiele an der ETH Zürich besteht darin, Umgebungsinformationen von am Körper getragenen Sensoren dem System zu übermitteln, damit es daraus die entsprechende Benachrichtigungsstrategie ermitteln kann. Es wird ein Modell mit zwei Achsen eingeführt, nämlich der persönlichen und sozialen Unterbrechbarkeit. Die Annahme, dass der Benutzerkon-

text durch die Variablen Beschleunigung, Befindlichkeit und Audio tendenziell richtig abgeleitet werden kann, wurde experimentell bestätigt.

Die Kosten einer Unterbrechung sind direkt abhängig von der Intensität beziehungsweise dem Grad der Unterbrechung. Der Grad der Unterbrechung ist von der Unterbrechbarkeit des Benutzers abhängig. Dieser ist in hohem Maße auf die aktuelle Situation und Beschäftigung der Person zurückzuführen. Zum Beispiel wäre der Anruf der letzten fehlenden Person bei einem Klassentreffen sehr erwünscht, beim Fahren im Auto ohne Freisprechanlage unter Umständen sehr teuer. Deshalb ist die Unterbrechbarkeit des Benutzers, im folgenden persönliche Unterbrechbarkeit genannt, der Schlüssel für die Kosten einer Benachrichtigung. Sie kann als kontinuierliche Variable betrachtet werden, die von *Nicht stören* bis *Unterbrechung kein Problem* reicht.

Eine Benachrichtigung erreicht nicht zwangsläufig den Benutzer allein. Ein akustischer Alarm kann genauso gut von der Umwelt wahrgenommen werden – und zu einer möglicherweise peinlichen Situation führen, beispielsweise im Vorlesungssaal. Daher unterscheiden wir zwischen der *persönlichen Unterbrechbarkeit* des Benutzers und der seiner Umwelt, der *sozialen Unterbrechbarkeit*. Dies ermöglicht es uns, eine Unterbrechungsstrategie zu verwenden, die entweder nur den Benutzer oder unter Umständen auch seine Umgebung ansprechen kann.

Das daraus resultierende Modell besteht aus einem zwei-dimensionalen Raum, der durch die persönliche und die soziale Unterbrechbarkeit einer Person aufgespannt wird. Die Abbildung 2 zeigt ein paar typische Situationen: Unter anderem erfordert das Lenken eines Autos eine hohe Aufmerksamkeit des Benutzers, hat jedoch eine niedrige Unterbrechbarkeitsstufe im sozialen Kontext, beispielsweise bezüglich lautem Handy-Klingeln.

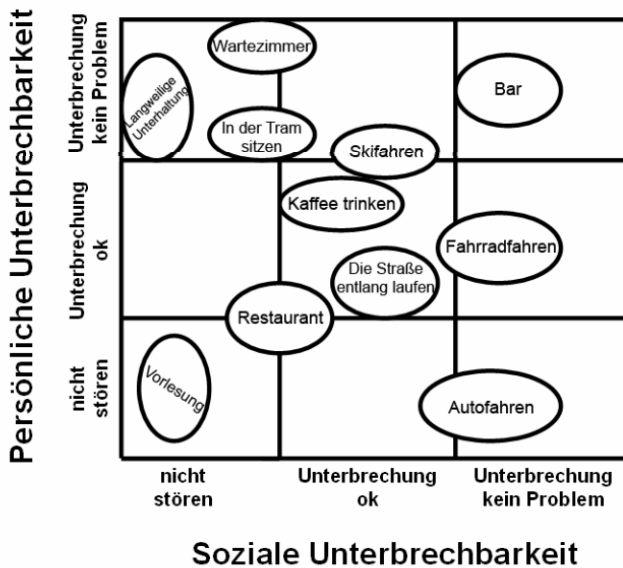


Abbildung 2. Der Darstellungsraum für Unterbrechbarkeit (aus dem Englischen übersetzt aus [3])

Um dieses Modell umzusetzen, wurden im Laborversuch Werte für die Befindlichkeit (location), die Beschleunigung (acceleration) und Töne (audio) von Testpersonen gemessen und daraus folgende vier Faktoren bezüglich der aktuellen Unterbrechbarkeit der Person abgeleitet: Die Aktivität des Benutzers, die soziale Aktivität, z.B. ob die Person sich gerade mit anderen Leuten unterhält, die soziale Situation, z.B. „im Restaurant“, und der Aufenthaltsort. Wie bereits erwähnt, konzentrieren wir uns auch in diesem Modell auf die Kosten der Unterbrechung, nicht auf die Wichtigkeit der Nachricht, obwohl letztere für die Wahl der Benachrichtigungsstrategie als fünfter Punkt natürlich miteinfließt.

Im Zusammenhang mit dem Beschleunigungssensor wurden die Klassifikationen Sitzen, Stehen, Gehen und Treppensteigen gemacht, Töne wurden in die Bereiche Straße, Unterhaltung, Vorlesung, Restaurant und Sonstige eingeteilt. Die Lokationsbestimmung umfasste die Kategorien Büro, Außen, Vorlesungshalle, Labor und Cafeteria.

Eine konkrete Situation wie *Arbeiten im Büro* wird vom System nicht erfasst, vielmehr betrachten wir Aktivitäten auf einem hohem Abstraktionsniveau wie zum Beispiel *Konversation*. Die Sensoren liefern lediglich eine Tendenz dafür, was die wahrscheinlichste Unterbrechbarkeit ist. Die Wahrscheinlichkeiten werden miteinander gewichtet und ergeben in ihrer Kombination einen Schätzwert (sensor estimate).

In einer 37minütigen Aufzeichnung des Testprogramms waren die geschätzten Werte in 88,5% der Zeit korrekt für die soziale Unterbrechbarkeit und 96,3% korrekt für die persönliche Unterbrechbarkeit [3].

2.2 Sensoren zur Bestimmung der mentalen Belastung (mental load)

Mehr in die Tiefe des menschlichen Organismus geht der Ansatz von Daniel Chen und Roel Vertegaal von der Queen's University in Kingston, Ontario, die mit Hilfe der Messung von Herzfrequenz und motorischer Aktivität die mentale Belastung einer Person und auf diese Weise die passende Benachrichtigungsstrategie herausfinden wollen. Es wird eine Benutzerschnittstelle entwickelt, die mit physiologischen Meßtechniken aktiv die Konzentrationsfähigkeit der Benutzer erfaßt und die Aufmerksamkeitsverteilung optimieren soll.

Sowohl das Herz als auch das Gehirn liefern Signale, die eine Modellierung der wahrscheinlichen Benutzeraktivität und –belastung erlauben. Um die Beschäftigung einer Person zu schätzen wurden daher Messungen des Elektrokardiogramms (EKG) und des Elektroenzephalogramms (EEG) untersucht. Das EKG ist die Registrierung der Summe der elektrischen Aktivitäten aller Herzmuskelfasern, das EEG gibt die gesamte elektrische Aktivität des Gehirns durch Aufzeichnung der Spannungsschwankungen an der Kopfoberfläche wieder. Die Messvorrichtungen sind in Abbildung 3 dargestellt.

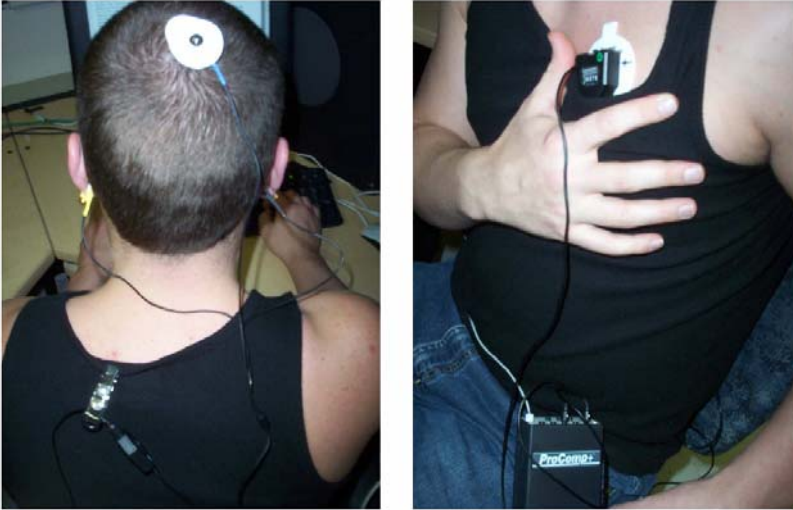


Abbildung 3. EEG-Sensoren (links) liefern Informationen zur motorischen Aktivität des Benutzers, während EKG-Sensoren (rechts) Informationen über die mentale Belastung ermitteln (aus [4])

Das elektrische Potenzial des Herzes produziert ein elektromagnetisches Feld, das 5000 mal größer ist als das des Gehirns. Damit sendet das Herz die höchste elektrische Aktivität in allen Körperorganen aus und liefert verlässliche physiologische Daten über die Belastung einer Person, die mit dem EEG allein viel schwieriger auszumachen wäre. Während das EKG Informationen über die mentale Last gibt, erlaubt das EEG robuste Aussagen über die motorische Aktivität. Die Kombination beider Signale ergibt eine gute Darstellung des geistigen Belastungszustands (attentional state) eines Menschen.

Das EKG-Signal wird vom sympathischen (SNS) und vom parasympathischen (PNS) Nervensystem reguliert. Biochemische Nachrichten vom Gehirn nutzen sowohl das SNS als auch das PNS, um das Herz oder auch andere Organe in bestimmten Situationen zu steuern. Das SNS wirkt anregend und erhöht den Herzschlag in Situationen von Ärger, Stress oder Frustration, während das PNS den Herzschlag verlangsamt und das Herz in normalen Situationen reguliert. Das EKG-Signal wird folgendermaßen analysiert: Die Herzfrequenz ist nach den Schlägen pro Minute definiert. Ermittelt wird sie durch die Zeitintervalle zwischen zwei aufeinander folgenden Maxima in der Herzkurve. Die Veränderung der Herzfrequenz (heart rate variability, HRV) wird mit der Standardabweichung der Herzfrequenz berechnet. Die Spektralanalyse der HRV führt zur Entscheidung über den Zustand des Benutzers. Schätzungen über die mentale Belastung erhält man durch Tiefpassfilterung der Spektralkomponenten der HRV, mit Messung von niedrigen Frequenzstärken (NF) im Bereich unter 0,1 Hz.

Um den Benutzerstatus noch genauer zu bestimmen, wird das EEG dazugezogen. Obwohl es das geringere elektrische Potenzial als das EKG hat, liefert es wertvolle Information über motorische Aktivitäten. Motorische Aktivität wird durch die Spektralanalyse des EEG-Signals ermittelt. Der Gebrauch einer einzelnen Elektrode reicht aus, um Informationen bezüglich der Motorik aus dem EEG-Signal zu erschließen. Dabei misst das System die ereignisbezogene Desynchronisierung des EEG im μ -Stärkenbereich, mit einem Frequenzbereich von 8 bis 30 Hz. Kurz bevor und während des Ausbruchs einer motorischen Aktivität kann ein Stärkeabfall in diesem Signal beobachtet werden. Diese Messung wird angewendet, um zwischen zwei Zuständen motorischer Aktivität beim Benutzer zu unterscheiden: die Ruhephase und das Durchführen einer motorisch bezogenen Aufgabe.

Sogenannte PAUIs (physiologically attentive user interfaces) unterscheiden durch Kombination mehrerer physiologischer Signale zwischen verschiedenen Graden von Aufmerksamkeitszuständen. Die Niedrigfrequenz-Spektralkomponenten des EKG liefern Indikatoren für die geistige Aktivität, aber weisen nicht zwangsläufig auf eine Aktion hin. Erst durch die Kombination von HRV-Informationen mit EEG-Signalen für motorische Aktivität können PAUIs entscheiden, ob der Benutzer aktiv an einer Sache teilnimmt oder eher passiv beteiligt ist. Obwohl wir in diesem Artikel nicht näher darauf eingehen, könnte die Benutzeraktivität durch Hinzunahme von Umgebungssensoren weiter präzisiert werden. Hierbei könnte man sich beispielsweise mit Eye-Tracking-Verfahren oder den in Abschnitt 2.1 beschriebenen Techniken behelfen.

Abbildung 4 zeigt die EKG- und EEG-basierte Klassifizierung, um den Benutzerstatus zu modellieren. Wir unterscheiden vier Zustände, die bei der Voraussage helfen, inwiefern der Benutzer für eventuelle Unterbrechungen verfügbar ist.

	Geringe motorische Aktivität (EEG)	Hohe motorische Aktivität (EEG)
Niedrige NF Stärke (EKG)	<p>STATUS 1</p> <ul style="list-style-type: none"> - Geringe mentale Aktivität - Ruhephase <p><i>Mögliche Aktivitäten:</i> Eine Pause machen, Relaxen</p>	<p>STATUS 2</p> <ul style="list-style-type: none"> - Geringe mentale Aktivität - Fortwährende Bewegung <p><i>Mögliche Aktivitäten:</i> Fortbewegung</p>
Hohe NF Stärke (EKG)	<p>STATUS 3</p> <ul style="list-style-type: none"> - Hohe mentale Belastung - Ruhephase <p><i>Mögliche Aktivitäten:</i> Autofahren, Lesen, Denken</p>	<p>STATUS 4</p> <ul style="list-style-type: none"> - Hohe mentale Belastung - Fortwährende Bewegung <p><i>Mögliche Aktivitäten:</i> Treffen, Schreiben, Vorlesung halten</p>

Abbildung 4. Klassifizierung von Aktivitäten mit Aufmerksamkeitszuständen (aus dem Englischen übersetzt aus [4])

In diesem Schema ist der niedrigste Grad der aktuell aufgewendeten Aufmerksamkeit in *Status 1* dargestellt, bei dem der Benutzer nicht aktiv in eine Aufgabe involviert ist. Im Zusammenhang mit Arbeit kann bei diesem Zustand typischerweise auf die niedrigst möglichen Unterbrechungskosten geschlossen werden. Diese Beobachtung ist jedoch nicht für alle Zusammenhänge verallgemeinerbar, denn ein Zustand des Ausruhens könnte in der Tat die höchsten Kosten an Unterbrechung hervorrufen. *Status 2* ist typisch für Benutzer, die unterwegs sind, zum Beispiel zu einer Verabredung. Dieser Zustand hat in der Regel geringe Konsequenzen für sprachbezogene Unterbrechungen wie Handyanrufe, aber höhere Kosten für Unterbrechungen, die es erforderlich machen, dass das motorische System in die Antwort einbezogen wird, so wie beim EMailen. *Status 3* weist auf geistige Aktivität während einer körperlichen Ruhephase hin, zum Beispiel beim Lesen, Autofahren oder Denken. In dieser Situation könnte man benachrichtigt werden wollen, aber nicht auf auditivem Wege, weil dies möglicherweise den aktuellen Denkprozess stören würde. Als letzter zeigt *Status 4* eine aktive Beteiligung bei einem Prozess an, der verfügbare mentale Ressourcen stark eingeschränkt hat und daher hohe Kosten im Falle einer Unterbrechung fordern würde. In diesem Fall würde man entweder eine Verzögerung der Benachrichtigung oder die Kommunizierung des Zustands *Beschäftigt* wünschen.

In Anlehnung an die vier Benutzerzustände von Abbildung 4 wurde ein Experiment zur Validierung der Praxistauglichkeit dieses Modells durchgeführt. Ein Handy mit den Unterbrechungsmodi *Klingeln*, *Vibrieren* und *Stumm* wurde an ein Sensorensystem angeschlossen, das nach dem oben beschriebenen Prinzip den Aufmerksamkeitszustand der Benutzer messen sollte. Für die einzelnen vier Zustände, die das System automatisch ermittelt, konnten die Benutzer selbst konfigurieren, auf welche Weise sie jeweils benachrichtigt werden wollten. Eine typische Präferenz war die folgende: Für Status 4 wurden Anrufe auf lautlos gestellt, der Instant-Messaging-Status (IMS) auf *Beschäftigt* und EMail auf *Stumm*, beim dritten Zustand alle Benachrichtigungen auf *Vibrieren*, der IMS auf *Verfügbar*. Im Falle von Zustand 2 entschieden sich die meisten für das akustische Klingeln bei Anrufen, für IMS *Beschäftigt* und Vibrieren bei EMail und Instant Messaging. Bei Level 1 schließlich war Klingeln akzeptiert, der IMS auf *Verfügbar* gesetzt und Email und IM – Benachrichtigungen auf *Vibrieren*. Erste Auswertungen des Systemtests waren ermutigend: So konnte in einem Sechs-Personen-Experiment mit einem initialen Prototypen in 83% der Fälle das passende Benachrichtigungslevel ermittelt werden [4].

3 Prinzipien von Benachrichtigungssystemen

Bisher wurde nur die Input-Seite eines Benachrichtigungssystems untersucht, nämlich die Erfassung des Benutzerkontextes, um die Dringlichkeit einer Benachrichtigung mit der aktuellen Situation einer Person in Einklang zu bringen. Wenn im weiteren Verlauf dieses Artikels von der Wichtigkeit einer Information gesprochen wird, ist damit immer die Gewichtung mit dem user context im Sinne von Abschnitt 2 gemeint, sofern ein System über Statuserfassung verfügt. Der Motion Monitor beispielsweise hat derartige Mechanismen nicht nötig, da die ausgesendeten Informationen kontinuierlich, hintergründig und in ihrer Darstellung eher angenehm

sind und es keine akuten Änderungen zu vermitteln gibt, die durch störende Unterbrechungen kommuniziert werden müssten.

Der Hauptteil dieses Artikel befasst sich nun mit Aspekten der tatsächlichen Darstellung von Benachrichtigungen und den Prinzipien, nach denen solche Systeme entworfen werden sollten. Es gibt bisher wenig Forschungsergebnisse darüber, warum ein Benachrichtigungssystem effektiver als ein anderes funktioniert, oder wie man derartige Anwendungen bewerten könnte. Untersuchungen bestehender Systeme identifizierten jedoch drei Schlüsselcharakteristika, die diese Systeme gemein hatten: Abstraktion, Benachrichtigungslevels und Transitionen.

3.1 Datenabstraktion

Datenabstraktion (data abstraction) bezeichnet die Feature-Extraktion aus Rohdaten beziehungsweise die Reduzierung von deren Umfang und Genauigkeit, sodass das Resultat weniger Details als das Original enthält und damit auf einen Blick oder gar aus dem Augenwinkel (im Falle visueller Benachrichtigung) zu erfassen ist. Dies ist wichtig, da Benutzer meist nur einen geringen Teil ihrer Aufmerksamkeit für Benachrichtigungen aufwenden wollen oder können. Graphik 5 zeigt ein Beispiel für die visualisierte, abstrahierte Darstellung eines komplexen Sachverhalts.

Experimente haben gezeigt, dass abstrahierte Daten genügend Informationen enthalten können, um dem Benutzer Informationen zu übermitteln und ihm es gleichzeitig erlauben, sich weiterhin auf seine Hauptaufgabe zu konzentrieren [2].

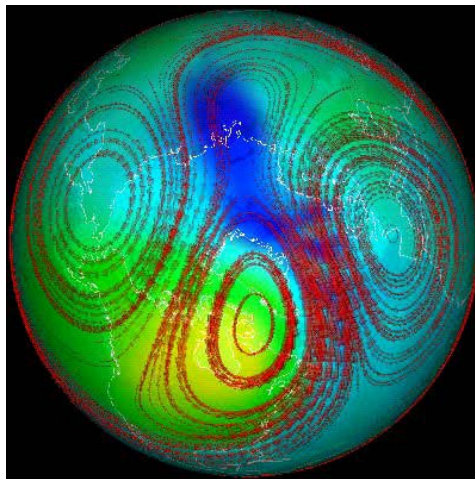


Abbildung 5. Visualisierung und abstrakte Darstellung von Temperaturverläufen und Windrichtung in der oberen Erdatmosphäre über der nördlichen Polarregion, basierend auf den Daten eines Wettersatelliten (aus [5])

3.2 Benachrichtigungslevels

Der Begriff *Benachrichtigungslevel* (notification level) wird benutzt, um die Wichtigkeit einer Nachricht zu kommunizieren. Höhere Levels korrespondieren mit kritischen Daten und werden in einer Weise dargestellt, dass sie schnell die Aufmerksamkeit des Benutzers auf sich ziehen. Niedrigere Benachrichtigungsstufen beziehen sich auf unkritische Daten und werden typischerweise so angezeigt, dass sie zwar keine Ablenkung hervorrufen, aber in der Peripherie für gelegentliche Informationsaufnahme zur Verfügung stehen. Die Entscheidung darüber, wie ein gegebenes Benachrichtigungslevel dargestellt werden soll, ist Bestandteil des nächsten Abschnitts. Jetzt geht es vielmehr um die Frage, welches abstrakte Darstellungsniveau ein Ereignis haben soll und mit welcher Intensität es die Aufmerksamkeit des Benutzers auf sich ziehen soll.

Wir unterteilen die menschliche Aufmerksamkeit in drei Typen: Unaufmerksamkeit (inattention), geteilte Aufmerksamkeit (divided attention) und volle oder zielgerichtete Aufmerksamkeit (focused attention). Im Falle von Unaufmerksamkeit sind Objekte nicht direkt für die bewusste Wahrnehmung verfügbar, könnten jedoch trotzdem das Handeln beeinflussen, zum Beispiel beim Abruf von Erinnerung. Geteilte und volle Aufmerksamkeit repräsentieren die beiden Möglichkeiten, wie Menschen bewusst Umweltreize wahrnehmen: Durch Verteilung der Aufmerksamkeit auf verschiedene Objekte oder durch Fokussierung der gesamten Aufmerksamkeit auf einen Stimulus.

Wie eingangs erwähnt, repräsentieren Benachrichtigungslevels eine relative Wichtigkeit von Informationen. Je wichtiger ein Ereignis, desto mehr Aufmerksamkeit sollte es auf sich ziehen. Demnach würde ein wichtiges Ereignis gerichtete Aufmerksamkeit erfordern, während ein weniger wichtiges für geteilte Aufmerksamkeit ausgelegt sein sollte. Wir definieren fünf Benachrichtigungslevels, basierend auf den gerade diskutierten Aufmerksamkeitsstypen: *Ignorieren* (ignore), *Kaum Änderung* (change blind), *Aufmerksam machen* (make aware), *Unterbrechen* (interrupt) und *Reaktion fordern* (demand action). *Kaum Änderung* repräsentiert Information von marginaler Wichtigkeit und wird derart angezeigt, dass sie nicht bewusst wahrgenommen wird, aber eventuell das Handeln beeinflusst, korrespondierend zu Unaufmerksamkeit. *Aufmerksam machen* repräsentiert Information von mittlerer Wichtigkeit, die etwas Aufmerksamkeit auf sich ziehen sollte, vergleichbar mit geteilter Aufmerksamkeit. Diese zwei Typen werden typischerweise bei ambient displays verwendet, wie zum Beispiel bei der Kugel des Motion Monitors. *Unterbrechen* und *Reaktion fordern* stehen für Informationen, die die volle Aufmerksamkeit einfordern und den Benutzer zeitweise von seiner primären Aufgabe ablenken sollen. Diese Levels werden typischerweise für Benachrichtigungsmonitore oder Alarmer verwendet. *Reaktion fordern* macht zusätzlich eine Aktion des Benutzers nötig, um den Alarm zu deaktivieren. Diese Levels sollten in peripheren Anzeigegeräten sehr sparsam benutzt werden, denn nur kritische Information sollte vom Benutzer fordern, alles liegen zu lassen und sich der angezeigten Information zuzuwenden. *Ignorieren* repräsentiert Information, die unwichtig ist und nicht angezeigt werden sollte. Somit spannen unsere Benachrichtigungslevels einen Bereich von geringster Wichtigkeit, der keine Aufmerksamkeit geschenkt wird, bis zu Informationen, denen volle Aufmerksamkeit zu Teil werden sollte, auf.

Obwohl es für periphere Displays widersprüchlich erscheint, ungeteilte Aufmerksamkeit zu erreichen, eröffnet dies einem Anzeigegerät jedoch die Möglichkeit im Hintergrund zu bleiben, wenn es keine relevanten Informationen anzeigt. Eine Armbanduhr braucht nur volle Aufmerksamkeit, wenn der Alarm ausgelöst wird, um der Person zu vermitteln, dass die Konferenz angeht. Wenn der Benutzer nicht vollends darauf vertrauen könnte, dass er den Alarm auch bei großer Ablenkung wahrnehmen wird, dann müsste er dauernd auf die Uhr sehen.

Die gängigsten Anzeigegeräte benutzen alle Benachrichtigungslevels außer dem *Fordere Reaktion. Aufmerksam machen* ist das bei weitem die gängigste Niveau, besonders bei Umgebungsanzeigen, obwohl viele Bildschirme *Kaum Änderung* verwenden und Benachrichtigungsmonitore zu *Unterbrechen* tendieren. Unser Motion Monitor ist ein typisches Beispiel. Er verwaltet verschiedene Benachrichtigungslevels für verschieden starke Bewegungen, basierend auf dem Aktivitätslevel der Zielperson. Niedrige und hohe Aktivitätsgrade werden auf *Kaum Änderung* und *Aufmerksam machen* Notifikationslevel gemappt. Weil hohe, entfernte Aktivität kein wirklich wichtiges Ereignis ist, benutzt der Motion Monitor nicht das *Unterbrechen* Level [2].

3.3 Transitionen

Es ist schwierig zu erklären, wie ein peripheres Display Benachrichtigungslevels verwalten soll, ohne dabei die Transitionen, die es bereitstellt, zu beschreiben. *Transitionen* sind ein Mechanismus zur Erstellung von Effekten, die eine angemessene Menge an Aufmerksamkeit vom Benutzer auf sich ziehen, basierend auf dem jeweiligen Benachrichtigungslevel des neuen Ereignisses.

Der Motion Monitor zum Beispiel verändert seine Farben mit unterschiedlicher Geschwindigkeit, abhängig vom Benachrichtigungslevel des eingehenden Ereignisses. Diese animierten Farbänderungen sind Transitionen.

Neue Studien geben Hinweise darauf, wie Informationen bezüglich des jeweiligen Benachrichtigungslevels angezeigt werden sollen, obwohl viel Forschungsarbeit bleibt um exakt bewerten zu können, wie fein oder abrupt Veränderungen auf Anzeigegeräten sein müssen, damit sie in korrekter Weise die Aufmerksamkeit des Benutzers erreichen. Ausgabegeräte, die schwerpunktmäßig für die Vermittlung wichtiger Informationen ausgelegt sind, machen typischerweise von abrupten Transitionen Gebrauch. Einige Studien haben gezeigt, dass signifikante Änderungen in einer Benutzerschnittstelle die Aufmerksamkeit von Personen auf sich ziehen. Displays, die auf geteilte Aufmerksamkeit ausgerichtet sind, verwenden eher unscheinbare Transitionen, wie zum Beispiel die Aktualisierung von kleinen Bildschirmausschnitten oder langsame Bewegungen. Für Monitore, die mit dem Wahrnehmungstyp *Unaufmerksamkeit* arbeiten, hat man herausgefunden, dass Animationen wie Überblenden, Rollen oder Laufschrift kaum von Hauptaufgaben ablenken, selbst wenn letztere nur mit geringer Konzentration ausgeführt werden. Dies impliziert, dass wiederholte und sich schrittweise verändernde Animationen in diesem Kontext für die sogenannte change blindness adäquate Lösungen darstellen. Basierend auf diesen Fakten wird deutlich, dass verschiedene Animationenstypen ein Schlüsselmechanismus sind für Transitionen in Anwendungen, die den Benutzer nicht verwirren wollen.

Änderungen in Displays, die explizit auf die Minimierung von Bewegungen achten, werden kaum wahrgenommen. Andere Systeme machen den Benutzer effektiv auf neue Information aufmerksam, indem sie sich stark ändern, Graphikobjekte unterschiedlich schnell bewegen oder andere Formen von Animation gebrauchen. Abrupte oder signifikante Bewegungen werden benutzt, um absichtlich eine Unterbrechung herbeizuführen [2].

3.4 Entwurfskriterien für Ambient Displays

Umgebungsanzeigen oder ambient displays wurden bereits in der Einleitung definiert. Obwohl sie nur einen Teilbereich der Benachrichtigungssysteme im allgemeinen abdecken, gewinnen sie immer mehr an Bedeutung, da sie oft als animierter Ziergegenstand für den kommerziellen Bereich entworfen werden. Deshalb seien an dieser Stelle die Nachforschungen von Mikko Laakso erwähnt, der für Umgebungsanzeigen die folgenden Entwurfsrichtlinien zusammengefasst hat:

Ambient displays können auf verschiedenen Ebenen analysiert werden. Eine ist der *Nutzen* (utility), was bedeutet, dass das Anzeigegerät Informationen darstellen soll, die für eine Person wirklich relevant sind, zum anderen die *Ästhetik* (aesthetics), also ob das Display attraktiv ist, sowie die *Ruhe* (calmness), was der Frage nachgehen soll: Erfolgt die Anzeige ausschließlich in der Peripherie der Benutzerwahrnehmung? Diese Aspekte können theoretisch individuell betrachtet werden, aber es ist eher wahrscheinlich, dass alle in irgendeiner Form voneinander abhängig sind und nur in ihrer Gesamtheit analysiert werden können. Möglicherweise ist dies auch der Grund, warum es in der Literatur kaum Evaluationen von Anzeigegeräten gibt. Mankoff definierte eine heuristische Analyse, die auf ambient displays angepasst ist. Verglichen mit den klassischen Heuristiken von Nielsen identifizierte man mit diesem Ansatz mehr und auch schwerwiegendere Probleme, was zeigt, dass die neue Methode bei der Klärung von Usability-Fragen im Bereich der Umgebungsanzeigen effektiver ist.

Im Folgenden seien nun die drei angesprochenen Entwurfskriterien näher erläutert. Die ästhetische Qualität spielt – wie bereits erwähnt – eine große Rolle, da ambient displays nicht nur als Informationsanzeige, sondern auch als dekoratives Objekt dienen können. Ein ansprechendes Design bietet darüber hinaus auch einige praktische Vorteile: Zum einen „verschwimmt“ es dadurch eher mit dem Hintergrund, was die gewünschte Eigenschaft als peripherer und unscheinbarer Gegenstand verbessert, zum anderen rechtfertigt es somit seine Existenz, auch wenn die angezeigte Information nur von marginaler Wichtigkeit ist. Letztlich gewinnt es damit auch die Akzeptanz des Benutzers auf längere Zeit. Eine mögliche Schwierigkeit ist natürlich, dass Ästhetik viel mit Geschmack zu tun hat und daher objektiv sicherlich nicht zu hundert Prozent bewertet werden kann.

Ruhe bedeutet minimale Unterbrechung des Benutzers. Die generelle Idee dabei ist, dass sich das Anzeigegerät völlig aus dem Vordergrund der Benutzerwahrnehmung ausblenden kann.

Nutzen schließlich heißt, dass tatsächlich etwas von Interesse angezeigt wird. Das klingt vielleicht trivial, ist aber nicht immer der Fall. Denn viele bisher entworfene Prototypen stellen oft bedeutungslose Information dar. Diese Tatsache könnte auf die

beschränkte Ausdruckstärke solcher Apparate zurückzuführen sein. Die meisten ambient displays wurden entworfen, um eine einzige skalare Größe anzuzeigen oder nur einen booleschen Wahrheitswert. Um eine höhere Vielseitigkeit für einen großen Bereich von Anwendungen zu erreichen, sollten Umgebungsanzeigen daher nicht auf eine Informationsquelle beschränkt sein, sondern auf andere Daten umprogrammierbar sein und auf diese Weise ihren relativen Nutzen erhöhen [1].

4 Multimodale Ausgabe

Bisher hat sich die Darstellung und Klassifizierung von Benachrichtigungssystemen implizit hauptsächlich auf den visuellen Kanal beschränkt. Der Motion Monitor ist ein solches Beispiel, das ausschließlich den Augen Informationen liefert. Benachrichtigungen können jedoch verschiedene Kanäle der menschlichen Sinneswahrnehmung ansprechen. Traditionelle Computer sind primär auf visuelle Benachrichtigung und den gelegentlichen Einsatz von Audiosignalen beschränkt, um dem Benutzer Informationen zu vermitteln. Aufgrund dessen werden Hintergrundereignisse gewöhnlich als optische oder auditive Reize überbracht. Dies allein beachtet jedoch nicht die Tatsache, dass der Mensch auch außerhalb dieser Kanäle außergewöhnliche Rezeptoren für Umweltsignale besitzt. Eine dieser zusätzlichen Kapazitäten ist der Geruchssinn, ein mächtiges und vielseitiges Medium, um Hintergrundinformationen jeder Art wahrzunehmen. Die effektive Nutzung dieses Kanals ist jedoch in der praktischen Anwendung deutlich komplizierter und steckt forschungstechnisch noch in den Kinderschuhen. Die Verwendung des Tastsinns findet beispielsweise beim Vibrationsalarm von Handys weite Verbreitung. Andere Ansätze binden die Möglichkeit der Wärmeempfindung des Menschen als interessante Alternative ein. Im Zusammenhang mit Multimedia gehen also die Darstellungsmöglichkeiten weit über den bekannten, visuellen Kanal hinaus: Der Trend, die fünf klassischen Sinne des Menschen, nämlich Sehen, Hören, Fühlen, Riechen und Schmecken, mit den Mitteln der Informationswissenschaft über alle Quellen zu integrieren und als Gesamtheit für die Kommunikation zu nutzen, eröffnet neue Möglichkeiten für Benachrichtigungssysteme, subtil oder auch vehement die Aufmerksamkeit zu verteilen.

Menschliche Sinne unterscheiden sich sowohl in ihrer Präzision als auch in ihrer Geschwindigkeit. Sehen und Fühlen sind für die Wahrnehmung von Objekteigenschaften wie Form, Oberfläche, Richtung, Distanz und Größe präziser und schneller als Hören. Hören erlaubt eine bessere Erfassung von zeitabhängigen Ereignissen wie Länge, Geschwindigkeit und Rythmus. Die allgemeinen und einzigartigen Charakteristiken der menschlichen Sinne ermöglichen den Entwurf von Benutzerschnittstellen, die mehrfache Ausgabemodalitäten benutzen und überdies auch zwischen diesen Möglichkeiten abwägen, je nach dem produzierten Unterbrechungseffekt [6].

Die Qualität der verschiedenen Ausgabe-Modalitäten wird hinsichtlich ihrer *Leistung* (performance), dem *Ablenkungsgrad* (disruptiveness), ihrer *Effizienz* (efficiency) und anderen Faktoren beurteilt. Die Leistung bezieht sich auf die *Zeit*, die nach einer Unterbrechung für die Durchführung einer Aktion gebraucht wurde. Der Ablenkungsgrad bezeichnet die Fehlerrate, die Testpersonen nach einer Unterbrechung produzierten, und die Effizienz die Reaktionszeit, bis ein Unterbrechungssignal wahrgenommen und anerkannt wurde. Weitere Faktoren sind beispielsweise die subjektive

Empfindung, ob Geruch oder Wärmeentwicklung in ihrer Form als Benachrichtigungsmittel angenehmer sind, oder gar Angstzustände von Testpersonen, als sich ihr Eingabegerät zwecks einer Nachrichtenübermittlung erwärmte. Diese Faktoren schwanken stark bei den verschiedenen Kanälen, und letztlich geben für die Auswahl der Output-Modalität auch Aspekte wie die Kosten der Implementierung oder die technische Umsetzbarkeit den Ausschlag. Nicht zu vergessen ist die Frage, welches Medium die Umwelt des Benutzers am meisten bei einer Benachrichtigung miteinbezieht. Dies hat wiederum mit der sozialen Unterbrechbarkeit zu tun. Während taktile Alarmer im Prinzip nur die betreffende Person ansprechen, haben auditive und olfaktorische Reize einen weiten Wahrnehmungsbereich. Neben der Motivation, möglichst komfortable und neuartige Anwendungen zu kreieren, bietet multimodales Output auch die Möglichkeit, hör- oder sehgeschädigten Menschen Informationen auf Kanälen, die für sie erfassbar sind, zu senden.

4.1 Visuelle Benachrichtigung

Die klassische Möglichkeit der Benachrichtigung scheint zunächst für jeden offensichtlich, so kennt doch jeder Computerbildschirm und die Laufschrift an Bushaltestellen. Und doch ist die ganze Bandbreite an Möglichkeiten, Gesetzmäßigkeiten und Barrieren der visuellen Benachrichtigung nicht jedem auf Anhieb bewusst. Neben den vielen verschiedenen Möglichkeiten der Darstellung wie Bewegungen, Blinken, Änderung von Farben, Formen und Objektgrößen sowie der Auswahl von Lichtstärke und Anzeigegeräten wie LEDs, Bildschirmen oder mechanischen Lösungen etc. sind vor allem auch physiologische Aspekte interessant. Das Wissen über die biologischen Funktionsweisen des menschlichen Sehapparats führt zu einem effektiven Einsatz von visuellen Sinnesreizen. Einfache Grundregeln, wie Benachrichtigungslevels graphisch auf Transitionen gebildet werden können, wurden bereits im Abschnitt 3.3 erwähnt. Die Absicht dieses Kapitels ist, einen Überblick über das menschliche Sehsystem zu geben und aufzuzeigen, wie dieses Wissen auf das Design von computerbasierender Benachrichtigung angewendet werden kann. Beispielhaft wird ein Modell erklärt, das sieben unterschiedliche Wahrnehmungsstufen implementiert.

Die Verarbeitung visueller Informationen beginnt auf der Rückseite unseres Gehirns am Hinterhauptslappen. Dort teilt sich das Sehsystem auf zwei verschiedene Wege auf. Beide Wege arbeiten getrennt für visuelle Erfassung und visuelles Verhalten, sie unterstützen sich aber auch gegenseitig. Der primäre visuelle Weg (auch der *Was-Weg* genannt) ist für die Wahrnehmung von Farbe, Helligkeit, Form, Kontrast und Tiefe zuständig. Der zweite visuelle Pfad (auch als *Wo-Weg* bezeichnet) sorgt für die Fähigkeit, Objekte im Raum zu lokalisieren und durch den Raum zu navigieren. Auf dem *Wo-Weg* wird visuelle Information mit auditiven und somato-sensorischen Informationen kombiniert.

Die beiden Wege unterscheiden sich in mehr Punkten als in den Informationen, die sie von der Umwelt erhalten. Erstens ist das *Wo-System* farbenblind, wohingegen das *Was-System* Farbinformationen benutzt. Zweitens ist das *Wo-System* zehn mal sensibler gegenüber Helligkeitskontrasten als das *Was-System*. Drittens bewegt sich die visuelle Information im *Wo-System* doppelt so schnell. Viertens hat das *Was-System* eine höhere Sehschärfe. Fünftens verarbeitet das *Was-System* nur Informationen aus

dem zentralen Sehbereich, während das Wo-System sowohl Informationen vom zentralen als auch vom peripheren Sehen verarbeitet. Die Unterschiede zwischen diesen Systemen sollten bei der Erstellung von Benutzeroberflächen in Erwägung gezogen werden, um die Aufmerksamkeit entsprechend zu lenken.

Korrespondierend zu diesen physiologischen Gegebenheiten gibt es zwei unterscheidbare und funktional unterschiedliche Formen von visueller Aufmerksamkeit. Eine Form ist präattentiv und die andere attentiv. Die präattentive Form hat unbegrenzte Kapazität und ist in der Lage, Informationen parallel zu verarbeiten. Die attentive Form hat eine begrenzte Kapazität und verarbeitet Informationen nacheinander. Im allgemeinen führt die menschliche visuelle Objekterfassung die beiden Formen sequentiell aus. Zuerst erstellt die präattentive Form eine Grobanalyse des gesamten Sichtfeldes, dann kümmert sich die attentive Form um ein oder mehrere Elemente gleichzeitig. Zusammengefasst ist präattentive Verarbeitung räumlich parallel und operiert simultan an verschiedenen Stellen über dem gesamten Sichtfeld. Attentive Verarbeitung operiert auf begrenzten Bereichen im Raum.

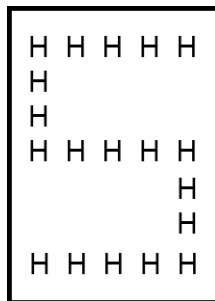


Abbildung 6. Globale und lokale Information: Ein S-Bild mit H-Detail (aus [7])

Abbildung 6 zeigt das Detail *H* als sogenannte lokale Information, während die globale Information der *S*-Umriss ist. Viele Experimente belegen, dass die Reaktionszeit für globale Information geringer ist als die für lokale Information. Das Detailniveau von globaler und lokaler Information („Ist die Aufmerksamkeit auf den Wald oder auf die Bäume gerichtet?“) spielt eine Schlüsselrolle in der selektiven Wahrnehmung. In Austens und Enns' Experiment konnten Testpersonen Änderungen auf lokalem Level besser erkennen als auf globalem Level, wenn sie ihre Aufmerksamkeit auf lokale Details gerichtet hatten. Hatte jemand seine Aufmerksamkeit auf die globale Information gerichtet, dann war er besser bei der Erkennung der globalen Informationsveränderung. Das Sehsystem benötigte dabei Zeit zur Anpassung, wenn sich die Aufmerksamkeitslevels änderten. Austen und Enns erwähnen desweiteren, dass die Reaktionszeit für die Erkennung des globalen Levels am kürzesten ist. Dies könnte den Grund haben, dass Informationserfassung auf globaler Ebene der Standard für das menschliche Sehsystem ist.

Mit dem Wissen über die zwei visuellen Wege und die selektive Wahrnehmung hat Taniya Pirapokin von der Columbia-Universität sieben Levels der visuellen Benachrichtigungsanzeige abgeleitet, die jeweils auf ein anderes Aufmerksamkeitsniveau

ausgerichtet sind. Diese vorgeschlagenen Alarmstufen sind im folgenden von der niedrigsten zur höchsten Stärke aufgezählt und wurden entworfen, um die Aufmerksamkeit des Benutzers ausschließlich auf dem präattentiven Level auf sich zu ziehen. Die Balance zwischen Informationsbereitstellung für das Präattentive als auch für das Attentive muss noch genauer erforscht werden. Voraussetzung für die Klassifizierung ist, dass der Benutzer hauptsächlich mit statischem Inhalt arbeitet.

Die erste Stufe ist *Abgrenzung* und spricht den Was-Weg an. Dieses Benachrichtigungsfeature hat die Eigenschaft, dass es sich vom umgebenden Inhalt in seiner graphischen Darstellung abgrenzt. Ein gutes Beispiel hierfür wäre der Cursor eines Texteditors, der als blinkender vertikaler Strich innerhalb der statischen Buchstaben steht und nicht besonders ablenkt. Bei der zweiten Methode wird das Zielobjekt einer Änderung in seinen Attributen wie Größe oder Farbe unterzogen. Bei dieser *Leichten Änderung* wird kein hohes Alarmlevel erzeugt, aber dennoch ein bemerkbarer Effekt. *Blinkende Bewegung* kann benutzt werden, um die sofortige Aufmerksamkeit des Benutzers zu erregen. Dennoch sind die beiden letztgenannten Effekte nur für kleine Informationsmengen relevant, weil sie aufgrund der Bewegung schwieriger zu lesen sind. *Bewegte Fenster* sind die vierte Stufe in der Hierarchie. Durch die Variation ihrer Größe und Geschwindigkeit könnte man sie noch in mehrere Sublevels unterteilen. Als nächstes kann durch das *Hinzufügen von Audiosignalen* die Aufmerksamkeit des Benutzers gesteuert werden. Auditive Reize sollten in diesem Abschnitt eigentlich nicht erwähnt werden und dennoch sind sie im gewissen Sinne auch ein Teil unserer visuellen Wahrnehmung: Das Wo-System benutzt sie – wie bereits erwähnt – in Kombination mit optischen Stimuli bei der Navigation durch den Raum. Als sechstes Level wird die Verwendung von *PopUp-Fenstern* beschrieben. Bei deren übermäßiger Verwendung besteht jedoch die Gefahr, dass sie vom Benutzer gewohnheitsmäßig ignoriert werden, vergleichbar mit der Situation bei Internet-Browsern. Als höchste Alarmstufe sind *PopUp-Fenster mit kurzen, schrillen Audio-Signalen* zu verstehen, wodurch das Wo-System effektiv auf zwei Wegen angesprochen wird.

Diese sieben Levels sind in ihrer Umsetzung nicht immer unterscheidbar. Attribute wie Größe, Bewegung, Geschwindigkeit oder das gewählte Geräusch können die Alarmwirkung zwischen den Leveln verschwimmen lassen. Wie in anderen Zusammenhängen gilt auch hier die Regel, einen Benachrichtigungstyp nicht im Übermaß zu verwenden, da er sonst nicht mehr als wichtig wahrgenommen wird. Außerdem sollte mit Alarmen sparsam umgegangen werden, da das System dem Benutzer sonst aufdringlich vorkommen könnte [7].

4.2 Auditive Benachrichtigung

Beim Thema Audio soll das Hauptaugenmerk in dieser Abhandlung nicht auf den Einsatz von möglichen Klangvariationen gerichtet werden. Eine adäquate Verwendung von Audiosignalen ist sehr intuitiv und kann in hohem Maße der Kreativität des Entwicklers der Benutzerschnittstelle überlassen werden. Töne können laut oder leise, lang oder kurz, alarmierend oder hintergründig, angenehm oder schrill eingesetzt werden, um ein entsprechendes Benachrichtigungslevel umzusetzen. Je wichtiger das Ereignis ist, desto lauter, schriller und unangenehmer sollte die Soundausgabe sein, da dies den Benutzer am meisten ablenkt. Für Umgebungsanzeigen würde man eher

ruhige, angenehme Klänge wie Wasserplätschern, Vogelzwitichern oder ähnliches auswählen.

Nun soll ein Thema abseits von Computern und Technik kurz beleuchtet werden, nämlich die Frage, welche spezifischen Probleme bei auditiven Nachrichten im sozialen Kontext auftreten können. Denn auch derer sollte man sich in Zukunft bewusst sein, gerade weil die Landschaft der mobilen Endgeräte immer dichter wird: Wir sind wieder bei der sozialen Unterbrechbarkeit angelangt. Rebecca Hansson, Peter Ljungstrand und Johan Redström erwähnen:

Derzeitige auditive Benachrichtigung kann aufmerksamkeitsfordernd, laut und aufdringlich sein und deswegen in vielen sozialen Situationen als unangemessen angesehen werden. Der individuelle Nutzer begegnet diesem Problem häufig mit der Abschaltung der Benachrichtigungsmodalität in manchen sozialen Kontexten, zum Beispiel im Kino oder in der Vorlesung. Es ist jedoch fraglich, ob dies eine zufriedenstellende Lösung ist, da eingehende Informationen wichtig und wertvoll für den Benutzer sein könnten.

Diese Frage ist mit dem Problem des Kommunikationsdefizits verbunden, wie von Ljungberg und Sörensen definiert: Das *Kommunikationsdefizit* charakterisiert Situationen, in denen Leute einer Kommunikationsform unterliegen, an der sie interessiert sind, aber wo die Art der Kommunikation unerwünscht ist. Audio-Benachrichtigung für mobile Geräte ist typischerweise darauf ausgelegt, die maximale Aufmerksamkeit auf sich zu ziehen und in der Lage zu sein, sogar sehr geräuschvolle Umgebungen zu durchdringen. Schon die Benachrichtigung an sich führt zu einer Ablenkung des Rezipienten.

Der Begriff *Aufmerksamkeits-Überladung* (attention overload) beschreibt das Phänomen, wenn Leute von aufdringlichen Ereignissen, die ihre Aufmerksamkeit erfordern, überflutet werden. Diese betrifft nicht nur den Besitzer eines mobilen Endgeräts, sondern auch die Leute in der Umgebung. Diese haben in der Regel kein Interesse an der Benachrichtigung einer anderen Person und werden trotzdem abgelenkt.

Der Begriff *Feinsinn* (subtlety) wird benutzt, um zu beschreiben, wie gut eine Benachrichtigung Information auf eine nicht-aufdringliche und subtile Art überträgt. Vibrationsalarm beispielsweise entspricht dieser Anforderung [8].

Das Design von auditiven Benutzerschnittstellen sollte also vor allem dem Feinsinn genügen.

4.3 Thermale und taktile Benachrichtigung

Dieser Abschnitt umfasst die taktilen und thermalen Benachrichtigungsmethoden. Sie werden deshalb zusammengefasst betrachtet, weil sie beide als Teilbereiche des menschlichen Sinnes *Fühlen* eng miteinander verwandt sind. Für den taktilen Kanal sei der allseits bekannte Vibrationsalarm beim Handy als ein selbsterklärendes Beispiel zu erwähnen. Taktile Reize werden jedoch in diesem Artikel nicht näher begutachtet. Interessanter erscheint ein Experiment von Ernesto Arroyo und Ted Selker vom MIT Media Lab zur Untersuchung der thermalen Modalität:

Das Ziel des Experimentes war es, eine neue Benachrichtigungsmethode für den thermalen Kanal einzuführen und den Effekt der Unterbrechungstechniken Hitze und Licht miteinander zu vergleichen. Neurophysiologische Studien zeigen, dass Finger

und Hände eine der sensibelsten Zonen des Körpers sind und sich exzellent als Ausgabekanäle eignen. Deswegen wurde Temperaturveränderung als eine neue Darstellungsmethode eingeführt, um Unterbrechungen zu generieren. Das Experiment zeigte, dass Hitze aufgrund ihrer Neuheit und des Gefühls von Unmittelbarkeit eine gute Unterbrechungsmodalität ist. Ein Vergleichsdisplay mit Licht bewies, dass Licht ebenso effektiv ist, weil es die anderen Sinne dominiert.

Das Experiment war wie folgt aufgebaut: Als Haupttätigkeit, von der die Testpersonen abgelenkt werden sollten, wurde eine Aufgabe mit hoher kognitiver Belastung gewählt. Die Personen sollten ein text-basiertes Rollenspiel mit graphischen Elementen steuern, wobei sie sich unter anderem eine Liste von Dingen merken, Anweisungen lesen, Gegenden erforschen und Objekte benutzen mussten. Die Unterbrechung erfolgte mittels zweier Umgebungsanzeigen, die die Aufmerksamkeit entweder durch Temperaturveränderung oder durch eine Änderung der Lichtintensität auf sich zogen. Nach einer Unterbrechung mussten diese Ereignisse bestätigt werden und es musste eine Sekundäraufgabe ausgeführt werden, die konzeptionell simpel, aber aufgrund der hohen mentalen Belastung schwierig durchzuführen war. Dabei wurden jeweils die Fehler, die Reaktionszeit und andere Daten aufgezeichnet und daraus die bereits erklärten Werte Ablenkungsgrad, Leistung und Effizienz ermittelt.

Als Testpersonen wurden 23 Männer und Frauen im Alter zwischen 22 und 34 Jahren ausgewählt. Als Anzeigegerät für den visuellen Kanal wurde eine Lampe installiert, die sich in der Peripherie des Sichtfeldes befand und deren Helligkeitslevel von 5 auf 95 Prozent gesteigert werden konnte. Als Apparat zur Wärmeübertragung wurde das *Thermo Mouse Pad* entwickelt, das die Fähigkeit hat, einen großen Bereich der Hand des Benutzers zu erwärmen. Abbildung 7 zeigt den eingesetzten Prototypen. Die Temperatur stieg dabei im Falle einer Benachrichtigung von den 22 Grad Raumtemperatur mit einem Grad pro Sekunde auf bis zu 40 Grad Celsius.



Abbildung 7. Das Thermo-Mauspad (aus [6])

Nach Auswertung der Daten kam man zu folgenden Ergebnissen: Auf die Frage nach der bevorzugten Modalität wählten 40 Prozent der Testpersonen die Temperatur und die restlichen 60 Prozent Licht. Ein überraschender Kommentar über das Thermo-Mauspad war die Angst, sich verbrennen zu können. So wurde Hitze oft als etwas Bedrohliches aufgefasst. Andere fanden Wärmeentwicklung angenehm, speziell für kalte Umgebungen. Interessanterweise war die Temperaturerhöhung, obwohl sie

anfänglich schwieriger zu erfassen war, während ihres Auftretens auch schwerer zu ignorieren, unter Umständen aufgrund der gefühlten Gefahrensituation. Licht führte zu keiner physischen Interaktion mit den Personen und wurde daher nicht als ein Eindringen in den persönlichen Bereich gewertet. 39 Prozent gaben an, dass Licht einfacher zu identifizieren sei als Hitze.

In puncto Leistung war die visuelle Ausgabe deutlich effektiver. Dies lag wohl zum einen an der verzögerten Wahrnehmung der Temperaturveränderung als auch an dem auftretenden Stress aufgrund der gefühlten Bedrohung. Beim Ablenkungsgrad lag in der subjektiven Benutzermeinung kein signifikanter Unterschied zwischen den beiden Unterbrechungstechniken vor, jedoch ließ die Fehlerrate auf einen größeren Ablenkungseffekt bei Hitze als bei Licht schließen. Ebenso war die Effizienz, also die Reaktionszeit, der visuellen Benachrichtigung um 42 Prozent besser als bei der thermalen, womit erstere deutlich effektiver darin ist, die Benutzeraufmerksamkeit auf sich zu ziehen.

Dieses Experiment zeigt also, dass die Ausgabe-Modalität einen gravierenden Einfluss auf die Leistung des Benutzers hat. Desweiteren lassen sich folgende Schlüsse für den Einsatz von thermaler Unterbrechung in Benachrichtigungssystemen ziehen: Zunächst sieht man, dass visuelle Reize im allgemeinen das effektivere Mittel sind, um Aufmerksamkeit zu erregen. Dennoch kann es manchmal sinnvoll und nützlich sein, die thermale Modalität zu wählen. Wenn der visuelle Kanal mit Informationen überflutet und gesättigt ist oder wenn der Benutzer gerade nicht seine Aufmerksamkeit auf den Bildschirm richtet, könnte Wärme das zuverlässigere Medium sein. Zusätzlich berührt die Unterbrechung dann nur eine Person und nicht ihre Umwelt [6].

4.4 Olfaktorische Benachrichtigung

Gehen wir nun noch näher auf die Nutzung des Geruchsinns als ein Benachrichtigungsmedium ein. Verglichen mit den anderen Sinnen wie Sehen, Hören und Fühlen ist der Geruchssinn nicht besonders gut erforscht, was seine Verwendung bei Benutzerschnittstellen angeht. Dafür gibt es einige Gründe: Olfaktorische Kommunikation ist in ihrer Bandbreite beschränkt und unpassend für schnell änderliche Information, die technische Beschreibung des Geruchspektrums sowie die Bereitstellung von olfaktorischen Reizen „on demand“ ist schwierig. Jedoch ist die Fähigkeit des Geruchsinns, sich von der Peripherie ins Zentrum der Benutzerwahrnehmung und wieder zurück zu bringen, höchst geeignet für die Darstellung von Information. Jüngste Forschungsarbeiten nutzen diese Eigenschaft aus und haben neue Formen an olfaktorischen Displays ins Leben gerufen. Dabei haben Erhebungen gezeigt, dass olfaktorische Reize, gepaart mit visuellen und auditiven Stimuli, ein Gefühl der Immersion in das System beim Benutzer ausgelöst haben. Diese Ergebnisse sind interessant und ermutigend, und dennoch scheint das volle Potenzial der olfaktorischen Ausgabemodalität noch nicht ausgeschöpft.

Mit dem AROMA-System der University of British Columbia wurde ein Experiment durchgeführt, das olfaktorische mit visuellen und auditiven Benachrichtigungsmethoden vergleicht, um neue Erkenntnisse über die Effektivität des Riechens im

Zusammenhang mit Benachrichtigungssystemen zu gewinnen. Dabei mussten Testpersonen so viele arithmetische Aufgaben wie möglich lösen und wurden in zufälligen Zeitabständen von den verschiedenen Modalitäten unterbrochen. Dabei wurden wiederum Daten über die Effizienz und den Ablenkungsgrad erhoben. Die Geräte, die für die Generierung von Gerüchen zuständig waren, sind in Abbildung 8 dargestellt.

Die Ergebnisse lassen sich mit folgenden, durch das Experiment verifizierten Hypothesen kurz zusammenfassen: Nach olfaktorischer Benachrichtigung wurden weniger Fragen gelöst als nach visueller und auditorischer Unterbrechung und sie führte nicht zu einer geringeren Fehlerrate bei den zu lösenden Aufgaben. Die Nutzung des Geruchssinns ist zudem weniger effektiv, aber auch weniger ablenkend als die beiden Alternativen.

Die Folgen für die Verwendung in Umgebungsanzeigen lassen sich daher so charakterisieren: Auch olfaktorische Reize sind wie die eben erwähnten thermalen Reize keine Konkurrenz für die klassischen Sinne Sehen und Hören, was die Performanz angeht. Jedoch könnte dies auch an der fehlenden Vertrautheit der Nutzer mit diesem Medium liegen. Aufgrund der mangelnden technischen Möglichkeiten, Gerüche zu generieren und vor allem diese bei Bedarf schnell zu beseitigen, wird sich dieser Kanal der menschlichen Sinneswahrnehmung in der Computertechnik so schnell nicht durchsetzen. Und trotzdem wäre er höchst geeignet für die Verwendung in Systemen, die ihr Ablenkungsniveau möglichst minimieren wollen, da der Unterbrechungseffekt sehr gering ist [9].

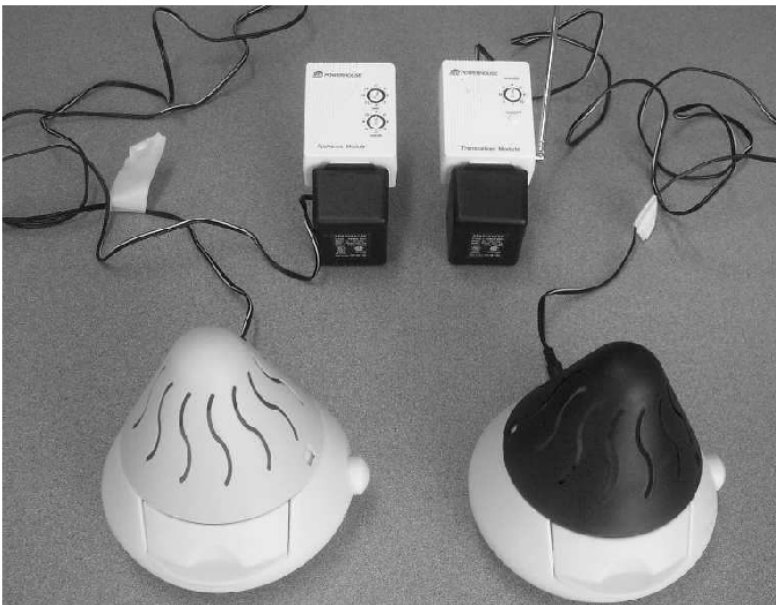


Abbildung 8. Geruchszerstäuber (aus [9])

5 Interessante Beispiele

Nach der Theorie sollen nun einige konkrete Beispiele für Benachrichtigungssysteme folgen. Die betrachteten Systeme stammen alle aus dem Bereich der ambient displays. Die hierbei rein zufällige Fokussierung auf Umgebungsanzeigen soll – wie bereits erklärt – nicht darüber hinwegtäuschen, dass sie nur eine Gruppe von Benachrichtigungssystemen mit speziellen Eigenschaften darstellen. Die zwei Hauptaspekte, die sie von klassischen Benachrichtigungssystemen unterscheiden, sind vor allem, dass sie keine kritischen Informationen kommunizieren und bezüglich ihrer Zweckmäßigkeit eher in der dekorativen Ecke anzusiedeln sind. Wenn man von Benachrichtigungssystemen spricht, denkt man zunächst an die Vermittlung wichtiger Information. Dafür wurden einige Beispiele bereits in der Einleitung genannt: Telefon, Handy, Funkwecker, Skype etc. Diese Anwendungen sind allseits bekannt und hinsichtlich ihrer Informationsübermittlung unspektakulär. Forschungsansätze in dieser Richtung wurden bereits im Verlauf dieses Artikels vorgestellt, wie zum Beispiel Handys mit Umgebungssensoren, das Thermomauspad oder das AROMA-System. Größter Innovationsfreudigkeit bezüglich ihrer Anwendung und ihres Designs erfreuen sich jedoch die ambient displays, deswegen nach dem Motion Monitor hier einige weitere Vertreter:

5.1 Der Dangling String

Der *Dangling String*, wörtlich übersetzt der *hängende Faden*, der im Jahre 1995 von Mark Weiser und John Seely Brown vorgestellt wurde, war eines der ersten und einfachsten ambient displays. Es handelt sich dabei um einen zweieinhalb Meter langen Plastikspaghetto, der von der Decke hängt und an einem kleinen Elektromotor befestigt ist (siehe Abbildung 9). Der Motor ist derart mit einem nahen Ethernet-Kabel verbunden, sodass jedes Bit, das durch die Leitung geschickt wird, eine leichte Drehung des Motors auslöst. Dadurch wedelt der Faden bei hoher Netzlast wie verrückt herum und erzeugt einen charakteristischen Lärm, während ein ruhiges Netzwerk den Faden nur sporadisch in Bewegung setzt [1].



Abbildung 9. Der Dangling String (aus [1])

5.2 Das Bus Mobile

Die Ambient Displays Research Group der Universität Berkeley hat einige interessante Prototypen für Umgebungsanzeigen entwickelt. Eines wurde entworfen, um Informationen über beliebte Buslinien darzustellen. Konkret zeigt das *Bus Mobile*, wie weit ein häufig genutzter Bus von der nächsten Haltestelle entfernt ist. „Nähe“ wird errechnet aus dem Busfahrplan und der aktuellen Tageszeit. Es gibt einige Tokens, die von einem weißen Bildschirm hängen und jeweils einen bestimmten Bus repräsentieren, wie in Abbildung 10 dargestellt. Wenn sich ein Bus einer Haltestelle nähert, bewegt sich der Token, der ihn repräsentiert, nach oben zur Spitze des Displays. Die Tokens bleiben außer Sichtweite (unter dem weißen Display), bis der Bus weniger als 25 Minuten entfernt ist. Danach bewegen sie sich zum tiefstmöglichen Punkt zurück, um ihre Annäherung erneut zu starten [1].



Abbildung 10. Bus Mobile (aus [1])

5.3 Das Daylight Display

Ein anderes Umgebungsdisplay der Ambient Displays Research Group wurde entworfen, um als ungefähre Anzeige der Tageszeit zu dienen und um allgemeine Informationen darüber zu geben, ob es draußen düster, hell oder dunkel ist. Das *Daylight Display* (gezeigt in Abbildung 11) flimmert, wenn der Abend beginnt (dies wird aus Beobachtungen erschlossen, wie lang vor Sonnenuntergang es anfängt zu dämmern), und geht dann langsam von hell nach dunkel über. Bei Tagesanbruch geschieht das Gegenteil. Das Gerät kann in einem Gebäude nützlich sein, in dem es Zimmer ohne Fenster gibt, obwohl die normale Beleuchtung, die in einem fensterlosen Raum benötigt wird, die Ausdrucksstärke und somit den Nutzen des Displays in der Regel deutlich mindert [1].



Abbildung 11. Das Daylight Display (aus [1])

5.4 Der Information Percolator

Der *Information Percolator* ist ein Medium von hoher Ausdrucksstärke und zugleich ein sehr ansprechendes, dekoratives Objekt. Es ist der Funktionsweise eines Bildschirms nachempfunden und benutzt Luftblasen, die in 32 transparenten Röhren voll Wasser aufsteigen. Durch das kontrollierte Auslassen der Luft wird ein Raster von Pixeln, die das Display hinaufwandern, erzeugt. Diese Eigenschaft ermöglicht es, kurzzeitig die Umrisse jedes kleinen Schwarzweiß-Bildes anzuzeigen (siehe Abbildung 12). Die Entwickler kreierten vier Testanwendungen für das System: eine Art personalisierte Uhr, eine Bewegungsmeldesystem, das Poetry Display und eine interaktive „Blasen- Malanwendung“ [1].



Abbildung 12. Der Information Percolator
(aus [1])

5.5 LumiTouch

Das *LumiTouch*-System wurde im MIT Media Lab erfunden und besteht aus einem Paar von interaktiven Bilderrahmen. Wenn ein Benutzer seinen Bilderrahmen berührt, wird diese Berührung in Licht übersetzt und über eine Internet-Verbindung an die andere Person geschickt, sodass deren Bilderrahmen aufleuchtet. Beispiele sind in Abbildung 13 dargestellt. Die Anzeige variiert in Abhängigkeit von den Druckattributen (wo, wie hart und wie lang). Damit wird beabsichtigt, Benutzern die Möglichkeit zu einer abstrakten Form von emotionaler Sprache zu geben; sie können in Echtzeit kommunizieren, indem sie sich gegenseitig Farbmixturen und Lichtmuster schicken [1].



Abbildung 13. LumiTouch (aus [1])

5.6 Die Oxygen Flute

Die Oxygen Flute (gezeigt in Abbildung 14) ist eine Gemeinschaftsinstallation von Chris Chafe und Greg Niemeyer. Die Arbeit erstellt eine aktive Beziehung zwischen dem Publikum und einer Gruppe von Bambuspflanzen, die in der Installation wachsen. Jeder Teilnehmer geht die Stufen hinauf in die Kammer, schließt die Tür und atmet. Mit jedem Atemzug inhalieren wir Menschen Sauerstoff und atmen Kohlenstoffdioxid aus. Zur selben Zeit absorbiert der Bambus Kohlenstoffdioxid und gibt Sauerstoff an die Umgebungsluft ab. Die Oxygen Flute stellt diesen Kreislauf durch Echtzeitmusik dar. Die Klänge sind computergeneriert, aber basieren auf traditionellen Bambusflöten [1].



Abbildung 14. Die Oxygen Flute (aus [1])

6 Zusammenfassung

In diesem Artikel wurde ein Überblick über Benachrichtigungssysteme und deren Unterbrechungstechniken gegeben. Die zwei grundlegenden Forschungsthemen hierbei berühren zum einen das Input, die automatische Erfassung der Benutzersituation, und zum anderen die Art der Informationsvermittlung, die stark an Multimedia-Techniken gebunden ist.

Mit dem Fortschreiten des Preisverfalls und der Miniaturisierung von Kleinstprozessoren sowie der stärkeren Ausweitung von drahtlosen Hochgeschwindigkeitsnetzen ist mit einer weiteren Zunahme von Benachrichtigungssystemen in unserer Informationsgesellschaft zu rechnen. Prototypen von Kleidung mit elektronischen Features sind längst keine Zukunftsvisionen mehr.

Doch gerade beim Paradebeispiel Handy sieht man, dass einige der angesprochenen Aspekte noch nicht hinreichend verwirklicht sind. Deren Umsetzung für Endgeräte des Massenmarktes scheint jedoch nur eine Frage der Zeit.

Literatur

1. Laakso, M., "Ambient Displays and Changing Information". In: Sitzung Interruptions and Intelligent Environments im Seminar User Interfaces and Usability HUT, SoberIT, Frühjahr 2004.
2. Matthews, T., Dey, A. K., Mankoff, J., Carter, S., Rattenbury, T., "A Toolkit for Managing User Attention in Peripheral Displays". In: UIST '04, Santa Fe, New Mexico, USA, Oktober 2004.
3. Kern N., Schiele B., "Context-Aware Multi-Modal Notification for Wearable Computing". In: 7th International Symposium on Wearable Computing, New York, USA, Oktober 2003
4. Chen D., Vertegaal R., "Using Mental Load for Managing Interruptions in Physiologically Attentive User Interfaces". In: CHI 2004, Wien, Österreich, April 2004
5. Lamond, B., "Computing/Visualisation", <http://www.brucelamond.net/visn.php>, Mai 2006
6. Arroyo E., Selker T., "Self-Adaptive Multimodal-Interruption Interfaces". In: Proceedings of the 8th International Conference on Intelligent User Interfaces, Miami, Florida, USA, 2003
7. Pirapokin, T., "Interfaces for Message Notification: Leveling Alert Magnitude". In: Final Project Report, Seminar "Interaction Design: A Perceptual Approach", 2003
8. Hansson R., Ljungstrand P., Redström J., "Subtle and Public Notification Cues for Mobile Devices". In: Proceedings of the 3rd International Conference on Ubiquitous Computing, Atlanta, Georgia, USA, 2001
9. Bodnar A., Corbett R., Nekrasovski D., "AROMA: Ambient awaReness through Olfaction in a Messaging Application". In: Proceedings of the 6th international conference on Multimodal Interfaces, State College, Pennsylvania, USA, Oktober 2004

Privatsphäre in der Multimediakommunikation am Beispiel Instant Messaging

Tobias Emrich

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`emrich@gmx.de`

Zusammenfassung Diese Arbeit beschäftigt sich mit dem komplexen Gebilde Privatsphäre in der Multimediakommunikation. Oftmals wird diese bei der Implementierung neuer Software außen vor gelassen oder als zweitrangig eingestuft. Damit wird allerdings keinesfalls den Wünschen und Erwartungen der Nutzer Rechnung getragen, welche bei Eingriffen in ihre Privatsphäre drastische Konsequenzen (wie z.B. die Ablehnung der Software) ziehen. Durch zwei Modelle im ersten Teil des Artikels, sollen Prinzipien und Konzepte des Themas Privatsphäre besprochen werden. In einem zweiten Teil werden die vier populärsten IM Systeme auf die Ausprägung verschiedener Features hin verglichen. Hier zeigt sich ein Fehlen verschiedener wichtiger Einstellungsmöglichkeiten. Zum Schluss wird versucht zu abstrahieren um Entwicklern von digitalen Systemen einen Einblick darüber zu geben, welche Faktoren bei der Nutzung auf den Nutzer wirken.

1 Einleitung

Die zunehmende Verbreitung von Multimediakommunikation in unserer Gesellschaft bringt viele Vorteile, ebenso wie Risiken mit sich. Die Beziehung zwischen Technologie und der eigenen Privatsphäre ist Vielseitig und komplex und wurde in den letzten Jahren oft diskutiert. In dieser Debatte werden vor allem folgende Standpunkte vertreten:

1. „Mit dem aufkommen neuer Technologie, werden Nutzer deren Privatsphäre-Einschränkungen übernehmen“.
2. „Privatsphäre ist ein komplexes Problem und löst sich nicht von selbst. Um erfolgreiche Anwendungen zu entwerfen, muss das Problem erkannt und beseitigt werden“.

Das erste Statement erinnert an frühere Aussagen der Computer-Industrie, schlecht-bedienbare Benutzeroberflächen seien nicht das Problem – die Nutzer würden sich daran gewöhnen. Das stetige Wachstum der Disziplin Mensch-Maschine-Interaktion zeigt das Gegenteil. Diese Arbeit, wird das Thema Privatsphäre und deren Wichtigkeit in der Multimedia-Kommunikation diskutieren.

1.1 Hintergrund

Große Teile der Literatur zum Thema Privatsphäre beschränken sich auf das Schützen von verschiedensten Arten von Daten ohne darauf zu achten, was Menschen als private Information betrachten. Hingegen ist es wichtig nicht Experten sondern die Nutzer selbst zu fragen, was sensible Informationen sein könnten.

Die meiste Forschungsarbeit zum Thema Privatsphäre konzentriert sich auf Mechanismen um „persönliche Information“ (Daten, die genutzt werden können um ein Individuum zu identifizieren) zu schützen. In dieser Arbeit wird davon ausgegangen, dass ein solcher „datenzentrierter“ Ansatz im Bereich der Multimediakommunikation nicht funktioniert. Ein Grossteil der Daten bei dieser Art der Kommunikation erlaubt die Identifizierung einer Person (z.B. Video, Stimme). Alle diese Daten als „persönliche Information“ anzusehen und somit schützenswert einzustufen, ist nur schwer zu realisieren. Vielmehr kommt es darauf an Individuen innerhalb der Gesellschaft in der sie sich befinden zu betrachten. Privatsphäre wird nur in einem öffentlichen Kontext benötigt. Somit hat z.B. die Unternehmenskultur oder die Wahrnehmung der Situation Einfluss darauf, was Nutzer von sich preisgeben wollen.

Um die vielen verschiedenen Einflussfaktoren zu verstehen wird in dieser Arbeit zunächst auf ein Modell („Multimedia Privacy“ Modell) eingegangen, welches das Problem der Privatsphäre in der Multimediakommunikation von der Nutzerperspektive aus betrachtet. Das Modell hilft zu erkennen welche Information Nutzer als privat einstufen, vor wem und in welchem Kontext. Außerdem wird der Kompromiss, zwischen Einschränkungen der Privatsphäre und dadurch möglichen Mehrnutzen der Multimediakommunikation, hervorgehoben.

2 Grundlegende Modelle

Privatsphäre ist ein komplexes Gebilde, auf das mehrere Faktoren Einfluss haben welche wiederum in wechselseitiger Beziehung stehen. Um diese komplizierten Zusammenhänge zu veranschaulichen sollen im Folgenden zwei Modelle [1] dargestellt und erläutert werden:

- Das „Multimedia Privacy“-Modell beinhaltet die Faktoren, die an Privatsphäreverletzungen beteiligt sind.
- Das „Privacy Invasion Cycle“-Modell, welches aufzeigt, wie diese Faktoren zu einer Verletzung der Privatsphäre führen können und was das für die zukünftige Nutzung des Kommunikationsdienstes für Auswirkungen hat.

2.1 Multimedia Privacy Modell

Mentale Modelle, welche sich Nutzer von komplexen Sachverhalten machen sind oft ungenau, unvollständig oder verzerrt. Ebenso muss die Wahrnehmung von Privatsphäre nicht unbedingt der Realität entsprechen. Das Modell (Abbildung 1) macht 3 Hauptfaktoren (Sensibilität der Information, Informationsempfänger und Informati-

onsnutzen) aus, die miteinander interagieren und somit den Gesamteindruck der Privatsphäre des Nutzers formen. Zusätzlich gibt es noch zwei wichtige aber nicht spezifische (für das Thema Privatsphäre) Faktoren (Kontext, Benutzer). Der Kontext in dem eine Interaktion stattfindet, beeinflusst die Wichtigkeit der anderen Faktoren. Situationsabhängig geschehen Kompromisse, in denen die verschiedenen Variablen gegeneinander abgewogen werden und eventuell Risiken (Verletzung der Privatsphäre betreffend) hingenommen werden. Das alles passiert auf Grundlage von Annahmen des Benutzers. Z.B. wird ein Nutzer auf eine Kommunikation eingehen, wenn der Informationsnutzen für ihn größer ist, als die Hemmschwelle Sensible Informationen preiszugeben. Umgekehrt werden sensible Informationen nicht an nicht vertrauenswürdige Empfänger versandt.

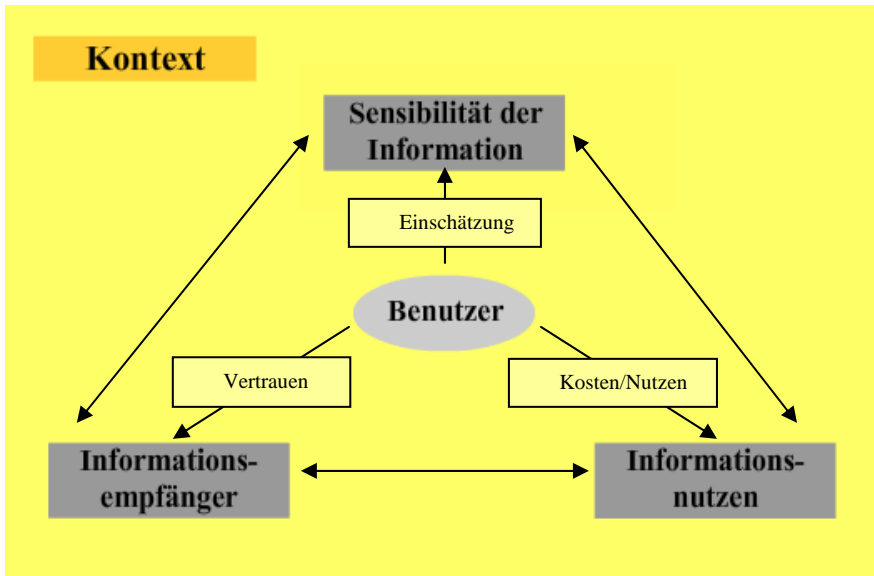


Abbildung 1: Multimedia Privacy Model [1]

2.2 Das „Privacy Invasion Cycle“-Modell (PIC-Modell)

Eine Verletzung der Privatsphäre wird meist dann wahrgenommen, wenn der Nutzer bemerkt, dass die Erwartung und die Realität die Kommunikation betreffend nicht übereinstimmen. Verletzungen der Privatsphäre haben einen Einfluss auf zukünftige Kommunikationen über die gleiche Technologie. Auf diese und weitere Wechselwirkungen geht das „Privacy Invasion Cycle“-Modell (Abbildung 2) ein. Folgende Punkte sind entscheidend für das Verständnis des Modells:

Vertrauen: Nutzer stellen nicht in jeder Situation sofort eine Kosten-Nutzen-Rechnung des Informationsaustausches auf. Abhängig vom Vertrauen, dass der Benutzer in den Empfänger und die verwendete Technologie hat, achtet er mehr oder weniger auf seine Privatsphäre.

Annahmen: Das gefühlte Vertrauen, das der Nutzer während eines bestimmten Informationsaustausches hat, hängt von mehreren Annahmen rund um die Interaktion ab:

- 1) Das Wissen, die Erfahrungen und die Rolle des Nutzers in der Interaktion
- 2) Empfundene Sensibilität der Information
- 3) Empfundener Empfänger der Information
- 4) Empfundener Nutzen der Information
- 5) Empfundener Kontext der Interaktion

Die vermittelnde Technologie kann allerdings all diese Annahmen nichtig machen.

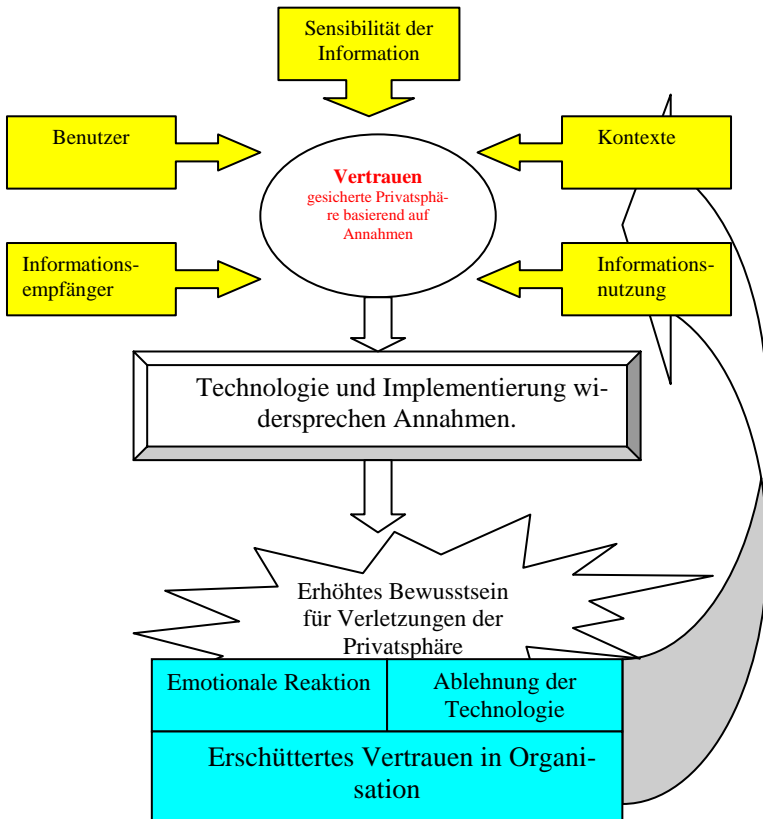


Abbildung 2: Privacy Invasion Cycle Modell [1]

Erkennen und Auswirkungen: Wenn Nutzer bemerken, dass ihre Annahmen falsch waren, empfinden sie eine Verletzung der Privatsphäre. Meist folgen emotionale Reaktionen woraus die Zurückweisung des benutzten Systems, vermindertes Vertrauen in den Empfänger der Information und in die Organisation, welche die Technologie implementiert hat.

Selbstverstärkung: Kommt der Nutzer in eine Situation, die ähnlich empfunden wird (z.B. gleicher Empfänger, Technologie oder Organisation welche die Technolo-

gie implementiert), so geht er mit einem niedrigerem Vertrauen in die Interaktion. Negative Annahmen über die Konversation verhindern eventuell, Erfahrungen zu machen, welche das Vertrauen wieder herstellen.

Das PIC-Modell geht also vor allem auf empfundene Verletzungen der Privatsphäre ein und wie sich diese auf zukünftige Interaktionen auswirken. Für Entwickler neuer Technologien, kann das Modell Aufschlüsse geben, welche Faktoren an gefühlten Verletzungen der Privatsphäre beteiligt sind und wie mögliche Lösungen aussehen könnten.

2.3 Komponenten der Modelle

Im Folgenden soll auf die 5 Komponenten (Sensibilität der Information, Empfänger der Information, Nutzen der Information, Benutzer und Kontext der Interaktion), welche die Basis der beiden Modelle bilden genauer eingegangen werden:

Sensibilität der Information: Die Sensibilität der übermittelten Informationen ist der Hauptfaktor in beiden Modellen. Sie hängt davon ab wie die Daten von Sender wahrgenommen und vom Empfänger interpretiert werden. Nutzer stufen Daten sehr flexibel in ihrer Sensibilität ein und nicht nur nach privat und öffentlich. Somit beeinflusst sowohl die Wahrnehmung der übertragenen Daten aber auch wie öffentlich oder privat die Situation ist, wie hoch die Sensibilität eingestuft wird:

Wahrnehmung der übertragenen Daten: Wie Multimediadaten wahrgenommen werden hängt zu einem großen Teil davon ab, inwiefern Information enthalten ist, die den Nutzer beschreibt. Meist können die übertragenen Daten in zwei Ebenen aufgeteilt werden:

- Primäre Ebene: Die Kernaussage der Übertragung/ Das Diskussionsthema z.B. die medizinischen Fakten, welche in einem computervermittelten Doktor-Patienten Gespräch diskutiert werden.
- Sekundäre Ebene: Soziale oder psychologische Merkmale des Senders z.B. Körpersprache oder Besonderheiten beim Sprechen

Viele Nutzer sind sich nicht bewusst, dass die zu beurteilenden Daten mehr als nur Informationen der primären Ebene enthalten können. Erst wenn Informationen der sekundären Ebene missbraucht worden sind, bemerken sie, dass ihre Privatsphäre verletzt worden ist. Z.B. könnte ein Verkäufer sich in seiner Privatsphäre verletzt fühlen, wenn die Sicherheitskameras im Laden dazu genutzt werden die Leistung zu beurteilen. Bei der Sensibilität von übermittelten Daten ist auch auf die Eigenheiten der Multimediakommunikation zu achten. Zum einen wird eine viel größere Menge an Daten generiert (was vor allem die Menge von sekundärlevel Informationen betrifft). Zum anderen können die Daten mit enormer Geschwindigkeit an ein riesiges Publikum weiterverbreitet werden womit die Gefahr unvorhersagbaren Verwendung steigt.

Situation: Wie andere uns sehen, hängt von der Situation ab, in der wir betrachtet werden. Manche Verhaltensweisen und Ansichten sind gesellschaftlich akzeptiert in

einer privaten Situation nicht dagegen in einer öffentlichen – und auch das umgekehrte kann der Fall sein.

Empfänger der Information: Mit Empfänger der Information ist die Vorstellung des Nutzers der Person welche die Daten empfängt (nicht unbedingt der tatsächliche Empfänger) und/oder manipuliert gemeint. Die Beziehung zwischen Sender und Empfänger kann vielschichtig und komplex sein, sodass an dieser Stelle nicht weiter darauf eingegangen werden kann. Wichtig für das Verständnis der Modelle ist, dass eine Bandbreite von Faktoren die Beurteilung des Empfängers beeinflusst, wobei Vertrauen wohl die wichtigste Rolle spielt.

Nutzung der Information: Beim Faktor Informationsnutzung geht es darum, wie nach der Auffassung des Nutzers seine Informationen im Moment bzw. zu einem späteren Zeitpunkt genutzt werden. Was der Nutzer darüber denkt, wie seine Daten im Moment genutzt werden, hängt vor allem von der Aufgabe ab, welche durch die Multimediakommunikation unterstützt werden soll. Weicht die tatsächliche Nutzung der Information davon ab, so kann eine Verletzung der Privatsphäre wahrgenommen werden. Die spätere Nutzung bezieht sich darauf, was passiert, wenn die Daten der Kommunikation aufgenommen, wieder verwendet und/oder Verändert werden. Nutzer haben in diesen Fällen Bedenken, dass diese Daten in unvorhersehbarer Weise verwendet werden könnten. Aufgenommene und später wiedergegebene Informationen verlieren zudem ihren Kontext. Eine 10 Jahre alte Videoaufzeichnung eines Professors beispielsweise, kann selbigen, ohne einen Zeitstempel, heute als veralteten Wissenschaftler aussehen lassen.

Auch der Verwendungszweck aufgenommener Daten darf sich nicht ändern. Die Daten einer Multimediakommunikation aufzunehmen um die Technologie zu analysieren, mag von den meisten Nutzern hingenommen werden. Werden allerdings die Daten dazu genutzt die Benutzer zu analysieren wird das nicht in Ordnung sein.

Benutzer: In der Multimediakommunikation wird der Ausdruck Benutzer für Sender und Empfänger genutzt. In den beschriebenen Modellen soll als Benutzer derjenige gelten, welcher direkt (primäre Information – Arbeitsergebnisse, Verbrauchsgewohnheiten, ...) oder indirekt (sekundäre Ebene – Persönlichkeit, Auffassungsgabe, Intelligenz, ...) Informationen von sich überträgt. Jeder Nutzer wird bei der Entscheidung ob er Risiken seine Privatsphäre betreffend hinnimmt, eine Abschätzung machen welche Gewinne ihn dafür erwarten.

Es kann sein, dass der Nutzer ein System nicht aktiv nutzt und somit nicht über die Verbreitung seiner Daten bescheid weiß. Ansonsten werden aber vorherige Erfahrungen – positive in Form von Sympathie, negative in Form von Ablehnung) mit einem System immer zukünftige Beeinflussen (siehe PIC).

Kontext der Interaktion: Der Kontext der Interaktion bezieht sich auf die Wahrnehmung der Technologie, soziale und unternehmerische Normen sowie Nationale und Internationale Beschränkungen.

Die Technologie muss den Nutzern die Kontrolle darüber geben wie sie sich selbst in Multimediakommunikationen präsentieren. Schon die Größe oder Qualität eines Übertragenen Gesichtes bei einer Videokonferenz, kann psychologische Effekte beim Empfänger hervorrufen. Nutzer machen oft falsche Annahmen darüber, wie sie dem Informationsempfänger gegenüber präsentiert werden. Deswegen ist ein Feedback darüber enorm wichtig, um die Möglichkeit zu haben gegebenenfalls entsprechend zu reagieren. Außerdem ist es wichtig, dass die Technologie Feedback darüber gibt, wer

alles möglicherweise an der Kommunikation teilnimmt und wer in der Lage ist Aufzeichnungen zu machen. Ungenügendes Feedback kann ebenfalls dazu führen dass die Technologie Abgelehnt wird.

Soziale und unternehmerische Normen sowie Nationale und Internationale Beschränkungen bilden immer den umschließenden Kontext einer Kommunikation und sind ebenfalls ein wichtiger Faktor.

3 Privatsphäre in Instant Messaging

Instant Messaging ist ein Paradebeispiel für eine Technologie, welche die Möglichkeit zur Multimediakommunikation bereitstellt. Deshalb soll im zweiten Teil dieser Arbeit Privatsphäre an diesem konkreten Beispiel und mithilfe mehrerer Studien, diskutiert werden.

3.1 Definition von „Instant Messaging“

Beim Instant Messaging ist es mithilfe einer Client-Software (dem Instant-Messenger) möglich, mit anderen Teilnehmern die Ebenfalls diese Software benutzen (und online sind), zu kommunizieren. Die Kommunikation kann dabei textbasiert (Chat), auditiv (Telefonie) oder audiovisuell (Video-Telefonie) erfolgen. Auch die Übertragung von Dateien ist möglich.

Die meisten Instant Messaging Systeme bieten zusätzlich die Möglichkeit so genannte Buddy-Listen zu führen. Ähnlich wie in einem Adressbuch können hier alle Kontakte abgespeichert werden, wodurch es möglich wird deren Stati (online, offline, nicht verfügbar,...) zu erfahren.

Instant Messaging ist in den vergangenen Jahren immer populärer geworden und hält inzwischen sogar Einzug in die Arbeitswelt. Dort vor allem um Teams die gemeinsam an einem Projekt arbeiten, die Planung und Koordination zu vereinfachen und beschleunigen.

3.2 Privatsphäre ⇔ Erkenntnis

Zusätzlich zur Übertragung des Inhaltes der eigentlichen Kommunikation werden noch andere Daten (sekundäre Informationen) über einen Nutzer preisgegeben. Die Information z.B. welchen Status ein Nutzer im Moment hat, kann ebenfalls nicht ganz unproblematisch für die eigene Privatsphäre sein. Hier kann es zu einem Konflikt kommen zwischen, dem was der Nutzer bereit ist ändern an Erkenntnissen über sich preiszugeben und dem Schutz seiner Privatsphäre.

Man könnte nun also Vermuten das zwischen Erkenntnispreisgabe und Privatsphäre allgemein ein umgekehrte Proportionalität besteht: Preisgabe von Erkenntnisinformation bedeutet weniger Privatsphäre und umgekehrt. Aber auch das Gegenteil kann der Fall sein. Z.B. kann das Bereitstellen einer Website mit allen Veröffentlichungen eines Professors, vor andauernden Störungen und Nachfragen von Studenten nach seinen Arbeiten schützen.

Das Verhältnis zwischen Privatsphäre und Erkenntnis ist also ebenfalls (wie die Privatsphäre für sich) hochgradig Situations- und Kontextabhängig.

3.3 Kausalfaktoren für Sorgen um die Privatsphäre

Zwar ist es vor allem interessant zu betrachten was für konkrete Sorgen Benutzer haben wenn es um ihre Privatsphäre geht, dennoch ist es in einem ersten Schritt entscheidend zu untersuchen, was die Gründe diese Sorgen haben und welche Faktoren diese unterstützen. Deswegen führten S. Patil und A. Kobsa im Jahr 2005 mit 622 Teilnehmern eine Befragung [3] durch. Mit einem detaillierten Onlinefragebogen wurde versucht Haltungen und Praktiken von Erwachsenen zum Thema Privatsphäre zu erfassen. Ebenfalls wurde nach der allgemeinen Nutzung von Instant Messaging gefragt. Im gesamten Fragebogen war allerdings keine Definition von Privatsphäre angegeben um zu verstehen wie Nutzer diese definieren und sie nicht zu beeinflussen. Die Testpersonen sollten alle Fragen anhand einer 7-Punkte Skala beantworten. Bei der Frage nach Angst vor Verletzungen der Privatsphäre im IM wurde die gesamte Skala von 1(niedrig) bis 7 (hoch) genutzt (Abbildung 3). Der Durchschnitt lag bei 3,34.

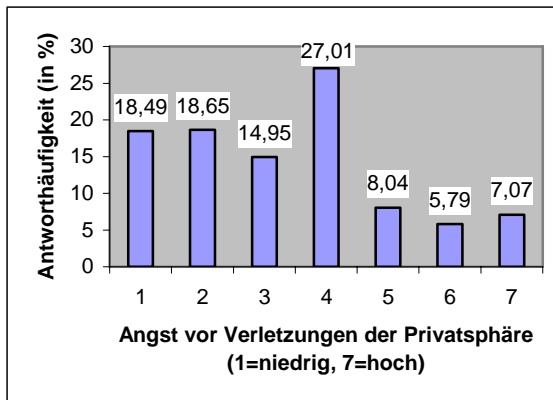


Abbildung 3: Ergebnisse der Umfrage: Angst vor Verletzungen der Privatsphäre in einer IM Konversation auf einer 7 Punkte Skala [3]

Folgende vier Haupteinflussfaktoren für eine Erhöhte Sorge um die eigene Privatsphäre traten bei der besagten Umfrage zu Tage:

- Sensibilität des Inhaltes
33% der Befragten gaben an, die Sensibilität des Inhaltes sei ein Grund für die erhöhte Angst um die eigene Privatsphäre. Ebenso korrelierte die Sensibilität der übertragenen Daten mit den Sorgen um Verletzungen der Privatsphäre (d.h. Unsensible Daten => keine Angst, Sensible Daten => hohe Angst).

- **Allgemeine Haltung zu Privatsphäre**
25% der Befragten, hielten die persönliche Haltung zu Privatsphäre Allgemein ausschlaggebend für Ihre Einstufung von Privatsphäre in IM. So zeigt sich auch in der Umfrage, dass 85% der Befragten, denen Privatsphäre „gleichgültig“ war, ihre Besorgnisse als gering (zwischen 1 und 3) einstufen. Auf der anderen Seite gaben 77% der Befragten welche Privatsphäre als wichtig empfinden an, dass Sie besorgt (zwischen 5 und 7) um Ihre Privatsphäre bei einer IM Konversation sind.
- **Verständnis der Technologie**
Für 22% der Befragten hing ihre Besorgnis um die Privatsphäre mit dem Verständnis für die Technologie zusammen. Bei der Auswertung zeigte sich, dass Missverständnis der Technologie zu einer falschen Einschätzung der Sicherheit und somit zu weniger Achtsamkeit auf die Privatsphäre führen.
- **Mögliche Speicherung von Unterhaltungen**
Schließlich gaben 21% der Befragten an, die Besorgnisse um ihre Privatsphäre gründen sich auf ihre Befürchtungen, die Konversation könnte aufgezeichnet werden.

3.4 Besorgnisse um die Privatsphäre

In einem nächsten Schritt sollen die konkreten Besorgnisse um die Privatsphäre diskutiert werden. Obige Autoren (S. Patil & A. Kobsa) führten ebenfalls zu diesem Thema eine Umfrage [2] durch. In dieser im Jahr 2004 vorgenommenen Studie wurden 7 Teilnehmer befragt, welche häufige Nutzer von Instant Messaging Systemen waren. Die Befragung der Teilnehmer erfolgte durch ein halbstrukturiertes Interview und ist somit als qualitative im Gegensatz zur vorherigen quantitativen Befragung anzusehen.

Obwohl die Befragten alle aus sehr unterschiedlichen Bereichen mit verschiedenen Hintergründen kamen, waren ihre Ansichten und Erwartungen zum das Thema Privatsphäre auffallend ähnlich. Allgemein fanden es die Befragten schwierig in Worte zu fassen, was Privatsphäre bedeutet. Sie fanden es einfacher den Begriff anhand von Situationen und Beispielen zu diskutieren. Als Begründung hierfür wird wiederum die hochgradige personen- und kontextabhängige Natur des Gebildes Privatsphäre gegeben.

Die Befragten gaben an bei der Nutzung von IM-Systemen nicht besonders besorgt um ihre Privatsphäre zu sein. Die meisten arbeiteten unter der Annahme, dass es online sowieso wenig Privatsphäre gibt. Trotzdem zeigen einige Handlungsabläufe und Praktiken der Testpersonen entgegen ihren Äußerungen, dass sie sehr wohl das Verlangen nach mehr Privatsphäre und Sorgen um die selbige haben. Insgesamt fanden Patil und Kobsa drei Haupt-Besorgnis-Faktoren um die Privatsphäre:

- **Privatsphäre vor Nicht-Kontakten**
Alle Befragten hatten das Verlangen nach einer hochgradig abgesicherten Privatsphäre vor Nicht-Kontakten (Personen, welche nicht in Ihrer „Buddy-

List“ enthalten waren). Nicht –Kontakte wurden als Fremde mit unbekanntem Absichten gesehen. Darum stellten sie sicher, dass keiner, der nicht auf ihrer Buddy-List war irgendeine Art von Informationen über sie sehen konnte. Nur einer der Befragten hatte ein öffentliches Profil.

Auf der anderen Seite, wurden die Personen in der Liste als vertrauenswürdige Bekannte gesehen. Da bei diesen Kontakten ein sehr viel niedrigerer Privatsphäreschutz vorgesehen ist, waren alle Befragten sehr vorsichtig, wer in diese Liste aufgenommen wird und wer nicht.

- Privatsphäre im Hinblick auf die Verfügbarkeit

Zu beobachten, war das die Befragten nicht zu jedem Zeitpunkt ihre Statusinformationen allen Kontakten zukommen lassen wollten. Vor allem war eine Trennung zwischen Arbeit und „Zuhause“ zu erkennen. In der Arbeit, wollten die Untersuchten so verfügbar wie möglich für Mitarbeiter sein. Ebenso galt deren Verfügbarkeit sehr große Aufmerksamkeit. Kontakte wurden schnellstmöglich informiert, wenn sich die Statusinformation änderte. Ebenfalls befanden sich auch teilweise Freundes- und Familienkontakte in den entsprechenden „Buddy-Listen“.

Umgekehrt hielt normalerweise die Arbeitswelt keinen Einzug in die Freizeit. Die Befragten stellten sicher, dass die Arbeit so keinen Einfluss auf ihr persönliches Leben hat. So hatten zwei der Untersuchten zwei getrennte Accounts (mit unterschiedlichen Kontaktlisten), auf denen sie sich je nach Situation (Arbeit oder Zuhause) einloggen.

Im Allgemeinen deuten die Praktiken der Probanden darauf hin, dass sie ein Verlangen nach verschiedenen Levels von Verfügbarkeit gegenüber verschiedener Gruppen (wie Freunde, Familie, Arbeit ...) haben, je nachdem an welchem Ort und in welchem Kontext sie sich befinden.

- Privatsphäre im Hinblick auf den Inhalt der Kommunikation

Zu einem großen Teil, wurde IM ähnlich genutzt wie Email oder andere geschriebene Kommunikation. Die Testpersonen waren sich der Sensibilität ihrer Konversationen bewusst. Sie wussten ebenfalls (und akzeptierten auch), dass ihre Konversationen von Systemadministratoren oder dem Netzwerk selbst beobachtet und/oder aufgezeichnet werden konnten. Trotzdem, wie beim versenden einer Email haben die Befragten die Erwartung, dass sie nur vom vorgesehenen Empfänger gelesen wird.

Außerdem wird von diesem Empfänger erwartet, dass dieser dieselben Regeln befolgt, wie sie auch für Email gelten, wenn Nachrichten an Dritte weitergegeben werden. Viele Probanden gaben Unbehagen an, bei der Vorstellung, ihre Konversation könnte vom Gegenüber aufgezeichnet werden. Allerdings war ihnen klar, dass keine Möglichkeit gibt dies entweder wahrzunehmen und noch weniger zu kontrollieren. Andersherum gaben alle Beispiele an, in denen eine aufgezeichnete Konversation hilfreich war. Generell wurde die Strategie verfolgt, Inhalte zu vermeiden, welche eventuell schädlich in der Zukunft sein könnten. Alle Probanden gaben an das Medium (z.B. Telefon oder persönliches Gespräch) zu wechseln, wenn die Konversation als zu sensibel für IM eingeschätzt wird.

Außerdem gaben alle Befragten an, darauf zu achten, wenn andere Ihnen bei der Konversation via IM über die Schulter schauen. Die meisten von ihnen minimierten dann die Fenster.

Zusammenfassen lässt sich festhalten, dass Verletzungen der Privatsphäre im Moment durch eine Kombination aus mehreren Handlungsweisen und Praktiken verhindert werden. Als Kernpunkt für das Verlangen nach Privatsphäre in IM geben die Autoren das Bedürfnis der Nutzer an, zu kontrollieren, wie sie gegenüber anderen erscheinen.

3.5 Designhinweise für IM Systeme

Ausgehend von obigen und weiteren Studien sollen hier Designhinweise gegeben um die Privatsphäre in IM Systeme zu schützen und auch deren Nutzern das Gefühl zu geben ihre Privatsphäre ist gesichert:

1. Sicherheit

“Man kann Sicherheit ohne Privatsphäre haben aber keine Privatsphäre ohne Sicherheit[8]“. An diesem Satz lässt sich erkennen, dass Sicherheit die Basis bildet ohne die keine Privatsphäre möglich ist. Kann die Übertragung beispielsweise ohne Probleme abgehört werden und ist nicht verschlüsselt, so helfen auch keine anderen Mechanismen um die Privatsphäre der Nutzer bei der Konversation zu schützen.

Ferner ist das Wissen um diese Sicherheit oder fehlende Sicherheit ist extrem wichtig. Auch wenn es unrealistisch ist zu erwarten, dass der Standardbenutzer weiß, wie Verschlüsselung funktioniert, so ist es doch ziemlich einfach ihn darüber zu informieren, dass ein unverschlüsselter Datenaustausch von Dritten gelesen werden kann.

2. Mitloggen von Informationen verhindern

Viele Nutzer haben bei einer IM-Konversation im Hinterkopf, dass ihre Gespräche mitgeloggt und gespeichert werden. Dies führt zu einer Art Selbstzensur, bei der man nicht alles über IM preisgibt und evtl. auf andere Medien umschwenkt. Gerade Teenager haben diese Angst um das Abspeichern und Weiterversenden ihrer „innersten Gefühle“. Um diese Ängste zu lindern und somit sensiblere Konversationen über IM zuzulassen, muss eine ausbalancierte Kontrolle über das Archivieren geschaffen werden (z.B. müssen alle Parteien zustimmen, damit eine Unterhaltung aufgezeichnet werden kann). Gerade dieses Problem wird wohl immer zu umgehen sein, aber für Normalnutzer kann man das Mitloggen erheblich erschweren.

3. Gruppierung

Nicht nur globale Privatsphäre Einstellungen sollen möglich sein, eine weitere Differenzierung in Untergruppen der Kontakte kann die Effizienz des Managements der eigenen Privatsphäre erheblich erhöhen:

Normalerweise lassen sich verschiedene Haltungen zur Privatsphäre auf Gruppen von Kontakten anwenden und nicht nur auf einzelne Kontakte. Je-

der Kontakt kann somit einer gewissen Gruppe (Familie, Arbeit, Freunde, Schule,...) zugeordnet werden. Zu den verschiedenen Gruppen wiederum können jeweils bestimmte Privatsphäre Einstellungen vorgenommen werden. Bei diesem Punkt ist besonders darauf zu achten, den Nutzer nicht mit einer Vielzahl von Optionen zu überfordern. Vielmehr muss jeder Benutzer (egal ob Anfänger oder Profi) seine spezifischen Einstellungen einfach vornehmen können (z.B. über intelligente Voreinstellungen, siehe Punkt 5).

4. Zugangskontrolle

Mit Zugangskontrolle ist zum einen die Möglichkeit den Zugang zu den eigenen Daten zu regulieren um „Fremde“ fernzuhalten gemeint und zum anderen die Kontrolle über die eigene Verfügbarkeit, wenn man gerade andere Aufgaben ungestört erledigen will.

In einigen Studien zur IM Nutzung wurde von den Probanden erwähnt, dass sie das Gefühl hatten andere Mitarbeiter würden ihnen auflauern. Dieses auflauern führte sogar dazu, dass Entscheidungen über das ob und wann online gehen davon beeinflusst wurden. Einige IM Systeme unterstützen dies sogar noch durch explizite Nachrichten, wenn ein Kontakt online geht.

Ein weiterer Punkt wurde in einer Studie zur IM Nutzung von Teenagern erwähnt. Freundeskreise von Jugendlichen ändern sich sehr schnell, so kann es sein, dass eine Gruppe von Teenagern nicht von einem ausgeschlossenen Freund online gesehen werden will, um sich befreit über diesen unterhalten zu können.

Somit sind Einstellungen, von wem man zu welcher Zeit gesehen werden will essentiell für den uneingeschränkten Schutz der Privatsphäre und die unbeschwerte Nutzung von IM Systemen.

5. Bessere Privatsphäre Voreinstellungen in den Systemen

Da Privatsphäre ein komplexes und kontextabhängiges Gebilde ist, ist die Anzahl der Optionen und Einstellungen um sie zu verwalten enorm groß. Darum ist es notwendig entweder Voreinstellungen bereitzustellen, die über weite Teile der Nutzer und ihrer Situationen anwendbar sind oder typische Profile (wie Arbeit, zu Hause, Student, Manager,...) mit verschiedenen Voreinstellungen anzubieten.

Die Voreinstellungen eines IM Systems werden selten betrachtet und erst recht nicht verändert. Das passiert nur, wenn das System als erstes installiert wird und dann erst wieder, wenn der Nutzer eine Verletzung seiner Privatsphäre erfahren hat oder sich zumindest Sorgen darum macht.

Das unterstreicht die Wichtigkeit der Einrichtung von passenden Voreinstellungen schon bei der Installation. Diese Voreinstellungen können natürlich auch dynamisch durch eine Betrachtung des Benutzers und seiner Nutzung des Systems erstellt werden. Auf keinen Fall aber dürfen die Voreinstellungen zu locker sein.

6. Transparenz des Systems

Abschließend ist einer der wichtigsten Punkte, die Transparenz des Instant Messaging Systems. Der Nutzer muss immer wissen was gerade passiert. Dazu zählt Feedback über aktuelle Ereignisse vor allem aber über die momentane Einstellung ihrer Parameter. Vor allem sollte das System dem Nutzer detaillierte Informationen darüber geben, wie dieser gegenüber anderen Nutzern präsentiert wird. Auch Statistiken und Historie sollten einsehbar sein. Das Feedback sollte sich andererseits auf ein gesundes Maß beschränken, da der Nutzer sonst in seiner eigentlichen Anwendung gestört wird.

3.6 Umsetzung in vorhandenen Systemen

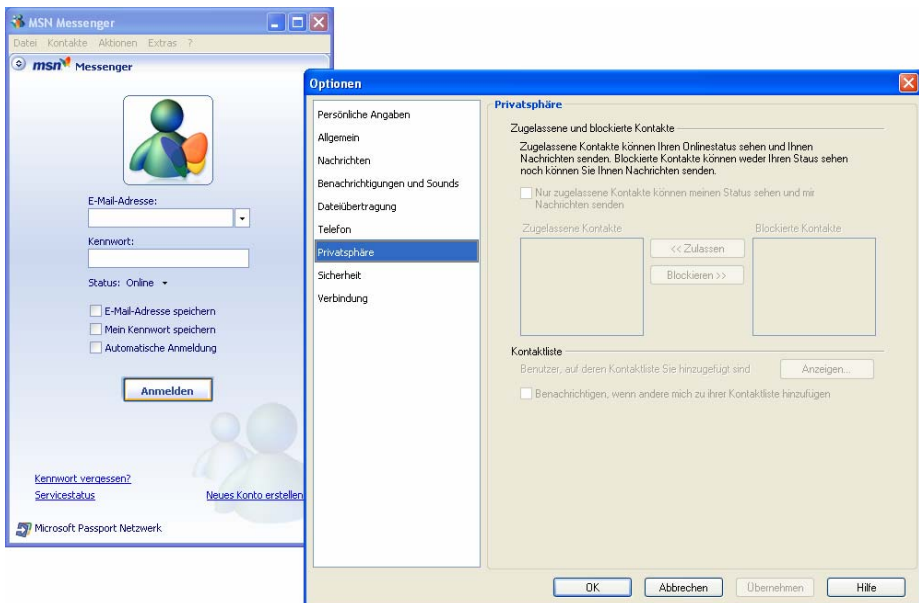


Abbildung 4: Privatsphäre Menü des MSN Messengers

Im Folgenden soll betrachtet werden, mit welchen Features Privatsphäre, das Konzept der Erkenntnispreisgabe und die dazu nötigen Konzepte in momentanen kommerziellen IM Systemen umgesetzt werden. In gängigen IM Systemen wird hierfür ein Menü teilweise mit einer Vielzahl von Optionen bereitgestellt (Abbildung 4 zeigt dieses vom MSN Messenger).

Hierzu sollen die vier bekanntesten Messenger - AOL Instant Messenger (<http://www.aol.com>), MSN Messenger (<http://www.messenger.msn.com>), Yahoo! Messenger (<http://messenger.yahoo.com>) und ICQ (<http://www.icq.com>) - auf folgende Merkmale hin verglichen werden[4]:

- Sound-Benachrichtigung: Bezieht sich auf die Möglichkeit verschiedenen Ereignissen (z.B. eingehende Nachrichten, Einloggen eines Freundes) unterschiedliche Sound-Signale zuzuordnen.
- Gruppierung: Gruppierung erlaubt es Kontakte in Gruppen, wie „Familie“, „Freunde“, „Mitarbeiter“, usw. zu organisieren.
- Privatsphäre Menüs: Privatsphäre Menüs erlauben dem Benutzer, verschiedene Einstellungen zu ändern und anzupassen um seine Privatsphäre zu verwalten. Beispiele wären ob eine Person wünscht ihre Statusinformationen preiszugeben oder ob ihre Telefonnummer von Freunden eingesehen werden kann.
- Blockierung: Die Blockierung eines Kontaktes verhindert, dass Erkenntnisinformationen über den Nutzer vom geblockten Kontakt eingesehen werden können. Der geblockte Kontakt wird, in diesem Fall den betreffenden Nutzer immer als „offline“ sehen.
- Angepasster Status: Die Möglichkeit seinen Status (bzw. den Status, den andere Nutzer zu sehen bekommen) an die momentane Situation anzupassen. Angepasste Stati (wie „Arbeite im Moment an Projekt X“) erhöhen die Flexibilität der vom System gegebenen Voreinstellungen und können somit anderen besser über die eigene Situation informieren, was evtl. zu weniger Störungen führen kann.
- Automatische Rückantwort: Erlaubt einem Nutzer auf eine eingehende Nachricht, eine automatische Antwort zurückzusenden wenn er gerade nicht am System ist. Die Personalisierung solcher Nachrichten hilft ebenfalls andere besser über den eigenen Status zu informieren.
- Popup Mitteilungen: Popup Mitteilungen geben die Möglichkeit den Nutzer (in einem kleinen Fenster in der Ecke) über aktuelle Ereignisse (wie „Einloggen eines Freundes“ oder „Anfrage auf einen Chat“) zu informieren.
- Individuelle Einstellungen: Dies Bezieht sich auf die Möglichkeit individuell für verschiedene Kontakte Einstellungen anzupassen. Zum Beispiel kann es sein, dass eine Person immer als Verfügbar für einen bestimmten Kontakt erscheinen will, egal welcher Status gegenüber anderen Personen angezeigt wird. Andersherum könnte er verhindern wollen, dass ein bestimmter Kontakt Zugriff auf seine Handynummer hat.

- Gruppen Einstellungen: Diese bieten die Möglichkeit bestimmte Einstellungen für ganze Gruppen festzulegen. Wird eine Einstellung der Gruppe geändert, so ändern sich die Individuellen Einstellungen aller Kontakte in der Gruppe. Wenn ein Nutzer immer als „Momentan nicht da“ für die Freundesgruppe erscheinen will, erscheint er für alle Kontakte in dieser Gruppe als selbiges.
- Reziprozität: Reziprozität bedeutet ob das System Wechselseitigkeit von bestimmten Einstellungen verlangt. Zum Beispiel wird, wenn Person A Person B blockt, ein reziprokes System veranlassen, dass Person B automatisch auch Person A blockt.
- Integration des Status auf Webseiten: Damit ist es möglich seinen Status auch auf Webseiten zu publizieren, sodass es für andere Möglich ist diesen Abzufragen, ohne sich in das IM System einloggen zu müssen.
- Erlaubnis zum Hinzufügen: Gibt an, ob ein Nutzer die explizite Erlaubnis benötigt, um andere zu seinen Kontakten hinzuzufügen und umgekehrt. Nutzer, welche eine Überflutung an anfragen zum Hinzufügen bekommen, können eine Option festlegen um dies jedem zu gestatten ohne explizite Erlaubnis.
- Sicherheit: Die Sicherheitsaspekte, sollen an dieser Stelle nicht verglichen werden, da sie nicht den Kernpunkt der Arbeit bilden. Bemerkt sei aber, dass ein grossteil der Kommunikation, vor allem mit dem IM Server von allen IM Systemen noch unverschlüsselt passiert [7].

Feature	AIM	MSN	Yahoo	ICQ
Sound-Benachrichtigung	Ja	Ja	Ja	Ja
Gruppierung	Ja	Ja	Ja	Ja
Privatsphäre Menüs	Ja	Ja	Ja	Ja
Blockierung	Ja	Ja	Ja	Ja
Angepasster Status	Ja	Nein	Ja	Nein
Automatische Rückantwort	Ja	Nein	Nein	Ja
Popup Mitteilungen	Nein	Ja	Ja	Ja
Individuelle Einstellungen	Nein	Nein	Teilweise	Ja
Gruppen Einstellungen	Nein	Nein	Nein	Nein
Reziprozität	Ja	Teilweise	Nein	Nein
Integration des Status auf Webseiten	Nein	Nein	Ja	Ja
Erlaubnis zum Hinzufügen	Nein	Ja	Ja	Ja

Tabelle 1: Vergleich der vier populärsten Instant Messenger [4]

Wie Tabelle 1 aufzeigt, gibt es viele Features, welche alle Instant Messenger implementieren. Einige jedoch werden gar nicht oder auf eine Andere Art und Weise in verschiedenen Systemen umgesetzt. Was vor allem auffällt, ist das nicht einmal die rudimentärsten Aspekte von Kontext in den Systemen umgesetzt sind. Dies wird im totalen Fehlen von Gruppen- und Individualeinstellungen sichtbar. Es ist daher nötig von den individuellen Features (bzw. deren Privatsphäre Management System) der Instant Messenger zu abstrahieren und allgemeine Prinzipien und Faktoren, welche auf diese Einfluss haben auszumachen.

3.7 Einflussfaktoren und -prinzipien auf das Privatsphäre Management

Patil und Kobsa machen in dieser Arbeit folgende Faktoren und Prinzipien aus, welche einen Einfluss auf das Privatsphäre Management haben:

- Reziprozität: Stellt sicher, dass ein Nutzer nur die Informationen erfragen kann, die er auch selbst bereit ist preiszugeben und umgekehrt.
- Feedback: Informiert Nutzer in einer angemessenen Weise welche Information von Ihnen von wem, wann, in welcher Form eingesehen wird.
- Kontext: Welche Information wem und in welcher Form verfügbar gemacht wird ist hochgradig Kontextabhängig und wechselt ständig.
- Kontrolle: Idealerweise, müssen alle Informationen vom entsprechenden Nutzer kontrolliert werden, damit dieser entscheiden kann, wer und wann die eigenen Informationen einsehen kann.
- Normen: Privatsphäre Normen können verschiedene Ursprünge haben (soziale, kulturelle, organisatorische,...) welche sich evtl. beeinflussen.
- Rückschluss: Nutzer managen ihre Privatsphäre auf dem Mikrolevel, bezogen auf eine konkrete Aufgabe. Hier kann es passieren, dass mehrere Mikrolevelinformationen gesammelt werden können um einen Einblick auf Makroebene zu ermöglichen, welcher vielleicht so nicht erwünscht ist.
Z.B. kann es sein, dass Nutzer in ihrem öffentlichen Kalender ihres Unternehmens angeben sie seien in einem bestimmten Meeting, aber nicht öffentlich angeben wollen, wer daran teilnimmt. Sieht jemand aber nun die Kalender aller Mitarbeiter durch so kann diese Information herausgefunden werden.
- Zeitaufwand: Die eigene Privatsphäre zu managen verschlingt Zeit, welche an der primären Aufgabe verloren geht. Dieses managen kann somit ablenkend und ermüdend sein.
- Anreize: Individual- oder Gruppenmotivation beeinflusst wie Nutzer ihre Privatsphäre Einstellungen vornehmen. Es kann sein, dass Nutzer bereit sind

etwas Privatsphäre zu opfern um draus auf der anderen Seite eventuell Vorteile zu ziehen.

- Konflikte: Möglicherweise kollidieren die Wünsche, Meinungen oder Erwartungen die Privatsphäre betreffend verschiedener Nutzer miteinander.
- Archivierung: Die Archivierung von Informationen in irgendeiner Form, konserviert diese über die Zeit. Sie kann dann allerdings zu einem späteren Zeitpunkt frei von dem Kontext in welchem sie ursprünglich stand, in einer anderen Art und Weise und vor nicht gedachten Empfängern wiedergegeben werden.

All die obigen Faktoren in einem System (speziell in einem IM System) unter einen Hut zu bekommen ist eine enorme Herausforderung.

Zum einen ist Privatsphäre ein Konzept aus der realen Welt, wobei versucht wird dieses in die digitale Welt zu übersetzen. Oft wird versucht Konzepte aus dem Alltagsleben nachzubilden. Bei dieser Transformation können aber die Grundannahmen von Privatsphäre total verloren gehen, woraus entweder zu starker oder zu schwacher Schutz der Privatsphäre resultiert.

Zum anderen hat die digitale Welt sehr spezifische Gegebenheiten. Es ist sehr viel einfacher, Informationen zu sammeln und auszuwerten, sie langfristig zu archivieren und schnell, überallhin weiter zu versenden. Dies alles stellt enorme Ansprüche an die Umsetzung der Privatsphäre in heutigen digitalen Systemen, wie z.B. Instant Messengern.

4 Zusammenfassung

Zum Grundverständnis des komplexen Gebildes Privatsphäre wurden am Anfang dieser Arbeit zwei Modelle erläutert und diskutiert. Sie zeigten zum einen welche Faktoren zusammenspielen damit ein Nutzer seine Privatsphäre als gesichert empfindet und zum anderen dass nur ein nutzerzentrierter Ansatz zum Schutz der Privatsphäre zum Erfolg führen kann. Es ist also essentiell vor dem Design einer Anwendung zu betrachten welche Vorstellungen und Erwartungen der Nutzer in Bezug auf die Privatsphäre hat. Das Privacy Invasion Cycle Modell zeigte was passiert, wenn die Vorstellungen der Nutzer in Bezug auf ihre Privatsphäre enttäuscht werden (weniger Vertrauen, evtl. Ablehnung der Technologie). Es untermauert, wie wichtig es ist, dass sich der Nutzer leicht ein mentales Modell des Systems schaffen kann, welches auch zu einem hohen Grad richtig ist.

Im zweiten Teil der Arbeit wurde dann spezieller auf die Privatsphäre in Instant Messaging eingegangen. Anhand mehrerer Studien wurde aufgezeigt:

- Was für Faktoren für die Besorgnis um die Privatsphäre von IM Nutzern eine Rolle spielen (Sensibilität des Inhaltes, Allgemeine Haltung zu Privatsphäre, Verständnis der Technologie, Mögliche Speicherung von Unterhaltungen)
- Welche diese konkreten Besorgnisse sind (Privatsphäre vor Nicht-Kontakten, Privatsphäre im Hinblick auf die Verfügbarkeit, Privatsphäre im Hinblick auf den Inhalt der Kommunikation)
- Welche Designhinweise beachtet werden sollten um diese Sorgen abzuschwächen

Um in einem nächsten Schritt zu sehen, wie tatsächlich all diese Konzepte in momentanen Instant Messaging Systemen umgesetzt werden, wurde ein Vergleich der bekanntesten vier Instant Messenger gemacht. Eine teilweise Umsetzung des Konzeptes der Privatsphäre war zwar in den meisten enthalten, jedoch fehlten oft noch essentielle Features (Gruppeneinstellungen, Individualeinstellungen).

Im letzten Abschnitt wurde dann versucht zu abstrahieren und somit Entwicklern einen Anhaltspunkt zu geben, auf welche Faktoren und Prinzipien beim Entwurf eines Privatsphäre Kritischen Systems, zu achten ist.

Diese Arbeit zeigt, dass Privatsphäre keineswegs ein zu vernachlässigendes Problem der digitalen Kommunikation ist. Für Nutzer kann dieser Punkt der entscheidende Faktor für die Nutzung oder Nichtnutzung einer Software sein. Denn auch wenn Nutzer sich zu Beginn keine großen Sorgen darum machen, so sind heftige Reaktionen zu erwarten, wenn sie einen Eingriff in ihre Privatsphäre erfahren. Diese zeigten sich sogar schon unter Versuchsbedingungen. So scheiterte die Einführung eines IM Systems zur Verbesserung der Zusammenarbeit an einem Arbeitsplatz unter anderem wegen der als zu gering empfundenen Privatsphäre der Nutzer [6].

Zwar wird hier teilweise sehr speziell auf die Umsetzung für IM Systeme eingegangen, da aber IM ein Paradebeispiel für die Multimediakommunikation ist, lassen sich die Prinzipien und Konzepte mit geringfügiger Mühe ebenso auf andere digitale Systeme übertragen.

Literatur

1. Anne Adams, Martina Angela Sasse. Privacy in multimedia communications: Protecting users, not just data. In A. Blandford and J. Vanderdonk, editors, People and Computers XV - Interaction without frontiers. Joint Proceedings of HCI2001, 2001
2. Sameer Patil, Alfred Kobsa. Instant Messaging and Privacy. In Proceedings of Human Computer Interaction 2004. Leeds, UK, 2004.
3. Sameer Patil, Alfred Kobsa. Uncovering privacy attitudes and practices in instant messaging. In Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work. Sanibel Island, Florida, USA, 2005.
4. Sameer Patil, Alfred Kobsa. The Challenges in Preserving Privacy in Awareness Systems. ISR Technical Report #UCI-ISR-03-3, Institute of Software Research, University of California, Irvine, 2003. http://www.isr.uci.edu/tech_reports/UCI-ISR-03-3.pdf
5. Rebecca E. Grinter, Leysia Palen. Instant messaging in teen life. In Proceedings of the 2002 ACM conference on Computer supported cooperative work. New Orleans, Louisiana, USA, 2002.
6. James D. Herbsleb , David L. Atkins , David G. Boyer , Mark Handel , Thomas A. Finholt. Introducing instant messaging and chat in the workplace. In Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves. Minneapolis, Minnesota, USA, 2002.
7. Mohammad Mannan. Secure Public Instant Messaging, Masters Thesis(ComputerScience). Carleton University, 2005. <http://dev.hil.unb.ca/Texts/PST/pdf/mannan.pdf>
8. Bruce J. Bakis. Privacy Fundamentals: What an Information Security Officer Needs to Know. 18. Okt. 2005. http://www.intellitactics.com/pdfs/privacy_final.pdf

Entwicklungsansätze für CSCW-Anwendungen und verteilte Benutzerschnittstellen

Andreas Attenberger

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
attenber [AT] cip.ifi.lmu.de

Zusammenfassung Diese Arbeit soll einen Überblick über die Forschungsergebnisse im Bereich der Entwicklung von Anwendungen für computerunterstützte Zusammenarbeit (Computer Supported Cooperative Work) geben. Nach einer kurzen Übersicht der Anforderungen an CSCW-Applikationen wird der Spezialfall der oft bei CSCW-Systemen eingesetzten verteilten Benutzerschnittstellen betrachtet. Hier wird auf Aspekte der Migration und der Entwicklung dafür passender Schnittstellen eingegangen. Anschließend werden die in diesem Kontext ebenso wichtigen Voraussetzungen für möglichst vielseitig anwendbare Toolkits betrachtet. Dabei wird der Augenmerk insbesondere auf die Forschungsergebnisse von Paul Dourish im Bereich der Open Implementation gelegt. Zum Abschluß erfolgt eine kurze Zusammenfassung sowie ein Ausblick in die Zukunft des Forschungsbereiches.

1 Einleitung

Die Abkürzung CSCW steht für den Begriff Computer Supported Cooperative Work und beschreibt das Forschungsfeld, das sich mit der Entwicklung von Anwendungen für die computergestützte Zusammenarbeit mehrerer Personen beschäftigt [1]. Dabei werden sowohl der technische Aspekt (Welche Funktionen muss eine Anwendung für das effektive Zusammenarbeiten beispielsweise bei einer Konferenz oder Projektplanung bieten?) als auch die sozialwissenschaftliche Seite (Welchen Einfluß hat Gruppendynamik? Ändern sich die Arbeitsstile im Verlauf der Bearbeitungszeit?) betrachtet. Die Forschungsdisziplin ist im Vergleich zu anderen Kerngebieten der Informatik noch relativ jung. Erst 1989 wurde mit der European Conference on Computer Supported Cooperative Work (ECSCW) eine europäische Konferenz zur Beschäftigung mit diesem Thema geschaffen [2]. Auch mit dem Anstieg der Leistungsfähigkeit von Computersystemen gewinnt dieses Gebiet immer weiter an Bedeutung und wird wohl auch in Zukunft mit der steigenden Verbreitung von Computern im Alltag interessante Entwicklungen unterlaufen. Das Ziel dieses Artikels ist es, den

momentanen Stand der Entwicklungsunterstützung für CSCW-Systeme zu erfassen, als auch einen Blick in die Zukunft zu geben, der die Bereiche beleuchtet, denen dabei besondere Aufmerksamkeit zukommen sollte. Zuerst werden die allgemeinen Anforderungen an CSCW-Systeme selbst betrachtet, um eine Basis für die Betrachtung von Entwicklungsansätzen zu schaffen, die eng mit diesen Bedingungen zusammenhängen. Danach folgt ein Abschnitt über die für CSCW-Systeme oft verwendeten verteilten Benutzerschnittstellen und deren für die Entwicklung wichtigen Merkmale. Diese gewinnen mit der Verbreitung des Ubiquitous Computing Prinzips immer mehr an Bedeutung. So wären bei CSCW-Systemen beispielsweise multimedial ausgestattete Seminarräume denkbar. Dadurch können sie für eine effektive Zusammenarbeit mehrerer Benutzer mit Hilfe verschiedenster Geräte sorgen. Die dabei auftretenden Probleme finden sich anschließend bei der Beschäftigung mit der Entwicklung von möglichst flexiblen, also vielseitig und universell einsetzbaren Toolkits wieder. In diesem Bereich hat Paul Dourish mit der Open Implementation angewendet auf Toolkits einen bemerkenswerten Ansatz geleistet, mit dem viele bisherigen Einschränkungen vermindert oder vermieden werden könnten. Auf bestimmte Aspekte seiner Lösung wird ein genauerer Blick geworfen. Zum Schluss erfolgt eine kurze Zusammenfassung und sowie eine Beurteilung der weiteren Entwicklung.

2 Anforderungen an CSCW-Systeme

Anders als Software für die Benutzung durch eine Person erfordern CSCW-Systeme für den Mehrbenutzerbetrieb verschiedene Funktionen, die für eine Zusammenarbeit erforderlich sind. Eine mögliche Aufteilung wäre in Applikationsanforderungen, sowie funktionale wie auch technische Anforderungen [3]:

Unter Applikationsanforderungen kann man die konzeptionellen Bedingungen für ein erfolgreiches Zusammenarbeiten einstufen. Bestimmte Tätigkeiten sind besonders oft Teil der computergestützten Zusammenarbeit. Deshalb sollten in einem vollständigen CSCW-System Funktionen wie Teamkalender oder gemeinsame Editoren zur Verfügung stehen.

Die funktionalen Anforderungen teilen sich auf mehrere Bereiche auf. Zum einen müssen vielfältige Arten von Interaktionen möglich sein. Darunter fällt beispielsweise synchrone, also gleichzeitige Interaktion von Benutzern, als auch asynchrone, etwa bei der zeitlich unabhängigen Bearbeitung desselben Dokuments. Weiterhin kann man auch von impliziter, vom Arbeitsobjekt bestimmter und expliziter, durch eine Echtzeitkonferenz

realisierte, Interaktion sprechen. Daneben gibt es noch die Unterscheidung zwischen formell und informell, die sich auf das Wissen und Können der beteiligten Nutzer bezieht. Dabei bedeutet formelle Interaktion, dass Regeln und Arbeitsabläufe im voraus festgelegt werden müssen, während informelle Interaktion die Zusammenarbeit zweier Benutzer mit gleichem Kenntnisstand bezeichnet. Dabei sind keine vorherigen Absprachen nötig. Ein weiterer Aspekt der funktionalen Anforderungen ist die Koordination zwischen den Nutzern. Diese hängt sowohl von der Gruppengröße als auch den spezifischen Tätigkeitsanforderungen ab. So ist es beim Brainstorming zum Beispiel erforderlich, dass alle Nutzer ihre Ideen möglichst ohne Verzögerung einbringen können. Dagegen wäre es beim gemeinsamen Durchsehen eines Dokuments eher kontraproduktiv, allen Nutzern die Möglichkeit zur Steuerung von Scrollfunktionen oder Ähnlichem zu geben. Eine andere funktionale Anforderung ist die Ermöglichung von Distribution, das heißt, Nutzer sollten von verschiedenen Orten aus zusammenarbeiten können. Ebenso müssen kulturelle und sprachliche Hürden berücksichtigt werden. Darüber hinaus gibt es die Unterscheidung zwischen "collaboration aware" und "collaboration transparent" für CSCW-Systeme, wobei letztere keine spezifischen Benutzerreaktionen ermöglichen, also in der Regel auf bestehende Anwendungen zurückgreifen. Die letzten zwei Punkte der funktionalen Anforderungen sind die Visualisierung und das "Data hiding". Programme sollten verschiedene Visualisierungsmodi für verschiedene Benutzer erlauben, so dass nicht alle Teilnehmer zwangweise an dieselbe Darstellung gebunden sind. Das "data hiding", also das Verstecken von Informationen für bestimmte Nutzergruppen, ist eine der wichtigsten Eigenschaften, die ein CSCW-System bieten muss. Damit hat ein Benutzer die Möglichkeit, Daten, an denen er gerade arbeitet, entweder öffentlich zu machen oder vor fremdem Zugriff zu schützen.

Neben den eben besprochenen Anforderungsbereichen gibt es auch noch technische Kriterien für CSCW-Systeme. Diese sind aufgeteilt in den Hardware-, Software- und Netzwerkbereich. Die Hardware steht für die Zielsysteme, Ein- und Ausgabegeräte, etc. Unter der Software fasst man das Betriebssystem und die graphische Oberfläche zusammen, während unter dem Punkt Netzwerke die Merkmale beziehungsweise Typen der Infrastruktur, wie LAN oder WLAN, bezeichnet werden.

All diese Anforderungen an CSCW-Systemen kann man hinsichtlich ihrer Flexibilität beurteilen. Manche Konzepte auf niedrigerem Level (z.B. der Hardware) können bestimmte Funktionen auf höherer Ebene unmöglich machen [3]. Deshalb stellt die Flexibilität einen so wichtigen Aspekt bei der Entwicklung von Frameworks und Toolkits, die am Ende dieser Ar-

beit genauer betrachtet wird, dar. Wenn auf niedrigerer Ebene bestimmte Vorgaben festgelegt werden, kann dies gewünschte Funktionen bei der späteren Handhabung unmöglich machen.

3 Entwicklungsansätze für verteilte Benutzerschnittstellen

Verteilte Benutzerschnittstellen bilden einen Spezialfall bei der Entwicklung von CSCW-Systemen. Mit ihrer steigenden Verbreitung und vielseitigen Einsetzbarkeit finden sie zunehmend Verwendung bei der Zusammenarbeit zwischen mehreren Benutzern. So können zum Beispiel spontan neue Mitglieder mit ihren Endgeräten wie PDAs oder Mobiltelefonen in ein Team integriert oder in der Umgebung verfügbare Projektionsflächen für die Darstellung einer Grafik benutzt werden.

3.1 Spezielle Anforderungen verteilter Benutzerschnittstellen

Zusätzlich zu den Anforderungen für CSCW-Systeme kommen in diesem Fall noch weitere Faktoren hinzu, die eine komplexere Situation und somit auch weitere Anforderungen an Toolkits entstehen lassen. Vandervelpen und Coninx machen in ihrer Arbeit dabei drei allgemeine Bereiche aus [4]. Der erste ist die Aufteilung einer Benutzerschnittstelle auf verschiedene Geräte. Mit konventionellen Schnittstellen ist dies nicht möglich, da diese üblicherweise eng mit der Implementierung verknüpft sind. Stattdessen muss eine unabhängige Beschreibung der Schnittstelle beispielsweise in einer XML-Sprache vorliegen. Die zweite Herausforderung ist der Erhalt der Benutzbarkeit. So muss es eine sinnvolle Aufteilung zwischen den Geräten geben. Beispielsweise sollten auf weiter entfernten Geräten nur graphische Ausgaben und keine Steuerungselemente ausgegeben werden. Auch sollte die Verteilung der Schnittstellen keine Arbeitspause beim Benutzer erzwingen. Daneben weisen verteilte Benutzerschnittstellen dieselben Voraussetzungen wie andere verteilte Systeme auf, was Datenaufteilung oder Steuerung betrifft. Insgesamt kann man für die Anforderungen sieben genauere Punkte definieren [4]:

Heterogenität: In einem System von verteilten Schnittstellen sind viele verschiedene Geräte vorhanden. Es müssen einerseits die technischen Voraussetzungen betrachtet werden, andererseits ist es nötig, die Benutzerschnittstelle an die unterschiedlichen Geräte auf die sie verteilt ist, anzupassen. Mit der zuletzt genannten Problemstellung haben sich Vandervelpen und Coninx in ihrer Arbeit intensiver beschäftigt [4].

Offenheit: Das System muss offen für Veränderungen und Erweiterungen sein. Zusätzliche Ein- und Ausgabegeräte sollten ohne größeren Aufwand unterstützt werden können.

Sicherheit: Eine fehlerfreie Funktionalität des Systems und Verfügbarkeit von Ein- und Ausgabemöglichkeiten muss sichergestellt sein. Darüber hinaus müssen persönliche Daten einzelner Benutzer vor dem Zugriff durch andere geschützt werden - sowohl was das Anzeigen von Informationen auf öffentlichen Displays betrifft, als auch während dem Zusammenarbeiten in einer kleineren Gruppe von Benutzern.

Skalierbarkeit: Auch wenn sich die Zahl der Ressourcen oder Benutzer erhöht, muss ein ungestörtes Arbeiten garantiert werden. Komplexe Situationen mit häufig wechselnden Akteuren sind bei verteilten Benutzerschnittstellen üblich.

Fehlerbehandlung: Wenn Geräte oder Ein- und Ausgabeschnittstellen versagen, muss das System über eine geeignete Strategie verfügen, so dass es nicht zu einem Absturz oder gar Datenverlust kommt.

Gleichzeitigkeit: Um Dateninkonsistenz zu vermeiden, muss der Zugriff auf Ressourcen geregelt werden. Das geschieht in der Regel mit Hilfe von Locking-Mechanismen. So soll sichergestellt werden, dass ein Benutzer beispielsweise eine Eingabe fertigstellen kann, bevor jemand anders die Kontrolle über das Eingabegerät erhält.

Transparenz: In verteilten Systemen gibt es vielfältige zu erfüllende Transparenzarten. Dazu zählen unter anderem Fehler-, Migrations-, Orts-, Performanz-, Replikations-, Skalierungs-, Transaktions- und Zugriffstransparenz. Der Ortstransparenz kommt bei verteilten Schnittstellen besondere Bedeutung zu. Einerseits sollten Ein- und Ausgabegeräte möglichst ohne Interaktion durch den Benutzer erkannt und zum Ressourcenpool hinzugefügt werden, andererseits müssen aber Benutzer und System wissen, an welcher Stelle sich eine Ressource befindet, um zu vermeiden, dass etwa ein Ausgabegerät aus einem anderen, nahe gelegenen Raum in die Liste der aktiven Geräte hinzugefügt wird.

Die genannten Punkte stellen also die Basis für die Entwicklung von Toolkits für verteilte Systeme dar. Im weiteren werden verschiedene Aspekte verteilter Benutzerschnittstellen mit Hilfe von modell-basierten Ansätzen betrachtet. Unter anderem wird die Interaktion mit verteilten Schnittstellen als auch die eng damit verknüpfte Migration zwischen Geräten betrachtet. Zum Schluß wird exemplarisch die Vorgehensweise beim Design der Schnittstellen beschrieben. Die einzelnen Beispiele beziehen sich in der Regel auf Einzelnutzer, aber die vorgestellten Konzepte kön-

nen ebenso auf den bei CSCW-Systemen vorliegenden Mehrbenutzerbetrieb übertragen werden.

3.2 Interaktion mit verteilten Benutzerschnittstellen

Besonders interessant bei verteilten Benutzerschnittstellen ist die gemeinsame Bearbeitung von Daten durch mehrere Benutzer mit Hilfe verschiedener Endgeräte. Mit der zunehmenden Bedeutung von Ubiquitous Computing wird es in Zukunft möglich sein, in Situationen, in denen ein Nutzer weitere Ein- oder Ausgabemöglichkeiten benötigt, von Hardware in der Umgebung Gebrauch zu machen. So könnten Räume für Besprechungen oder Teamarbeit so ausgestattet werden, dass sie für eine Vielzahl von verschiedenen Situationen entsprechende Möglichkeiten wie Projektionsflächen oder Smartboards bereithalten. Im Notfall würde in der Regel aber zumindest ein PDA oder Mobiltelefon für den jeweiligen Benutzer verfügbar sein. Darauf kann zurückgegriffen werden, wenn es nicht möglich ist, andere Geräte aus der Umgebung in den Arbeitsprozess mit einzubinden. Da solch ein "Fallback" allerdings eine Einschränkung der Nutzungsmöglichkeiten darstellt, ist dieser - wenn möglich - zu vermeiden. Im folgenden soll also besprochen werden, wie man einem Nutzer ermöglichen kann, die Infrastruktur zusammen mit den eigenen Geräten zu benutzen, anstatt eine Entweder-Oder-Entscheidung treffen zu müssen. Braun und Mühlhäuser haben in ihrer Arbeit zunächst einmal die Einschränkungen von Ubiquitous Computing Umgebungen bestimmt [5]. Hierbei kann unterschieden werden zwischen räumlichen Einschränkungen (zum Beispiel kann mit Touchscreens nur interagiert werden, wenn diese nicht weiter als Armlänge entfernt sind), Einschränkungen durch die Geräte (nicht alle Bildschirme sind auch Touchscreens, lassen also auch eine Eingabe zu) und soziale Einschränkungen (beispielsweise sollten persönliche Informationen nicht ohne weiteres auf öffentlich sichtbare Bildflächen abgebildet werden).

Wenn ein persönliches Mobiltelefon oder ein PDA zusammen mit der umgebenden Infrastruktur genutzt werden kann, ist es möglich, alle drei eben aufgeführten Einschränkungsarten zu überwinden. Beispielsweise können weiter entfernte Anzeigeflächen mit dem eigenen Eingabegerät gesteuert werden. Wenn in der Umgebung keine Audioausgabefunktionen vorhanden sind, könnte ein Mobiltelefon diese technische Einschränkung aufheben. Und auch persönliche Informationen würden in einem sicheren Modus nur auf dem eigenen, persönlichen Gerät mit kleinerem Bildschirm gezeigt werden. Somit können in Kombination mit einer Ubiquitous Computing Umgebung alle Einschränkungen sowohl der äußeren Infrastruk-

tur als auch des persönlichen Handhelds kompensiert werden. Braun und Mühlhäuser sprechen von sogenannten "federated user interfaces" [5].

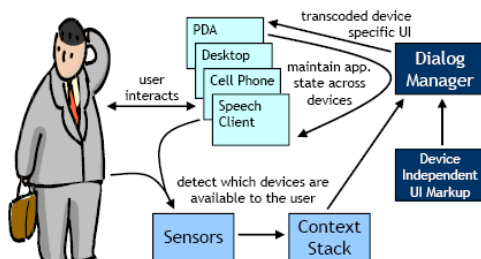


Abbildung 1. Interaktionsmanagement in Ubiquitous Computing Umgebungen [5]

In ihrer Arbeit schildern sie auch eine von ihnen entwickelte Laufzeitumgebung, die den Anforderungen entspricht [5]. Sie basiert auf einem Model-View-Controller und wird verglichen mit einer Art Webserver, der je nach Gerät eine passende Schnittstelle generiert. Die Vorgehensweise bei der Benutzung ist wie folgt: Zunächst wird ein einzelnes XML-Dokument, das die Benutzerschnittstelle beschreibt, erstellt. Es enthält Informationen über private Bereiche der Anzeige, Bedienelemente, die Eingaben erfordern, etc. Wenn ein Benutzer die Anwendung startet, wird eine Verbindung zu einem Dialog Manager gestartet. Dieser generiert aus den Daten im XML-Dokument eine Benutzerschnittstelle. Wenn der Benutzer sich in der Nähe einer Projektionsfläche oder Ähnlichem befindet, wird ein Event ausgelöst, durch das der Dialog Manager erneut eine Benutzerschnittstelle erstellt, die jeweils auf die Eigenschaften der benutzten Geräte eingeht. Anschließend werden alle Geräte synchronisiert und die Benutzeroberfläche übertragen. Falls der Benutzer jetzt eine Eingabe macht, wird diese über ein Event dem Dialog Manager gemeldet und wiederum die Anzeige auf den verschiedenen Clients aktualisiert. Abbildung 1 zeigt eine graphische Darstellung dieser Mechanismen.

3.3 Migration zwischen Ein- und Ausgabegeräten

Nach der Interaktion zwischen einem Nutzer und verteilten Benutzerschnittstellen soll nun genauer auf die Migration, also den direkten Wechsel zwischen einzelnen Geräten eingegangen werden. Dass dies neue Probleme mit sich bringt, zeigt der folgende Abschnitt. Anschließend werden

die verschiedenen Migrationsarten sowie das sinnvolle Design dafür geeigneter Benutzerschnittstellen betrachtet.

Probleme bisheriger Toolkits Mittlerweile gibt es mit BEACH, Aura und anderen schon einige Ansätze für Toolkits, die verteilte Benutzerschnittstellen unterstützen. Allerdings haben sie noch mit diversen Problemen zu kämpfen, die besonders Umgebungen betreffen, in denen viele verschiedene Ein- und Ausgabegeräte vorhanden sind [6]. In Zukunft muss damit gerechnet werden, dass - vor allem in Firmen - solch heterogene Umgebungen immer öfter anzutreffen sein werden, da die Verfügbarkeit verschiedener Geräte besonders für die Zusammenarbeit mehrerer Benutzer von hoher Relevanz ist. Deshalb sollte folgenden Problemfeldern besondere Aufmerksamkeit geschenkt werden [6]:

Zum einen ist das Fenstermodell, das sich in Anlehnung an bisherige Darstellungsformen mit rechteckigen Bildschirmflächen etabliert hat, nicht mehr zeitgemäß. Stattdessen sollte mit einer beliebigen Form und Beschaffenheit von Flächen gerechnet werden können, auf die beispielsweise etwas projiziert werden kann oder die mit entsprechenden Anzeigeeinheiten ausgestattet sind. Auch Situationen, die eine drehbare Anzeige ermöglichen - beispielsweise für Projektionen auf runde Tische - sollten einkalkuliert werden. Bisher mussten für solche Gegebenheiten jedes Mal Schnittstellen von Grund auf neu erstellt werden. Für eine solide Lösung ist also sowohl eine flexible Ausgabe als auch eine Reihe von Sensoren nötig, um die jeweiligen Eigenschaften einer Fläche zu erfassen.

Die geometrischen Unterschiede zwischen den Flächen werden bisher nicht miteinbezogen. So sollten in Zukunft auch 3D-Daten erfassbar sein, um von den physischen Gegebenheiten optimal profitieren zu können. So könnten zum Beispiel 3D-Präsentationen auf vertikalen Flächen und 2D-Präsentationen auf horizontalen Flächen möglich sein.

Noch gibt es kaum Unterstützung für mehrere Eingabegeräte. Bei konventionellen Desktop-Systemen wird beispielsweise jeweils nur mit einer Maus und einem Keyboard gerechnet. Bei mehreren Benutzern ist es allerdings wünschenswert, eine größere Anzahl von Eingabegeräten nutzen zu können. Dafür müssen Toolkits noch weiter optimiert werden.

Die Interaktion bezieht sich meist auf die Ressourcen von einzelnen Workstations. Es sollte aber die Möglichkeit geben, Eingabegeräte von mehreren Computern oder anderen Geräten zusammen zu nutzen. Dabei müssen ebenfalls die Vielzahl von Eingabegeräten und ihre jeweiligen Eigenheiten berücksichtigt werden. Dadurch könnte dann ein Computer

gleichzeitig auch mit einem PDA oder ähnlichen Geräten gesteuert werden.

Oft gibt es keine dynamische Erkennung von neu hinzukommenden oder aus der Umgebung entfernten Geräten. Diese muss aber konstant und mit möglichst geringen Intervallen erfolgen, um einer Ubiquitous Computing Umgebung gerecht zu werden. Beispielsweise könnten Benutzer spontan in eine Situation kommen, in der sie Informationen an eine öffentliche Anzeigetafel oder Ähnliches schicken möchten. Diese Art der Interaktion ist im Moment noch ungenügend oder ohne Einbeziehung der Benutzerschnittstellenentwicklung unterstützt.

Jeweilige Beispiele für Toolkits sind in der Arbeit von Coutaz et al. aufgeführt [6]. Die Autoren stellen dort ebenfalls ein Toolkit namens IAM vor, das eine Lösung dafür anbietet. Diese ist eng mit dem Prinzip der Flexibilität verknüpft, die im nächsten Kapitel dieser Arbeit erläutert wird.

Migrationsarten und Durchführung Nach Betrachtung der Probleme bei einem Wechsel zwischen Geräten folgt nun eine Analyse der verschiedenen Migrationsarten, die es zwischen Geräten geben kann. Von Bandelloni und Paternò wurden drei unterschiedlichen Typen bestimmt [7]: vollständige, partielle und gemischte Migration. Unter der vollständigen Migration wird der komplexe Wechsel der Benutzerschnittstelle von einem Gerät auf ein anderes verstanden. Bei der partiellen Migration werden Kontroll- und Anzeigeteil der Benutzeroberfläche getrennt und jeweils auf ein eigenes Gerät verteilt. Bei der gemischten Migration schließlich können auch Kontroll- und Visualisierungsbereich noch einmal in kleinere Einheiten zerlegt und an verschiedene Geräte geschickt werden.

In ihrer Arbeit stellen Bandelloni und Paternò für die vollständige und partielle Migration jeweils einen Lösungsansatz mit Hilfe des von ihnen entwickelten Tools TERESA vor [7]. Das Ziel war, eine effektive Migration während der Laufzeit zu ermöglichen. In ihrem Beispiel betrachten sie eine Webseite mit einem Darstellungs- und einem Kontrollbereich.

Bei der vollständigen Migration wird zuerst ermittelt, welche Aktionen auf der neuen Geräteplattform möglich sind, anschließend wird eine passende Aktionsrepräsentation, also eine Sammlung von Aktionen, um ein Ziel zu erreichen, ausgewählt. Dabei sind mehrere Fälle möglich. Im einfachsten Fall ist die Ausgangspräsentation auf dem bisherigen Gerät gleich der Zielpräsentation auf der neuen Plattform (one-to-one). Dann kann sofort die entsprechende Webseite geladen werden. Im nächsten Fall (one-to-many) gibt es verschiedene Alternativen für die Zielseite, die ver-

schiedene Aktionen unterstützen. In solchen Situationen muss eine davon ausgewählt werden. Dies geschieht in der Regel mit einem Vergleich, welche Repräsentationen am ähnlichsten sind, also dieselben Tätigkeiten ermöglichen. Falls damit kein eindeutiges Ergebnis ermittelt werden kann, wird noch die Reihenfolge der möglichen Operationen als Kriterium herangezogen. Zuletzt gibt es noch die many-to-one Kombination bei der verschiedene Quellrepräsentationen auf eine Zielrepräsentation passen. Dann wird das "task set" ausgewählt, das der Zielrepräsentation entspricht. In TERESA sind diese Prozeduren mit einem Client-Server-Modell gelöst. Nach einer Analyse wird für die Zielplattform jeweils die passende Benutzerschnittstelle generiert. Abbildung 2 zeigt eine graphische Darstellung des Prinzips.

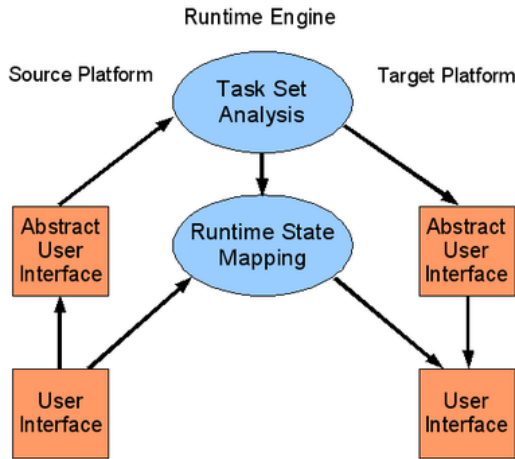


Abbildung 2. Analyse der Task Sets und Zuordnung von Schnittstellen nach [7].

Prinzipiell hat die partielle Migration ähnliche Anforderungen wie die vollständige Migration. Das Ziel ist es, ein Gerät mit Hilfe eines anderen steuern zu können (beispielsweise eine Videowand mit Hilfe eines PDAs). Die Herausforderung besteht dabei darin, eine sinnvolle Aufteilung in Kontroll- und Visualisierungsbereich einer Anwendung zu finden. In TERESA selbst wurden, um die Komplexität zu senken, nur selbst entwickelte Anwendungen und nicht generell auffindbare Webseiten betrachtet. Trotzdem bleibt das Problem, dass unter Umständen aufgrund einer Benutzereingabe weitere Kontrollelemente erzeugt werden müssen. Dabei muss verhindert werden, dass es zu einer unsinnigen Migration von Kontrollelementen auf das mit der Visualisierung belegte Gerät kommt.

Bandelloni und Paternò haben deshalb zwei verschiedene Szenarien definiert, die sich für eine partielle Migration eignen. Im ersten Fall kann man eine Seite einfach aufteilen, Kontrollelemente und eigentlicher Seiteninhalt bleiben immer getrennt. Im zweiten Fall bleibt die Seite weiterhin komplett auf dem Steuergerät sichtbar, während auf dem Präsentationsmedium eine speziell angepasste Version dargestellt wird. So kann der Benutzer weiter frei navigieren, hat aber zusätzlich eine größere Anzeigefläche zur Verfügung.

In Zukunft wird es noch nötig sein, ein Werkzeug zu entwickeln, mit dem auch aus herkömmlichen Webseiten Navigations- und Steuerfunktionen extrahiert werden können. Ansonsten ist keine angepasste Migration möglich. Dies kann dann nur realisiert werden, wenn Benutzerschnittstellen nach speziellen Kriterien erstellt werden. Eine Herangehensweise dafür wird im nächsten Unterpunkt besprochen.

Design von verteilten Benutzerschnittstellen Ein Vorschlag für ein grundlegendes Konzept zur Entwicklung und zum Design von verteilten Benutzerschnittstellen wird von Luyten und Coninx vorgestellt [8]:

Für das Erstellen von Schnittstellenelementen benutzen sie UIML, die User Interface Modelling Language. Dabei wird jedes Element einer Tätigkeit zugeordnet und kann mehrere UIML-Dokumente als Alternativen enthalten. Je nachdem, welches der UIML-Dokumente am geeignetsten für das gewählte Gerät ist, wird eines davon für die Generierung der Schnittstelle genutzt. Aufgrund der Nutzung von UIML sind Inhalte, Funktionszuordnungen (also welchen Bedienelementen welche Funktionen zukommen) und Applikationslogik konzeptionell voneinander getrennt. So ist es auch möglich, dass Personen aus unterschiedlichen Fachbereichen (graphischem Design, Software-Entwicklung, etc.) zusammen ein Schnittstellendesign erstellen können. Mit Hilfe der multiplen Dokumentinstanzen können auf diesem Weg passende Lösungen für die verschiedensten Schnittstellen vereint werden.

Bei der Entwicklung von Benutzerschnittstellen müssen Überlegungen in drei unterschiedlichen Dimensionen angestellt werden: Handlungs-, Geräte- und Ortsdimension. Luyten und Coninx sprechen von einem "Distributed Interaction Space", der Benutzerschnittstellen klassifiziert [8]. Die Handlungsdimension beschreibt die Tätigkeiten, die von den verschiedenen Benutzern ausgeführt werden müssen, um ein bestimmtes Ziel zu erreichen. Durch die Erstellung von "task sets" kann im voraus entschieden werden, wie die Elemente der Benutzerschnittstelle über die verfügbaren Ressourcen verteilt werden können. Abbildung 3 zeigt eine sol-

che Zuordnung von Ressourcen zu den jeweiligen, je nach Zeitpunkt auszuführenden Aufgaben. Die Gerätedimension wird von den technischen Eigenschaften der genutzten Geräten bestimmt. Viele von diesen können mittlerweile Informationen über sich selbst übertragen, wenn sie in ein Netzwerk eingebracht werden. Aus diesen kann ebenfalls eine Grundlage zur Verteilung der Schnittstellenelemente gewonnen werden. Die letzte Dimension, die Ortsdimension, steht für die momentane Position im Raum, an der sich ein Gerät befindet. Je nachdem, um welchen Gerätetyp es sich handelt und wie weit sich der Benutzer im Raum bewegt, sollte eine Benutzerschnittstelle angepasst werden.

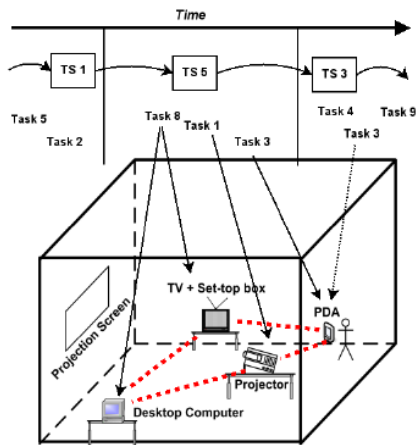


Abbildung 3. Task Sets im Interaktionsraum [8]

Neben dem Interaktionsraum ("interaction space") und dem Speicherformat muss auch die Distribution der Benutzerschnittstellen in die Überlegungen mit einbezogen werden. Man kann zwischen drei Methoden unterscheiden: vom Benutzer ausgelöste, vom System ausgelöste und kontinuierliche Verteilung. Beim ersten Fall geht die Initiative vom Benutzer aus. Er meldet sich an einem "distribution daemon" an und legt manuell fest, welche Geräte er in seinem persönlichen Interaktionsraum benutzen möchte. Oft ist dies die präferierte Methode. Ansonsten kann die Verteilung auch automatisch durch das System erfolgen. Nach der Anmeldung beim "distribution daemon" teilt dieser nach einem Scan der verfügbaren Geräte die Benutzerschnittstellen entsprechend zu. Dabei kann beispielsweise mit Least Cost Routinen gearbeitet werden, um Bandbreitenverbrauch, Latenz oder beliebige andere Werte möglichst niedrig zu halten.

Der letzte Fall der kontinuierlichen Verteilung beschreibt die Situation, dass eine Interaktion schon im Gange ist, aber Geräte wegfallen oder hinzukommen. Dann können oder müssen Schnittstellenelemente unter Umständen anders verteilt werden. Man kann diesen Prozess ebenfalls mit Kostenberechnungen kombinieren. Dabei muss allerdings darauf geachtet werden, dass die Benutzertätigkeit nicht unterbrochen oder erschwert wird.

Abschließend sollte noch die Benutzbarkeit von verteilten Benutzerschnittstellen bewertet werden. Luyten und Coninx sehen hier zwei Metriken vor: Vollständigkeit und Kontinuität. Vollständigkeit bedeutet, dass mit den bestehenden Schnittstellenelementen alle zur Erreichung eines gesetzten Ziels nötigen Tätigkeiten erledigt werden können. Für traditionelle Computersysteme ist dies leicht durchzusetzen, bei verteilten Geräten stellt dies aber eine große Herausforderung dar. Nicht immer ist es möglich, sicherzustellen, dass alle nötigen Aktionen durchgeführt werden können. Beispielsweise könnte für eine Aufgabe das Aufnehmen von Audiodaten erforderlich sein aber zum gegenwärtigen Zeitpunkt kein Gerät mit Mikrophon vorhanden sein. Um solche Fälle zu kompensieren kann bei der Festlegung der vorher besprochenen "task sets" eine alternative Vorgehensweise bestimmt werden. Trotzdem kann damit nicht für alle Situationen eine entsprechende Lösung gefunden werden. Deshalb besteht hier noch Forschungsbedarf. Die Kontinuität stellt die Forderung, dass Interaktionen nicht unterbrochen werden sollen, auch wenn sich die Zusammensetzung der benutzten Geräte ändert. Dies ist natürlich bei verteilten Schnittstellen ein erhebliches Problem. Ein Ansatz ist die Spezifikation eines Regelsatzes für bestimmte Tätigkeiten, die beispielsweise besagen, dass es in bestimmten Situationen zu keiner Redistribution von Schnittstellenelementen kommen darf. Da dies offensichtlich nicht immer möglich ist, sollten zusätzlich Alternativen für die Weiterführung einer Aktion definiert werden.

4 Flexibilität von CSCW-Toolkits

Nach dem vorangegangenen Kapitel über verteilte Benutzerschnittstellen in denen die Probleme bei der Entwicklung erläutert wurden, stellt sich nun die Frage, wie man ein Toolkit entwickeln kann, das nicht nur für eine spezielle Nutzungssituation geeignet ist. Die ersten CSCW-Systeme wurden noch als alleinstehende Systeme entwickelt und jedesmal von Grund auf neu konstruiert. Mit der zunehmenden Verbreitung dieser Systeme schien es bald wünschenswert, gewisse Elemente wiederzuverwerten, anstatt im-

mer wieder neu zu erstellen [2]. Auch hinsichtlich der späteren Zahl und Zusammensetzung der Benutzergruppen ist es notwendig, einen möglichst flexiblen Ansatz zu schaffen, da jeder Benutzer unterschiedliche Anforderungen hat und sich die Arbeitsweisen sogar während der Zusammenarbeit ändern können. Beispielsweise könnte ein Benutzer nach einiger Zeit eine andere Visualisierung für in einem Projekt angefallene Daten wünschen. Die Zielsetzung ist also, Toolkits beziehungsweise Frameworks zu schaffen, die höchste Flexibilität bieten. In seiner Dissertation hat Paul Dourish verschiedene etablierte Toolkits hinsichtlich dieser Eigenschaft analysiert und als Basis für seine weiteren Überlegungen genutzt [2].

4.1 Vielseitigkeit und Erweiterbarkeit

Um Flexibilität erreichen zu können, muss zuerst darüber nachgedacht werden, mit Hilfe welcher Ansätze man dies bewältigen kann. Paul Dourish macht zwei Merkmale aus, die Grundsteine für die Flexibilität sein können: "Genericity", Vielseitigkeit und "Extensibility", Erweiterbarkeit [2]. Die Vielseitigkeit beschreibt die Palette an Verhaltensweisen und Unterstützung von verschiedenen Aktionen, die ein Toolkit bietet, während die Erweiterbarkeit für das Prinzip der Analyse und Reflexion steht, mit dem das genannte Verhalten verändert werden kann. Mit Hilfe dieser beiden Prinzipien können die Funktionen und Komponenten von Toolkits bewertet werden. Als Komponenten bezeichnet Paul Dourish die dem Nutzer von CSCW-Systemen angebotenen Objekte, wie z.B. Teile einer graphischen Oberfläche. Als Funktionen sieht er die direkten Funktionalitäten eines Systems, also das Zusammenwirken der verschiedenen Komponenten.

Die Vielseitigkeit sieht Dourish als wichtigste Voraussetzung für flexible Toolkits. Sie kann sowohl auf Komponenten als auch auf Funktionen angewendet werden. Das Ziel ist es, ähnliche Verhaltensweisen verschiedener Applikationen zu erfassen und mit Hilfe der gewonnenen Informationen Komponenten und Funktionen so zu erstellen, dass sie in einer Vielzahl von verschiedenen Anwendungsbereichen eingesetzt werden können. Sie charakterisiert also die grundsätzliche Einsetzbarkeit eines Toolkits in den unterschiedlichsten Arbeitssituationen (z.B. Internetkonferenzen, gemeinsames Bearbeiten von Dokumenten, Brainstorming, etc.)

Die Erweiterbarkeit wiederum soll sicherstellen, dass Toolkits auch vom Anwendungsentwickler erweitert werden können, um spezifische Bedingungen zu erfüllen. Damit bildet ein Toolkit praktisch ein Framework für Toolkiterweiterungen mit dem wiederum neue Toolkit Funktionen erstellt werden können. Dabei können ganz verschiedene Strategien zum

Einsatz kommen. Einerseits bietet sich die Erweiterung bestehender Funktionalitäten an, andererseits sollte es auch möglich sein, komplett neue Toolkitbestandteile zu erstellen. Diesem Aspekt kommt eine hohe Bedeutung zu, da es einem Toolkit-Entwickler nicht möglich ist, alle in Frage kommenden Anwendungsfälle von vornherein zu berücksichtigen.

Bisher kamen die genannten Konzepte schon in Toolkits zum Einsatz, allerdings in der Regel nur eines von beiden. So sind beispielsweise Toolkits wie Oval und COLA ausschließlich auf Vielseitigkeit ausgerichtet und bieten keine Möglichkeit, tatsächliche Erweiterungen zu erstellen [2]. In den Toolkits MEAD, Rendezvous and Suite wurde dem Programmierer schon mehr Kontrolle über die Funktionen eingeräumt, aber erst GroupKit war ein Ansatz der auch die Erweiterbarkeit mit ins Spiel brachte [2]. Allerdings bleiben selbst bei Einsatz beider Prinzipien noch zwei weitere Probleme bestehen [2]. So ist es zum einen für den Programmierer nicht möglich, die Semantik des Toolkits, also das Datenmodell zu verändern - in der Regel können nur strukturelle Änderungen eingebracht werden. Zum anderen können, was die Erweiterbarkeit betrifft, bestehende interne Strukturen meist nicht selbst angepasst werden, sondern müssen neu implementiert werden. Dadurch kann aber kein Einfluß auf das Verhalten zwischen den einzelnen Komponenten erreicht werden. Dazu wäre ein tieferes Eingreifen in das Toolkit-System mit seinen internen Aspekten nötig. Wie man einen solchen Ansatz für Toolkits realisieren kann, wird im nächsten Abschnitt besprochen.

4.2 Das Prinzip der Open Implementation

Wie schon bei den Anforderungen von CSCW-Systemen angedeutet und in der vorhergehenden Betrachtung der Flexibilität in Toolkits weitergeführt, rühren viele der Einschränkungen für den Programmierer von Entscheidungen auf "tieferer" Ebene her. Bisherige Abstraktionsmodelle sind nicht darauf ausgerichtet, Kontrolle über niedrigere Abstraktionsebenen zu geben [9]. In der Regel liegt eine Abstraktionsbarriere vor, die Implementierungsdetails "versteckt" [2]. Abbildung 4 zeigt eine vereinfachte Darstellung dieses Konzeptes.

In der Vergangenheit gab es vor allem zwei Methoden, um dieses Problem zu umgehen [2]. Zum einen die sogenannten "Hematomas", siehe Abbildung 5a, bei denen über der eigentlichen Abstraktionsbarriere noch einmal Code produziert wird, der die ursprüngliche Abstraktion erweitert - aber in der Regel viele Funktionen noch einmal implementieren muss. Ein Beispiel wäre die Berechnung von Daten, die aus der graphischen Oberfläche eines Programms, beispielsweise einer Tabellenkalkulation, ausgelesen

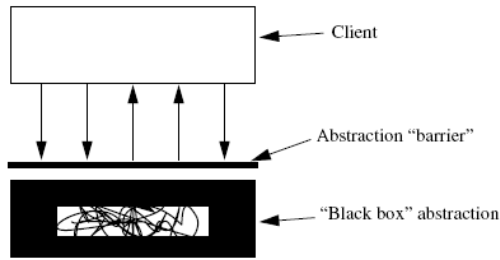


Abbildung 4. Die Sicht auf niedrigere Ebenen wird durch die Abstraktionsbarriere verhindert [2]

werden. Daneben gibt es das "Coding between the lines" (Abbildung 5b), dabei wird mit Wissen über die eigentliche Implementierung zusätzlicher Code eingebaut, der darauf Bezug nimmt. Hierbei besteht das Problem, dass der resultierende Code unübersichtlicher und schwerer zu verstehen ist. So könnte ein Entwickler die Verwaltung von virtuellem Speicher aufgrund von Performanzproblemen im Betriebssystem selbst anpassen, indem er auf systemspezifische Eigenheiten eingeht. Dadurch ist allerdings eine Portierung auf andere Systeme wiederum mit erheblichen Schwierigkeiten verbunden. Aufgrund der genannten Probleme wird klar, dass beide Lösungen keine zufriedenstellenden Ansätze darstellen.

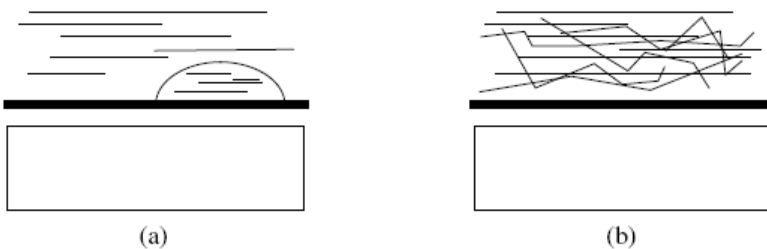


Abbildung 5. Bisherige Lösungen des Abstraktionsproblems [2]

An dieser Stelle greift die "Open Implementation". Bei der Erstellung einer Implementierung müssen verschiedene "mapping decisions" getroffen werden, die beschreiben, wie Anforderungen von höheren Ebenen den jeweiligen Einheiten aus niedrigeren Ebenen zugeordnet werden sollen.

Das erfordert aber eine Beurteilung der zukünftigen Einsatzgebiete der Anwendungen. Da diese aber sehr stark variieren können, ist eine Einschränkung im voraus nicht wünschenswert [9]. Deshalb setzt die "Open Implementation" auf die sogenannte "computational reflection", die ein System beschreibt, das mit Hilfe einer Repräsentation der eigenen Aktivitäten Reflexionen über sich selbst anstellen kann und sein Verhalten - wenn nötig - durch Veränderungen dieser Repräsentation beeinflussen kann. Konkret bedeutet das, dass innerhalb der Implementierung eines Toolkits gleichzeitig auch eine Repräsentation der Toolkit-Struktur und des Verhaltens vorhanden ist. Dabei erfolgt eine Aufteilung in ein "base interface" und ein "metalevel interface", wobei das "base interface" die traditionelle Abstraktionsbarriere darstellt, während das "metalevel interface" eine Spezialisierung der Implementierung auf die jeweiligen Bedürfnisse ermöglicht. Als Konsequenz kann man auch den Code in "base code" und "metalevel code" einteilen [9]. Vereinfacht betrachtet kann man sagen, dass der "base code" bestimmt, was zu tun ist, während der "metalevel code" festlegt, wie dies auszuführen ist. Durch diese Mechanismen sind die Metalevel-Techniken besonders gut für die Entwicklung von CSCW-Systemen geeignet, da mit entsprechenden Toolkits Anwendungen für ein sehr großes Einsatzgebiet erstellt werden können. Auf jeweilige Anforderungen kann durch Einbeziehung der Implementierungsart flexibel reagiert werden.

4.3 Metalevel-Techniken für flexible Toolkits

Um die Wirkungsweise zu demonstrieren, hat Dourish das Prinzip der Open Implementation in einem von ihm entwickelten CSCW-Toolkit, Prospero, verwirklicht. Dabei war sein Ziel, Applikationsentwicklern nicht nur möglichst viele verschiedene Anwendungsverhaltensweisen zu ermöglichen - anstatt einer vorgegebenen Optionspalette - sondern auch die Möglichkeit, bestehende Komponenten anwendungsspezifisch zu kontrollieren und diese Kontrollmöglichkeiten über ein "metalevel interface" verfügbar zu machen [9]. Er spricht von einem vielseitigen und spezifizierbaren Anwendungsmodell als Ausgangsbasis. Im folgenden werden zwei solche Modelle für die Datendistribution und das Konsistenzmanagement genauer beschrieben. Beide sind auch in Prospero implementiert.

Divergenz und Synchronisation Eines der größten Probleme bei der Entwicklung von CSCW-Systemen stellt sicherlich die Verteilung der Daten, die von den Benutzern bearbeitet werden, dar. Zum einen müssen

diese jederzeit möglichst ohne Verzögerung zur Verfügung stehen, andererseits muss gewährleistet werden, dass die von unterschiedlichen Personen bearbeiteten Dateien anschließend wieder zusammengefügt werden können. Dabei gibt es die Möglichkeit, Daten nur an einem Ort zu speichern oder - die in der Regel häufigere Variante - Kopien der Daten an verschiedenen Orten abzulegen. Bei der zuletzt genannten Methode kommt es jedoch meistens zu Inkonsistenzen. Entscheidungen über das Datenhaltungsverfahren einer Applikation stellen einen elementaren Schritt bei der Entwicklung dar. Gerade bei CSCW-Systemen werden dadurch Verfügbarkeit, Transparenz, Inkonsistenzen und Antwortzeit in erheblichem Maße kontrolliert [9]. Deshalb sollte ein Toolkit nach Möglichkeit keine Einschränkungen der Optionen forcieren, sondern diese dem Applikationsentwickler überlassen. Im Prospero Toolkit ist das Management von verteilten Daten deshalb mit einem Divergenz- und Synchronisationsmodell verwirklicht. Der erste grundlegende Unterschied zwischen diesem Modell und den üblichen Herangehensweisen in anderen Toolkits besteht darin, dass letztere normalerweise versuchen, Inkonsistenzen zu vermeiden [9]. Gerade diese Methodik kann aber die Freiheit von möglichen Arbeitsstilen bei der Benutzung erheblich einschränken. Bei synchronem Arbeiten wird dabei in der Regel ein einziger Strom von Benutzeraktionen mit Hilfe von Methoden wie Locking oder Floor Policies erzeugt. Prospero weist dagegen eine vollkommen andere Art der Datenverwaltung auf. Anstatt einen einzelnen "Stream" aus den Aktionen von mehreren Benutzern zu generieren werden mehrere simultane Ströme zugelassen. Divergenz tritt dann auf, sobald unterschiedliche Datensichten bei den Clients vorliegen. Dies kann beispielsweise durch in Konflikt tretende Aktionen auf den Daten geschehen. Das Beseitigen der Divergenz erfolgt anschließend an bestimmten Synchronisationspunkten, die individuell festgelegt werden können. So ist es möglich, sowohl asynchrones Arbeiten mit langen Synchronisationsintervallen als auch synchrones Arbeiten mit häufiger Divergenzbeseitigung zu modellieren.

Ströme, Aktionen auf den Daten und Synchronisationsbedingungen bilden die Elemente für ein Metaobjekt Protokoll mit dem der Anwendungsentwickler letztlich die Toolkit-Infrastruktur an verschiedene, gewünschte Szenarien anpassen kann.

Einschränkungen der Divergenz Beim Prinzip der Divergenz und Synchronisation handelt es sich in erster Linie um eine sogenannte "optimistische Strategie" [9]. Normalerweise wird angenommen, dass es nicht zu Konflikten kommt - falls doch, wird anschließend versucht, diese zu

beseitigen. Bei den üblichen Locking-Strategien bei denen es darum geht - wie im vorhergehenden Abschnitt besprochen - Inkonsistenzen zu vermeiden, handelt es sich um "pessimistische Strategien". Sie verhindern aber im Gegenzug bestimmte Arten der Zusammenarbeit über CSCW-Systeme, wie das flexible gemeinsame Bearbeiten eines Dokuments. Um allerdings in manchen Situation (wie Datenbanken mit wichtigen Finanzdaten) trotzdem garantieren zu können, dass es nicht zu Inkonsistenzen kommt (da diese nicht immer beseitigt werden können), muss das Toolkit dafür explizite Mechanismen zur Konsistenzkontrolle bereitstellen [9]. Hierbei gibt es eine Palette von mehreren einsetzbaren Techniken.

Es existiert beispielsweise die Möglichkeit, mit variabler Konsistenz zu arbeiten [9]. Dabei muss zwischen syntaktischer und semantischer Konsistenz unterschieden werden. Die letztere ist nötig, damit eine Anwendung fehlerfrei Daten verarbeiten kann. Bei der syntaktischen Konsistenz ist es möglich, schwächere Anforderungen zu stellen. Wenn zwei Autoren denselben Text bearbeiten, können größere auftretende Divergenzen nicht mehr ohne weiteres beseitigt werden. Dazu kommt es beispielsweise, wenn ganze Paragraphen geändert werden. Ein Ansatz zur Beseitigung ist das Verfügbarmachen beider Versionen für die Autoren, die sich dann erst später auf die gewünschte Version einigen können.

Auch für die traditionellen Locking-Verfahren bietet Prospero eine Lösung an. Es wird dafür mit Versprechen und Garantien gearbeitet [9]. Der Begriff Garantie rührt daher, dass man beim Locking davon ausgehen kann, dass eine Systemkomponente die Garantie über den alleinigen Zugriff auf Daten im Austausch gegen ein Versprechen über ihr zukünftiges Verhalten erhält. In Anlehnung an die Stärke und Granularität von Sperrungen kann man analog auch von verschiedenen Stufen der Garantien sprechen - nicht alle Verarbeitungsprozesse erfordern einen alleinigen Datenzugriff im selben Maße. Auch die korrespondierenden Versprechen über zukünftiges Verhalten können in ihren Eigenschaften variieren. Das eröffnet die Möglichkeit, von Systemkomponenten im Ausgleich für eine höherstufige Garantie ebenso rigorose Versprechen einzufordern. Als Konsequenz könnte man natürlich hier ebenfalls wieder eine Einschränkung der Toolkit-Flexibilität sehen - besonders was das opportunistische Arbeiten angeht, wenn sich zu anfang festgelegte Vorgehensweisen im weiteren Verlauf unvorhergesehen ändern. Prospero begegnet diesem Problem damit, dass es einer Anwendungskomponente auch gestattet ist, ein Versprechen wieder zu brechen. Allerdings führt das wiederum dazu, dass eventuell gegebene Garantien im Gegenzug nicht mehr eingehalten werden müssen. So können auch unvorhergesehene Aktionen mithilfe entsprechender Toolkit-

Optionen kompensiert werden. Es ist aber genauso für einen Entwickler möglich, auf den Mechanismus der Garantien und Versprechen zu verzichten. Dies kann zum Beispiel in Applikationen mit hohen Sicherheitsvorschriften von Vorteil sein. Prospero kann also sowohl für Anwendungen mit maximaler Flexibilität, als auch für Situation mit traditioneller, "pessimistischer" Inkonsistenzvermeidung eingesetzt werden.

Mit Hilfe der soeben vorgestellten Methoden kann erfolgreich das Auftreten einer nicht mehr behebbaren Inkonsistenz verhindert werden. Es kann also einerseits eine maximale Flexibilität, was die Datendistribution betrifft, gewonnen werden, andererseits sind ebenso Szenarien realisierbar, die bisher mit traditionellen Lockingverfahren gelöst wurden. Für eine genauere technische Analyse sowie Implementierungsdetails sei auf die am Ende dieser Arbeit aufgeführte Literatur verwiesen.

5 Zusammenfassung und Ausblick

Ausgehend von allgemeinen Konzepten wurden zuerst die grundlegenden Anforderungen von CSCW-Systemen selbst, die für eine erfolgreiche Entwicklung nötig sind, betrachtet. Daraufhin folgte eine Beschreibung der für CSCW-Systeme besonders wichtigen, verteilten Benutzerschnittstellen. Hierbei wurde eine genaueren Analyse der in diesem Spezialfall zusätzlich anfallenden Bedingungen durchgeführt. So stellen verteilte Benutzerschnittstellen, was die Bewegung des Nutzers zwischen einem Netzwerk aus Geräten betrifft, besonders hohe Anforderungen. Die Eigenheiten der unterschiedlichen Komponenten müssen schon beim Design der Schnittstellen beachtet werden. Dabei hat sich gezeigt, dass viele Überlegungen während der Entwicklung die spätere Funktionalität einer Anwendung erheblich beeinflussen. Mit diesem Problem hat sich insbesondere Paul Dourish eingehend befasst. Seine Überlegungen zu flexiblen Toolkits stellen ein wichtiges Konzept im Forschungsbereich dar. Aus den Erläuterungen zu verschiedenen, bisherigen Lösungen für verteilte Benutzerschnittstellen wurde deutlich, wie notwendig Ansätze sind, die nicht von vornherein auf spezifische Anwendungsbereiche festgelegt sind. Da die Bedeutung von Ubiquitous Computing Umgebungen in naher Zukunft sicherlich noch steigen wird, sollte eine Zielsetzung sein, die Forschung in diesem Bereich näher zusammenzubringen. Flexible Toolkits, die eine solide Grundlage bilden, könnten ein erster Schritt sein. Mit ihnen wäre es dann auch möglich, wiederum Komponenten für speziellere Situationen zu schaffen.

Literatur

1. Brinck, T.: Tom's csw and groupware page. 13.jun. 2006.
<http://www.infres.enst.fr/~vercken/multicast/csw.html>
2. Dourish, P.: Open implementation and flexibility in CSCW toolkits. PhD thesis, University College London (1996)
3. Reinhard, W., Schweitzer, J., Völksen, G., Weber, M.: Csw tools: Concepts and architectures. *Computer* **27** (1994) 28–36
4. Vandervelpen, C., Coninx, K.: Towards model-based design support for distributed user interfaces. In: *NordiCHI '04: Proceedings of the third Nordic conference on Human-computer interaction*, New York, NY, USA, ACM Press (2004) 61–70
5. Braun, E., Mühlhäuser, M.: Interacting with federated devices. In: *Advances in Pervasive Computing: Adjunct Proceedings of the Third International Conference on Pervasive Computing*, Vol. 191, Austrian Computer Society (OCG) (2005) 153–160
6. Coutaz, J., Balme, L., Lachenal, C., Barralon, N.: Software infrastructure for distributed migratable user interfaces. In: *UbiHCISys 2003*. (2003)
7. Bandelloni, R., Paternò, F.: Flexible interface migration. In: *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, New York, NY, USA, ACM Press (2004) 148–155
8. Luyten, K., Coninx, K.: Distributed user interface elements to support smart interaction spaces. *ism* **0** (2005) 277–286
9. Dourish, P.: Using metalevel techniques in a flexible toolkit for csw applications. *ACM Trans. Comput.-Hum. Interact.* **5** (1998) 109–155

Collaborative Filtering

Daniel Hilkert

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
hilkert@cip.ifi.lmu.de

Zusammenfassung Der vorliegende Artikel beschreibt zunächst allgemeine Eigenschaften und Ausprägungen von Empfehlungssystemen und geht dann detailliert auf das „collaborative filtering“ ein. Neben der Darstellung der genauen Funktionsweise wird auch die technische Umsetzung bei Amazon analysiert.

1 Einleitung

E-Commerce und elektronischer Versandhandel erzielen jedes Jahr enorme Zuwachsraten [11]. Für die Kunden wird es zunehmend schwieriger, aus dem steigenden Angebot die für sie interessantesten Artikel auszuwählen.

Im Gegenzug sehen sich auch die Anbieter von Einkaufsplattformen im Internet einer immer größeren Konkurrenz ausgesetzt. Dabei entsteht eine hohe Markttransparenz, d.h. Kunden haben die Möglichkeit mit geringem Aufwand viele Angebote zu vergleichen. Um in diesem schwierigen Marktumfeld dennoch erfolgreich sein zu können, sind innovative Konzepte zur Kundengewinnung und Bindung erforderlich.

Der Trend geht dabei zur Massen-Personalisierung. Internetseiten großer E-Commerce Portale wie Amazon oder E-Bay werden auf die Bedürfnisse jedes einzelnen Kunden angepasst. Ein zentrales Element solcher individuellen Angebote sind Empfehlungssysteme.

In der vorliegenden Arbeit wird daher zunächst auf Empfehlungssysteme im Allgemeinen eingegangen. Es soll erläutert werden, warum und wo Empfehlungssysteme eingesetzt werden, welche Daten dabei verwendet werden und welche Ausprägungen solcher Systeme existieren.

Im Anschluss erfolgt eine detaillierte Betrachtung des „collaborative filtering“¹ als eine spezielle Form von Empfehlungssystemen. Es werden die historischen Ansätze aufgezeigt und es erfolgt eine Klassifizierung der unterschiedlichen Formen des CF. Auf die Darstellung der allgemeinen Funktionsweise folgt eine detaillierte Betrachtung eines speziellen Algorithmus. Vor- und Nachteile des Ansatzes werden diskutiert und es erfolgt ein Vergleich mit „content based filtering“. Den Abschluss

¹ Die Abkürzung CF sowie die deutsche Übersetzung kollaboratives Filtern sind im Folgenden als äquivalente Begriffe zu „collaborative filtering“ zu sehen.

dieses Kapitels bildet eine Betrachtung möglicher Kenngrößen um die Leistungsfähigkeit der Methode zu evaluieren.

Abschließend wird der Webshop Amazon mit Blick auf die verwendeten Empfehlungssysteme betrachtet. Dabei wird auch auf die speziellen Algorithmen eingegangen, die notwendig sind um trotz der großen Anzahl an Kunden und Produkten relevante Empfehlungen in Echtzeit geben zu können.

2 Empfehlungssysteme (Recommendation Systems)

Collaborative Filtering stellt einen Mechanismus zur Auswahl von zu empfehlenden Elementen dar. Er ist also i.A. Teil eines Empfehlungssystems (Recommendation System). Daher soll im folgenden Abschnitt kurz erläutert werden, was Empfehlungssysteme sind, warum sie verwendet werden und welche Möglichkeiten es gibt, solche Systeme zu realisieren.

2.1 Bedarf und Einsatz von Empfehlungssystemen

In einer Meldung vom November 2005 schreibt der Bundesverband des deutschen Versandhandels e.V., dass durch elektronischen Handel in Deutschland mittlerweile 6,1 Milliarden Euro pro Jahr umgesetzt werden. Dies sind 24% mehr als noch im Jahr zuvor [11]. Der Markt für E-Commerce hat damit bereits eine enorme wirtschaftliche Bedeutung entwickelt.

Die Ansprüche der Kunden steigen dabei weiter. In Anlehnung an die Bedürfnispyramide nach Maslow [9] stellt Selbstverwirklichung die höchste Stufe des menschlichen Strebens dar. Übertragen auf elektronischen Handel bedeutet dies, dass immer mehr Kunden sich möglichst individuelle Produkte oder Produkt-Bündel wünschen. Diese sollen dabei jedoch zum Preis eines Massenproduktes angeboten werden.

Für die Betreiber entsteht die Notwendigkeit die Zahlungsbereitschaft der Kunden maximal auszunutzen. Eine solche Preisdifferenzierung 1. Grades (bei der die individuelle Zahlungsbereitschaft ausgenutzt wird) kann dabei nur durch umfassende Informationen über die jeweiligen Kunden und deren Kaufgewohnheiten umgesetzt werden [10].

Die logische Folge dieser Anforderungen ist, dass für jeden Kunden automatisch eine personalisierte Oberfläche des Shops zur Verfügung gestellt wird. Ein bekanntes Zitat von Jeff Bezos dem Gründer und CEO von Amazon verdeutlicht diese Aussage [6]:

“If I have 3 million customers on the Web, I should have 3 million stores on the Web.”

E-Commerce stellt somit die wichtigste Anwendung für Empfehlungssysteme dar. Die automatisierte Analyse und Auswertung von allen Aktivitäten eines Kunden innerhalb eines Shops ermöglicht dabei persönliche Empfehlungen mit geringem individuellem Aufwand zu erstellen.

Die Vorteile sind sowohl für Anbieter als auch für Kunden offensichtlich. Für Anbieter erhöht sich die Chance aus Interessenten Käufer zu machen und für Kunden reduzieren sich die Suchkosten, da ihnen schneller die gewünschten Artikel präsentiert werden können.

2.2 Aufstellungen der Input-Faktoren

Um die Funktionsweise solcher Empfehlungssysteme näher zu beschreiben sollen hier zunächst die verwendeten Daten betrachtet werden.

Grundsätzlich muss zwischen Daten aller Nutzer und Daten des aktiven Nutzers unterschieden werden. Um nicht nur allgemeine Empfehlungen geben zu können sondern eine Personalisierung im eigentlichen Wortsinn zu erreichen, kommt dabei letzteren eine besondere Bedeutung zu.

Ist dem Nutzer bewusst, dass seine Angaben zur Erstellung einer Empfehlung führen, sind diese explizit. Beispiele dafür sind Begriffe, die in ein Suchfeld eingegeben werden, oder die Auswahl bestimmter Attribute bei der Erstellung eines persönlichen Profils. Eine weitere Möglichkeit besteht darin Bewertungen des Nutzers zu verwenden. In einer Abfrage können dabei entweder bereits gekaufte oder repräsentativ ausgewählte Artikel vorgestellt und durch den Nutzer bewertet werden. Diese Methode ermöglicht den Geschmack des Kunden relativ genau zu erkennen. Mancher Kunde wird jedoch vor dem verhältnismäßig hohen Aufwand zurückschrecken, zudem bestehen nach wie vor erhebliche Bedenken der Verbraucher zu viele persönliche Informationen preiszugeben [5].

Alternativ können die Daten auch implizit, also ohne dass sich der Nutzer darüber bewusst ist, gewonnen werden. Aufschluss über die Präferenzen kann z.B. der Verlauf der betrachteten Elemente innerhalb des Shops liefern („Click-Stream-Analyse“ [8]). Rückschlüsse lassen sich auch aus den bereits gekauften Artikeln ziehen.

Um aus den Angaben des Nutzers sinnvolle Empfehlungen zu erzeugen, müssen diese mit den Angaben aller Nutzer verglichen werden.

Hierfür können zunächst globale Verkaufsstatistiken wie Bestsellerlisten herangezogen werden. Diese dienen jedoch meist nur dazu den „Mainstream“ Geschmack abzubilden und eignen sich daher nicht, persönlich abgestimmte Angebote zu erstellen.

Besser geeignet sind dazu von Nutzern verfasste Kommentare zu bestimmten Artikeln die dann auf den Produktseiten angezeigt werden.

Zudem können die Informationen des aktiven Nutzers, wie Verkaufshistorie und Bewertungen, mit denen anderer Nutzer verglichen werden. Bei entsprechenden Überschneidungen kann dann auf einen ähnlichen Geschmack geschlossen werden.

Folgende Grafik stellt die aufgezeigten „Input Faktoren“ in einer Übersicht zusammen:

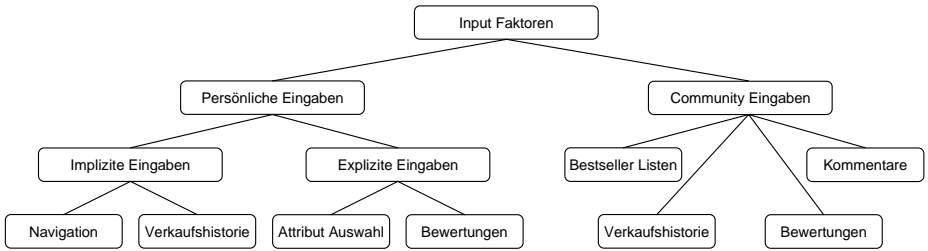


Abbildung 1. Übersicht der Inputfaktoren für Empfehlungssysteme.

2.3 Formen von Empfehlungssystemen

Eine Möglichkeit ein solches Empfehlungssystem zu implementieren ist „collaborative filtering“. Daneben existieren jedoch noch zahlreiche andere und teilweise wesentlich einfachere Methoden, die hier kurz vorgestellt werden ².

Die trivialste Form eines solchen Systems stellt eine simple **Suchfunktion** dar. Diese kann zwar im technischen Sinn nicht als Empfehlungssystem gewertet werden, Nutzer nehmen sie jedoch als solche wahr, denn das Ergebnis einer Suche stellt eine auf den aktuellen Geschmack (Suchbegriff) abgestimmte Liste von Empfehlungen dar. Da aber die wenigsten Suchfunktionen semantische Verknüpfungen erkennen und darstellen, ist das Ergebnis der „Empfehlung“ zumeist nicht besonders hilfreich.

Eine weitere einfache Möglichkeit zur Umsetzung sind **manuell erstellte Listen** von Empfehlungen. Diese können von Experten oder Nutzern zusammengefügt und mit einer Beschreibung versehen werden. Alle Nutzer haben so die Möglichkeit, sich speziell auf ihre Bedürfnisse zugeschnittene Listen herauszusuchen. Ein Beispiel für eine solche Liste, könnte zum Beispiel die Zusammenstellung „Informatikbücher fürs Grundstudium“ sein, die von einer Fachschaft in einem Online-Buchversand hinterlegt wird.

Sollen in das Ergebnis der Empfehlungen die Meinung aller Nutzer eingehen, so stellen **statistische Zusammenfassungen** das einfachste Mittel dar. Durch einfache Hitlisten bzw. Charts kann sich der Nutzer einen schnellen Überblick verschaffen, was aktuell besonders gefragt ist.

Besteht die Möglichkeit Profile der jeweiligen Nutzer anzulegen, kann eine **Attribut-basierte Empfehlung** erfolgen. Ist beispielsweise bekannt, dass ein bestimmter Nutzer bestimmte Genres, Artisten, Preisklassen, etc. bevorzugt, kann anhand dieser Informationen ein individuelles Angebot erstellt werden.

Bessere Ergebnisse können durch eine „item-to-item“ Korrelation erreicht werden. Bei der auch als „**content based filtering**“ bekannten Methode werden Artikel gefunden, die besonders häufig in Assoziation mit dem aktuell ausgewählten Artikel gebracht werden können. In komplexeren Systemen können dabei nicht nur Verknüpfungen zu einzelnen Elementen hergestellt werden. Es kann auch die semantische Nähe von mehreren Artikeln, wie z.B. denen der Einkaufshistorie oder dem Inhalt des Warenkorbs eines Nutzers, zu anderen Artikeln dargestellt werden. In

² Die folgende Aufstellung der Möglichkeiten der Umsetzung von Empfehlungssystemen entstammt den Quellen [6], [12] und [13]

vielen Shops wird so z.B. zu einer Digitalkamera automatisch ein passendes Speichermodul angeboten. Diese Herangehensweise führt einerseits zu einem hohen Aufwand, ermöglicht aber andererseits wesentlich bessere und vielfältigere Empfehlungen als die zuvor beschriebenen Methoden.

Nicht nur die Verbindung zwischen Elementen, sondern auch die Beziehung von Nutzern zueinander kann zur Erstellung von Empfehlungen genutzt werden. In der als „user-to-user“ Korrelation oder „**collaborative filtering**“ bekannten Methode wird durch verschiedene Attribute eine Nähe zwischen Benutzern festgestellt. In die Empfehlung gehen dann Artikel ein, die besonders ähnliche Kunden hoch bewertet bzw. gekauft haben. Häufige Beschreibung des Ergebnisses dieser Methode sind dabei der Art „*andere Kunden die ... gekauft haben, haben auch ... gekauft*“. Im Rahmen des gestellten Themas wird die Funktionsweise dieser Methode im weiteren Verlauf noch genauer erläutert werden.

3 Collaborative Filtering

Als eine der am häufigsten verwendeten Methoden um Empfehlungssysteme zu realisieren soll das „collaborative filtering“ im Folgenden näher erläutert werden.

3.1 Geschichte des CF³

Der Ausdruck „collaborative filtering“ wurde erstmals 1992 von Goldberg et Al erwähnt [17]. Das Xerox Palo Alto Research Center sah sich einer ansteigenden Informationsflut durch E-Mails gegenüber, da Rundmails grundsätzlich an alle Mitarbeiter geschickt wurden. Die in der Folge eingerichteten Mailinglisten für verschiedene Funktionsbereiche bzw. Interessensgebiete konnten jedoch nicht sicherstellen dass jeder Mitarbeiter alle relevanten E-Mails erhält. Goldberg entwickelte daher zusammen mit seinem Team das so genannte „Tapestry“ System. Dadurch konnte jeder Mitarbeiter nach dem lesen einer E-Mail diese entsprechend ihrer Relevanz bewerten. Andere Mitarbeiter hatten dann die Möglichkeit eigene Filter zu definieren, durch die ihnen nur E-Mails zugestellt wurden, die von ihnen bekannten Mitarbeitern als relevant eingestuft wurden. Somit konnten sie sich gegenseitig bei der Informationsfilterung unterstützen. Diese Methode wurde folglich in dem von Goldberg veröffentlichten Artikel als „collaborative filtering“ bezeichnet.

Ein weiterer historisch wichtiger Schritt war die „SIGCHI“ Konferenz 1995, in deren Rahmen weitere wichtige Artikel veröffentlicht wurden. Im Paper von Maltz und Ehrlich wurden dabei vor allem die Begriffe des aktiven und passiven „collaborative filtering“ definiert [18]. Bei der aktiven Kollaboration werden die abgegebenen Empfehlungen direkt an alle Nutzer gesendet, in der passiven Variante werden sie gespeichert und stehen dem Nutzer „on Demand“ zur Verfügung.

Eine 1997 erschienene Ausgabe von „Communications of the ACM“ enthielt vornehmlich Artikel über Empfehlungssystemen und „collaborative filtering“. Hier wurde unter anderem das in der Literatur bekannte System „GroupLens“ (Konstan et

³ Die Auswahl der historischen Beispiele ist den Beiträgen [5] und [8] entnommen.

AI) vorgestellt [19]. Analog zu den Problemen die zur Entwicklung von „Tapestry“ geführt hatten, sollten relevante UseNet Nachrichten gefiltert werden. Dies wurde dabei ebenfalls durch ein Bewertungssystem und einen „collaborative filtering“ Algorithmus gelöst. Die genaue Funktionsweise der verwendeten Algorithmen dieses Systems wird im Folgenden noch näher erläutert werden.

3.2 Einteilung in verschiedene Klassen

Grundsätzlich kann zwischen verschiedenen Klassen von „collaborative filtering“ unterschieden werden, primär erfolgt die Einteilung in aktives und automatisiertes CF [8].

Das historische System „Tapestry“ bildet hierbei ein klassisches Beispiel für aktives Filtern, da die Auswahl der Nutzer mit ähnlichen Präferenzen manuell erfolgt. Konkret müssen Nutzer diejenigen Mitarbeiter selbst auswählen, von denen Sie glauben E-Mails mit ähnlicher Relevanz zu bewerten wie Sie selbst.

Werden diese relevanten Nutzer anhand mathematischer Modelle durch das System ausgewählt, so spricht man von „automatic collaborative filtering“. Im Gegensatz zum aktiven Filtern müssen sich die Nutzer nicht gegenseitig kennen. Diese Methode kann also auch für Systeme mit einer großen Anzahl von Nutzern eingesetzt werden. Für das eingangs beschriebene Aufgabenfeld des E-Commerce ist dies somit die relevante Variante.

Aktive System werden nochmals in zwei Kategorien unterteilt [20]. Beim „memory based automatic collaborative filtering“ werden die Daten aller Nutzer direkt dazu verwendet die Bewertungen eines Nutzers für einen bestimmten Artikel vorherzusagen.

Beim „model based automatic collaborative filtering“ hingegen werden die Daten dazu genutzt, Modelle zu erzeugen. Diese werden in der Folge ebenfalls genutzt um Vorhersagen bezüglich des Geschmacks des Nutzers zu machen. Breese et Al (1998) schlägt dabei zwei verschiedene modellbasierte Algorithmen vor. Einerseits kann eine Zusammenfassung von Nutzern mit ähnlichem Geschmack zu Clustern vorgenommen werden. Andererseits können bayessche Netze [24] erstellt werden, die dann das schließen bzw. vorhersagen unbekannter Werte erlauben.

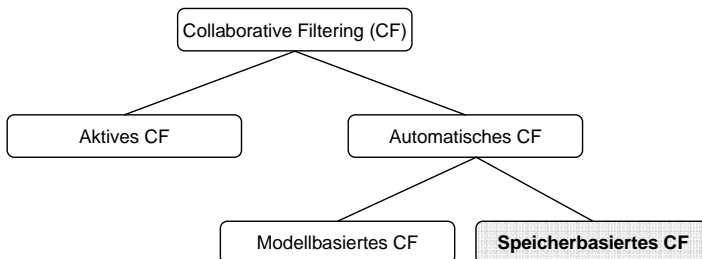


Abbildung 2. Einteilung der verschiedenen Klassen von „collaborative filtering“ [8]

3.3 Allgemeine Funktionsweise

Die Erstellung einer Empfehlung erfolgt in vier Schritten [8].

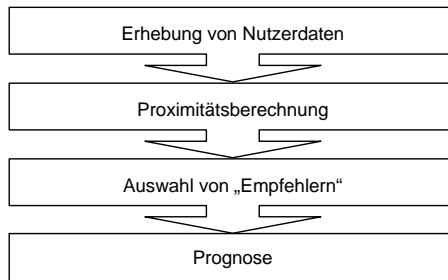


Abbildung 3. Schritte einer Empfehlung [8]

Daten, welche die Vorlieben der Nutzer repräsentieren, werden zunächst gesammelt und in Profilen gespeichert. Durch einen “collaborative filtering” Algorithmus wird dann die Ähnlichkeit zwischen den Profilen des aktiven Nutzers und denen aller anderen Nutzer erhoben [14]. Diejenigen Nutzer mit einer besonders hohen Proximität werden dann als so genannte „Empfehler“ ausgewählt. Im letzten Schritt werden Artikel ausgewählt die von den „Empfehlern“ besonders hoch bewertet wurden und die der aktive Nutzer noch nicht kennt bzw. gekauft hat.

3.4 Algorithmen

Ein in der Literatur häufig zitierter Ansatz um die relevante Nutzer zu identifizieren und daraus Empfehlungen abzuleiten, wurde 1994 von Resnick et al veröffentlicht.

Bei dem Vergleich verschiedener Nutzerprofile wird davon ausgegangen, dass alle Bewertungen, die von Nutzern abgegeben wurden (in einer Datenbank) gespeichert sind [14]. Zwei verschiedene **Nutzerprofile A** und **B** repräsentieren dann jeweils einen n-dimensionalen Vektor, wobei der Wert A_i der **Bewertung des i-ten Artikels** von Nutzer A entspricht. Die Dimension **n** wird durch die absolute **Anzahl aller verfügbaren Artikel** bestimmt. Liegt für einen Artikel j keine Bewertung eines Nutzers vor (was z.B. in einem Web-Shop häufig der Fall sein wird) so ist der Wert $A_j = 0$.

Um nun einen Vergleich dieser Vektoren vorzunehmen, kann der **Korrelationskoeffizient $Kor(A,B)$** nach Pearson eingesetzt werden [15]. Der Wertebereich von $Kor(A,B)$ ist dabei $[-1, 1]$ wobei 1 einem vollständig positiven und -1 einem vollständig negativen Zusammenhang entspricht.

$Kor(A,B)$ ist definiert durch:

$$Kor(A, B) := \frac{\sum_{i=1}^n (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2 \sum_{i=1}^n (B_i - \bar{B})^2}}$$

Abbildung 4. Korrelationskoeffizient nach Pearson.

Dabei entsprechen \bar{A} und \bar{B} dem empirischen Erwartungswert, werden also durch folgende Formel bestimmt (analog dazu \bar{B}):

$$\bar{A} := \frac{1}{n} \sum_{i=1}^n A_i$$

Abbildung 5. Empirischer Erwartungswert

Auf Basis der so errechneten Nähe zweier Profile kann nun eine Wertung A_x des Nutzers A für den Artikel X vorhergesagt werden, indem die Bewertungen von X aller m anderen Nutzer J_m mit dem entsprechenden Korrelationskoeffizienten ($Kor(A, J_i)$) gewichtet werden [16]:

$$A_x := \bar{A} + \frac{\sum_{i=1}^m (J_{ix} - \bar{J}) Kor(A, J_i)}{\sum_{i=1}^m |Kor(A, J_i)|}$$

Abbildung 6. Vorhergesagte Bewertung des Artikels x für Nutzer A

3.5 Vor- und Nachteile des Ansatzes

Die Vorteile dieses Ansatzes liegen auf der Hand. Im Gegensatz zu vielen anderen Systemen können mit CF Inhalte beliebiger Art, wie z.B. auch Filme oder Restaurants, empfohlen werden. Dies ist dabei auch möglich, wenn keine textbasierten Beschreibungen vorliegen, da Empfehlungen auf Basis der „Erfahrungen“ anderer Nutzer erstellt werden. Da nicht die inhaltliche Nähe sondern die Nähe verschiedener Nutzer im Vordergrund steht, können auch vermeintlich „unerwartete“ Elemente empfohlen werden.

Dieser Ansatz weist jedoch trotz seiner guten Performance einige Schwächen auf [14]. Die Korrelation zweier Nutzer kann nur auf Basis von Artikeln festgestellt werden, die von beiden bewertet wurden. Existiert also eine große Anzahl von

Artikeln sinkt die Wahrscheinlichkeit für eine Überlappung. Es können nur wenige ähnliche Nutzer gefunden werden, deren Bewertungen in die Vorhersage eingehen können.

Existieren z.B. 3 Nutzer A, B und C, so kann sowohl zwischen A und B als auch zwischen B und C eine absolut positive Korrelation ($Kor(A,B) = 1$ und $Kor(B,C) = 1$) existieren. Dieser Fall würde implizieren, dass auch zwischen A und C eine positive Korrelation existiert. Wird jedoch der vorgestellte Algorithmus verwendet, so ist dies nicht zwingend erforderlich. Hat nämlich B eine sehr hohe Anzahl von Artikeln bewertet, so dass diese Menge die Mengen der Bewertungen von A und C vollständig einschließt, jedoch keine Schnittmenge zwischen den Bewertungen von A und C existiert, wäre die Korrelation von A und C dennoch gleich 0 ($Kor(A,C) = 0$). Dieses, wenn auch konstruierte Beispiel zeigt, dass Bewertungen von Nutzern „zu unrecht“ niedriger gewichtet werden können, und somit relevante Informationen verloren gehen.

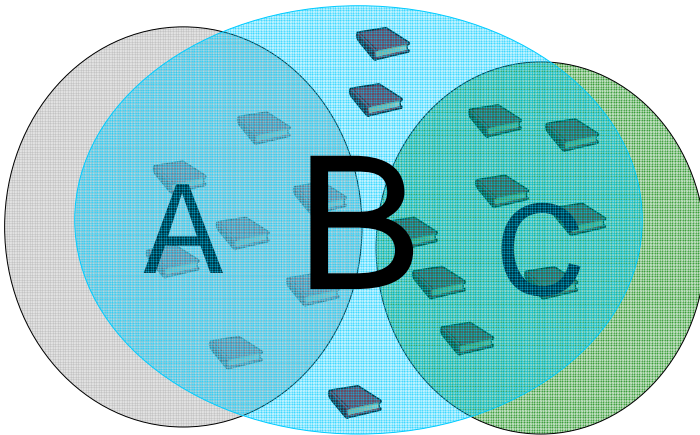


Abbildung 7. Jeweils absolut positive Korrelation zwischen A und B sowie zwischen B und C führt nicht zwingend zu einer Korrelation zwischen A und C.

Insbesondere in Bereichen, in denen häufig neue Elemente zur Auswahl stehen, eignen sich „collaborative filtering“ nur sehr schlecht [1]. Da die neuen Artikel noch nicht bewertet wurden, können sie in den Empfehlungen noch nicht auftauchen. Sie werden also schwerer gefunden was zu einem Kreislauf führt. Beispiele für solche Bereiche mit schnell wechselnden Angeboten sind Webseiten oder Newsfeeds.

3.6 Vergleich mit Content-Based-Filtering

Bei der Vorstellung verschiedener Empfehlungssysteme im Abschnitt 2.3 wurden bereits das „content based filtering“⁴ erwähnt. In vielen Systemen ist CBF jedoch nur

⁴ Die Abkürzung CBF und die Begriffe „item-to-item collaboration“ sowie inhaltsbasierte Empfehlungen sind im Folgenden als äquivalente Begriffe zu „content based filtering“ zu sehen.

ein Teil des Konzepts, daher soll zunächst abgegrenzt werden welche Eigenschaften ein „reines“ CBF System besitzt.

Inhaltsbasierte Empfehlungen basieren auf Profilen die ausschließlich durch die Analyse des Inhalts entstehen, die der Nutzer in der Vergangenheit bewertet hat [2], wie es z.B. bei dem System NewsWeeder [21] der Fall war. Diese Systeme weisen jedoch einige Schwächen auf [2].

Zunächst beschränkt sich die Analyse eines Elements im Normalfall auf textbasierte Inhalte, was es erheblich erschwert multimediale Inhalte wie Musik oder Filme in Relation zueinander zu setzen.

Ein zweites Problem stellt die „Überspezialisierung“ dar. Es werden nur Elemente in die Empfehlung aufgenommen, die denen die der Nutzer bereits kennt, ähnlich sind. Diesem Problem wird zumeist nur durch das hinzufügen zufällig ausgewählter Elemente begegnet.

Der kollaborative Ansatz unterscheidet sich dabei erheblich von CBF. Wie oben ausführlich beschrieben wurde, werden Bewertungen für Artikel die der Nutzer noch nicht kennt vorhergesagt, indem die Bewertungen ähnliche Nutzer herangezogen werden.

Wichtig ist jedoch, dass (automatisches) kollaboratives Filtern und inhaltsbasiertes Filtern nicht notwendigerweise in Konkurrenz zueinander stehen [3]. Das von Balabanovic vorgestellte hybride System „Fab“ verbindet die Vorteile beider Ansätze [1].

Dieses System sammelt dabei Websites durch so genannte „Collection Agents“ in Bereichen die bei den Nutzern aktuell im Trend sind [2]. Durch einen Abgleich mit den Nutzerprofilen werden diese bei Übereinstimmung dann direkt passenden Nutzern zugeordnet. Zusätzlich werden Webseiten empfohlen, die von anderen „ähnlichen“ Nutzern hoch bewertet wurden. Somit können Webseiten auf Basis der Erfahrung anderer bewertet werden, andererseits werden auch neue, bisher unbewertete Seiten in den Kreislauf einbezogen. Zudem kann so auch ein Bezug zwischen Nutzern hergestellt werden, die noch kein Element gemeinsam bewertet haben; es genügt wenn sie ähnliche Elemente bewertet haben.

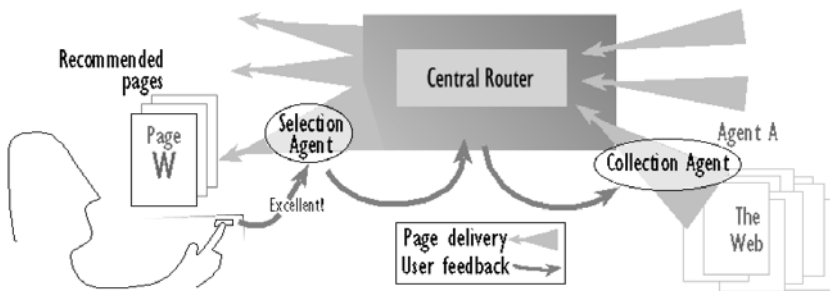


Abbildung 8. Darstellung der Architektur des „FAB“ Systems [2]

3.7 Leistungsmessung

Da aktuell verwendete Systeme vor allem im Bereich des E-Commerce häufig eine Kombination verschiedener Ansätze verwenden (siehe dazu z.B. Abschnitt 3.6) fehlt im Rahmen dieser Arbeit die Grundlage für eine konkrete Performancemessung eines bestimmten Algorithmus. Daher wird an dieser Stelle das Konzept der „Receiver Operating Characteristic“ (ROC) Kurve hergeleitet, das als Maß für die Qualität eines Filters verwendet werden kann.

Will man die Performance von Empfehlungssystemen betrachten, muss man sich zunächst vor Augen führen, in welcher Form Ergebnisse ausgegeben werden können.

Durch CF soll primär die „binäre“ Entscheidung getroffen werden ob ein Element im aktuellen Kontext relevant ist, oder nicht. Eine Anzahl von insgesamt N Elementen kann also einerseits in tatsächlich relevante und nicht relevante Elemente und andererseits in durch CF ausgewählte bzw. nicht ausgewählte Elemente eingeteilt werden. Einen Überblick über alle verwendeten Kenngrößen soll folgende Matrix geben (Die verwendeten Variablen werden noch erläutert werden):

	relevant	irrelevant	
ausgewählt	N_{RS}	$N_S - N_{RS}$	N_S
nicht ausgewählt	$N_R - N_{RS}$	$N - N_R - N_S + N_{RS}$	$N - N_S$
	N_R	$N - N_R$	N

Abbildung 9. Klassifikation (in Anlehnung an [8])

Aus dieser Einteilung lassen sich wichtige Kenngrößen ableiten. Der Wert der Präzision (P) beschreibt das Verhältnis zwischen der Anzahl tatsächlich relevanten ausgewählten Elemente (N_{RS}) und der Anzahl der insgesamt ausgewählten Elemente (N_S):

$$P := \frac{N_{RS}}{N_S}$$

Abbildung 10. Präzision [7]

“Recall” (R) oder auch Sensitivität ($SENS$) [20] ist definiert durch das Verhältnis der ausgewählten relevanten Elemente (N_{RS}) im Vergleich zu der Gesamtzahl der tatsächlich relevanten Elemente (N_R):

$$R = SENS = \frac{N_{RS}}{N_R}$$

Abbildung 11. Recall [7] (entspricht SENS [8])

Die Spezifität (*SPEC*) gibt die Wahrscheinlichkeit wieder, dass ein nicht relevantes Element auch nicht ausgewählt wird.

$$SPEC = 1 - \frac{N_S - N_{RS}}{N - N_R}$$

Abbildung 12. Spezifität (SPEC) [20]

Diese Kennzahlen berücksichtigen nur binäre Entscheidungen bzgl. der Relevanz. In realen Systemen wird jedoch häufig eine numerische Skala verwendet, so dass ein bestimmter Schwellenwert definiert werden muss. Dieses Problem ergibt sich nicht nur bei der Messung, sondern auch in den Systemen selbst. Auch diese müssen entscheiden, ab welchem Wert auf der Skala ein bestimmtes Objekt in die Auswahl aufgenommen wird.

Genau diese Entscheidung beeinflusst die Werte aller aufgezeigten Kenngrößen dabei erheblich. Wird der Schwellenwert niedrig angesetzt, so werden mehr der tatsächlich relevanten Elemente ausgewählt, die Sensitivität steigt also an. Gleichzeitig sinkt jedoch die Spezifität, da die Wahrscheinlichkeit sinkt ein irrelevantes Objekt nicht auszuwählen. Bei einem hohen Schwellenwert ist dieser Effekt genau entgegengesetzt, die Sensitivität sinkt und die Spezifität steigt.

Daher ist besonders die gegenseitige Beeinflussung dieser beiden Größen relevant. Zu diesem Zweck können daher die beiden Werte in einer so genannten „Receiver Operating Characteristic“ (ROC) Kurve gegeneinander angetragen werden [7]. Je größer dabei die Fläche unter dem Graphen ist, desto größer ist die Qualität der gemessenen Methode.

In folgendem Beispiel ist eine exemplarische sehr gute ROC Kurve (blauer Graph) gegen die Kurve eines Zufallsalgorithmus (orangene Diagonale) angetragen:

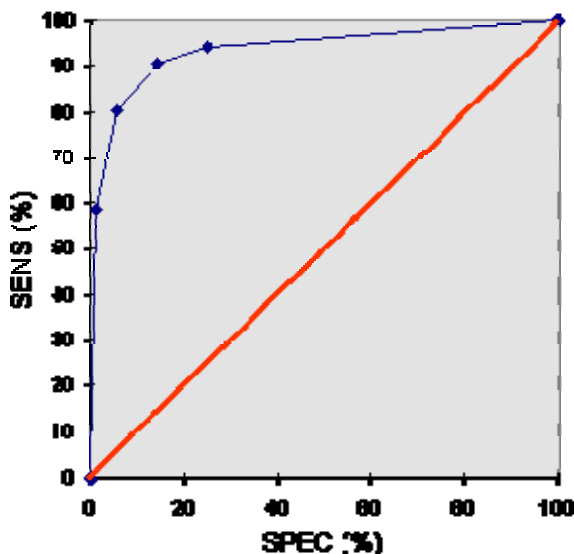


Abbildung 13. Beispiel einer ROC Kurve

Die Bewertung des Filters ist jedoch nicht die einzige Kenngröße. Evaluiert werden können zudem die Übereinstimmung der Reihenfolge der ausgewählten Ergebnisse oder allgemein die Güte der Vorhersage unbekannter Bewertungen des Nutzers.

4 Aktuelles Praxisbeispiel: Amazon

Das mit Sicherheit populärste Praxisbeispiel für die Verwendung von Empfehlungen ist Amazon. Die folgenden vier Screenshots sollen dabei die Bandbreite der individuell erstellten Empfehlungen bei amazon.de wiedergeben:

9.



Taschenbuch der Wirtschaftsinformatik.

von Georg Distler, und andere

Erscheinungsdatum: 1. Oktober 2002

Amazon-Preis: **EUR 10,00**

Gebraucht & neu ab EUR 10,00

In den Einkaufswagen

Auf meinen Wunschzettel

Gehört mir Kein Interesse Artikel bewerten

Diesen Artikel haben wir empfohlen, weil Sie u.a. **Taschenbuch Multimedia** gekauft haben.
(Ändern)

Unser Vorschlag

Kaufen Sie Taschenbuch der Wirtschaftsinformatik. und Einführung in die Wirtschaftsinformatik.



**Amazon-Preis: EUR
29,95**



Beide jetzt kaufen!

Kunden, die dieses Buch gekauft haben, haben auch diese Bücher gekauft:

- Taschenbuch Rechnernetze und Internet. von Erich Stein
- Taschenbuch Multimedia. von Peter A. Henning
- Einführung in die Wirtschaftsinformatik. von Peter Stahlmecht, Ulrich Hasenkamp
- Arbeitsbuch Wirtschaftsinformatik I. Lexikon, Aufgaben und Lösungen. von Hans Robert Hansen, Gustaf Neumann

► **Entdecken Sie verwandte Produkte**

Lieblingslisten

- Notwendige Bücher für das BWL Grundstudium: Eine Liste von wohebo, BWL Student
- Prozessmanagement: Eine Liste von mihofi, Simulant
- Gute IT- Bücher: Eine Liste von deltababy, Netzwerktechniker

Abbildung 14. Empfehlungen bei Amazon.de

Die Anforderungen an die technische Plattform sind dabei enorm. Eine Vielzahl von personalisierten Elementen in Kombination mit der extrem hohen Anzahl an Produkten und Kunden führt dazu, dass klassische CF Mechanismen nicht mehr den Effizienzansprüchen einer Echtzeitdarstellung der Portale genügen würden [4].

Im worst-case wäre eine Komplexität von $O(N*J)$ (N entspricht der Gesamtzahl der Produkte, J der Anzahl der Kunden) zu erwarten. Bei einer Zahl von mehr als 29 Mio. Kunden und mehreren Millionen Produkten (Zahlen von 2003) wäre dies selbst mit aktueller Hard- und Software nicht mehr zu bewältigen.

Amazon setzt daher auf eine Methode die als „item-to-item collaborative filtering“ bezeichnet wird. In einer Offline-Phase wird ein Katalog von ähnlichen Elementen erstellt indem diejenigen Artikel miteinander in Verbindung gebracht werden, die besonders häufig zusammen gekauft werden. Eine Vereinfachung des zugrunde liegende Algorithmus lässt sich wie folgt darstellen:

```

For each item in product catalog, I1
  For each customer C who purchased I1
    For each item I2 purchased by customer C
      Record that a customer purchased
      I1 and I2
    End For
  End For
End For

For each item I2
  Compute the similarity between I1 and I2
End For

```

Abbildung 15. Algorithmus zur Offline-Berechnung

Die Erstellung dieses Katalogs hat zwar eine worst-case Komplexität von $O(N^2 \cdot J)$, diese Operation muss jedoch nicht für jeden Kunden einzeln und in Echtzeit ausgeführt werden. Die Auswahl der zu empfehlenden Artikel erfolgt dann nur noch in jener Teilmenge der Artikel, die als ähnlich zu denen meines „Einkaufskorb“ identifiziert werden können.

Offline Berechnungen waren in den ursprünglichen Empfehlungssystemen nicht vorgesehen. Solche Ansätze sind jedoch für aktuelle E-Commerce Anwendungen unabdingbar, denn nur so können die enormen Datenmengen effizient bewältigt werden.

5 Zusammenfassung

Die Verbreitung des elektronischen Versandhandels in Deutschland wird weiter zunehmen. Da dies vor allem der Kosteneinsparung dient, kann nur noch selten eine persönliche Beratung erfolgen. Entsprechende Hotlines sind zwar zumeist eingerichtet, diese sind jedoch häufig chronisch überlastet oder es handelt sich um teure Mehrwertdienste.

Um den Kunden dennoch einen Wegweiser durch den „Produkt Dschungel“ bieten zu können, wird häufig auf Empfehlungssysteme zurückgegriffen. Besonders Produkte, deren Popularität auch in der realen Welt durch „Mund zu Mund Propaganda“ entsteht, wie dies beispielsweise bei Büchern oder Filmen der Fall ist, werden bevorzugt durch kollaborative Systeme empfohlen.

Die hohe Marktdurchsetzung zeigt dabei aber auch, dass die Technologie bereits sehr weit entwickelt ist. Sowohl die Kombination verschiedener Filtermechanismen als auch die notwendig gewordene Skalierbarkeit entsprechender Systeme sind weitgehend gelöste Probleme. Neue Forschungsansätze beziehen sich vor allem auf die Wahrung der Privatsphäre [22] oder die Verwendung neuer Input Faktoren wie z.B. Augenbewegungen [23]. Die kommerzielle Nutzung und die zugehörigen ökonomischen Auswirkungen werden primär in wissenschaftlichen Beiträgen aus der Betriebswirtschaft diskutiert.

Abschließend bleibt festzuhalten, dass Empfehlungssysteme und damit insbesondere auch „collaborative filtering“ ein wichtiges und etabliertes Werkzeug zur Filterung und Zusammenstellung von Informationen ist. Aus wissenschaftlicher Sicht sind hingegen keine großen Innovationen auf diesem Gebiet mehr zu erwarten.

Literatur

1. Marko Balabanovic. An Adaptive Web Page Recommendation Service. In: Proceedings of the first international conference on Autonomous agents, S. 378 - 385, ACM Press, 1997
2. Marko Balabanovic, Yoav Shoham. Content-Based, Collaborative Recommendation. In: Communications of the ACM, Vol. 40, No. 3, S. 66-72, ACM Press, 1997
3. Jonathan Herlocker, Joseph Konstan, John Riedl. Explaining Collaborative Filtering Recommendations. In: Proceedings of the 2000 ACM conference on Computer supported cooperative work, ACM Press, 2000
4. Greg Linden, Brent Smith, Jeremy York. Amazon.com recommendations: item-to-item collaborative filtering. In: Internet Computing, IEEE, Volume 7, Issue 1, IEEE Computer Society, 2003
5. Paul Resnick, Hal R. Varian, Guest Editors. Recommender Systems. In: Communications of the ACM, Volume 40, Issue 3, ACM Press, March 1997
6. Ben Schafer, Joseph Konstan, John Riedl. E-Commerce Recommendation Applications. In: Data Mining and Knowledge Discovery, Volume 5, Numbers 1-2, Springer, 2001
7. Jonathan Herlocker, Joseph Konstan, Loren Terveen, John Riedl. Evaluating Collaborative Filtering Recommender Systems. In: ACM Transactions on Information Systems, Volume 22, Number 1, S. 5 - 53, ACM Press, 2004
8. Matthias Runte. Personalisierung im Internet: Individualisierte Angebote mit Collaborative Filtering. Dissertation, Christian-Albrechts-Universität zu Kiel, Kiel, 2000
9. Abraham Maslow. A Theory of Human Motivation. In: Psychological Review, Volume 50, S. 370-396, American Psychological Association, 1943
10. Bernd Skiera, Martin Spann. Preisdifferenzierung im Internet. In: M. Schögel, T. Tomczak, C. Belz (Hrsg.), "Roadmap to E-Business - Wie Unternehmen das Internet erfolgreich nutzen", S. 270-284, St. Gallen, 2002
11. Bundesverband des Deutschen Versandhandels e.V.. Pressemitteilung: "10 Jahre E-Commerce in Deutschland - 2005 bringt elektronischem Versandhandel Rekordumsätze". 21.11.2005 <http://www.versandhandel.org/uploads/media/bvh-Pressemitteilung_E_Commerce_21112005.doc>
12. John Breese, David Heckerman, Karl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98), S. 43 - 52, Morgan Kaufmann Publishers, 1998
13. Jonathan Herlocker, Joseph Konstan, Al Borchers, John Riedl. An algorithmic framework for performing collaborative filtering. In: Proceedings of SIGIR'99, S 230 - 237, ACM Press, 1999
14. Daniel Billsus, Michael J. Pazzani. Learning collaborative information filters. In: Proceedings of the Fifteenth International Conference on Machine Learning, S. 46 - 54, Morgan Kaufmann Publishers, 1998
15. Upendra Shardanand, Pattie Maes. Social Information Filtering: Algorithms for Automating "Word of Mouth". In: Proceedings of the SIGCHI conference on Human factors in computing systems, S. 210 - 217, ACM Press/Addison-Wesley Publishing, 1995
16. Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, John Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In: Proceedings of the 1994 ACM conference on Computer supported cooperative work, S. 175 - 186, ACM Press, 1994
17. David Goldberg, David Nichols, Brian M. Oki, Douglas Terry. Using collaborative filtering to weave an information tapestry. In: Communications of the ACM archive, Volume 35, Issue 12, S. 61 - 70, ACM Press, 1992

18. David Maltz, Kate Ehrlich. Pointing the way: active collaborative filtering. In: Proceedings of the SIGCHI conference on Human factors in computing systems, S. 202 – 209, ACM Press/Addison-Wesley Publishing, 1995
19. Joseph Konstan, Eden Prairie, Bradley Miller, David Maltz, Jonathan Herlocker, Lee Gordon, John Riedl. GroupLens: applying collaborative filtering to Usenet news. In: Communications of the ACM archive, Volume 40, Issue 3, S. 77 – 87, ACM Press, 1997
20. John Breese, David Heckerman, Carl Kadie. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, S. 43 – 52, Morgan Kaufmann Publishers, 1998.
21. Ken Lang. Newsweeder: Learning to filter netnews. In: Proceedings of the 12th International Conference on Machine Learning, Tahoe City, 1995
22. John O'Donovan, Barry Smyth. Trust in recommender systems. In: Proceedings of the 10th international conference on Intelligent user interfaces, S. 167 – 174, ACM Press, 2005
23. Kai Puolamäki, Jarkko Salojärvi, Eerika Savia, Jaana Simola, Samuel Kaski. Combining eye movements and collaborative filtering for proactive information retrieval. In: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, S. 146 – 153, ACM Press, 2005
24. David Heckerman. Bayesian Networks for Data Mining. In: Data Mining and Knowledge Discovery, Volume 1, Issue 1, S. 79 – 119, Springer, 1997

Kollaboration mit Hilfe von physikalischen Objekten

Johannes Jüngst

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
Johannes.Juengst@ifi.lmu.de

Zusammenfassung Diese Ausarbeitung befasst sich mit den Auswirkungen von Tangible User Interfaces (TUI) auf die Teamarbeit und Kooperation in Gruppen. TUIs und Tangible Interaction sind demnach in besonderem Maße geeignet, die Kooperation und Kommunikation im Team und in der Gruppe zu unterstützen und zu festigen. TUIs stellen eine physische Schnittstelle zwischen Mensch und Maschine dar. Gegenständliche Objekte dienen dabei zur Interaktion und Manipulation virtueller, digitaler Objekte. Die Kooperation kann parallel über verschiedene physische Objekte stattfinden und mehrere Personen in den Interaktionsprozess einbeziehen. Diese Vergegenständlichung und ihre Auswirkungen auf das Verhalten und die Kommunikation in Gruppen werden in dieser Arbeit im Einzelnen analysiert und erläutert. Im Zusammenhang mit der Vorstellung einzelner praktischer Anwendungsbeispiele werden die Effekte dargestellt.

1 Einführung

Graphical User Interfaces (GUI) waren über viele Jahre mit die einzige Möglichkeit, mit einem Computer zu interagieren. Der Bildschirm war und ist dabei das Fenster, um digitale Informationen wahrzunehmen, und Maus und Tastatur geben uns die Möglichkeit, digitale Objekte zu manipulieren. Eine Steuerung des Computers, die die Wahrnehmungsfähigkeit des Menschen auf Finger und Augen reduziert [21], wurde in einigen Bereichen zunehmend als unbefriedigend empfunden. So sind in manchen Anwendungsgebieten Maus und Tastatur nicht mehr oder nur sehr umständlich einsetzbar, wie z.B. im Grafikdesign, wo das Phantom¹ als Eingabemedium im Arbeitsprozess viele Vorteile bringt. Angesichts der kontinuierlich steigenden Komplexität von Software insbesondere im professionellen Anwendungsbereich (Desktop Publishing, Bildbearbeitung, CAD, usw.) werden die beschränkten Einsatzmöglichkeiten dieses generischen Werkzeuges als zunehmend frustrierend wahrgenommen. Darüber hinaus gestaltet sich vor allem auch die Zusammenarbeit mit mehreren Nutzern auf Basis von GUIs als höchst restriktiv und schwierig.

¹ Das Phantom ist ein haptisches Interface, bei dem der Nutzer seine Fingerspitze in einen speziellen Fingerhut steckt, der auf die Fingerspitze gezielt Kräfte ausüben kann. Somit entsteht eine weitere Möglichkeit der Kontrolle und Wahrnehmung virtueller Objekte [18].

Diese Restriktionen begründeten die Suche nach Alternativen, die es ermöglichen zur verbesserten Erfassung und Bearbeitung der digitalen Welt, weitere Fähigkeiten des Menschen in den Interaktionsprozess einzubeziehen und somit die menschliche Wahrnehmungsfähigkeit umfassender auszuschöpfen. Weiterhin ausgehend von der physischen Welt bzw. Alltags-Umwelt, wurde mithilfe „gegenständlicher Schnittstellen“ bzw. „Graspable / Tangible User Interfaces“ (TUI) nach direkten Kommunikationsformen gesucht, die einerseits eine Verknüpfung bzw. Erweiterung um die digitalen Fähigkeiten / Eigenschaften erlauben und es andererseits ermöglichen, gewohnte Umgangsformen mit Objekten und Informationen beizubehalten [12, 16, 22]. Die Interaktion soll weg von den virtuellen GUIs zurück in die physikalische Welt, in unsere Hände geführt werden. Der Handlungs- und Gestaltungsspielraum für die Eingabe und Repräsentation von Informationen wird durch diese Entwicklung so signifikant erweitert, dass man sich von TUIs große Fortschritte bezüglich Intuitivität, Benutzbarkeit und Effizienz verspricht [8].

Vor diesem Hintergrund hat neben der Idee des „Rechners im Rücken“ das „Konzept der Übergänge, d.h. das Schaffen von Übersetzungen und einer engen Koppelung zwischen Realem und Virtuellem in der Interaktion“ [16] erheblich an Bedeutung gewonnen. Dies gilt auch und gerade für die Problematik der kooperativen Nutzung und Kommunikation in der Gruppe, die in früheren Studien zumeist weitgehend unberücksichtigt blieb.

2 Grundlagen und Grundbegriffe

Sehen, Berühren und materielles Manipulieren spielen in der realen Welt eine wichtige Rolle zur Wahrnehmung und zum Verständnis unserer Umwelt und Umgebung [4]. „Tangible“ heißt ins Deutsche übersetzt „erfassbar“ oder auch „greifbar“, was bedeutet, dass ein tangibles Medium zum einen als ein eigenständiges Objekt wahrgenommen werden kann, man bestimmte Funktionalitäten damit verbinden kann, aber auch, dass man es physisch greifen und es manipulieren kann.

In Verbindung mit einem Eingabemedium für digitale Medien entsteht ein Tangible User Interface, früher eher bekannt als Graspable User Interface [15, 17]. TUIs ermöglichen es, digitale Informationen mit physischen Gegenständen zu koppeln. Diese können dabei sowohl zur Repräsentation der Informationen, als auch als Interface zur Manipulation dieser Informationen dienen [20].

TUIs gewinnen in den letzten Jahren immer mehr an Bedeutung für den Bereich der Mensch-Maschine-Interaktion (MMI). Durch diese Entwicklung sind - im Gegensatz zu den traditionellen Graphical User Interfaces - komplett neue Möglichkeiten der Interaktion und Kommunikation entstanden. TUIs haben sich für die Mensch-Maschine-Interaktion zwischenzeitlich als ein neues, eigenständiges Modell entwickelt [15].

In der Literatur lassen sich drei Sichtweisen für dieses Modells finden, die die Möglichkeiten ihres Einsatzes definieren. Zum Teil gibt es unterschiedliche Bezeichnungen, sie stellen jedoch grundsätzlich die gleichen Erkenntnisse dar.

Data-centered view (datenorientierte Sicht): Diese Sichtweise kommt vorwiegend aus der Mensch-Maschine-Interaktion und beschreibt die physische Darstellung digitaler Informationen und die damit gekoppelte physische Handlung und digitale Funktion. Digitale Zustandsänderungen können durch materielle Zustandsmanipulationen hervorgerufen werden. Digitaler Zustand und physischer Zustand bedingen sich dabei, so dass eine direkte Rückkoppelung der jeweiligen Manipulationen erforderlich ist. [11, 15, 20] Diese Sichtweise wird auch mit „Tangible Interfaces“ bezeichnet.

Expressive-Movement-centered view (ausdrucksorientierte Sicht): Diese Sichtweise stammt aus der Industrie und dem Produkt-Design. Der Nutzen und die Interaktion stehen dabei im Vordergrund. Durch das Objekt / Material soll eine intuitive Verständlichkeit und Anwendbarkeit geschaffen werden. Somit soll durch die Interaktion und die Handhabung die Funktion verdeutlicht werden [11, 15]. Meist auch „Tangible Interaction“ genannt.

Space-centered view (raumorientierte Sicht): Diese Sicht hat sich insbesondere in der Architektur und Kunst herausgebildet. Dabei handelt es sich um die Kombination von realem Umfeld und realen Objekten mit digitalen Displays und Sounds. Sensoren unterstützen diese Umgebungen. Dabei wird ermöglicht, mit greifbaren Objekten in Interaktion mit digitalen Umgebungen zu arbeiten [11, 15]. Diese Betrachtungsweise, oft auch „Interaction Space“ genannt, spielt im Rahmen der Themenstellung eine untergeordnete Rolle.

2.1 Tangible User Interfaces

Um zu verstehen, was Tangible User Interface bedeutet, ist zunächst zu klären, was es für ein Objekt bedeutet, ‚tangible‘ oder ‚graspable‘ zu sein. Im engen Sinne ist damit gemeint, dass ein Objekt greifbarer und stofflicher Natur ist, physikalischen Gesetzen unterliegt, in einer Umgebung situiert ist und körperlich erfahrbar ist [16]. Diese Eigenschaft von Objekten beschrieb James J. Gibson 1977 schon mit dem Begriff „affordance“ [9], die Fähigkeit von Objekten sich selbst zu erklären und ihr Verständnis zu implizieren. Anfassen und Fühlen sind entscheidend für die Interaktion.

TUIs bieten eine physikalische Schnittstelle zu digitalen Informationen. Anstelle der Manipulation von virtuellen Objekten auf einem Bildschirm werden physische, berührbare Objekte zur Interaktion eingesetzt [8]. Die Benutzerschnittstellen erweitern die reale Welt, indem sie Alltagsgegenstände und Umgebung mit digitalen Informationen koppeln.

Tangible User Interfaces zeichnen sich durch bestimmte Merkmale und Eigenschaften aus. In der Wissenschaft gibt es hierzu verschiedene Definitionen. Eine der überzeugendsten und der am meisten verwendeten ist die von Ullmer & Ishii [20] aus dem Jahre 2000. Vier Merkmale sind demnach für Tangible User Interfaces charakteristisch:

- „1. *Physical representations are computationally coupled to underlying digital information.*
2. *Physical representations embody mechanisms for interactive control.*
3. *Physical representations are perceptually coupled to actively mediated digital representations.*
4. *Physical state of tangibles embodies key aspects of the digital state of a system.*” [20]

TUIs verbinden mithin Physikalität mit Digitalität. Digitale Informationen nehmen eine physische Form an, die zum einen zur Repräsentation und zum anderen gleichzeitig auch als Medium zur Steuerung derselben dient. Technisch basiert das Medium auf Sensor- und Video-Systemen (Augmented Reality). Mit unterschiedlichen Techniken - wie z.B. Bilderkennung, Datenhandschuhen, Sensorisierung der Objekte und Umgebungen und berührungssensitiven Oberflächen - erfolgt eine direkte Kopplung der physischen Objekte mit den digitalen Repräsentationen. Die Objekte dienen somit sowohl als Eingabe- als auch als Ausgabeschnittstelle [16]. Bewegungen und Manipulationen können registriert werden und zusätzlich kann das Medium auch ein direktes Feedback auf Aktion und Interaktion geben. Dabei unterscheiden sich diese Interfaces von den traditionellen Schnittstellen dadurch, dass sie keinen sichtbaren direkten Zusammenhang zum Computer haben müssen [20].

2.2 Tangible User Interaction

Tangible User Interaction beschreibt die Aktion und Interaktion mit Tangible User Interfaces im Raum. Die Interaktion findet in der Regel verteilt über verschiedene Objekte statt und kann auch von mehreren Personen zum selben Zeitpunkt ausgeführt werden. Hornecker und Buur [11, 15] unterscheiden dabei zwischen vier Themenstellungen bzw. Konzepten, die wechselseitig untereinander verbunden sind. Diese Differenzierung spielt insbesondere für die später näher zu erläuternde Kooperation und Zusammenarbeit eine wichtige Rolle [15].

Tangible Manipulation (greifbare Manipulation). Bedeutet die körperliche Manipulation von physischen Objekten mit materiellen Handlungen, bei denen Körper und Hände eingesetzt werden. Dabei sind Material, Form und Verhalten der physischen Objekte entscheidend für die Interaktion und Einsetzbarkeit.

Spatial Interaction (räumliche Interaktion). Die Interaktion bezieht sich in diesem Rahmen auf den realen Raum und den Menschen. Körper und Raum stehen in Relation zueinander. Die Interaktion folgt deshalb durch die Bewegung im Raum.

Embodied Facilitation (körperliche Erleichterung). Tangible Interaction findet hierbei im realen und virtuellen Raum statt und ist darauf fokussiert, wie die Konstellation von physischen Objekten im Raum das entstehende Gruppenverhältnis beeinflusst und steuert.

Expressive Representation (Ausdrucksfähigkeit). Material, Form und digitale Funktionalität spielen in diesem Rahmenwerk die wichtigste Rolle. Expressive Representation konzentriert sich auf das Material bzw. physische Objekt und die digitale

Darstellung, die von Tangible Interaction Systems benutzt werden, auf ihre Ausdruckskraft und Lesbarkeit.

In Bezug auf die Themenstellung dieser Arbeit sind in besonderer Weise Spatial Interaction und Embodied Facilitation von Bedeutung. Beide Konzepte beziehen sich auf den Raum und insbesondere auf die Aktionen und Bewegungen von Objekten und Körpern im Raum und haben somit einen direkten Bezug zur Kooperation und Kommunikation.

2.3 Collaboration

Das englische „Collaboration“ sollte nicht mit dem deutschen Begriff „Kollaboration“, der meist mit der Zusammenarbeit mit dem Feind [6] interpretiert wird, verwechselt werden.

"Collaboration is broadly defined as the interaction among two or more individuals and can encompass a variety of behaviors, including communication, information sharing, coordination, cooperation, problem solving, and negotiation" [10]

Diese Definition von Collaboration ist sehr allgemein und wurde von der Intelligence Community im Rahmen ihrer „Collaboration Baseline Studie“ gegeben. Sie beschreibt eine Zusammenarbeit von mehreren Personen mit Hilfe diverser Verhaltensmuster, wie z.B. Kommunikation, Koordination und Austausch von Informationen, um ein gemeinsames Ziel zu erreichen.

Der deutsche Begriff „Kooperation“ greift im Verhältnis zu dieser Definition zu kurz, da er in Bezug auf die Themenstellung allzu unspezifisch die unterschiedlichsten Formen der menschlichen Zusammenarbeit bezeichnet. Insoweit überzeugt die Begriffsbildung und -wahl von Hornecker [14], die zur Bezeichnung des spezifischen Kooperationstypus in diesen Szenarien von „kooperativem und erfahrungsorientiertem Modellieren“ spricht. Hierunter versteht sie „die gemeinsame Produktion von etwas Neuem und insofern eine nicht-routinehafte Situation.“ Dies umfasse auch Lernprozesse, bei denen etwas subjektiv Neues erzeugt werde. Es sei eine kreativ-konstruktive Tätigkeit, für deren Verlauf und Ergebnis es meist keine Vorschrift oder Vorgehensregel gäbe (sog. „open-ended design issues“). Diese Tätigkeit sei potentiell konflikthaft wegen der differierenden Sichtweisen und Interessen der Teilnehmer.

Diskussion und Konflikte sind mithin ein grundlegender Bestandteil des Kooperations- und Kommunikationsprozesses. Das Wissen und die Zusammenarbeit aller sind gefordert.

3 Vereinigung physischer Objekte mit digitaler Information

Tangible User Interaction/Interfaces verbinden zwei Welten, die Digitale und die Physikalische, zu einer neuen Welt und eröffnen damit eine neue Dimension für die Interaktion zwischen Mensch und Maschine [2, 15], aber auch zwischen Menschen

untereinander. Im folgenden Kapitel sollen daher - in Anlehnung an die grundlegenden Ausführungen von Arias, Eden und Fischer [2] - die Eigenschaften und Merkmale materieller Objekte und digitaler Informationen im Einzelnen dargestellt werden, aber auch die Chancen und Möglichkeiten der Ausnutzung von Synergien aus dem Digitalen und Materiellen, insbesondere in welcher Weise sich jeweils beide Welten vervollständigen können und ihre Nachteile ausgleichen können.

3.1 Eigenschaften und Merkmale physischer Umgebung/Objekte

Die Realität und unser materielles Umfeld sind uns schon von Kind auf bekannt. Durch Berühren, Greifen, Fühlen und Ausprobieren haben wir gelernt Dinge, zu verstehen und zu begreifen. Dies ist ein entscheidender Vorteil bei der Verwendung physischer Objekte als Interfaces, den jeder Mensch beherrscht und der intuitives Handeln hervorrufen kann. Arias, Eden und Fischer [2] haben sich darüber hinaus auch mit den Stärken und Schwächen von Materialität befasst. Die grundlegenden Stärken physischer Medien sind:

- *Direkte, einfache Anwendbarkeit und intuitives Verständnis.* Es liegt in der Natur des Menschen, Objekte anzufassen und zu benutzen. Mit Form, Größe und Gewicht können zudem Einsatzmöglichkeit und Verwendungszweck verdeutlicht werden. Wenn ein Objekt z.B. eine gewisse Größe überschreitet, muss es mit zwei Händen bewegt werden oder ist nicht dazu gedacht hochgehoben zu werden.
- *Fühlbare Interaktion.* Der Einsatz eines weiteren Sinnes, nämlich des Fühlens, erweitert die Wahrnehmung unserer Umwelt. Fühlen kann helfen, Objekte besser zu verstehen. Material und Form (z.B. Oberfläche, Widerstand, spitz / stumpf oder fest / weich) sind entscheidende Merkmale von Objekten.
- *Vermittlung von Kommunikation und sozialer Interaktion.* Physische Gegenstände können zu zentralen Bezugspunkten in Kommunikation und Interaktion zwischen Menschen werden. Zudem können Objekte auch ausgetauscht oder weitergegeben werden, was einen entscheidenden sozialen Faktor für die zwischenmenschliche Kommunikation darstellt. Gruppen können dadurch miteinander und untereinander agieren, wobei die physischen Medien die Kommunikation nicht behindern, sondern eher unterstützen und optimieren können.
- *Realitätsgebundenheit.* Physische Artefakte unterliegen den Gegebenheiten der Realität und den physikalischen Gesetzen. Das bedeutet, sie können einfach im Raum bewegt und platziert werden, aber sie können nicht dupliziert oder vervielfacht werden.

Entsprechend den Vorteilen, zeigen sich auch die Nachteile und Einschränkungen durch das Physikalische. Physische Medien sind passiv und nicht reaktiv. Das Objekt kann sich nicht selbst verändern, das bedeutet, dass jeweils eine dritte Person Änderungen vornehmen muss. Weiterhin können sich Objekte nur begrenzt beschreiben und erklären. Es obliegt dem Nutzer, wie er ein Objekt interpretiert. Auch dadurch, dass sich physische Objekte der Realität unterordnen müssen, können diese Objekte

nicht beliebig verändert werden oder bestimmten Gegebenheiten angepasst werden. Eine virtuelle Realität bleibt dem Medium vorbehalten. Es ist nur möglich, was real möglich ist. Ein weiterer Nachteil liegt gegebenenfalls in der Veralterung der Informationen und Objekte. Neue Informationen können nur sehr schwer und nicht zeitgleich ins Physikalische transferiert werden. Physische Objekte sind träge und können meist nur einen Zustand darstellen [2].

3.2 Eigenschaften und Merkmale digitaler Umgebung/Informationen

Die digitale bzw. virtuelle Welt ist erst mit der Entwicklung der modernen Computertechnologie entstanden. Nach Arias, Eden und Fischer [2] kann Digitalität die physikalische Welt unterstützen und ergänzen. Sie bietet die Möglichkeit Informationen zu bearbeiten, zur Verfügung zu stellen und wirkt sich somit positiv auf das Verständnis neuer Sachverhalte aus. Weiterhin unterstützen die digitalen Medien eine direkte Rückkopplung, d.h. man kann einen interaktiven Prozess mit diesen Medien eingehen. Zudem hat man leichten und schnellen Zugang zu weiterführenden Informationen. Die digitalen Daten können unter optimalen Bedingungen gesammelt, archiviert und geordnet werden und erlauben damit ein bislang nicht gekanntes Maß an zeitlicher und räumlicher Verfügbarkeit.

Nachteile der digitalen Welt sind ihre Virtualität und damit Undurchsichtigkeit. Auch ist der Umgang mit digitalen Medien nicht für jeden sofort verständlich, sondern muss erst erlernt werden. Diese Medien sind ein reaktives System, das permanent volle Aufmerksamkeit verlangt. Sie unterliegen einer virtuellen Realität und somit Gesetzmäßigkeiten, die für den normalen Nutzer kaum, bzw. nicht mehr nachvollziehbar sind.

3.3 Nutzen der Integration von physischen und digitalen Medien

Der besondere Nutzen von TUIs ist, dass sie sowohl die digitale als auch die physikalische Welt in ihr Konzept mit einbeziehen. Somit können sich beide Welten ergänzen und ihre Stärken optimieren. Genauer betrachtet sind die Eigenschaften beider komplementär, d.h. die Schwächen physischer Medien können durch die Stärken der digitalen ausgeglichen werden [2].

Neben dieser Ergänzungsfunktion sind als weitere Vorteile durch die Integration und Kombination der digitalen mit den physischen Medien zu benennen:

- *Erweiterte Möglichkeiten.* Durch die Vereinigung entstehen neue Möglichkeiten, um Ziele verständlicher zu machen und mehr Freiheit bei der Gestaltung zu erhalten.
- *Kontinuität der Argumentation.* Bei der Argumentation und Veranschaulichung kann durch die Integration von physischen und digitalen Modellen die Argumentation unterstützt werden. Sie dient als Bezugspunkt und es kann durch digitales Verhalten zudem noch die Funktionalität veranschaulichen.

- *Transparenz.* Physische Objekte selbst können ein intuitives Verständnis durch Material und Form hervorrufen, jedoch kein Feedback auf bestimmte Situationen geben. Durch die Digitalität kann dieses Problem gelöst werden und ein höheres Maß an Transparenz geboten werden.
- *Verständnis.* Greifbare, materielle Objekte sind bekannt und leicht erfassbar. Innerhalb von Gruppen kann man sich leicht auf das Verständnis dieser Objekte einigen. In der digitalen Welt ist dies nur schwer möglich. Beide Eigenschaften, physische Merkmale und digitaler Sinn, können zusammen jedoch das Verständnis und die Anwendbarkeit verbessern und vereinfachen. Auch Hornecker verweist darauf, dass durch die Gestaltung von Objekten Bedeutung, Kontext und Umfang der digitalen Information suggeriert werden kann und bezeichnet diese Eigenschaft als Mapping [13].
- *Verbesserte Konfliktlösung und geteiltes Verständnis.* Die Vorteile für soziale Interaktion in Gruppen durch den Bezug auf physische Objekte verhelfen digitalen Informationen zu einem besseren Verständnis und dienen somit einer verbesserten Interaktion in der Gruppe. Konflikte können somit anschaulich besprochen und verständlich für alle Parteien gelöst werden.

4 Tangible Interaction und soziale Effekte

Tangible User Interfaces beschreiben eine neue Art von Schnittstelle für die Mensch-Maschine-Interaktion und sind gerade damit auch für die soziale Interaktion - wie sie in Abschnitt 2.3 dargestellt und erläutert wurde - von besonderem Interesse. Dieser Themenstellung haben sich in der wissenschaftlichen Literatur bislang allerdings nur wenige Autoren angenommen. In diesem Zusammenhang sind insbesondere die Studien von Hornecker und Buur [11, 12, 13, 15, 16] zu benennen, die die Bedeutung dieser Problematik früh erkannt und die sozialen Aspekte und Auswirkungen von Tangible Interaction untersucht haben.

Die von Hornecker und Buur vorgenommene Differenzierung zwischen vier verschiedenen Themenstellungen der Tangible Interaction (Abbildung 1) wurde eingangs bereits dargestellt und kurz erläutert (siehe Abschnitt 2.2).

Es gilt nunmehr, insbesondere in Anlehnung an die Dissertation von Hornecker [13] die sozialen Effekte der verschiedenen Modelle der Tangible Interaction in ihrer spezifischen Ausprägung im Einzelnen aufzuzeigen.

In der nachfolgenden Grafik sind die Rahmenwerke von Tangible Interaction von links nach rechts vom Spezifischen zum Allgemeinen geordnet, wobei Tangible Manipulation das jeweils spezifischste und Expressive Repräsentation das allgemeine Konzept darstellt [15].

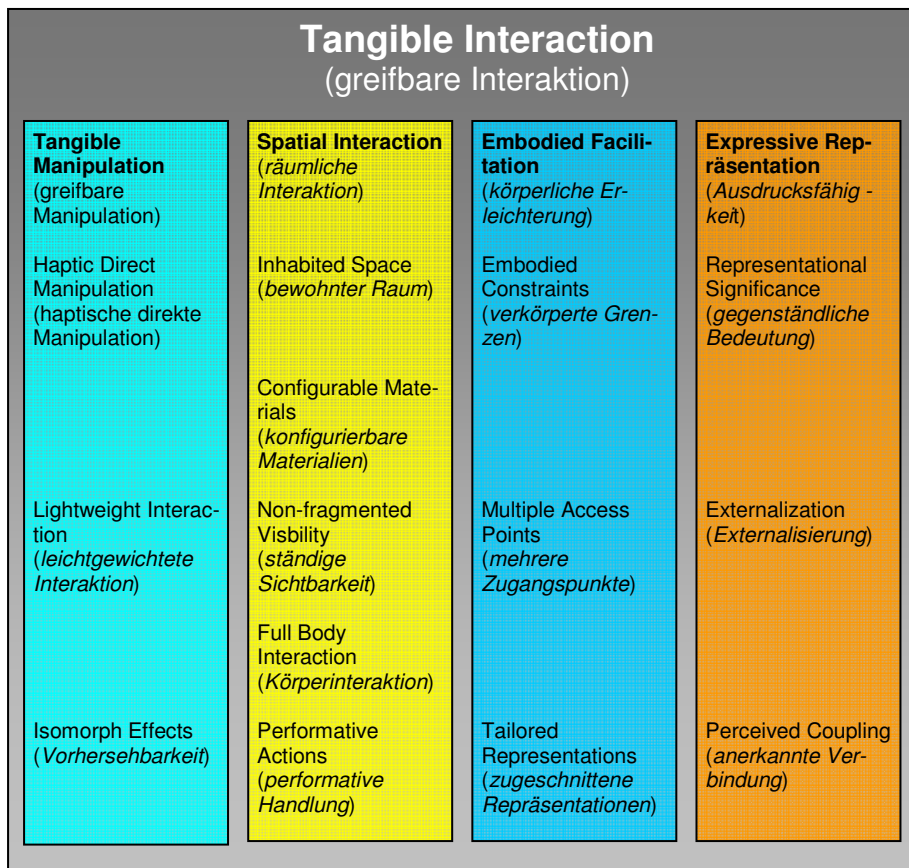


Abbildung 1: Untergliederung von Tangible Interaction nach Hornecker und Buur [15].

4.1 Tangible Manipulation (greifbare Manipulation)

Die greifbare Manipulation (*Tangible Manipulation*) ist die spezifischste Betrachtung und umfasst eine haptische direkte Manipulation mit dem Medium selbst. Derjenige, der mit dem Interaktionsobjekt arbeitet, hat direkten Kontakt zum Interface und kann dadurch ein direktes Feedback über Verhalten und Material des Objektes bekommen. Greifbare Objekte laden dazu ein, durch Ausprobieren, Anfassen und Fühlen mit diesem Interface zu interagieren [15]. Diese intuitive Manipulierbarkeit kann auch die „Hemmschwelle zur aktiven Beteiligung herabsetzen und den Zugang zum System erleichtern“ [13]. Menschen die bereits den Umgang mit physischen Objekten kennen, haben somit eine gute Basis für den Umgang mit einem Tangible User Interface.

Eine weitere Eigenschaft der haptischen direkten Manipulation ist die verbesserte Nachvollziehbarkeit und Vorhersehbarkeit (*Isomorph Effects*). Wenn Informationen durch physische Objekte repräsentiert und manipuliert werden können, resultiert daraus eine bessere Verständlichkeit. Eine sichtbare Verbindung kann zwischen Handlungen und Wirkungen hergestellt werden. Diese isomorphen Effekte erhöhen die Transparenz in der Interaktion [13].

Lightweight Interactions erlauben dem Nutzer eine konstante Rückkoppelung, kleine Schritte und bessere Ausdrucksmöglichkeiten für seine Aktionen und Ideen und schaffen somit eine Grundlage für eine direkte Konversation in der Gruppe [15]. Handlungen und Argumentation werden somit fassbar und ermöglichen eine Intervention, Klarstellung und Positionierung.

4.2 Spatial Interaction (räumliche Interaktion)

Bei der räumlichen Interaktion (*Spatial Interaction*) wird mehr Aufmerksamkeit auf den Raum und damit auch auf die Positionierung und die Bewegung im Raum gelegt. Zudem spielt in diesem Kontext der soziale Aspekt eine entscheidende Rolle. *Inhabited Space* bedeutet, dass die Handlung in einem realen, physischen Raum stattfindet [15]. Die Sichtbarkeit der Aktion im erfassbaren Umfeld erleichtert die Wahrnehmung und das Begreifen. Die Objekte und die Umgebung stehen in einer Relation zueinander, die entweder durch die eigene Bewegung oder durch das Positionieren der Gegenstände verändert werden kann (*Configurable Materials*) [15]. Der Nutzer kann den Raum selbst kontrollieren oder erforschen. Die ständige Sichtbarkeit von Objekten und Umfeld macht die eigentliche Handlung im Kontext und mit Objekten überschaubar (*Non-fragmented Visibility*). Diese Sichtbarkeit ermöglicht erst eine performative Handlung (*Performative Action*) und die gestische Kommunikation untereinander [13]. Handlung und Gesten können somit zum Bezugspunkt in der materiellen Umgebung für Interaktion und Kommunikation werden, wobei dies die Kooperation mit Interaktionspartnern und Interaktionsgegenständen voraussetzt. „Full-body interaction, ... , acquires communicative and performative function“ [15] Durch den Einsatz der Umgebung und des Körpers werden Gestik, Raum und Sprache zu entscheidenden Faktoren. Soziale Aspekte gewinnen mehr an Bedeutung und können somit Kooperation und Interaktion unterstützen [13].

Performative Bedeutung von Handlung wird erst durch die Co-Präsenz von Akteur und Interaktionsobjekt ermöglicht. Der Nutzer kann mit Körper und Handlung gleichzeitig argumentieren [15]. Durch die körperliche Interaktion im Raum kann damit die symbolische und emotionale Bedeutsamkeit der Interaktion erhöht werden [13].

4.3 Embodied Facilitation (körperliche Erleichterung)

Das Konzept von *Embodied Facilitation* umfasst die gegenständliche Interaktion im realen und zugleich virtuellen Raum. Größe, Position und Form von Repräsentations-

flächen können die Aufmerksamkeit der Nutzer lenken und das soziale Gefüge verändern [13].

Embodied Constraints stellen eine bestimmte Konfiguration oder Begrenzung des Raumes und der Objekte dar. Sichtbarkeit und Erreichbarkeit können Handlungen und Verhaltensweisen z. T. vereinfachen oder aber auch begrenzen [15]. Anhand der Konstellation und Konfiguration des Umfeldes können unterschiedlichen Möglichkeiten des Zugangs und des Mitwirkens entstehen. (*Multiple Access Points*) Parallele Manipulierbarkeit vereinfacht somit sowohl den Zugriff, wie auch die Aufteilung der Rollen in der Gruppe. Jeder hat eine „Stimme“ und kann aktiv partizipieren [13]. *Tailored Representation* bezieht sich auf eine maßgeschneiderte Repräsentation für eine Verbesserung des Arbeitsprozesses durch Einbeziehung von kognitiven und emotionalen Zugängen im Umgang mit den Objekten [15]. Neue und unerfahrene Nutzer können durch eine maßgeschneiderte Interaktionsumgebung schneller den Umgang mit den Objekten erlernen und den Nutzen verstehen.

4.4 Expressive Representation (Ausdrucksfähigkeit)

Das umfassendste Konzept ist das der *Expressive Representation*. Sie legt wert auf die direkte physische Repräsentation von digitalen Daten. Dabei kann das Objekt auch zur Externalisierung (*Externalization*) dienen. Diese Funktion beruht vor allem auf der Sichtbarkeit und haptisch direkten Manipulierbarkeit und kann somit die Ausdrucksfähigkeit und das Denken unterstützen [13, 15].

Diese Auslagerung wird auch mit *Representational Significance* verglichen. Objekte zeigen hiermit ohne digitale Repräsentation, in welchem Zustand sie sich befinden [15]. Relation, Position und Zustand haben dabei die bedeutendsten Rollen und bieten zu dem digitalen auch einen realen Kontext, der vom Nutzer interpretiert werden kann.

Perceived Coupling kann im Sinne von Nachvollziehbarkeit verstanden werden. Der Nutzer kann eine wahrgenommene Verbindung zwischen Handlung und Wirkung sowie physischem Objekt und digitaler Repräsentation herstellen. Durch die Erfahrung mit dem System und die Verständlichkeit von Systemreaktionen kann eine gewisse Vorhersehbarkeit geschaffen und folglich die Interaktion verbessert werden [13, 15].

5 Chancen und Grenzen von Tangible Collaboration

Die Ausführungen zur Kopplung physischer Objekte mit digitalen Informationen (siehe 3. Kapitel) belegen den Nutzen und die vielfältigen Vorteile einer integrierten Anwendung von physischen und digitalen Medien. Dies gilt mit allen aufgezählten positiven Eigenschaften (siehe Abschnitt 3.3) gerade auch zur Unterstützung und Optimierung von Kooperation und Kommunikation im Team und in der Gruppe. Im nachfolgenden 4. Kapitel wurden - in Anlehnung an Hornecker [13] und unter Be-

rücksichtigung ihrer Differenzierung zu den Modellen / Rahmenwerken der greifbaren Interaktion - die diversen Effekte für die soziale Interaktion dargestellt. Die besondere Bedeutung von Spatial Interaction und insbesondere von Embodied Interaction für die Kooperation und das menschliche Zusammenwirken in der Gruppe wird dabei deutlich und hervorgehoben.

Es ist daher festzuhalten, dass durch den Einsatz physischer Objekte/Modelle die normale Struktur der Kommunikation und Kooperation gefördert werden kann. Sie sind nicht so aufdringlich, einfacher zu verwenden und können soziale Aspekte unterstützen [12]. Die Externalisierung spielt dabei eine bedeutende Rolle. Die realen Objekte verkörpern konkrete Problemstellungen und sind als solche immer sichtbar und gegenwärtig. Dies unterstützt generell den pragmatischen Umgang mit Konflikten und die Konsensbildung zu ihrer Lösung. Durch Multiple Access Points können von mehreren Kommunikationspartnern parallel berührbare Objekte ergriffen und manipuliert werden, die es somit ermöglichen, dass die geübte traditionelle Kommunikationsstruktur beibehalten werden kann [16]. Eingabe- und Steuerungsmittel sind in gewissem Maße frei verfügbar, ohne dass sich Rollen- oder Arbeitsverteilungen bilden müssen [13]. Anschauliche Modelle, Prototypen und praktische Erprobung sind ein entscheidender Faktor für die Interaktion in Gruppen, um ein gemeinsames Ziel zu erreichen [5].

Debatten und Brainstorming gehören zu den wichtigsten und interessantesten Bestandteilen in der Teamarbeit. Dabei steuern oder behindern TUIs diesen Part der Diskussion nicht, sondern können viel eher ein neues Verständnis mit einbringen [12]. Tendenziell vermögen sie es, die Qualität der Zusammenarbeit zu verbessern.

Der Einsatz von TUIs bietet andererseits aber auch keine Garantie für eine gute und erfolgreiche Kommunikation. Solche realen Objekte sind lediglich ein Werkzeug beim kooperativen Modellieren bzw. ein Medium, das Selbst- und Welterfahrung und Kommunikation unterstützt [19]. Dies wird insbesondere dadurch deutlich, dass die physischen Objekte, die sowohl als Eingabe- als auch als Ausgabeschchnittstellen dienen, nicht in beliebiger Anzahl zur Verfügung stehen und dass ihre Verwendung zumeist eine Verständigung voraussetzt, wenn sie nicht gleich als unfreundlicher bzw. gar aggressiver Akt gedeutet werden soll. Dies gilt insbesondere dann, wenn das Objekt/Modell nur als einziges Exemplar vorhanden ist. Soziale Protokolle zur Steuerung der Interaktion kommen deshalb immer mehr zum Tragen. Soziale und organisierende Faktoren sind entscheidend für die zwischenmenschliche Kommunikation und Interaktion und für den Umgang mit greifbaren Objekten [5].

Entsprechendes gilt auch für die diversen anderen Problemstellungen bzw. Konfliktsituationen im Zusammenhang mit der Nutzung von TUIs, z.B. aufgrund der vorgegebenen Größe, Gestaltung oder der Handhabung der Objekte, der Größe der Interaktionsfläche, der emotionalen Spannungen oder der Verständigungswilligkeit/-fähigkeit der Beteiligten oder auch anderer vorgegebener Bedingungen für das wechselseitige Zusammenwirken. Gerade vorgegebene restriktive Bedingungen, die auf den ersten Blick die Kooperation scheinbar behindern und erschweren, erweisen sich in der Anwendung von TUIs zumeist als Wegbereiter für eine vernünftige und erfolgreiche Verständigung unter den Beteiligten [11].

Selbstverständlich gilt diese These nicht unabhängig von der Möglichkeit, die gegebenen Beschränkungen und Restriktionen auch letztlich überwinden zu können. Dort wo es praktisch keine Chance zur Verständigung gibt und sich die Konfliktsituation als quasi unüberwindbar darstellt, kann bei realistischer Betrachtung auch kein fördernder Einfluss auf das Kooperationsverhalten erwartet werden. Für bereits festgefahrene, sehr schwierige Diskussionen oder bei Problemen mit von vornherein hohem Konfliktpotential dürfte es sich immer als sinnvoll und zweckmäßig erweisen, einen Moderator bzw. Mediator einzuschalten.

Dadurch, dass die Interaktionen durch Tangible Interfaces in einen sozialen und kulturellen Kontext eingebettet werden, steht bei Tangible Interaction eher die Art und Weise der Verständigung als die Schnittstelle selbst im Vordergrund [13]. Das Verständnis von Medien und Handlungen spielt deshalb eine entscheidende Rolle beim Aufbau von Interaktion und Kommunikation [5]. Was jedoch auch bedeutet, dass sich die Beteiligten auf ein Verständnis einigen müssen, da kulturelle und soziale Hintergründe meist verschieden sind.

Durch das Zusammenwirken im realen Raum und die Bewegung im Raum werden die Handlungen und auch das soziale Verhalten für alle Beteiligten jeweils sichtbar und gewinnen somit performative Bedeutung. Gestik, Sprache und Bewegung im Umfeld bzw. „full body interaction“ werden also zu maßgeblichen Einflussfaktoren. Soziale Aspekte kommen hinzu. Im Rahmen des gesamten Spektrums der menschlichen Wahrnehmung findet bewusst oder auch unbewusst soziale Interaktion statt. Aus dem sich hieraus ergebenden Gefühl sozialer Nähe kann für die zusammenarbeitende Gruppe eine tendenziell erhöhte Kooperationsbereitschaft abgeleitet werden.

In der Integration von physischen und digitalen Medien bzw. Realem und Virtuellem liegt darüber hinausgehend eine Chance, durch wechselseitige Ergänzung in Argumentation und Veranschaulichung zu einem besseren Verständnis und zu erhöhter Transparenz im Kommunikationsprozess beizutragen. Dies zeigt sich vor allem, wenn neben den realen und virtuellen Objekten auch noch zusätzliche und verschiedenartige Informationen und Darstellungen angeboten werden können. Tangible Collaboration eröffnet somit neue Handlungs- und Gestaltungsmöglichkeiten, die in Verbindung mit konkreten gegenständlichen Objekten / Modellen in der zwischenmenschlichen Kommunikation und Kooperation - oder besser beim „Modellieren“ - neue, intuitive und erfahrungsorientierte Vorgehensweisen ermöglichen. Die Beteiligten erhalten subjektive Spielräume und können dort neue schaffen, wo diese durch den Einsatz von Computern verloren gegangen sind [16].

Weiterhin werden durch die digitale Ergänzung allen Beteiligten Möglichkeiten der Archivierung, der Simulation und auch der Rückverfolgung (Playback) des kooperativen Modellierungsprozesses angeboten, die jederzeit eine gemeinsame Beobachtung und kritische Analyse des Verfahrens- und Handlungsablaufs gestatten.

6 Anwendungsbeispiele

Nach den bislang weitgehend theoretischen Ausführungen zur Tangible Collaboration werden im Folgenden einzelne Anwendungsbeispiele und aktuelle Entwicklungen in diesem Bereich vorgestellt. Verschiedene Ansätze verbinden die Vorteile von Tangible User Interfaces mit der Kooperation und Teamarbeit in Gruppen, um Lernprozesse oder die Konsensbildung für Lösungen zu unterstützen. Auf die jeweils im Vordergrund stehenden Aspekte zur Unterstützung von Kommunikation und Kooperation wird hingewiesen.

6.1 EDC Environment

Das „Envisionment and Discovery Collaboratory“ (EDC) wurde am „Center for Lifelong Learning and Design“ entwickelt, um kooperative und partizipative Städteplanung zu unterstützen [2, 3].

Der Vorteil der Planungsumgebung ist, dass physische Spiele, Computersimulationen und dynamische Informationen in dem Projekt miteinander gekoppelt sind. Dabei werden diese Systeme in zwei wechselseitig in Beziehung stehende Räume aufgeteilt, zum einen dem „Action Space“ und zum anderen dem „Reflection Space“ (Abbildung 2).

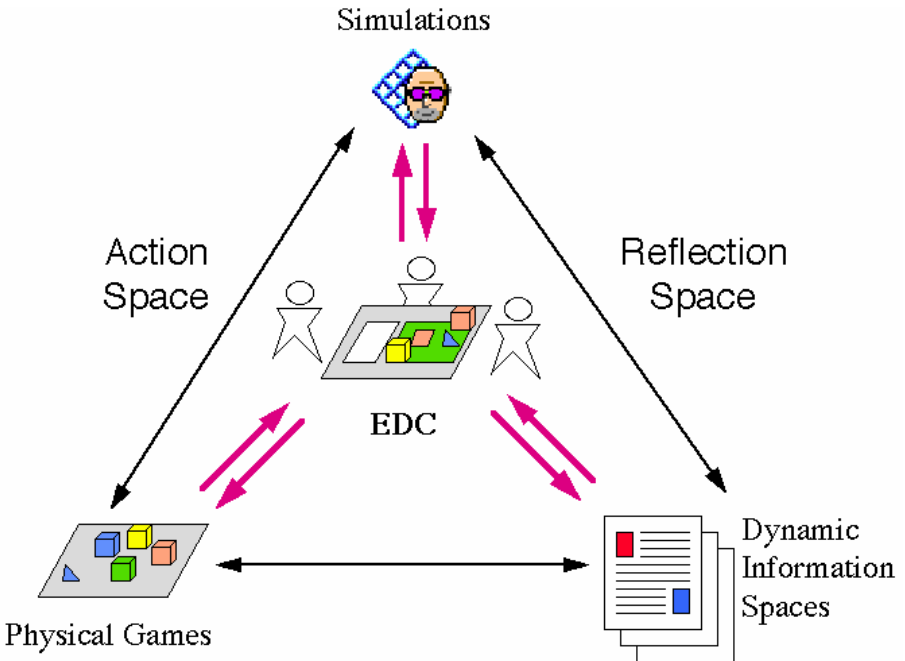


Abbildung 2: Aufbau der EDC Umgebung [3].

Nutzer des EDC können so an einem rechnerisch unterstützten Modelliertisch zusammen kommen und experimentieren (Abbildung 3). Dieser Bereich dient als „Action Space“, der zusätzlich zur Verdeutlichung durch Simulationen unterstützt wird. Technisch basiert der Modelliertisch auf einer berührungssensitiven Fläche, die gleichzeitig auch der Bildschirm ist. Die Nutzer können die digitale Simulation, die auf dem Tischdisplay angezeigt wird, durch das Bewegen von physischen Gegenständen auf dem Modelliertisch manipulieren. Diese Gegenstände haben jeweils bestimmte Bedeutungen, z.B. Häuser, Schulen, Parks und Bushaltestellen, und dienen in diesem Beispiel zum gemeinschaftlichen Erstellen eines Nachbarschaftsmodells für die Städteplanung. Zusätzlich können mit einem Stift Strassen und Arbeitsrouten eingezeichnet werden. Der Rechner kann mit diesen Daten z.B. eine optimale Busroute errechnen und darstellen.

Im „Reflection Space“ befindet sich ein weiteres Smart-Board, das zur zusätzlichen Ausgabe von Informationen dient. Fakten, Meinungen und frühere Diskussionen können somit betrachtet werden und in den Erstellungs- und Gestaltungsprozess mit einbezogen werden.

Insgesamt wird durch die Koppelung dieser Geräte ein Interaktions-System zwischen Mensch und Maschine entworfen, das auf der physikalischen Welt basiert, jedoch virtuell verschiedene Aspekte beleuchten kann. Es bietet somit eine Grundlage für eine sinnvolle interaktive Städteplanung für jedermann und mit jedermann, um durch experimentelle Erprobung und Diskussion eine gemeinsame und sozialverträgliche Lösung zu finden.



Abbildung 3: Experimentelles Arbeiten in der EDC Umgebung [3].

6.2 Ely The Explorer

Ely The Explorer ist ein Konzept, das speziell für das interaktive und kooperative Lernen in der Schule von Kindern entwickelt wurde [1].

Das System besteht aus Ely der Puppe, einem Touch Screen auf dem Teleporter, aus diversen Steuerungsmöglichkeiten direkt am Teleporter, aus einem radio frequency identification (RFID) System mit Sendern (RFID tagged cards) und RFID Lesern, einem PDA und einer Digital-Kamera (Abbildung 4). Mit diesen Interaktionsmöglichkeiten erlaubt das System den Kindern sowohl individuelles als auch gruppenorientiertes Arbeiten und Lernen. Kinder werden dazu aufgefordert durch Ausprobieren und Manipulieren zu einem gemeinsamen Ergebnis zu kommen.

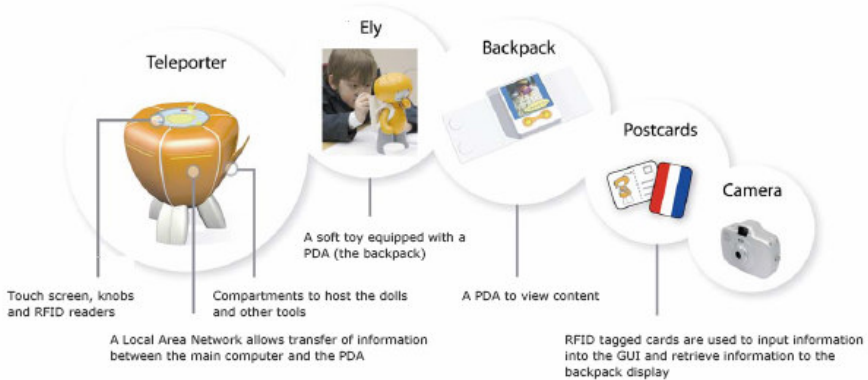


Abbildung 4: Systemkomponenten von Ely the explorer [1].

Das System zeigt insbesondere, dass die Kinder durch die Veranschaulichung und die Möglichkeit zum physischen Manipulieren und Greifen, zum spielerischen Experimentieren und Lernen motiviert wurden [1].

6.3 Tangible Video Editor (TVE)

Der „greifbare“ Video Editor bildet ein System, das aus mehreren Handheld Computern mit eigens implementierter Funktion besteht, um physisch einen Film zusammen zu setzen [23]. Durch diese Externalisierung verbindet der Editor die Vorteile physischen Editierens mit denen des nicht-linearen Editierens.

Die Handhelds dienen dabei sowohl zur Visualisierung des Films als auch zur Repräsentation von einzelnen Filmausschnitten. Durch diese Repräsentation bekommen die Filmstücke einen physischen Charakter und können somit von mehreren Personen parallel bearbeitet und kombiniert werden (Abbildung 5).



Abbildung 5: Arbeit mit dem Tangible Video Editor [23].

Im Ergebnis hat sich gezeigt, dass durch die physische Visualisierung die Arbeit effizienter wurde und dass innerhalb kurzer Zeit Diskussionen ausgelöst wurden. Die Nutzer teilten sich die Interfaces und Aufgaben, um zu ihrem Ergebnis zu kommen [23].

6.4 Build-It

Das *Build-It* System [7] dient zur Planung und basiert auf Computer-Visualisierungen zur Verdeutlichung von komplizierten Modellen und Anordnungen. Dabei kann eine Gruppe von Nutzern, gruppiert um einen Tisch mit physischen Objekten ein virtuelles Modell erstellen und manipulieren (Abbildung 6). Eine Plan-Ansicht wird dabei auf den Tisch und eine Seiten- oder auch Perspektivansicht wird auf die Wand projiziert.

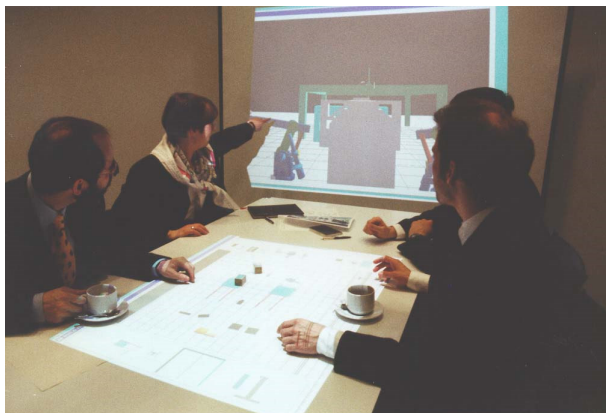


Abbildung 6: Teamarbeit mit *Build-It* [7].

Reale Gegenstände werden, im Gegensatz zu den meisten anderen TUI Systemen, nur temporär an auf den Tisch projizierte Objekte gebunden (Abbildung 7). Am Ran-

de gibt es eine Menüleiste mit Symbolen für verschiedene Werkzeuge zur Manipulation. Durch das Verschieben des realen Gegenstands auf das Symbol wird dieses ausgewählt und der reale Gegenstand dient nun auch zur Steuerung des Werkzeuges.

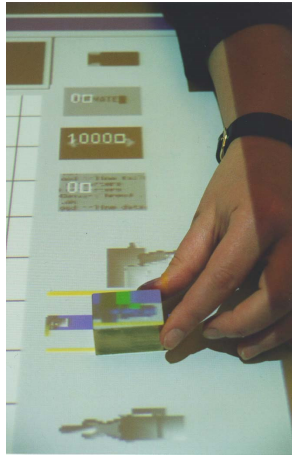


Abbildung 7: Auswahl eines Menüpunktes [7].

Genau betrachtet ist das *Build-It* System eigentlich kein Tangible User Interface, da die realen Gegenstände nur zeitweise an digitale Objekte gebunden werden. Jedoch verdeutlicht diese Art der Anwendung gut die intuitive Manipulation und Steuerung über physische Objekte [16].

7 Ausblick

Im Ergebnis ist festzuhalten, dass Tangible User Interfaces und Tangible Interaction in besonderer Weise geeignet sind, die Kooperation und Kommunikation im Team und in der Gruppe zu verbessern und zu festigen. Zahlreiche Anwendungsbeispiele und die Entwicklung immer neuer Modelle und Techniken belegen diese Effekte des „Collaboration Support“ durchaus überzeugend. Andererseits ist auch festzustellen, dass gerade diesen hoch interessanten und für das menschliche Zusammenwirken in fast allen Lebensbereichen äußerst bedeutsamen Auswirkungen in der wissenschaftlichen Forschung nicht die Rolle bzw. Bedeutung eingeräumt wird, die ihnen eigentlich zukommt. Viel zu sehr beschäftigen sich die verfügbaren Studien und Veröffentlichungen zumeist nur mit Begriffsbestimmungen, diversen Versuchen der Kategorisierung sowie den modelltheoretischen Ansätzen und konzeptionellen Grundlagen von TUIs.

Für die künftige Entwicklung und Gestaltung von „gegenständlichen Schnittstellen“ ist der weitere Bedarf an zusätzlichen Informationen und Erkenntnissen über die tatsächlichen Auswirkungen auf die Interaktion im Team und in der Gruppe sowie über die Integration von realen und virtuellen Medien im menschlichen Zusammenwirken offenkundig. Nur mit Hilfe weiterer empirischer Studien zu den verschiedenen

Anwendungsfällen und -methoden sowie im direkten Abgleich mit den traditionellen Formen der Kooperation und Kommunikation in der Gruppe können hier die Effekte konkreter eingeschätzt und ein höheres Maß an Klarheit erlangt werden. Insbesondere die verstärkte Zusammenarbeit von Wissenschaft und Praxis bzw. von Informatikern in Forschung und Entwicklung und den Anwendern in der praktischen Nutzung ist dringend nötig, um die Erkenntnisse über die zwischenmenschliche Interaktion im Umgang mit TUIs zu sammeln und eine breitere und erfolgreiche Anwendung in vielen Lebensbereichen wie in Wirtschaft und Verwaltung, Ausbildung und Lernen, Freizeit und Unterhaltung und auch Politik zu ermöglichen .

Literatur

1. Africano, D., Berg, S., Lindbergh, K., Lundholm, P., Nilbrink, F. und Persson, A.: Designing Tangible Interfaces for Children's Collaboration. *Proc. of CHI 2004*, (2004).
2. Arias, E., Eden, H. und Fischer, G.: Enhancing communication, Facilitating Shared Understanding, and Creating Better Artifacts by Integrating Physical and Computational Media for Design. *Proc. of DIS '97*, ACM (1997), S. 1-12.
3. Arias, E. G., Eden, H., Fischer, G., Gorman, A. und Scharff, E.: *The Envisionment and Discovery Collaboratory*. Internet: <http://13d.cs.colorado.edu/systems/EDC/index.html>. Letztes Update: 15.4.1999.
4. Brave, S., Ishii, H. und Dahley, A.: Tangible Interfaces for Remote Collaboration and Communication. *Proc. of CSCW '98*, (1998).
5. Dourish P.: *Where the action is. The Foundations of Embodied Interaction*. MIT Press (2001).
6. DUDEN. 2000. Die deutsche Rechtschreibung. Duden Verlag: Mannheim, Leipzig, Wien, Zürich.
7. Fjeld, M., Voorhorst, F., Bichsel, M., Lauche, K., Rauterberg, M. und Krueger, H.: Exploring Brick-Based Navigation and Composition in an Augmented Reality. In H.-W. Gellersen (ed.) *Handheld and Ubiquitous Computing (HUC 99), Lecture Notes in Computer Science, Vol. 1707*. Berlin/Heidelberg: Springer-Verlag, (1999), S. 102-116.
8. Frigg, P.: *Tangible User Interfaces – Beispiele und eine Taxonomie*. ETHZürich, Department für Informatik, Seminar Verteilte Systeme SS2005 zum Thema „Smarte Objekte und smarte Umgebungen, (2005).
9. Gibson, J. J.: The Theory of Affordance. In: Shaw, R., Bransford, J. (Eds.), Lawrence, E. *Perceiving, Acting and Knowing: toward an Ecological Psychology*. NF (1977), S. 67-82.
10. Hall, T.: *Intelligence Community Collaboration Base Line Study – Final Report*. Internet: „http://collaboration.mitre.org/prail/IC_Collaboration_Baseline_Study_Final_Report/toc.htm“, (1999).
11. Hornecker, E.: A Design Theme for Tangible Interaction: Embodied Facilitation. *Proc. of ECSCW '05*, Springer (2005), S. 23-43.
12. Hornecker, E.: Understanding the Benefits of graspable interfaces for cooperative use. *Proc. of CSD '02* (2002), S. 71-87.
13. Hornecker, E.: *Tangible User Interfaces als kooperationsunterstützendes Medium*. PhD-thesis. University of Bremen (2004).
14. Hornecker, E.: *Ein-Jahres-Bericht über das Promotionsprojekt „Gegenständliche Benutzungsschnittstellen und kooperative Modellierung“*, artec, Universität Bremen, (2000).
15. Hornecker, E., Buur, J.: Getting a Grip on Tangible Interaction: A Framework in physical Space and Social Interaction. *Proc. of CHI '06*. ACM (2006).

16. Hornecker, E., Robben, B. und Bruns, F.W.: Technische Spielräume: Gegenständliche Computerschnittstellen als Werkzeug für erfahrungsorientiertes, kooperatives Modellieren. In: Matushek, I., Henniger, A. und Kleemann, F. (eds.): *Neue Medien im Arbeitsalltag*. Westdeutscher Verlag (2001), S. 15-34.
17. Ishii, H. und Ullmer, B.: Tangible Bits: Towards seamless interfaces between people, bits, and atoms. *Proc. of CHI '97*, (1997), S.234-241.
18. Massie, T. H. und Salisbury, J. K.: The PHANTOM Haptic Interface: A Devise for Probing Virtual Objects. *Proc of the ASME '94*, (1994).
19. Schelhowe, H.: *Das Medium aus der Maschine – Zur Metamorphose des Computers*. Frankfurt/New York: Campus, (1997).
20. Ullmer B. und Ishii H.: Emerging frameworks for tangible user interfaces. *IBM Systems Journal* 39 (3-4,) (2000), S. 915-931.
21. Weiser, M.: Some Computer Science Issues in ubiquitous computing. *Communications of the ACM*, Vol. 36, Nr. 7, (1993), S. 74-78.
22. Wellner, P., Mackay, W. und Gold, R.: Computer-Augmented Environments. Back to the Real World. *Communications of the ACM* Vol. 36, Nr. 7, (1993), S. 24-26.
23. Zigelbaum, J., Horn, M., Shaer, O. und Jacob, R. J. K.: *Tangible Video Editor: Designing for Collaboration, Exploration and Engagement*. Report, (2005).

Multi-user Collaboration with Large Displays

Torsten Fritsche

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
fritsche@cip.ifi.lmu.de

Zusammenfassung Diese Arbeit gibt einen Überblick, wie der Einsatz von Großformatdisplays eine computergestützte Teamarbeit forcieren und unterstützen kann. Ich gehe dabei nicht nur auf Anwenderstudien über die kooperative Eignung von Großformatdisplays ein, sondern möchte vor allem aktuelle Forschungen bezüglich der Interaktionsmöglichkeiten und dem Interaktionsdesign anführen. Ein kurzer Ausblick unter Einbeziehung der eigenen Meinung bilden den Abschluss der Arbeit.

1 Einleitung

Wenn man sich Teamarbeit vorstellt, denkt man zum Beispiel an Verhandlungen, Brainstormings, Diskussionsrunden, Vorträge oder Konferenzen. Üblicherweise werden, je nach Art der Zusammenarbeit, dafür traditionelle Tische bevorzugt [1], wodurch sich auch der Begriff des "Runden Tisches" prägte. Für Vorträge oder Konferenzen ersetzen meist vertikale Präsentationsflächen die horizontalen Arbeitsflächen der "Tischrunde".

In den meisten Fällen solch einer Zusammenarbeit gibt es zentrale Oberflächen für das Zeichnen und Aufnehmen von Ideen oder für das Darstellen von Inhalten. Diese Aufgaben erledigt heutzutage ein Computer, mit dessen Hilfe man Präsentationen steuern oder relevante Dokumente besprechen bzw. gemeinsam bearbeiten kann [2]. Gewöhnliche Computer eignen sich jedoch nicht sonderlich für eine solche Teamarbeit. Der Einsatz von Großformatdisplays bietet in diesem Zusammenhang ein deutlich höheres Potential. Daher liegt es nahe diese Flächen durch großformatige Displays zu ersetzen. Dieser Ansatz birgt den Vorteil, dass die Flächen als Informationsdarstellung, wie auch als Interaktionsschnittstelle Verwendung finden können. So hätte theoretisch jeder Benutzer bzw. Teilnehmer der Diskussionsrunde den Zugriff auf die gleichen Daten, was im Sinne einer Zusammenarbeit wünschenswert ist.

Aus technischer Sicht kann die Darstellung solcher interaktiven Arbeitsflächen über eine Projektion mittels Beamer oder über entsprechend

großformatige LCD-, TFT- oder Plasmadisplays realisiert werden. Gekoppelt mit kooperativen Interaktionsmöglichkeiten in Soft- und Hardware werden zukünftig solche Systeme in einer effektiven Teamarbeit nicht mehr wegzudenken sein.

2 Großformatdisplays für kooperative Aufgaben

In einer Teamarbeit können neben bzw. anstelle des PC's Großformatdisplays zum Einsatz kommen. Ob sich diese nun besser für die Teamarbeit eignen, oder welche Szenarien und Anforderungen ein solcher Einsatz bedingt, wird in diesem Kapitel dargelegt.

2.1 PC versus Großformatdisplay

Die Nutzung von Großformatdisplays ist im abstrakten Sinne nicht auf ein Display beschränkt, denn auch mehrere, semantisch miteinander kombinierte Displays können eine großformatige Präsentationsfläche bilden. Somit lässt sich ein Trend zu solchen, großformatigen Pendanten der gängigen CRT-Monitore an aktuellen Desktop-Arbeitsplätzen feststellen. Hier beobachtet man oft den Einsatz von zwei Monitoren an einem Computer – insbesondere dann, wenn es um komplexe Aufgaben mit mehreren Programmen und Programmfenstern geht. "DualHead" bzw. die Möglichkeit mehr als einen Monitor zu nutzen, verbreitet sich mit zunehmendem Maße [3]. Zum einen ist dies den rapide gesunkenen Preisen für Monitore bzw. für Flachbildschirme zuzuschreiben. Zum anderen sind moderne Betriebssysteme in der Lage diese Möglichkeiten entsprechend auszuschöpfen, da auch die Leistungsfähigkeit aktueller Hardware diese Varianten offeriert. Die meisten User haben Displays, deren Größe nicht mehr als 10% ihres physischen Arbeitsplatzes entsprechen [3]. Wie agieren bzw. reagieren User, wenn sie 25% bis 30% ihrer Arbeitsfläche als Bildschirmfläche zur Verfügung hätten?

Entgegen dem Trend und der höheren Verfügbarkeit ist jedoch noch wenig über das Potential und den Nutzen im Vergleich zu gewöhnlichen Monitoren bekannt [3]. Warum eignet sich ein großformatiges Display als Darstellungs-, und möglicherweise Eingabemedium für die Zusammenarbeit? Was für Vorteile bietet die Verwendung dieser Displaytypen im Vergleich zu "gewöhnlichen" PCs oder Laptops?

2.2 Szenarien der Teamarbeit

Verglichen mit einer räumlich getrennten Gruppenarbeit an mehreren Single-User-PCs können Großformatdisplays direkter von mehreren An-

wendern genutzt werden. Diese können auf dargestellte Inhalte weisen bzw. sie konkret berühren, während alle Teammitglieder die Interaktion verfolgen können [4]. Es gibt viele Display-Varianten bezüglich Größe, Form, Orientierung, Anzahl oder Auflösung. Die meisten kommerziellen Ansätze lehnen sich jedoch an das vertikale Whiteboard-Prinzip an [4]. Dieser, in Abbildung 1 dargestellte, häufig anzutreffende Ansatz bildet nicht nur die Grundlage für sämtliche Vortragszenarien, sondern damit auch das erste, zu betrachtende Szenarium einer computergestützten Teamarbeit.



Abbildung 1. Vertikales "Whiteboard-Prinzip" [5]

Bei der gewöhnlichen und traditionellen Zusammenarbeit sitzen alle Beteiligten an einem Tisch und kommunizieren miteinander. Die Interaktion geschieht dabei direkt mit, auf dem Tisch befindlichen, Gegenständen [4][6]. Dieser Ansatz einer horizontalen Displayverwendung, wie in Abbildung 2 dargestellt, soll im Weiteren das zweite, zu betrachtende Szenarium darstellen.

2.3 Eignung für die Teamarbeit

Zu dem Thema, ob und in welchem Maße sich großformatige Displays für die Teamarbeit eignen, existieren einige Untersuchungen. Dabei lässt sich die Eignung von Großformatdisplays nicht nur im Hinblick auf den Einsatz in kooperativen Aufgaben untersuchen. Auch einzelne Anwender sind in der Nutzung von großformatigen Displays zufriedener, produktiver und sogar signifikant schneller [3]. Diese Ergebnisse beziehen sich

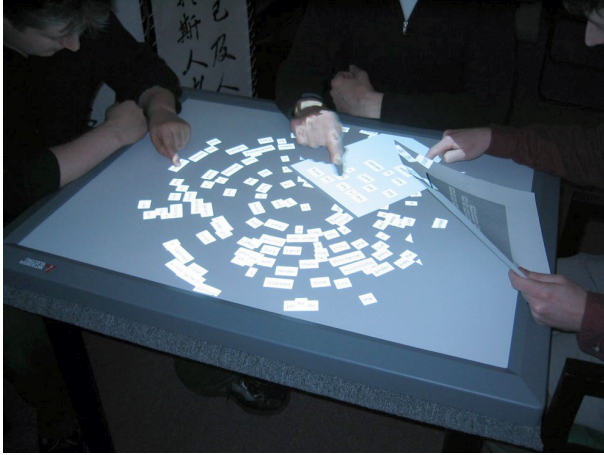


Abbildung 2. Horizontales "Table-Top-Prinzip" [7]

auf komplizierte und komplexe Officeanwendungen, die eine Interaktion mit mehreren Fenstern erfordern. Die Zeitersparnis bei den zu erledigenden Aufgaben lag in dieser Untersuchung bei ca. 9%. Das klingt zwar im ersten Moment nicht übermäßig viel. Doch eine Reihe von Studien zum Interaktionsdesign an Großformatdisplays kommen zu der Ansicht, dass existierende Userinterfaces nicht geeignet sind und sogar fundamentale Bedienkonzepte überdacht werden müssen [3][2][1]. Auf diesen interessanten Teil gehe ich ausführlicher anhand aktueller Forschungen in Kapitel 4 bzw. 4.2 ein. Das bedeutet, dass ausgehend von dieser These zusammen mit den in Kapitel 4 vorgestellten Forschungen, eine deutlich höhere Performanz im Einsatz von Großformatdisplays in Verbindung mit geeigneten Bedienkonzepten zu erwarten ist.

Bezogen auf die Teamarbeit hängt die Eignung der Displays von der Art des Einsatzes ab. Betrachtet man das erste Szenarium (aus Abschnitt 2.2) so ist ein, wie auch immer technisch geartetes, Wall-Display denkbar, welches sich vor einem Publikum an der Seite des Vortragenden befindet. Allein die Sichtbarkeit und Größe der dargestellten Inhalte verlangt nach einer entsprechend großen Darstellung, damit diese von allen Personen im Publikum aufgenommen werden können. Dementsprechend müssen die Bedienelemente möglichst groß dargestellt werden, damit sie leicht und schnell erreicht werden können. Auch die Schriftgröße wäre auf "normalen" PCs viel zu klein für einen solchen Einsatz [2]. Das zweite Szenario (aus Abschnitt 2.2) verlangt nach einem großflächigen Display, das horizontal verwendet eine "Table-Top-Situation" erzeugt. In diesem Umfeld ist

eine direkte Kommunikation der Teammitglieder, wie auch die individuelle und gemeinsame Interaktion mit dem Display möglich. Beide Szenarien haben ihre Berechtigung, sind jedoch für unterschiedliche Situationen geeignet, denn die Ausrichtung des Displays hat einen großen Einfluss auf die Zusammenarbeit.

Generell lässt sich zusammenfassen, dass ein horizontales Table-Top-Display Personen mehr koragiert. Sie sind eher bereit Ideen zu erforschen, sind aufmerksamer und näher am Geschehen [4]. Im Gegensatz dazu fühlen sich Personen, wenn sie an einem Wall-Display stehen bezüglich einer Gruppenarbeit eher unbeholfen. Dies reduziert die Chancen einer effektiven Zusammenarbeit. Das Aufstehen von Personen, um vor den anderen an dem Display zu interagieren, trennt sie vom Rest der Gruppe, was eine Koordination der Gruppenmitglieder erschwert. Sollen bei einer Teamarbeit mehrere Mitglieder interagieren und zusammenarbeiten, dann ist ein horizontales Display geeigneter als ein vertikales. Jedoch haben vertikale Großformatdisplays durchaus ihre Berechtigung. Sie garantieren für alle Teammitglieder eine gute Sichtbarkeit und haben damit einen guten Zugang zu den Inhalten, denn alle Informationen auf dem Display sind für jeden zugänglich – unabhängig von deren Sitzposition. (siehe Kapitel 4.2) Damit eignen sie sich gut für ein gemeinsames Betrachten und kommentieren von Inhalten, wie beispielsweise von Architekturplänen [4].

2.4 Allgemeine Anforderungen an Bedienkonzepte

Soziologische Aspekte "Eine neue Technologie muss die Reflexionen einer Gesellschaft beinhalten, in welcher sie integriert ist und der Gesellschaft das Gefühl geben, dass etwas anerkannt wird, das sie benötigt, möchte oder verlangt." [8]

An ein Userinterface, das für die kooperative Nutzung entworfen ist, werden hohe Anforderungen gestellt. Es muss die Möglichkeit bieten, mit ihm auch komplexe Aufgabenstellungen zu lösen, bei denen mehrere Nutzer gleichberechtigt interagieren können. Trotz dieses Anspruches soll es aber auch schnell und einfach zu bedienen sein, damit die Hauptaufgabe und die Zusammenarbeit der Mittelpunkt bleibt.

Menschen, die in Gruppen arbeiten, neigen nämlich dazu, nicht so geduldig mit "Softwaremysterien" oder komplizierten Bedienungen umzugehen. Sie geben in der Gruppe eher auf, solche Probleme zu lösen, als wenn sie alleine im Büro vor ihrem PC sitzen [2].

Betrachtet man nun das Table-Top-Szenarium, das nach Kapitel 2.3 eher für die Teamarbeit geeignet erscheint, so lassen sich interessante

Verhaltensmuster von Probanden feststellen, die an einem Table-Top-Prototyp untersucht wurden [6]. So tendieren die Probanden dazu – ob alleine oder in der Gruppe – ihren "digitalen Arbeitsplatz" in Bereiche zu unterteilen, um ihre Arbeit besser zu organisieren [6][1]. Bei der Untersuchung kristallisierten sich typische Bereiche für die persönliche Verwendung, die öffentliche Verwendung von allen Teammitgliedern und Bereichen zum Ablegen und Speichern von Inhalten heraus. Die Begrenzung dieser Bereiche wurde von den Probanden meist flexibel gewählt. Wo ein Bereich endet, und wo der nächste anfang hing von der Position der Inhalte und der durchgeführten Aktionen ab. Ist zum Beispiel ein Bereich mit Inhalten gefüllt und wird von außen ein neues Element auf den Rand gezogen, so muss sich dieser Bereich um den benötigten Platz vergrößern.

Die Unterscheidung dieser "semantischen Bereiche" geschah durch ihre Position. Bereiche, die unmittelbar vor einem Proband platziert wurden, wurden als privat und nur durch diesen Proband nutzbar angesehen, wenn sie sich in einer Gruppenarbeit befanden [6][1]. Agierten die Probanden allein, so wurde dieser Bereich für die Hauptaufgabe genutzt. Allgemein gilt für die Abbildung der Position eines Bereiches auf die Funktion desselben, wie im realen Umfeld: "Wenn man etwas nicht erreichen kann, kann man damit auch nicht interagieren" [7]. In dieser, oben schon erwähnten Studie [6] stellte man auch fest, dass Gruppenmitglieder eher zu kleineren Bereichen für ihre eigene Arbeit tendieren, als Single-User. Auch die Größe der zuvor beschriebenen persönlichen Bereiche hängt in gleichem Maße von der Anzahl der mitarbeitenden Personen bzw. davon ab, ob die Person allein arbeitet. Je mehr Personen an dem Table-Top-Display mitarbeiteten, je kleiner wurden diese Bereiche und waren am größten, wenn sie allein agierten.

Bei der Gruppenarbeit wurde das Zentrum des "Tisches" für die Haupt- bzw. Gruppenarbeit genutzt. Die Gruppenbereiche wurden auch für Hilfestellungen unter den einzelnen Mitgliedern genutzt. Zum Beispiel räumte jemand seine Inhalte beiseite, damit ein anderer dort Platz fand. Bestimmte Bereiche wurden auch für Untergruppen gebildet, wie zum Beispiel an einer Tischseite, an der sich diese Gruppe befand.

Interessant für die Ableitung von Bedienkonzepten für solche Displayszenarien ist, dass Inhalte, die in den "Speicherbereichen" abgelegt wurden, um so häufiger genutzt wurden, je näher sich dieser Bereich am "Gruppen-geschehen" befand. Innerhalb dieser "semantischen Bereiche" wurden die Inhalte intuitiv locker organisiert und zu Stapeln zusammengefasst. Wenn man in einem User-Interface diese "semantischen Bereiche" beweglich und verschiebbar konstruiert, beobachtet man einen interessanten Effekt. Die

Probanden platzierten die Bereiche, aus denen sie Informationen holen wollten oder in die sie Inhalte ablegen wollten direkt über ihrer Arbeit, um sie im Anschluss wieder zurück zu platzieren [6].

"Die Möglichkeit Stapel auf dem Tisch verschieben zu können bildet eine Schlüsselrolle in Koordinationsaufgaben und Gruppeninteraktionen." [1] Generell gilt, dass Personen dazu tendieren Stapel oder mehrere Inhalte auf einfach zu verschiebenden Oberflächen zu platzieren, wie zum Beispiel Puzzle-Teile in dem entsprechenden Deckel der Verpackung. So können sie leicht alle verschieben, haben Zugriff auf alle darin befindlichen Objekte und beeinflussen nicht die anderen, auf dem Tisch befindlichen, Inhalte. Auch die Bildung "semantischer Bereiche" bzw. überhaupt die Gruppierung von kausal zusammenhängenden Inhalten bildet meiner Meinung nach eine entscheidende Schlüsselrolle für die Teamarbeit, aber auch für die individuelle Tätigkeit. Diese Möglichkeit bietet eine Strukturierung und Organisation der Arbeit, erinnert an Arbeit, die getan werden muss, bieten einen Mechanismus um Dinge abzulegen, die schwer klassifizierbar sind und bieten während der Zusammenarbeit so die Chance sich auf die "Hauptaufgabe" zu konzentrieren. Der Nachteil ist allerdings, dass die Inhalte oft schwer überschaubar sind und Koordinationsaufgaben erschwert [1].

Abgeleitete Anforderungen Welche allgemein abgeleiteten Anforderungen könnten sich an ein solches Großformat-Display-System für den kooperativen Einsatz ergeben? Unabhängig von der aktuellen Forschung zu der Art und Weise einer geeigneten Interaktion, oder deren Technik lassen sich allgemeine Anforderungen ableiten, die ein Großformatdisplay für den Teameinsatz geeigneter erscheinen lassen, als einen "gewöhnlichen" PC.

Ein vertikales Display in einer Vortragssituation eignet sich dann gut, wenn die Darstellung der Inhalte groß genug und auch qualitativ gut genug ist, dass sie von jeder, im Raum befindlichen, Person aufgenommen werden können. Für dieses, wie auch andere Displayszenarien gilt, dass die Bedienung leicht und schnell möglich sein muss. "Userinterfaces für Gruppenszenarien müssen robust und intuitiv sein." [2] Horizontale Displays müssen zusätzlich auf ein User-Interface setzen, dass das "Erreichbarkeitsproblem" und das "Orientierungsproblem" geeignet umgeht. Darauf werde ich im Kapitel 4.2 näher eingehen.

Zusätzlich zu diesen genannten allgemeinen Anforderungen ergeben sich, entsprechend dem gewähltem Bedienkonzept, oder -technik, unter-

schiedliche und individuelle Anforderungen, die eine intuitive Interaktion ermöglichen. (mehr dazu im Kapitel 4.2)

3 Interaktionsmöglichkeiten

Was haben die traditionelle Teamarbeit an einem Tisch und die computer-gestützten Teamarbeit gemeinsam? In beiden Fällen können die Teammit-glieder direkt miteinander kommunizieren und zusammen oder individuell Aufgaben bearbeiten. In beiden Fällen sind dazu Informationen, die Auf-gaben betreffend, auf dem Tisch bzw. dem Display sichtbar. Es ist auch in beiden Fällen möglich direkt mit diesen Inhalten zu agieren. Diese direk-te Interaktion ist mit Händen oder auch mit Hilfsmitteln möglich. Diese Hilfsmittel können in beiden Fällen Stifte oder ähnliches darstellen.

Das Potential in der Nutzung großformatiger Displays für eine solche Teamarbeit liegt jedoch meiner Meinung nach in den Unterschieden un-ter Beibehaltung der Gemeinsamkeiten. Man könnte eine solche digitale Tischvariante mit mehreren Displays oder Kameras koppeln und gleich-zeitig eine gemeinsame Datenbasis behalten. Es können so alle möglichen realen Objekte einbezogen werden, um inhaltliche Informationen an phy-sische Objekte zu koppeln (zum Beispiel könnten persönliche Notizen oder Adressen an einen Notizblock gekoppelt werden, den man auch an einem anderen Display mit den gleichen Informationen weiterverwenden könn-te). Genauso können alle Teilnehmer einer Teamsitzung immer das gleiche Dokument sehen oder schnell permanente Änderungen durchführen.

Dem Design einer solchen geeigneten Interaktion sind kaum Grenzen gesetzt, sofern sie sich an aktuellen technischen Gegebenheiten orientie-ren. Im Allgemeinen kann man zwischen direkter und indirekter Interak-tion unterscheiden. Bei der direkten Interaktion kann eine Person mit der Hand auf der Oberfläche des Displays interagieren – unabhängig von des-sen Ausrichtung. Eine indirekte Interaktion bedient sich dabei bestimmter Hilfsmittel. So wäre zum Beispiel eine Interaktion über einen Laserpoin-ter oder ein Handy denkbar. Auf die verschiedenen Möglichkeiten oder Techniken soll an dieser Stelle nicht eingegangen werden, es lassen sich jedoch Thesen aufstellen, welche Art der beiden Interaktionen für welches konstruierte Teamarbeitsszenarium aus 2.2 geeignet erscheint.

3.1 Displayinteraktionen und Anwendungen

Walldisplays sind, entsprechend der beschriebenen Anforderungen und so-ziologicalischen Aspekte gut für Vortragssituationen, und besonders auch für

öffentliche Präsentationen oder interaktive Systeme verwendbar. Dabei sind direkte oder indirekte Interaktionen denkbar. Meiner Meinung nach sind direkte Interaktionen bei diesen Anwendungen eher geeignet, wenn wenige oder nur eine Person zur gleichen Zeit mit dem System interagieren bzw. interagiert. Wie oben schon beschrieben sind vertikale Displays ideal, wenn es um die Präsentation vor einem größeren Publikum geht oder um das gemeinsame Betrachten von Plänen, um nur einige Beispiele zu nennen. Eine direkte Interaktion schon von einer Person würde die Inhalte für andere verdecken, weswegen mir eine indirekte Interaktion in dieser Konstellation sinnvoller erscheint. Bei den horizontalen Displays, wie es in den oben beschriebenen Table-Top-Szenarien vorgestellt wurde, sind meiner Meinung nach direkte bzw. in so weit indirekte Interaktionen sinnvoll, als das Stiftinteraktionen oder ähnliche Varianten dazu zählen. Den Grund sehe ich hier in einer aktiven Arbeit bzw. Zusammenarbeit, die intuitiver Weise – unabhängig von der durch das Display initiierten Digitalisierung des Arbeitsplatzes – mit den Händen oder Werkzeugen, wie der Stift eines ist, durchgeführt wird.

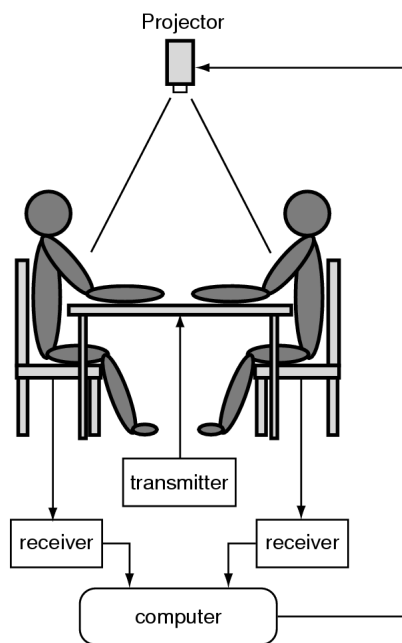


Abbildung 3. DiamondTouch: direkte, simultane Interaktion mehrerer Anwender mit einem horizontalen Display [9]

Als Anwendungsbeispiele für eine solche Interaktion beschäftigen sich die Artikel [9] und [10]. In der ersten Studie – in Abbildung 3 zu sehen – wurde ein System konstruiert, das aus einem weißen, durch einen Projektor beleuchteten, Tisch besteht, der simultane Berührungen lokalisieren und dem jeweiligen User als zugehörig erkennen kann. Somit sind beidseitige Handinteraktionen auf einem großflächigen Table-Top-Display möglich, an dem mehrere Anwender simultan zusammen oder individuell arbeiten können.

In der zweiten Studie – in Abbildung 4 zu sehen – wurde ein System konstruiert, das dem Vortragenden eine direkte Interaktion mit dem Großformatdisplay erlaubt. Dieses System hat den oben beschriebenen Vorteil, dass der Anwender bei der Interaktion keine Inhalte verdeckt, aber auch sehr intuitiv damit interagieren kann.

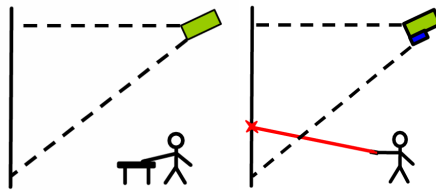


Abbildung 4. links "indirekte Laptop-Interaktion" – rechts "direktere Laserpointer Interaktion" (Bild aus [10])

4 Aktuelle Forschung

An dieser Stelle möchte ich einige aktuelle Forschungen vorstellen, die ich aus meiner Sicht für eine computergestützte Teamarbeit für wichtig erachte. Im Anschluss werde ich einen subjektiven Vergleich unter Einbeziehung der vorher schon genannten kooperativen Systemanforderungen darlegen. Dabei möchte ich nicht auf konkrete technische Lösungen eingehen, sondern die Art und Weise der Interaktion darlegen und im Kontext der kooperativen Arbeit vergleichen.

4.1 Studien und Konzepte für die Teamarbeit

Bulletin Boards Am Rande des Themengebietes möchte ich hier eine Studie erwähnen, die keine direkte, aber eine indirekte Zusammenarbeit un-

terstützt. Dennoch deckt sie meiner Meinung nach ein sehr wichtiges Einsatzgebiet bei einer Teamarbeit ab. In der Studie wurden digitale "Bulletin Boards" entworfen [11] und in einer weiteren Anwenderstudie [12] auf deren Akzeptanz sowie auf die Nutzungsgewohnheiten der Anwender hin untersucht. Wie in Abbildung 5 dargestellt hatten Testpersonen die Möglichkeit Nachrichten ähnlich den aus dem Internet bekannten Foren zu hinterlassen, in den vorhandenen "Postings", wie sie genannt wurden, zu stöbern und sie zu lesen. Die Anwenderstudie zeigte den Einsatz dieser "digitalen Schwarzen Bretter" über einen Zeitraum von ca. eineinhalb Jahren. Je nach eingesetzter Umgebung stagnierte die Anzahl der Autoren, die dort Beiträge veröffentlichten. Bemerkenswert ist jedoch die kontinuierliche Nutzung insbesondere auch im Lesen der Beiträge anderer Autoren [12].



Abbildung 5. Das "Bulletin Board" als digitale Kommunikationsplattform [11]

iRoom Ein weiteres interessantes Projekt ist der "iRoom" [13], dargestellt als eine Draufsicht auf diesen instrumentierten Raum in Abbildung 6. Dieses "Interaktiver-Arbeitsplatz-Projekt" basiert auf der Idee einen (oder mehrere) Räume mit u.a. großformatigen Displays auszustatten, und gleichzeitig mobile Computer einzubinden. Über die mobilen Geräte sollte jeder Nutzer die Möglichkeit haben eigene Inhalte zu integrieren. Der Gedanke einer unkomplizierten und einfachen Bedienung jedes Nutzers von jedem Standort sollte die Zusammenarbeit unterstützen, da so Unterbrechungen vermieden werden. Jedes Gerät im Raum (Großformat-

displays, Laptops, PDAs) war über eine Software (PointRight) an den Kanten des Bildschirmes mit den anderen Geräten assoziiert. So konnte man, neben einer direkten Interaktion an den Großformatdisplays, auf diese auch von jedem anderen Gerät zugreifen, als wäre die gesamte Umgebung ein virtueller Desktop.

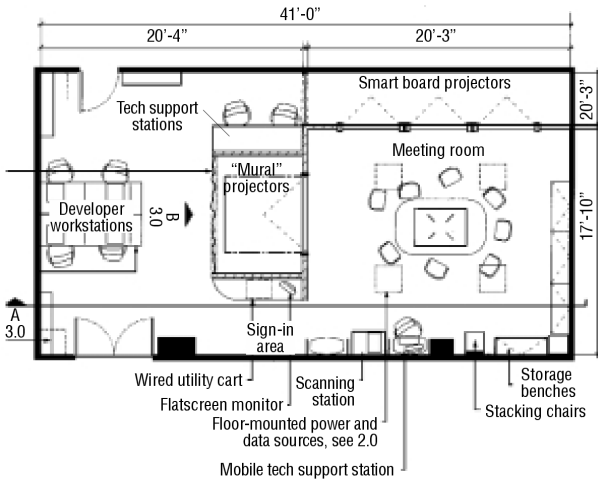


Abbildung 6. Draufsicht eines instrumentierten Raumes "iRoom2" (Bild aus [13])

Roomware Auch mit dem Roomware-Projekt [14] werden verschiedene Komponenten miteinander kombiniert, so dass Informationen, von einem Arbeitsplatz zum nächsten "verschoben" werden können. Ich möchte dieses Projekt an dieser Stelle mit aufzählen, weil es insofern interessant ist, als das es die verschiedenen Interaktionskomponenten konkret in Möbelstücke und Handlungszenarien integriert. Man kann sich – auf Abbildung 7 dargestellt – greifbar vorstellen, wie innerhalb eines Meetings eine Präsentation oder Abbildung von einem CommChair (ein instrumentierter Stuhl) zu dem Wanddisplay (DynaWall) geschickt wird, um diese im Team zu besprechen. Zwar ist das Großformatdisplay als Bestandteil der DynaWall oder des InteracTable nur "Teil des Ganzen", aber dieses Projekt, wie auch der iRoom zeigen damit eine Kombination aus direkter und indirekter Interaktion mit diesen.



Abbildung 7. Roomware: instrumentierte Möbelstücke zur computergestützten Zusammenarbeit [14]

4.2 Intuitives Interaktionsdesign

Ein Interface bezeichnet man als intuitiv, wenn es für alle Nutzer unmittelbar verständlich ist, ohne dass diese über einen bestimmten Bildungsstand oder über spezielles Wissen verfügen [15]. Ein intuitives Userinterface ist für die Zusammenarbeit von besonders herausragender Bedeutung, denn bei einer Gruppe von Personen kann nicht von gleichen Vorkenntnissen, Wissen oder Fähigkeiten ausgegangen werden – das würde dem Sinn der Zusammenarbeit teilweise widerstreben. Trotz dieser Vorbedingung soll eine Gruppenarbeit möglich sein, ohne durch technische Interaktionsprobleme einzelner beeinflusst zu werden. Um die Teamarbeit an Großformatdisplays zu forcieren und die Akzeptanz der Anwender zu wahren, sollte eine geeignete Interaktion möglichst intuitiv gestaltet werden.

Je nach Art des Displays (Bezug zu Displayszenarien aus Kapitel 2.2) stellen sich dem Designer unterschiedliche Probleme bei der Gestaltung eines intuitiven Userinterfaces. Betrachtet man erneut die Szenarien des vertikalen Großformatdisplays und des Table-Top-Displays, so stellt sich das schon vorher erwähnte "Erreichbarkeitsproblem". Das User-Interface muss also Techniken oder Konzepte bereitstellen, dass eine am Tisch bzw. am Display befindliche Person nicht mit Gegenständen, Menüs oder Inhalten interagieren muss, die sie physisch nicht erreichen kann. Als Beispiel möchte ich hier Menüs zum Auslösen von Aktionen anführen, auf die der Anwender mehrmals während der Arbeit zugreifen müsste. Bei einem vertikalen Display wäre eine indirekte Interaktion geeignet, weil der Anwender dann nicht auf die physischen Displayausmaße angewiesen wäre. Bei einer direkten Interaktion jedoch kann er nur innerhalb seiner

physischen Reichweite interagieren. Dieses Problem wurde bei dem sogenannten "Tracking Menu" aufgegriffen [16]. Abbildung 8 zeigt ein solches Tracking Menu, welches sich jeder Anwender frei auf dem Tisch positionieren und verschieben kann. Somit sind die benötigten Menüs nicht nur innerhalb der physischen Reichweite des Anwenders, sondern lassen auch flexible Positionsänderungen zu. Es spielt keine Rolle, ob ein Vortragender beispielsweise links oder rechts vom Display steht, wenn sich dieser "sein" Menü mitnehmen kann.

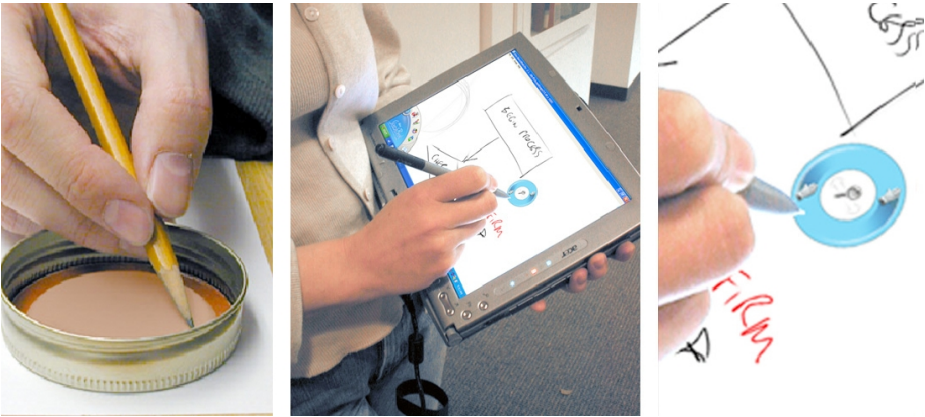


Abbildung 8. Mobile Menüs, die so dem Anwender "folgen" können [16]

Bei der Interaktion mit horizontalen Displays, wie sie in der Table-Top-Situation vorkommen, kommt ein weiteres Problem zum Vorschein: Das "Orientierungsproblem" – das in den soziologischen Aspekten 2.4 angedeutet wurde – entsteht, weil bei einer solchen Teamarbeit am Tisch befindliche Personen, Inhalte von verschiedenen Seiten betrachten. So ist zum Einen Text schwer von anderen Seiten lesbar, zum Anderen können aber auch Standard-Interface-Komponenten mehrdeutig erscheinen [1]. So sieht man zum Beispiel auf Abbildung 9, dass User-Interface-Komponenten, wie zum Beispiel Buttons, je nach Betrachtungswinkel gedrückt oder ungedrückt erscheinen können. Auch für dieses Problem gibt es verschiedene Ideen, das Userinterface-Design betreffend. Eine Idee besteht darin, die Inhalte dynamisch den jeweiligen Anwendern "entgegen" zu drehen [17]. Abbildung 10 zeigt einen solchen Prototyp, in dem jedes Bild dem jeweiligen Anwender zugewandt ist.

Stellt man sich ein Publikum vor einem sehr großen, kinoartigen, vertikalen Display vor, so kann es je nach Art der Präsentation dazu kommen,



Abbildung 9. Standard-Interface-Komponenten erscheinen je nach Betrachtungswinkel gedrückt oder ungedrückt [1]



Abbildung 10. Rotation dargestellter Inhalte in Richtung des jeweiligen Anwenders [17]

dass sich die Aufmerksamkeit des Publikums nicht eindeutig steuern lässt. Das ist auch eine der Thesen, warum sich im Kino 3D- und IMAX-Filme nicht in dem Maße durchgesetzt haben, wie es erwartet wurde. Das Problem liegt darin, dass bei einer solchen Informationsvielfalt jeder Zuschauer andere Bildbereiche wichtig oder interessant finden würde. Für viele Präsentationstypen – oder Filme – spielt es eine entscheidende Rolle die Aufmerksamkeit auf bestimmte Bereiche zu lenken. Aufgegriffen aus dem klassischen Theater und der gestalterisch geschickten Lichtsetzung beim Film hat man eine "Spotlight-Technik" benutzt, um bestimmte Bildbereiche zu einem bestimmten Zeitpunkt in einer Präsentation hervorzuheben [18]. Abbildung 11 zeigt, dass alle Teilnehmer einer Teamarbeit so auf einen bestimmten Bereich aufmerksam gemacht werden können.

Wie in dem Beispiel mit der "Spotlight-Technik" muss ein Großformatdisplay nicht in jeder Situation auch viele, hochaufgelöste Informationen anzeigen. Nun fragt man sich vielleicht, warum sind Großformatdisplays sinnvoll, wenn man nur einen bestimmten Bereich mit der Spotlight-Technik hervorhebt, oder gar Informationen nur gering aufgelöst darstellt? Es gibt viele Situationen, in denen es eher um den Überblick geht. Es reicht, die eigentlichen Detailinformationen durch Licht oder hohe Auflösung angemessen darzustellen. Diesen Überblick zu geben ist eines der vielleicht wichtigsten Argumente für den Einsatz solcher Displays. Um



Abbildung 11. Spotlight-Technik lenkt die Aufmerksamkeit der Betrachter [18]

diese Idee plakativer darzustellen, stelle man sich nur ein Navigationssystem vor bzw. einen solchen Ausschnitt einer Wegbeschreibung auf einem "gewöhnlichen" Monitor. Der Ausschnitt wird hinreichend genau dargestellt, doch hat man Probleme sich in diesem zu orientieren, wenn man nicht genau weiß, wo man sich befindet. Nun wäre es vielleicht nicht gleich nötig, einen größeren Ausschnitt bzw. den ganzen Weg auf einem großformatigen Wanddisplay darzustellen (da bliebe auch das Problem, den eigenen Standort zu finden). Dagegen könnte aber eine kombinierte Displaytechnologie Anwendung finden. Ein niedrig aufgelöstes Bild des Weges könnte an die Wand projiziert werden, während ein hochauflöster Ausschnitt den Standort, sowie alle "primär" wichtigen Informationen liefert [19]. So oder ähnlich könnte eine Motivation für eine Studie gewesen sein, die genau diese zwei Darstellungen auf einem Table-Top-Display (Foveal Display) kombiniert [20]. Abbildung 12 zeigt links den Aufbau dieser Technologie, die zwei Projektoren über ein Spiegelsystem auf unterschiedlich große Bereiche eines Tisches projiziert. Eine geeignete direkte Interaktionstechnologie machte es möglich Dokumente aus dem niedrig aufgelösten Bereich in den hochauflösten Bereich zu holen und umgedreht. Somit konnte man an einem Dokument mit der nötigen Darstellungsqualität arbeiten, verlor aber gleichzeitig nie den oben beschriebenen Überblick. Das Anwendungsgebiet ist weitreichend, denn gerade in redaktionellen Arbeiten, in denen man viele Dokumente oder Forschungen einbezieht, ist diese Technologie sehr intuitiv und vorteilhaft.

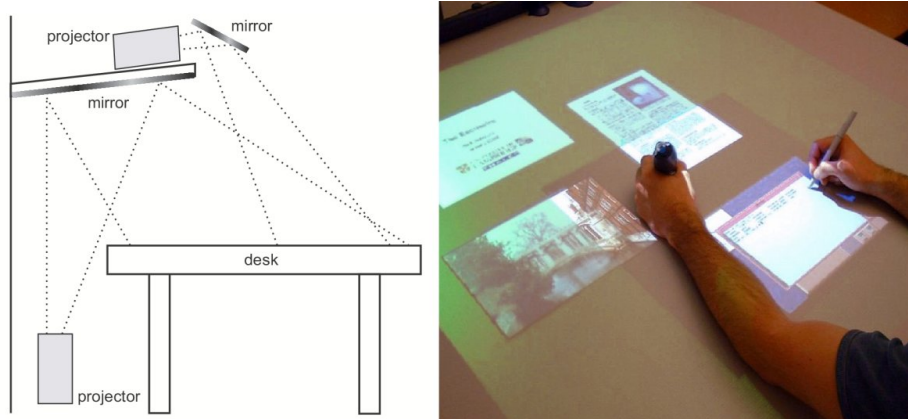


Abbildung 12. Kombinierte Auflösungen durch mehrere Projektoren [20]

4.3 Zusammenfassung und Wertung

Ich möchte in diesem Abschnitt versuchen, die vorgestellten Arbeiten mit den zuvor zusammengetragenen Hintergründen gegeneinander zu stellen und entscheidende Aspekte herausstellen, die meiner Meinung nach eine Zusammenarbeit an großformatigen Displays unterstützen und forcieren können. Diese Gegenüberstellung soll nicht in Form eines Vergleiches erfolgen, denn dazu sind die Projekte, wie deren betreffende Anwendungsbereiche zu verschieden. Da sie jedoch alle wichtige Bereiche einer Zusammenarbeit einbeziehen, lassen sie sich in diesem Kontext zumindest subjektiv werten.

Viele Ideen, Forschungen oder auch Entwürfe im kreativen Bereich kommen nicht von einem Moment zum Nächsten. Auch denke ich, dass ein Meeting nie alle Problemstellungen auf einmal lösen kann. Meiner Meinung nach müssen Ideen oft "wachsen", indem man sie sich erneut zu anderen Zeitpunkten ansehen kann, und auch andere Ideen betrachten, beurteilen oder kommentieren kann. Aus diesem Grund habe ich das Projekt der "Bulletin Boards" aufgenommen. Diese eher indirekte Technik bietet genau diesen Ansatz. So könnten während oder außerhalb eines Meetings Ideen von einzelnen Personen notiert werden, die die Personen sich selbst oder andere Personen, unabhängig von der Zeit, ansehen und kommentieren können. So könnte diesem "kooperativen Ideenwachstum" Rechnung getragen werden.

Die Teamarbeit an sich – ob herkömmlich oder mit instrumentierter Unterstützung – lebt von unterschiedlichen Ideen, Gedanken und Ansätzen, die meiner Meinung nach am geeignetsten durch kombinierte Techno-

logien (wie im iRoom oder mittels Roomware) präsentiert und bearbeitet werden können. Ich denke zwar, dass die Kombination mit "persönlichen" Geräten, wie dies im iRoom-Projekt vorgestellt wurde sinnvoller ist, als ausschließlich, für die Teamarbeit optimierte Geräte. Den meisten Teamarbeiten geht eine individuelle Arbeit voraus, die oft in diese Teamarbeit eingebracht bzw. dort präsentiert wird. Diese Vorarbeiten kommen dann meist von solchen "persönlichen" Geräten, wie Laptops oder PDAs. Auch fühlen sich Anwender mit bekannten bzw. eigenen Geräten vertrauter. Insofern halte ich eine geeignete Integration eigener "Ressourcen" in eine instrumentierte Umgebung für ideal. Die reine Teamarbeit betreffend, sind die Spotlight-Technik oder die Kombination unterschiedlicher Auflösungen auf einem Display eher für spezielle Einsatzgebiete wichtig, obwohl sie interessante Problemlösungen bieten.

5 Fazit

Dieser Artikel stellt einen Überblick über die computergestützte Zusammenarbeit an Großformatdisplays dar. Es wurden, ausgehend von soziologischen Aspekten und Displayvarianten Anforderungen konstruiert, die eine solche Zusammenarbeit unterstützen oder forcieren können. Eine Auswahl aktueller Studien, Interaktionstechniken und Interaktionsdesign betreffend zeigt, dass dieses Thema eine Veränderung typischer Interaktionstechniken und -gewohnheiten mit sich bringen muss, um attraktiv zu sein. Sieht man sich nun die fiktiv konstruierten Anforderungen an Großformat-Display-Systeme an, die eine Zusammenarbeit in geeigneter Weise unterstützen; bezieht man soziologische Aspekte der handelnden Personen, sowie prinzipbedingte "Userinterface-Design-Probleme" mit ein, möchte ich folgenden Schluss unter Einbeziehung meiner eigenen Meinung fassen: Großformatdisplays eignen sich bedeutend besser als "herkömmliche" Single-User-Displays für die Zusammenarbeit, wenn gleich sie nur Bestandteil eines Systems sein können, das auf diese Belange optimiert ist. Meiner Meinung nach haben vertikale wie horizontale Setups Vorteile, die in Kombination mit geeigneten intuitiven Interaktionsmöglichkeiten ein Optimum für eine computergestützte Zusammenarbeit bilden können. Ich denke, dass das zuvor beschriebene Projekt des iRooms genau den Kern dieser Forschungsrichtung trifft und umfassend einbezieht, denn die Kombination verteilter, unterschiedlicher Systeme und Technologien gliedert sich am Besten in eine Teamarbeit von genauso "unterschiedlichen" Personen, deren gleichartige "Interessenkombination" das Wesen der Kooperation beschreibt.

Literatur

1. Scott, S.D., Carpendale, M.S.T., Habelski, S.: Storage bins: Mobile storage for collaborative tabletop displays. Published by the IEEE Computer Society (2005)
2. Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J., Welch, B.: Liveboard: A large interactive display supporting group meeting, presentations and remote collaboration. *Proceedings of ACM* (1992)
3. Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., Starkweather, G.: Toward characterizing the productivity benefits of very large displays. *Proc. INTERACT 2003* (2003)
4. Rogers, Y., Lindley, S.: Collaborating around large interactive displays: which way is best to meet? *Interacting with Computers* (2003)
5. Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N., Kurtenbach, G.: A remote control interface for large displays. *Proceedings of ACM* (2004)
6. Scott, S.D., Carpendale, M.S.T., Inkpen, K.M.: Territoriality in collaborative tabletop workspaces. *Proceedings of ACM* (2004)
7. Ryall, K., Forlines, C., Shen, C., Morris, M.R.: Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. *Proceedings of ACM* (2004)
8. Arvola, M.: Shades of use: The dynamics of interaction design for sociable use. Dissertation No. 900. Department of Computer and Information Science at Linköping University (2005)
9. Dietz, P., Leigh, D.: Diamondtouch: A multi-user touch technology. *Proceedings of ACM* (2003)
10. Cheng, K., Pulo, K.: Direct interaction with large-scale display systems using infrared laser tracking devices. *Proceedings of the Australian symposium on Information visualisation - Volume 24* (2006)
11. Churchill, E.F., Nelson, L., Denoue, L.: Multimedia fliers: Information sharing with digital community bulletin boards. *Proceedings of Communities and Technologies* (2003)
12. Churchill, E.F., Nelson, L., Denoue, L., Helfman, J., Murphy, P.: Sharing multimedia content with interactive public displays: A case study. *Proceedings of Designing Interactive Systems (DIS2004)*. ACM Press (2004)
13. Johanson, B., Fox, A., Winograd, T.: The interactive workspaces project: Experiences with ubiquitous computing rooms. *IEEE Pervasive Computing* 1:2 (April-June 2002), 67-75 (2002)
14. Prante, T., Streitz, N.A., Tandler, P.: Roomware: Computers disappear and interaction evolves. *IEEE Computer*, December, 2004. pp. 47-54 (2004)
15. Bærentsen, K.B.: Intuitive user interfaces. *Scandinavian Journal of Information Systems*, 2000, 12: 29-60 (2000)
16. Fitzmaurice, G., Khan, A., Pieké, R., Buxton, B., Kurtenbach, G.: Tracking menus. *Proceedings of ACM* (2003)
17. Hancock, M.S., Vernier, F.D., Wigdor, D., Carpendale, S., Shen, C.: Rotation and translation mechanisms for tabletop interaction. *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)* (2006)
18. Fitzmaurice, G., Khan, A., Kurtenbach, G., Binks, G.: Cinematic meeting facilities using large displays. *IEEE CG&A*, 25(4) (2005)

19. Baudisch, P., Good, N., Stewart, P.: Focus plus context screens: Combining display technology with visualization techniques. *Proceedings of ACM* (2001)
20. Ashdown, M., Robinson, P.: The escritorio: A personal projected display. *11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 03)*, pages 33?40 (2003)

Kollaboration mittels direkter Interaktion mit großen Displays

Hendrik Richter

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`richter.hendrik@tiscali.de`

Zusammenfassung Diese Arbeit gibt einen Überblick über Möglichkeiten zur gemeinsamen direkten Interaktion mit großen Displays bzw. Bildschirmen. Anforderungen an die Interaktionstechniken werden anhand von Beispiel-Szenarien verdeutlicht, Technologien werden genannt und erläutert. Zusätzlich werden einige sinnvolle Ansätze zur Unterstützung der gemeinsamen Arbeit am großen Display vorgestellt. Im Ausblick werden schließlich mögliche zukünftige Entwicklungen im Bereich der kollaborativen Systeme aufgezeigt.

1 Einleitung

Um gemeinsam mit einer großen Menge von Informationen zu arbeiten, ist es oft sinnvoll, diese Informationen auf einer für alle sichtbaren Ebene oder Fläche zu platzieren. Die Informationen bzw. Daten sind dadurch einfacher zu strukturieren, zu organisieren und zu bearbeiten. Möglich ist das auf einem großen, interaktiven und hochauflösenden Screen mit dem Vorteil der digitalen Datenverarbeitung.

Durch sinkende Preise und steigende Rechenleistung sind große Displays bzw. Multi-Monitor-Systeme in immer mehr Arbeitsumgebungen anzutreffen. Studien zeigen, dass die Verwendung von größeren Bildschirmen Produktivität und Arbeitsgeschwindigkeit steigern kann [1]. Speziell beim gemeinsamen kreativen Arbeiten wie z.B. beim Brainstorming, strategischen Planen oder der Entscheidungsfindung verwenden die Nutzer aber noch immer eher traditionelle Medien wie z.B. Tafelmarker, farbige Zettel, Post-Its oder Karteikarten. Die Interaktion mit physischen Objekten scheint in diesen Situationen leichter zu sein als mit Computermonitoren. Leider kann die in solchen Szenarien gesammelte Information nicht sofort digital gespeichert werden und somit sogar komplett verloren gehen [2].

Mit Displays, die viel größer als ein Computermonitor sind, muss man aus verschiedenen Gründen anders interagieren. Das große Display kann dazu genutzt werden, mit einer riesigen Menge von gleichzeitig sichtbarer Information zu arbeiten, die Interaktion findet auf dem Bildschirm statt, direkt mit der Hand oder einem Stift-ähnlichen Eingabegerät, eher nicht mit der Tastatur oder einem indirekten Zeigegerät. Außerdem arbeiten die Nutzer oft zusammen an einem Display, wobei technische und soziale Interaktion gleichzeitig stattfindet.

Die Techniken und Technologien zur Interaktion mit diesen Displays müssen also einige Anforderungen erfüllen, die hier genannt werden sollen. Diese Arbeit beschränkt sich auf die direkte Interaktion und versucht einen Überblick über ihre Möglichkeiten zu geben.

2 Anforderungen

In einem realistischen Arbeitsumfeld wie z.B. einem Meeting-Room führen die Nutzer von großen Interaktiven Displays verschiedene Tätigkeiten aus. Daten werden gesammelt, präsentiert und bearbeitet, um eine Aufgabe zu lösen. Der Screen wird nicht ständig von derselben Person bzw. Personengruppe verwendet. Die soziale Interaktion spielt eine viel größere Rolle als bei der Arbeit alleine an einem Desktoprechner, die Person am Screen präsentiert Themen, interagiert und spricht mit den Zuhörern, die Aufmerksamkeit ist also nicht permanent auf das Display gerichtet. Das Display integriert sich in die Umgebung [2].

Eines der ersten großen interaktiven Anzeigesystemen ist das Liveboard aus dem Xerox Palo Alto Research Center [4]. Es stellt eine Grundlage zur Forschung über Benutzer-Schnittstellen bei Gruppen-Meetings, Präsentationen und verteilter Zusammenarbeit dar.

Interaktive Bildschirme wie das Liveboard sollen nicht nur Whiteboards sein, auf denen man zeichnen und Ideen anzeigen kann, sondern sollen in eine computerisierte Umgebung eingebunden werden.

2.1 Anforderungen an die Interaktionstechnik mit dem Bildschirm

Die direkte Interaktion mit dem an der Wand befestigten Screen geschieht in kurzem Abstand (max. Armlänge). Das angezeigte Material kann bunt und komplex sein. Mehrere Nutzer arbeiten in Echtzeit gleichzeitig am Display (siehe Abb.1). Die Möglichkeit, die (z.B. bei einem Brainstorming) gesammelten Daten hinterher leicht in digitaler Form zu katalogisieren bzw. zu speichern, sollte gegeben sein.

Alle Anforderungen zielen darauf ab, eine flüssige Interaktion sowohl zwischen Nutzern als auch zwischen Gerät und Nutzer zu gewährleisten. Der Workflow in einem kreativen Meeting soll nicht unterbrochen werden.

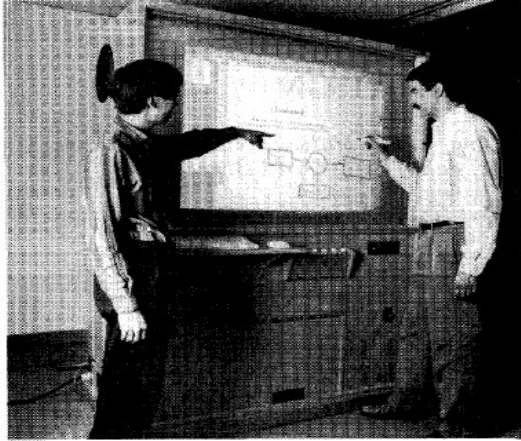


Abbildung 1: Direkte Interaktion bei der gemeinsamen Arbeit an einem großen Interaktiven Display

Die Techniken zur Interaktion mit diesen Displays sollten also im besten Fall folgende Anforderungen erfüllen:

- **Interaktion in Echtzeit.** Die Interaktion sollte in Echtzeit stattfinden, um eine dynamische Gruppenarbeit zu gewährleisten und den Workflow nicht zu unterbrechen.
- **Interaktionstechnik sollte schnell und einfach zu erlernen sein.** Die zu verwendende Interaktionsform kann sich fundamental von der Interaktion mit einem Desktoprechner unterscheiden. Ein Nutzer, der die Verwendung einer Maus gewohnt ist sollte aber keine Schwierigkeiten haben, die Eingabe am Screen zu erlernen und nach kurzer Zeit flüssig anzuwenden. Das flüssig-aktive Arbeiten am großen Screen wird so ermöglicht.
- **Freihändige Interaktion.** Eine direkte Interaktion, also eine Interaktion die direkt am Screen stattfindet sollte im besten Fall freihändig möglich sein. Für den Nutzer interessante Objekte können so mit dem Finger (oder einem Stift) "auf natürliche Weise" manipuliert werden. Verkabelte oder unhandliche Eingabegeräte erschweren die Zusammenarbeit direkt am Screen, da Platz verschwendet wird und die Verkabelung beim Positionswechsel störend wirkt.
- **Keine Eingabegeräte, die den Screen verdecken.** Geräte zur direkten Interaktion am Screen sollten nicht zu groß und zu unhandlich sein. Bei Verwendung z.B. eines Eingabestifts ist es sinnvoll, zu verhindern, dass Stift oder Stiftschatten die freie Sicht auf das aktuell selektierte oder bearbeitete Objekt erschweren.

- **Möglichst keine Moduswechsel.** Moduswechsel (wie z.B. das Wechseln vom Zoom- zum Edit-Modus) sollten vermieden werden. Die Aufmerksamkeit des Nutzers ist normalerweise nicht ständig dem interaktiven Display zugewandt, er führt evtl. gleichzeitig eine Unterhaltung mit anderen Leuten im Raum. Gemeinsame Arbeit am Screen, besonders am Anfang eines Projekts, benötigt freie Interaktion zwischen Nutzern und Geräten. In kreativen Meetings gibt es sowohl Phasen der Einzelarbeit als auch Phasen der Kollaboration [2]. Zwischen beiden Phasen wird häufig hin- und hergewechselt. Teilnehmer tauschen Informationen und Zwischenergebnisse aus, indem sie Daten auf dem großen Screen präsentieren. Ein Moduswechsel, bei dem der Nutzer sich den aktuellen Stand einer komplizierten Kommandoabfolge merken muss, verhindert den Wechsel zwischen diesen Phasen und unterbricht den Workflow.
- **Geordneter Screen.** Es besteht ein Unterschied zwischen dem Aussehen einer Wand, an der mit Hilfe von PostIts™ oder Zetteln ein Projekt geplant wird und einem Monitor oder Screen mit Graphical User Interface (GUI). Bei einer Planungswand besteht der relevante Inhalt nur aus den vom Nutzer befestigten und beschrifteten Zetteln. Bei einer GUI werden zusätzlich noch viele verschiedene “widgets” angezeigt. Widgets sind Fensterumrandungen, Titel-Anzeigen, Scrollleisten, Werkzeugleisten, Buttons zu Fenstersteuerung, etc. Es ist also sinnvoll, den sichtbaren Bereich des Displays für Inhalte zu reservieren und die Ablenkung durch sichtbare Interaktionshilfen zu minimieren. Eine Lösung ist es, die sichtbare Oberfläche mit Hilfe von halb- und undurchsichtigen “Zetteln” zu organisieren. Diese können unabhängig voneinander erzeugt, bearbeitet und bewegt werden. Diese “Zettel” oder “Sheets” enthalten nur Informationen, keine zusätzlichen Widgets sind notwendig [3].

3 Technologien

Im folgenden werden die verschiedenen technologischen Möglichkeiten zur direkten Interaktion bei besonderer Beachtung ihrer Eignung zur gemeinsamen Arbeit am Display dargestellt. Die Arbeit trennt zwischen direkter Interaktion mit Finger, Hand und Armen und direkter Interaktion mit Eingabegeräten wie einem Stift.

Für Ben Shneiderman bedingt direkte Interaktion [15] kontinuierliche Darstellung des aktuellen Objekts von Interesse, physikalische Aktionen oder das Drücken eines beschrifteten Schalters anstelle einer komplexen Eingabe-Syntax sowie schnell und schrittweise umkehrbare Operationen, deren Auswirkungen auf das Objekt von Interesse sofort sichtbar sind.

3.1 Interaktion mit der Hand bzw. dem Finger

Die Eingabe, bei der die Hand bzw. der Finger direkt auf ein Objekt auf dem Screen zeigt oder es berührt, ist die natürlichste und direkteste Form der Interaktion.

Die Bewegungen der Hand bzw. des Fingers auf dem Screen können auf verschiedene Arten elektronisch registriert werden.

Resistive (widerstandsgesteuerte) Systeme: Ein Touchscreen, Tastschirm bzw. Sensorbildschirm ist ein Computerbildschirm mit Bereichen, durch deren Berührung der Programmablauf gesteuert werden kann. Touchscreens finden als Info-Monitore, z.B. auf Messen, zur Orientierung in großen Kaufhäusern oder für die Fahrplanauskunft auf Bahnhöfen Verwendung. [5]

Bei Berührung einer Stelle des Touchscreens kommen sich zwei leitfähige Schichten im Touchscreen, die mit konstanter Gleichspannung angesteuert werden, näher. Dadurch entsteht ein Kontakt; durch den Widerstand dieses Kontakts entsteht an jeder Stelle eine unterschiedliche Spannung. Die Spannungsänderung kann dann zur Bestimmung der x- und y-Koordinate bestimmt werden. (siehe Abb.2)

Touchscreens haben den Vorteil, die Interaktion sehr natürlich erscheinen zu lassen, da der Nutzer das Objekt, das er bearbeiten oder selektieren möchte, direkt auf dem Screen berührt. Die Interaktion per Touchscreen ist einfach und schnell zu erlernen, ein Touchscreen verbraucht keinen zusätzlichen Platz für Eingabegeräte, es gibt keine beweglichen Teile und die Technik ist sehr widerstandsfähig..

Eine Studie von Sears und Shneiderman [6] zeigt, dass die Eingabe per Touchscreen zwar schneller als mit der Maus, aber auch weniger genau ist. Bei der Selektion von besonders kleinen Objekten können Techniken wie "Pulling" (der Cursor wird zum Objekt gezogen, wenn er nah genug daneben platziert wurde) oder "Zooming" (kleine Objekte werden bei Annäherung des Cursors vergrößert) helfen.

Die Ermüdung des Arms stellt bei allen Displays, die an der Wand montiert sind und mit dem Finger bedient werden, ein Problem dar. Kleinere Bildschirme können in einem flacheren Winkel montiert werden. Die Befürchtung, dass die Verschmutzung des Screens zu Problemen führen könnte, hat sich in der Studie [6] nicht bewahrheitet.

Beim herkömmlichen Touchscreen ist die Eingabe mit mehr als einem Finger nicht möglich, für die Arbeit in der Gruppe an nur einem Screen ist er also ungeeignet.

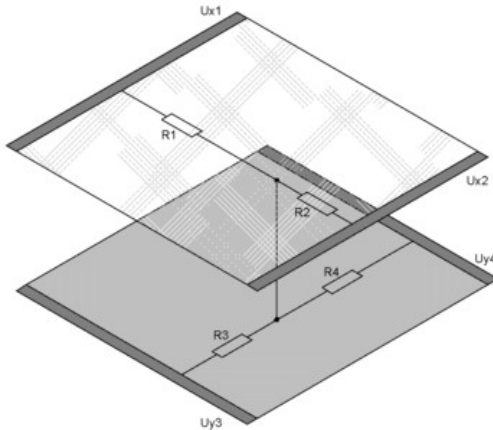


Abbildung 2: Beispiel eines resistiven Touchscreens: Bei Berührung einer Stelle des Touchscreens kommen sich zwei leitfähige Schichten im Touchscreen, die mit konstanter Gleichspannung angesteuert werden, näher. Dadurch entsteht ein Kontakt; durch den Widerstand dieses Kontakts entsteht an jeder Stelle eine unterschiedliche Spannung. Die Spannungsänderung kann dann zur Bestimmung der x- und y-Koordinate verwendet werden.

Kapazitive Systeme: Kapazitive Sensoren arbeiten auf Basis der Veränderung der Kapazität eines einzelnen Kondensators oder eines ganzen Kondensatorsystemes. Die Kapazität ist ein Maß für das Ladungs-Fassungsvermögen eines Körpers.

Eine Variante des kapazitiven Sensings stellt die Sensor-Architektur SmartSkin [7] dar. Hierbei wird die Position der Hand mit einem Netz aus Sender- bzw. Empfänger-Elektroden gemessen. (siehe Abb.3) Jeder Schnittpunkt zwischen den Elektroden wirkt als (sehr schwacher) Kondensator bzw. Ladungsspeicher. Kommt ein (auch schwach) leitfähiger und geerdeter Gegenstand (z.B. eine Hand) in die unmittelbare Nähe eines solchen Schnittpunkts, werden die Elektroden gekoppelt und die gespeicherte Ladung kann abfließen. Das von der Empfänger-Elektrode gemessene Signal wird schwächer. Durch diesen Effekt kann man die Annäherung eines oder mehrerer Objekte registrieren.

Anders als beim Touchscreen ist es beim SmartSkin möglich, zwischen den Zuständen “mouse-over” und “mouse-press” zu unterscheiden, die Sensoren registrieren auch den Abstand des Fingers. Die Interaktionstechnik “Shape Based Manipulation” führt dazu, dass Objekte von Hand und Arm “abgestoßen” werden, man kann also einen Stapel Fotos vom Tisch “wischen“. Die Interaktion erscheint natürlicher als die mit der Maus.(siehe Abb.4) Auch diverse Gesten werden verwendet.

Der größte Vorteil von SmartSkin zu herkömmlichen Maus-basierten Systemen ist aber die Möglichkeit der Bedienung mit mehreren Händen bzw. mehreren Nutzern.

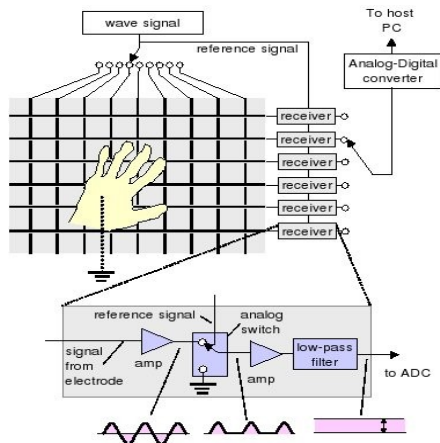


Abbildung 3: SmartSkin Sensortechnik: Netz aus Empfängern und Sendern wird verwendet, um die Handposition zu berechnen

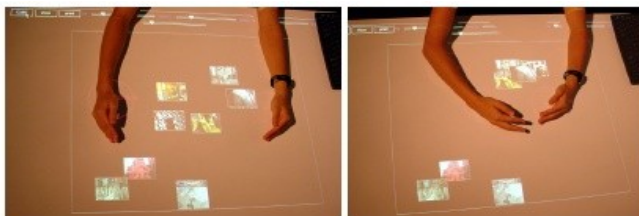


Abbildung 4: "Shape Based Manipulation" erlaubt es, mehrere Objekte ohne direkte Berührung zu manipulieren.

Das SmartSkin ist unempfindlich gegen schlechte Lichtverhältnisse und verbraucht, ähnlich wie ein Touchscreen, kaum zusätzlichen Platz. SmartSkin funktioniert ohne zusätzliche Eingabegeräte. Die Genauigkeit der Positionserkennung ist aber (in der Bauweise des Prototypen 1) relativ gering.

Ein weiteres Multi-User-System, das kapazitiv arbeitet, ist das DiamondTouch [8] der Mitsubishi Electric Research Laboratories. Beim DiamondTouch werden bei Berührung schwache elektrische Signale von der Oberfläche abgegeben. Die Impulse wandern durch den Körper des Nutzers und werden von einer Empfängereinheit aufgenommen, die dem entsprechenden Nutzer zugeordnet ist. (siehe Abb.5) Diese eindeutige Zuordnung ist beim oben genannten SmartSkin nicht möglich.

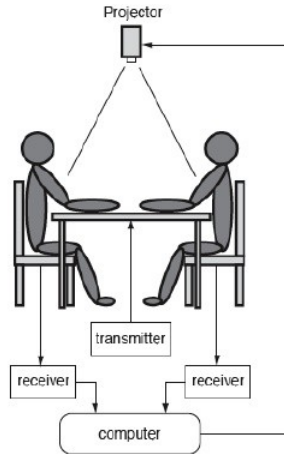


Abbildung 5: Signale gehen über den Körper zu einem Empfänger, der dem jeweiligen Nutzer eindeutig zugeordnet ist.

Das DiamondTouch-System versucht folgende Vorgaben zu erfüllen: mehrere gleichzeitige Berührungen sollen zuverlässig erkannt und den jeweiligen Benutzern eindeutig zugeordnet werden. Objekte, die auf der Oberfläche stehen, sollen die Eingabe nicht beeinträchtigen. Das System soll günstig in der Herstellung und sehr haltbar sein. Keine zusätzlichen Eingabegeräte sind nötig.

Optische Systeme: Eine relativ neue Entwicklung im Bereich der Touchscreen-Technologie stellt die Positionsbestimmung mit Hilfe von Kameras dar. Zwei oder mehr Bild-Sensoren sind in den Ecken des Screens platziert. Auf den gegenüberliegenden Bildschirmseiten befinden sich Infrarot-LEDs. Berührt ein Gegenstand den Screen, wird das als Schatten oder dunkle Stelle im Infrarotbild der Kamera deutlich. Durch Triangulation kann dann die Position des Fingers oder der Hand bestimmt werden. [5] Durch die Größe der LEDs ist nur eine begrenzte Auflösung möglich.

Eine kommerzielle Anwendung dieses Prinzips steckt im Digital Vision Touch (DViT) der Firma SmartTech. In allen vier Ecken des rechteckigen Rahmens befindet sich je eine Kamera, die ständig die Oberfläche aufnimmt. (siehe Abb.6)

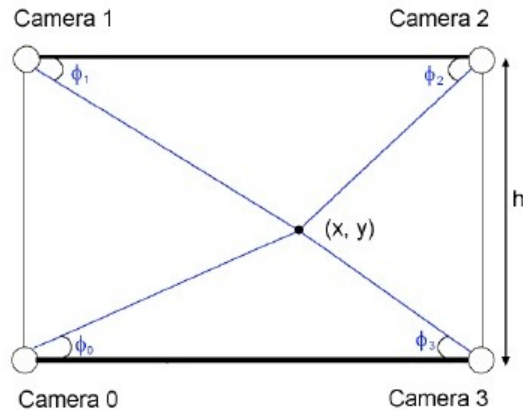


Abbildung 6: Der Abstand zwischen zwei Kameras und ihr jeweiliger Blickwinkel aufs Objekt ermöglichen eine Berechnung des Kontaktpunkts.

Die Framerate der Kameras beträgt hier 100fps. Eine genaue Positionsbestimmung sowie die Erkennung von mehr als einem Kontaktpunkt ist durch die Verwendung von 4 (statt 2) Kameras möglich. [9]

SAW (Surface Acoustic Wave): Die so genannten (Schall)wellen-gesteuerten Systeme basieren auf der Ausbreitungsgeschwindigkeit von Schall. Hier werden piezoelektrische Mikrofone an den Ecken einer (meist rechteckigen) Oberfläche positioniert.

Die Mikrofone messen die auf der Oberfläche auftretenden Schallwellen, die durch Klopfen o.ä. entstehen. Ein digitaler Signal-Prozessor extrahiert relevante Charakteristika wie Amplitude, Frequenzanteile und unterschiedliche Ankunftszeiten des Schalls an den Mikrofonen. So ist es möglich, z.B. die Position des Klopfens zu bestimmen.[10] Diese Technik bietet *nicht* die Möglichkeit der kollaborativen Arbeit am Display, die relativ ungenaue Positionsbestimmung ist nicht für die Arbeit mit einer grossen, gleichzeitig sichtbaren Menge von Daten geeignet.

3.2 Interaktion durch Eingabegeräte

Stifteingabe am Beispiel des Liveboard-Pens: (siehe Abb. 7) Das Design des Liveboard-Stiftes sollte verschiedene Anforderungen erfüllen: er sollte klein, leicht und gut ausbalanciert sein. Der Stift enthält eine LED und gibt so optische Strahlung ab, die von einem Detector Module bzw. einer Photo-Diode hinter der Bildfläche aufgenommen wird. Dabei beträgt die Abweichungsungenauigkeit weniger als 1mm. Die LED-Leuchtstärke wird moduliert, so ist eine Unterscheidung von mehreren Stiften durch den Computer möglich. Der Stift ist kabellos, bei simultaner Verwendung mehrerer Stiftes kommt es dadurch nicht zu Problemen. Mit dem Stift ist es auch möglich,

eine Eingabe in einigem Abstand zum Screen zu tätigen. Dadurch kann der Nutzer aus der Entfernung auf Objekte auf dem Screen zeigen oder Gesten auszuführen. Der Liveboard-Pen hat drei Knöpfe, um mit bereits vorhandener Software, die durch eine Drei-Tasten-Maus gesteuert wird, arbeiten zu können. [4]

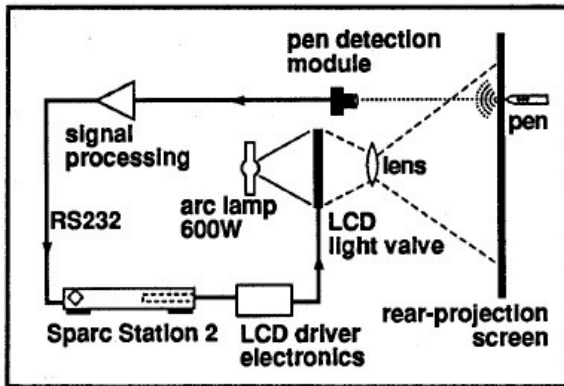


Abbildung 7: Der Stift des Liveboards gibt optische Strahlung ab, die von einem Gerät hinter dem Screen aufgenommen wird. Ein Computer berechnet die Stiftposition.

Ein stiftähnliches Eingabegerät hat den Vorteil, dass die Verwendung jedem Nutzer bekannt ist. Ein Stift ist klein, leicht und bietet durch die Spitze eine hohe Genauigkeit.

Nachteilig ist, dass der Stift verloren gehen kann, im Raum mit dem Screen ständig vor Ort sein muss und evtl. in einem ungünstigen Moment ein Batteriewechsel notwendig wird. Die LiveBoard-Studie zeigt, dass zusätzliche Knöpfe oder Buttons direkt am Stiftkörper ungern und selten genutzt werden.[4]

Beim LumiPoint-System der Stanford University [13] wird ein kabelloser, stiftförmiger Laserpointer direkt auf den Screen aufgesetzt, um Programmverhalten zu steuern. Er kann aber auch aus der Entfernung als Zeige- oder Eingabegerät verwendet werden. (siehe Abb.8)



Abbildung 8: Interaktion mit Laserpointern am von hinten beleuchteten Screen

Der Screen ist 1,8m x 0,6m groß und wird von hinten durch acht Projektoren angestrahlt, so dass die Oberfläche wie ein einziges großes Display erscheint. Kameras filmen die Screen-Fläche und nehmen die Bewegung der hellen Laser-Punkte wahr, wenn ein Benutzer am Bildschirm arbeitet. Durch einen zentralen Rechner werden die Punktbewegungen anhand von zeitlichen und dynamischen Zusammenhängen als Pfade oder Strichbewegungen analysiert.[Abb.9]

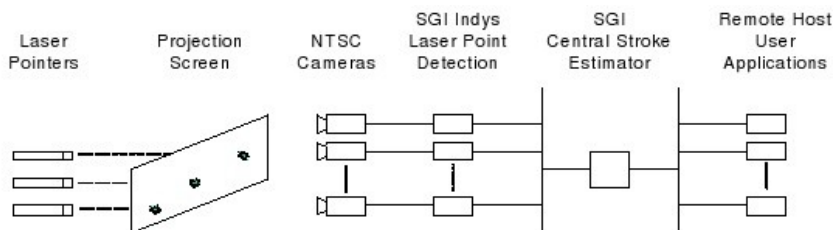


Abbildung 9: Laser-Punkte als Eingabegeräte an der Bildschirm-Wand. Kameras nehmen die Laserpunkte auf. Diese Informationen werden gefiltert und vom Estimator in stroke-events umgerechnet.

Das System ist sehr gut skalierbar, wenn die zu bearbeitende Screenfläche wächst, werden mehr Kameras zur Abdeckung verwendet. Wenn die Screenfläche für die direkte Interaktion zu groß wird, kann der Stift auch aus der Entfernung zur Eingabe verwendet werden.

3.3 Gesten

Gesten sind unverzichtbar und allgegenwärtig in der zwischenmenschlichen Kommunikation. Gesten können verschiedene Funktionen haben: sie unterstreichen das Gesagte, können aber auch ausdrücken, was alleine durch Sprache nicht erklärbar ist. Nicht unmittelbar präsente Objekte oder Ereignisse können erläutert und referenziert werden. Gesten können sich auf das gleiche Thema beziehen wie der gleichzeitig gesprochene Satz, aber einen komplett anderen Aspekt davon repräsentieren. Auch abstrakte Sachverhalte können mit Gesten erläutert werden.

Während im Bereich der Spracherkennung große Fortschritte erzielt worden sind, fehlt es bei der multimodalen Mensch-Computer-Interaktion immer noch an Erkenntnissen und geeigneter Sensing-Technologie, um *natürliche* Gestiken zu erkennen. Die Verwendung von Geräten wie einem Datenhandschuh oder einem Stift zur Eingabe von bereits vordefinierten Gesten führt zu unnatürlichen und zwanghaften Bewegungen. [11]

Schon beim Liveboard [4], bei dem der Nutzer einen kabellosen Stift zur Eingabe benutzt, wurden bestimmte vordefinierte Gesten wie eine Aufwärts-Abwärts-Bewegung verwendet um verschiedene Funktionen einer Präsentation aufzurufen wie z.B. das Scrolling. Im Roomware-Projekt [2] werden (vordefinierte) Gesten zum Beispiel dafür verwendet, einen Arbeitsplatz oder einen Hyperlink zu erzeugen. Durch Gestiken wie "Dragging" von Objekten wird ein Bezug zur echten Welt hergestellt (Objekte werden gezogen) und außerdem ein flüssigeres Arbeiten als im point-and-click Interface gewährleistet.

Eine Weiterentwicklung des SmartSkin [7] mit dichter gelegtem Sensorgitter zeigt die Möglichkeit einer interaktiven Oberfläche zur Manipulation virtueller Objekte. Die Erkennung der Handposition ist hier viel genauer, mehr als eine Fingerposition wird erkannt. Dadurch sind einige neue Interaktionstechniken möglich. (vgl. Abb. 10) Ein gleichzeitiges Manipulieren mehrere Punkte (wie zum Beispiel das Bewegen von Kurven-Kontroll-Punkten in der 3D-Gestaltung) wird einfacher und effizienter. Als Geste wurde hier das Zusammenführen zweier Finger ("Greifen") interpretiert. Die gegenteilige Bewegung lässt ein virtuelles Objekt wieder fallen. Eine andere Möglichkeit ist ein Vergrößern von Bildern relativ zum Abstand zwischen Daumen- und Zeigefinger.

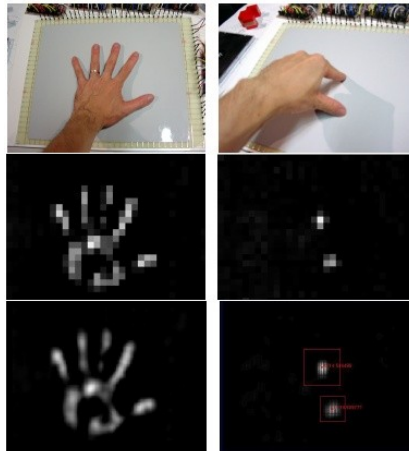


Abbildung 10: SmartSkin - Gesten und zugehörige Sensorwerte

Beim Flatland-Interface [12] besteht die Möglichkeit, mit dem Stift “Metastrokes” zu zeichnen und somit Gesten zu erzeugen, die sowohl das visuelle Layout des Whiteboards als auch Objektverhalten definieren. Auch Menüs werden durch Gesten aufgerufen. (vgl. Abb.11)

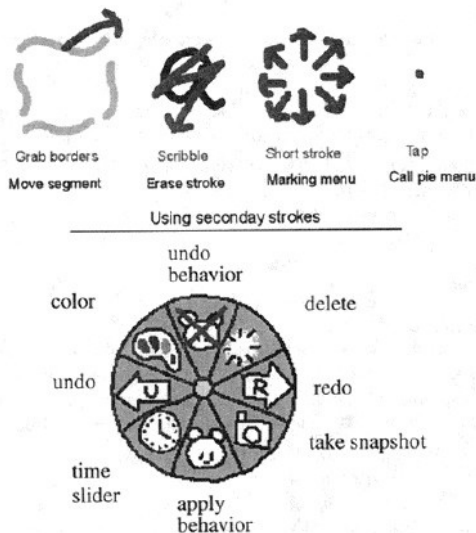


Abbildung 11: Gesten und Menüs bei Flatland

Gesten werden also hauptsächlich mit vordefinierter und eingeschränkter Syntax in der Mensch-Computer-Interaktion verwendet, sind in dieser Form aber sinnvoll, um längere Arbeitsschritte abzukürzen oder komplizierte, mehrteilige Eingabe-Operationen zusammenzufassen bzw. zu vereinfachen.

3.4 Spezielle Widgets

Widgets sind Fensterumrandungen, Titel-Anzeigen, Scrollleisten, Werkzeugleisten, Buttons zu Fenstersteuerung, etc. In [3] wird als grundlegendes Designziel der Clean-Screen genannt. Control-Widgets, die auf dem Screen eines Desktoprechners verwendet werden, um komplizierte, mehrteilige Kommandos auszuführen, haben auf einem Whiteboard-ähnlichen Screen, der im Team und mit weniger fokussierter Aufmerksamkeit genutzt wird, keinen Platz. Der sichtbare Bereich sollte für Inhalte reserviert bleiben, der Nutzer sollte nicht abgelenkt werden.

Bei dieser Studie wird das FlowMenu verwendet, um eine klare Trennung zwischen Kommandoingabe und Inhaltserzeugung zu gewährleisten (siehe Abb.12). Umgeschaltet wird zwischen diesen beiden Modi durch Druck auf einen Meta-Button direkt am Stift.

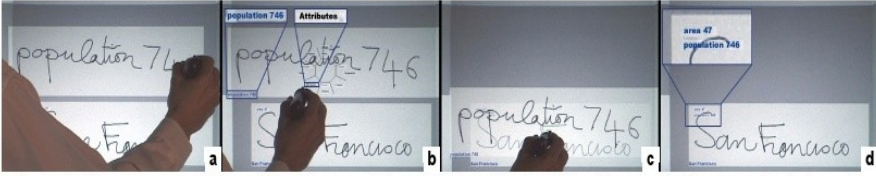


Abbildung 12: Eine Folge von Aktionen, um dem Zettel "San Francisco" ein Attribut zuzuweisen: Parameter/Wert-Paar schreiben

- a) Flow-Menü aufrufen und "item...->attribut auswählen" (Handschriftenerkennung)
- b) Drag+Drop des neuen Zettels aufs Ziel
- c) Attribut des Ziel-Zettels wird aktualisiert

Widgets werden bei der gemeinsamen Arbeit an einem großen Screen also anders genutzt als Widgets am Desktoprechner. Auch das LiveBoard-Projekt [4] zeigte, dass feste Buttons auf einem großen Bildschirm schwer zu lokalisieren und oft nicht leicht zu erreichen sind, da mit dem Arm bzw. der Hand "größere" Entfernungen zurückgelegt werden müssen, um das Ziel, den zu drückenden Knopf, zu erreichen. Es zeigte sich, dass hier entweder die Eingabe durch Gesten oder eben spezielle Widgets wie bewegliche oder sich selbst bewegende Menüs (ähnlich wie in [2]) benötigt werden. Es wurde auch erkannt, dass die Schrift auf dem Screen für die im Raum sitzenden Personen oft zu klein und daher schwer zu erkennen ist.

3.5 Weitere Ansätze

Im Roomware-Projekt des Fraunhofer Instituts Darmstadt [2] wurde u.a. ein Lokalisationsprinzip beim gemeinsamen Arbeiten am großen Screen definiert: Je lokaler die Einflussnahme eines Benutzers ist, desto unwahrscheinlicher ist ein Konflikt mit einer Aktion eines anderen, zur gleichen Zeit am Bildschirm arbeitenden Nutzers.

Die User-Interface-Software bietet die Möglichkeit evtl. auftretende Lokalisierungsprobleme zu verhindern. Eine besondere Schwierigkeit entsteht hierbei aus dem Prinzip der „shared documents“ oder einem verteilten Dateisystem. Die Anzeige muss auch bei gleichzeitigen gegensätzlichen Änderungen stets konsistent sein. Ermöglicht wird das durch die softwareseitige Unterstützung von kombinierten Screens, die durch verschiedene Rechner angesteuert werden. (siehe Abb.13)

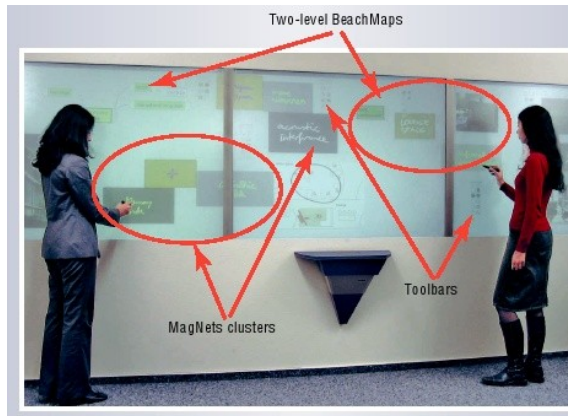


Abbildung 13: Nutzung der Roomware DynaWall. Nutzer arbeiten parallel am selben Dokument. Die Screenfläche besteht aus drei Teilen, die jeweils durch separate Rechner angesteuert werden.

„Scribble Groups“ sind ein Beispiel für lokalisierte Interaktionsmodi, da Notizen und Eingaben, die zusammen erzeugt wurden, auch zur späteren gemeinsamen Bearbeitung verbunden bleiben. Außerdem wird dadurch ein globaler Moduswechsel unterbunden bzw. vermieden, da in einer „Scribble Group“ getätigte Stifteingaben nicht als Gesten sondern als Zeichnung interpretiert werden. „Löschen“-Gesten z.B. werden in einer „Scribble-Group“ nicht als Befehl zum Löschen, sondern als Scriblings/Kritzeleien gedeutet.

Das Roomware-Projekt führt auch interessante Interaktionsformen und Möglichkeiten zum Tauschen und Verschieben von Informationseinheiten (auch über Systemgrenzen hinweg) ein. Objekte können nicht nur gezogen, sondern auch „geworfen“ sowie „gegriffen“ und auf einem anderen Screen wieder „abgelegt“ werden. Die Software des Roomware-Projekts unterstützt auch das Drehen des Arbeitsplatzes auf dem InteracTable (berührungsempfindliches Plasma-Display in Tischform). Mehrere Nutzer haben so einen guten Blick auf den Screeninhalt. Eine Analogie zum einfachen Drehen eines Papierzettels wird so hergestellt. Wieder wird ein globaler Moduswechsel vermieden, jeder Nutzer kann in einem jeweils eigenen Interaktionsmodus am Tisch arbeiten.

Als Beispiel für kreative Möglichkeiten, Ideen interaktiv zu strukturieren, können die drei „Creativity Tools“ des Roomware Projekts genannt werden: MagNets, Beach-Map und PalmBeach. MagNets basiert auf der Idee von virtuellen magnetischen Kärtchen, die sich entweder abstoßen, um so Überlappung auf dem Screen zu verhindern oder sich zusammenhängen, wenn der Nutzer sie übereinanderzieht, um eine thematische Zusammengehörigkeit zu definieren.

BeachMap erweitert diese Idee hin zu einer hierarchischen Ansicht von Informationen mithilfe von Graphen. MagNets können mit Hilfe von Gesten in die BeachMap-Strukturen eingebunden werden. Diese beiden Features sind für die Ideen-Strukturierung gedacht und dienen dazu, große Mengen von Informationen sinnvoll zu strukturieren. (siehe Abb.14)



Abbildung 14: Interaktion mit Hilfe von MagNets und BeachMaps auf der DynaWall des RoomwareProjekts.

4 Zusammenfassung

Diese Arbeit versucht, Anforderungen an direkte Interaktionstechniken für große interaktive Displays zu nennen. Techniken zur Interaktion werden genannt und beschrieben und auf ihre Eignung für die gemeinsame Interaktion am Screen untersucht.

Direkte Interaktionstechniken kann man in zwei Kategorien einteilen: aktive und passive Techniken. Passive Systeme beruhen auf Technik in der Oberfläche des Screens, die auf Berührung reagiert.

Ich habe in diesem Bereich mehrere resistive (Touchscreen) und kapazitive (SmartSkin, DiamondTouch) Systeme genannt. Besonders die Touchscreens sind mittlerweile weit verbreitet, eignen sich aber nicht für die gemeinsame Arbeit am Screen, da sie nur eine Fingerberührung erkennen.

Mit dem SmartSkin[7] sind neue mehrhändige Eingabeformen wie das "Wischen" von Objekten möglich. Auch Arme und andere Körperteile können in Echtzeit zur Eingabe von Gesten genutzt werden. Interessant ist hier auch die Fähigkeit zur Entfernungsmessung zur Oberfläche. Die Interaktionstechnik ist schnell und einfach zu erlernen. Eine Anwendung in einer realistischen Arbeitsumgebung mit entsprechender Software wäre hier sehr interessant.

Das DiamondTouch[8] hat die Möglichkeit, mit mehreren Personen in Echtzeit an einem Screen zu arbeiten. Der Vorteil ist hier, dass die Fingerberührung den Nutzern jeweils eindeutig zugeordnet werden kann, das ist sehr sinnvoll für die gemeinsame Arbeit an einem Projekt oder die personenspezifische Aufzeichnung von Arbeitsabläufen.

Auch optische Interaktionstechnologien wie das SmartTech DViT erlauben die gleichzeitige Eingabe von mehreren Benutzern in Echtzeit mit guter Genauigkeit.

Bei aktiven Systemen verwendet der Nutzer ein Eingabegerät, wie z.B. einen Stift, als Beispiel sind hier ultraschallgesteuerte, lasergesteuerte (LumiPoint[13]) und elektromagnetische Systeme zu nennen, bei denen die Eingabe mit einem speziellen Gerät, meist mit einem Stift mit Tasten erfolgt.

Diverse neue Widgets wurden für den großen Screen entwickelt, so zum Beispiel sich bewegende oder nur auf Aufruf erscheinende Menüs, allein schon um dem Arm weite Wege zu einem feststehenden Menü zu ersparen.

Als Beispiele für kreative Möglichkeiten, Ideen zu strukturieren und den großen Screen effektiv zu nutzen, habe ich MagNets und BeachMaps von Roomware[2] genannt.

5 Ausblick

Die Entwicklung hin zu immer kleineren (und immer größeren) Computern ist abzusehen, es könnte sein, dass der Rechner in Desktop-Größe mehr und mehr verschwindet. Neue Entwicklungen auf dem Gebiet der Interaktionstechniken entstehen. Größere Bildschirme werden günstiger und bieten viele neue Möglichkeiten der Nutzung und Interaktion. Bei der kollaborativen Arbeit steht der große Screen an einer festen Stelle, die Nutzer aber bewegen sich im Raum, interagieren miteinander und mit kleinen, mobilen Eingabegeräten. [16] Es wird versucht, die Interaktion mit Rechnern natürlicher und der echten Welt ähnlicher zu machen, auch um bereits existierende menschliche Fähigkeiten und Gesten (wie zeigen, greifen, Objekte bewegen) auszunutzen.

Besonders bei der gemeinsamen Arbeit am großen interaktiven Screen in der Brainstorming- bzw. Meetingsituation - hier steht die soziale Zusammenarbeit im Mittelpunkt - macht es Sinn, die Eingabetechnik natürlich zu gestalten, um die zwischenmenschliche Interaktion und den Workflow nicht zu unterbrechen.

Interessante Wege, direkt mit dem Körper virtuelle Objekte zu manipulieren und gemeinsam kreative Arbeit zu leisten, werden die direkte Interaktion mit großen interaktiven Bildschirmen auch in Zukunft zu einem faszinierenden Thema machen. Große Displays werden dann evtl. nicht mehr als Rechner wahrgenommen, sondern als Whiteboards mit fast unendlichen Möglichkeiten.

Literatur

1. Mary Czerwinski, Greg Smith, Tim Regan, Brian Meyers, George Robertson and Gary Starkweather, Toward Characterizing the Produktivity Benefits of Very Large Displays, Human-Computer Interaction – INTERACT'03, M.Rautenberg et al.(Eds.), Published by IOS Press,(c) IFIP, pp. 9-16, 2003.
2. Thorsten Prante, Norbert A. Streitz, Peter Tandler, Roomware: Computers Disappear and Interaction Evolves, Computer Vol. 37, No.12, pp. 47-54, December 2004
3. Guimbretire, F. Stone, M. and Winograd, T. Fluid Interaction with High-resolution Wall-size Displays, in Proceedings of UIST 2001, 21-30, ACM Press, 2001
4. Elrod, S., et. al. Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In Proc. CHI'92, pp. 599--607.
5. Wikipedia, the free Encyclopedia: Touchscreen (2006)
6. A. Sears and B. Shneiderman, High Precision Touchscreens: Design Strategies and Comparison with a Mouse, International Journal of Man-Machine Studies, vol. 43, no. 4, pp. 593-613, April 1991.
7. J. Rekimoto, SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces, Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 113-120, Minneapolis, USA, 2002.
8. P. Dietz, D. Leigh, DiamondTouch: a Multi-User Touch Technology, Proceedings of the 14th annual ACM symposium on User interface software and technology, Orlando, Florida, Tactile user interface, pp. 219 – 226, 2001.
9. http://www.smarttech.com/dvit/DViT_white_paper.pdf
10. Joseph A. Paradiso, Che King Leo, Tracking and characterizing knocks atop large interactive displays, Sensor Review, Emerald Group Publishing Limited, pp.134-143, 2005
11. S. Kettebekov, R. Sharma, Toward Natural Gesture/Speech Control of a Large Display, Engineering for Human-Computer Interaction (EHCI'01),Toronto, Canada. May 11-14, 2001. Lecture Notes in Computer Science, Springer Verlag. 14 pages
12. E. Mynatt et al, Flatland: New Dimensions in Office Whiteboards, Conference on Human Factors in Computing Systems, Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, Pittsburgh, Pennsylvania, United States, pp.346 – 353, 1999
13. X. Chen and J. Davis. LumiPoint: Multi-User LaserBased Interaction on Large Tiled Displays. Technical report, Stanford University, 2001.
14. The Mimio electronic white board <http://www.mimio.com>
15. Ben Shneiderman, Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces, Proc. ACM International Workshop on Intelligent User Interfaces '97, ACM, New York, NY, pp33-39. 1997
16. Robert J.K. Jacob, Human-Computer Interaction: Input Devices, ACM Computing Surveys, Vol.28, No.1, 1996

Kollaboration mittels entfernter Interaktion mit großen Displays

Stefan Seitz

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
stefan@seitz-muenchen.de

Zusammenfassung In der vorliegenden Arbeit werden zunächst unterschiedlichste Merkmale für indirekte Interaktionstechniken genannt, insbesondere die verwendeten Geräte und Funktionen, Verfahren und Technologien sowie die Synchronität der Interaktion. Anschließend werden Ergebnisse ausgewählter, relevanter Forschungsprojekte vorgestellt und erläutert. Die Arbeit schließt mit einem Ausblick.

1 Einleitung

Anschließend an die zwei vorausgegangenen Arbeiten, die sich mit der Zusammenarbeit mehrerer Personen über große Displays sowie mit direkter Display-Interaktion befassen, konzentriert sich der vorliegende Text auf indirekte Interaktionstechniken für die Zusammenarbeit mittels großer Displays. Dabei wird nicht eigens untersucht, welche Arten von Displays es gibt und auch nicht, welche Möglichkeiten zur Zusammenarbeit generell existieren. Vielmehr werden relevante Interaktionstechniken aus Forschung und Entwicklung vorgestellt, die für die Zusammenarbeit an großen Displays entworfen wurden oder für eine solche zumindest nicht von vorne herein ungeeignet erscheinen.

Bei der Vorstellung von Forschungsprojekten und -ergebnissen wird keinesfalls Anspruch auf Vollständigkeit erhoben. Das gesteckte Ziel besteht im Aufzeigen der Vielfalt der aus der Forschung hervorgegangenen Arten von Interaktionstechniken, wofür in Kapitel 3 einige herausragende Beispiele herausgegriffen und näher erläutert werden. Eine Grundlage für die Betrachtung dieser Beispiele soll der zunächst in Kapitel 2 folgende Überblick schaffen.

2 Überblick über Merkmale indirekter Interaktionstechniken

Der an dieser Stelle gegebene Überblick über Geräte, Funktionen, Verfahren, Technologien und weitere Charakteristika soll in erster Linie einen Eindruck von den wichtigsten Blickwinkeln und Betrachtungsweisen für indirekte Interaktionstechniken vermitteln. Die gefundenen Merkmale und Kategorien stellen zudem ein nützliches

Hintergrundwissen bei der Betrachtung konkreter Forschungsprojekte in Kapitel 3 dar.

2.1 Nutzerseitig eingesetzte Geräte und Funktionen

Aus Sicht des Nutzers könnte das Hauptaugenmerk zunächst weniger auf die zugrundeliegenden Interaktionskonzepte, die zur Realisierung verwendete Display-seitige Hardware oder die Software-Protokolle fallen. Stattdessen wird er vermutlich als erstes bemerken, dass die neue Form der Interaktion den Gebrauch seines Handys erfordert, oder dass das Absenden einer SMS eine Reaktion auf dem Display zur Folge hat.

Als besonders vielseitiges und mobiles Gerät, das zudem fast jeder ständig mit sich führt, ist zunächst das Mobiltelefon zu nennen, das in zahlreichen Projekten (vgl. [1, 2, 6, 8]) auf unterschiedliche Weise eingesetzt wird. Daneben steht eine Vielzahl weiterer mobiler Eingabegeräte zur Verfügung: Das Dynamo-Display [10] beispielsweise erlaubt die Verbindung mit PDAs, Notebook-Computern und USB-Sticks. Andere Systeme greifen auf Laserpointer als "verlängerten Arm" des Benutzers zurück (vgl. [13, 14], erwähnt auch in [15]). Über diese mobilen Interaktionswerkzeuge hinaus gibt es natürlich immer auch Maus und Tastatur (ebenfalls verwendet im Dynamo Project [10]).

Der Vollständigkeit halber sei auch auf die zahlreichen Erfindungen neuer Eingabegeräte hingewiesen, wie sie u. a. in der Forschungsgruppe „Embedded Interaction“ [22] an der Ludwig-Maximilians-Universität München entstehen. In Abschnitt 3.3 wird beispielhaft ein intuitives Steuerungsgerät vorgestellt, das drahtlos mit einem Display kommuniziert, während der Benutzer lediglich die Lage des Gerätes verändert [12]. Zu den Vorteilen physisch greifbarer Interaktions-Medien sei das kurze Überblicks-Paper von Rowanne Fleck zur Lektüre empfohlen [19].

In manchen Fällen kann man gar nicht mehr von Eingabegeräten sprechen. Dies ist etwa der Fall, wenn menschliche Gesten durch das optische Verfolgen von Reflektoren am Körper erkannt werden [17, 18], oder wenn diese Erkennung sogar ohne Verwendung jeglicher Gegenstände möglich ist [16].

Aus der gewöhnlichen Nutzung von (nicht nur mobilen) Endgeräten wie Handy, PDA oder PC ist der Anwender typischerweise mit bestimmten Funktionen zur Telekommunikation vertraut, die auch zur Interaktion mit Displays eingesetzt werden können. Durch den Einsatz von SMS/MMS, E-Mail, Sofortnachrichten (Instant Messages) oder Web-Technologie können – in beiden Richtungen – Nachrichten und Inhalte übermittelt werden, ohne dass das Erlernen neuer Spezialsoftware erforderlich wird.

Einige Beispiele hierfür werden in [1] genannt, wie der mit einem Display ausgestattete große Würfel in der Eingangshalle der Lissabonner Vodafone-Niederlassung, wo Besucher nach dem Anwählen einer Telefonnummer Spiele auf dem Bildschirm spielen können – alleine oder gegeneinander. Der BBC-Ableger BBCi, ein Anbieter interaktiver Dienste, ermöglicht es Passanten im Zentrum von London, per SMS Fragen an ein im Studio befindliches (und durchs Fenster sichtbares) Display zu senden und damit Einfluss auf den weiteren Verlauf eines Interviews zu nehmen. Ich selbst habe Ende vergangenen bis Anfang diesen Jahres an der Lancaster University an einer

Software zur Steuerung des Hermes Photo Display (Teil des CASIDE Projekts [23]; siehe auch Text [2] sowie Abbildung 1) gearbeitet, die Client-seitig vor allem auf der Nutzung von statischen und dynamischen Webseiten in einem Browser beruht. Eine Veröffentlichung hierzu ist für die kommenden Monate vorgesehen. Das Photo Display erlaubt darüber hinaus in seiner heutigen Konfiguration die Übermittlung von Fotos u. a. per MMS und E-Mail.

Neben der Nutzung bekannter Softwarefunktionen kann bei den heutigen Handys auch der Umgang mit einer eingebauten Digitalkamera und die Navigation mittels eines Mini-Joysticks bei Anwendern vielfach als bekannt vorausgesetzt werden. Beim



Abbildung 1. Testperson vor dem Hermes Photo Display (Foto: Keith Cheverst)

PDA kommt zu den Handy-ähnlichen Funktionen noch das berührungssensitive Display, das zusammen mit dem zugehörigen Stift als Mausersatz dienen kann. Benutzer von PDAs vieler Hersteller sind darüber hinaus normalerweise auch mit „Graffiti“ (vgl. [3]) vertraut, einem System zur Erkennung von handgeschriebenen Buchstaben, das auch in anderen Kontexten, etwa mit Laserpointern wie in [14], eingesetzt werden kann.

2.2 Verfahren und Technologien

Die digitalen Übertragungstechniken bei mobilen Eingabegeräten interessieren insofern, als dass man vorhandene Protokolle einsetzen möchte, insbesondere dann, wenn sie bereits ausgiebig (von anderen) getestet und für tauglich befunden wurden.

GPRS und UMTS sind, wie die Mobiltelefonie generell, abhängig von der Netzabdeckung und werden eingesetzt, um Internet-Verbindungen über den jeweiligen Netzbetreiber herzustellen. Beispielsweise verwenden die Testpersonen in [6] ein Mobiltelefon mit einer Software, die eine TCP/IP-Verbindung über GPRS zum Display-Server herstellt, um bei einer Live-Abstimmung in einem Lokal ihren Wunsch zu übermitteln. Ein wesentlicher Nachteil dieser Verbindungsart ist, dass der Benutzer normalerweise für das übertragene Datenvolumen beim Netzbetreiber bezahlen muss, und dass die hierdurch anfallenden Kosten auch nicht von vorne herein genau beziffert werden können. Auch wird manchen Benutzern nur schwer vermittelbar sein, warum sie für eine Verbindung zu einem Server, der wie in [6] im selben Raum steht, überhaupt Gebühren zahlen sollen.

Für den Benutzer kostenneutral ist der Aufbau einer Bluetooth-Verbindung. Bei meiner (bereits oben unter 2.1 erwähnten) Arbeit am Hermes Photo Display habe ich diese Technologie in zwei Konstellationen getestet: Einmal wurden HTTP-Verbindungen über einen Bluetooth-Proxy aufgebaut, der auf dem Handy lokal im Hintergrund lief. Die Übertragung von Webseiten an den installierten Web-Browser funktionierte ohne auffällige Verzögerungen – Surfen ohne teuren Mobil-Provider ist also zumindest technisch möglich. Die Steuerung des Displays war damit jedoch nicht schnell genug: Das Zwischenschalten des Servlet-Containers auf dem Server hatte zur Folge, dass das Display gelegentlich erst mit über 20 Sekunden langer Verzögerung auf übertragene Befehle reagierte. Probleme dieser Art tauchten nicht mehr auf, als in der zweiten Versuchsanordnung ein Java-MIDlet [4] verwendet wurde, das per OBEX [5] über Bluetooth mit dem Server kommunizierte – hier erfolgte die Reaktion des Displays stets sofort. Nachteile von Bluetooth nach heutigem Stand sind (sofern notwendig) das umständliche Pairing (Prozedur beim Verbindungsaufbau), die mitunter geringe Reichweite (in ungünstigen Fällen nur etwa zehn Meter innerhalb von Gebäuden) sowie die auf sechs Geräte begrenzte Anzahl aktiver Clients, die in einem sogenannten „Pico-Netz“ mit einem Server verbunden sein können.

In Konkurrenz zu Bluetooth befindet sich mittlerweile die drahtlose Netzwerk-Technologie WLAN, die seit einiger Zeit auch in Mobiltelefonen zu finden ist (vgl. bereits Artikel der Zeitschrift *c't* anlässlich der CeBIT 2004 [7] oder aktuelle Nokia-Modellübersicht unter [25]). Auch Scheible und Ojala [6] erwägen deren Einsatz.

Die annähernde Bedeutungslosigkeit der IrDA-Infrarot-Schnittstelle ist neben der Notwendigkeit einer Sichtverbindung vermutlich der Begrenzung auf einen geringen Einsatzradius sowie der mangelnden Mehrbenutzer-Fähigkeit geschuldet.

Optische Verfahren können im Endgerät des Anwenders oder auf Seiten des Displays bzw. verbundener Computer zum Einsatz kommen. Soll die Optik durch den Anwender selbst benutzt werden, so greift man normalerweise auf eine im Mobiltelefon eingebaute Digitalkamera zurück. Mit dieser können – trotz schwacher Rechenleistung der mobilen Geräte – schon Bewegungen des Handys wahrgenommen oder im Bild befindliche Marker erkannt werden (vgl. hierzu [9, 10], Erläuterung in Abschnitt 3.1). Andere Verfahren kommen zum Einsatz, wenn die Erkennung Displayseitig von einer fest installierten Kamera aus erfolgt. In Kapitel 3 finden sich gleich mehrere Beispiele, abhängig von dem Ziel, das erkannt werden soll:

- Schatten oder Laserpunkte/-bewegungen auf dem Display
- Farbige und/oder sich bewegende Punkte im Raum (Paddel, Handy mit blinkendem Display)
- Körperliche Bewegungen (grobe Bewegung einer Menge von Menschen, Bewegungen einer einzelnen Person oder von Körperteilen)

Auf heimischen und Büro-PCs sind schon seit mehreren Jahren ausgereifte kommerzielle Spracherkennungssysteme anzutreffen, die vorwiegend zum Diktat von Texten verwendet werden, so beispielsweise die Produkte IBM ViaVoice und Dragon NaturallySpeaking. Auch funktionieren selbst einfache Spracherkennungswerkzeuge wie das Microsoft Speech SDK für Visual Basic sehr robust, sofern ein begrenzter, definierter Wortschatz zugrundegelegt wird, wie er zur Befehlserteilung an den Computer denkbar ist. Dennoch ist zu beobachten, dass in den meisten Projekten zur Display-Interaktion die Sprache allenfalls ergänzend zum Einsatz kommt. Dies könnte an verschiedenen Einschränkungen liegen: So kann nie mehr als eine Person sprechen,

die gleichzeitige Befehlerteilung durch mehrere Benutzer ist also nicht möglich. Auch sorgen Hintergrundgeräusche ab einer gewissen Lautstärke für fehlerhafte und dadurch unbrauchbare Resultate, wodurch ein Einsatz an belebten Orten problematisch wird. Sieht man von Schwächen seitens des Systems ab, so bleibt noch der Nachteil, dass etwa in Großraum-Büros die Bedienung per Sprache alle übrigen Anwesenden in ihrer Arbeit stören könnte. Immerhin wird die Spracherkennung als alternative Eingabemethode von Sang Chul Ahn et al. [18] bzw. zur zukünftigen Erweiterung des Systems von Sparacino et al. [16] aufgegriffen.

2.3 Weitere Aspekte

Ein weiteres wichtiges Kriterium ist die Synchronität der Interaktion. Synchrone Interaktion bedeutet, dass auf die Aktion (Eingabe) des Benutzers innerhalb einer vorgegebenen, in der Regel sehr kurzen Zeitspanne eine Reaktion des Gerätes – hier: des Displays – erfolgt, so dass beide Vorgänge scheinbar gleichzeitig stattfinden. Ein Beispiel hierfür ist die Computermaus: Der Benutzer bewegt sie und kann zugleich die entsprechende Bewegung des Zeigers auf dem Bildschirm verfolgen.

Existiert eine solche zeitliche Bindung zwischen Aktion und Reaktion nicht, dann spricht man von asynchroner Interaktion. Diese findet statt, wenn man etwa eine E-Mail an das System schickt, die verarbeitet und deren Inhalt am Ende irgendwann auf dem Bildschirm angezeigt wird.

Manchmal ist eine exakte Zuordnung nicht möglich. Von einem Instant Messenger wird zwar erwartet, dass er die Daten in sehr kurzer Zeit überträgt. Dennoch lässt sich eine garantierte Obergrenze für die Übermittlung – und damit auch für die frühestmögliche Reaktion des Systems – nicht angeben, da es stets zu (nahezu beliebigen) Verzögerungen kommen kann, die beim Benutzer das Gefühl der Gleichzeitigkeit aufheben.

Einige Interaktionsmodelle sehen die Client-seitige Verwendung von Spezialsoftware vor. Dies muss schon deshalb erwähnt werden, weil sich in so einem Fall eine Interaktion nicht mehr vollkommen spontan ergeben kann, sondern der Anwender zunächst für den Download, die Installation und den Programmstart sorgen muss. Zudem muss beim Nutzer erst einmal die Bereitschaft hierzu bestehen.

Der Verzicht auf zusätzliche Software und ausschließliche Rückgriff auf bekannte Funktionen (siehe oben in Abschnitt 2.1) erweitert demgegenüber den potenziellen Nutzerkreis (z. B. „jeder mit einem Handy“ anstatt „jeder mit einem Handy, auf dem Software XY lauffähig und installiert ist“). Evtl. vorhandene Vorbehalte der Anwender etwa aufgrund des mit Fremdsoftware verbundenen Risikos von Computerviren oder wegen möglicher Datenschutzprobleme können abgemildert werden, und zusätzliche Hürden bei der Kompatibilität (Systemvoraussetzungen wie etwa Symbian OS, Java-Unterstützung, Implementierung optionaler Java-Packages auf dem Handy) entfallen.

Der Vollständigkeit halber sei auch noch auf folgende Kriterien hingewiesen:

- Zahl der (gleichzeitig möglichen) Nutzer des Systems,
- Skalierbarkeit,
- anvisierte Zielgruppe und
- Kontext der Nutzung.

Meiner Beobachtung nach liegt die Zahl der Anwender, die sinnvollerweise gleichzeitig an einem System arbeiten, in den mir bekannten Tests zwischen 1 und 10 Personen. Diese Zahl variiert jedoch stark mit der verwendeten Technologie und auch mit der Frage, ob einzelne Personen oder Personengruppen als Ganzes mit dem Display interagieren sollen (vgl. hierzu Abschnitt 3.5 weiter unten im Kontrast zu den übrigen Abschnitten von Kapitel 3).

Einige Projekte weisen auf die Möglichkeit zur Erweiterung der Nutzerzahl explizit hin, andere konzentrieren sich auf den Test mit einem Benutzer, ohne den Einsatz mit mehreren Personen gleichzeitig an einem oder mehreren Displays näher zu untersuchen (wenngleich zu erwarten ist, dass auch solche Tests folgen werden). Hier liegt es augenblicklich noch am Leser, Ideen zum Mehrbenutzer-Einsatz zu entwickeln (oder natürlich weitere Tests abzuwarten).

Die vorgestellten Projekte, die man der Grundlagenforschung zurechnen kann, definieren ihre Aufgabe nicht statisch über eine Zielgruppe, die ein bestimmtes Anwendungsfeld im Auge hat. Dennoch ergeben sich natürlich Unterschiede für die Eignung in bestimmten Einsatzfeldern, die zu diesem Zeitpunkt jedoch noch nicht bzw. nicht vollständig erforscht sind.

3 Ausgewählte Forschungsprojekte

Vor dem Hintergrund des in Kapitel 2 gegebenen Überblicks werden nun einzelne relevante Projekte vorgestellt, die sich – orientiert man sich an den oben genannten Merkmalen – zum Teil sehr deutlich unterscheiden. Wenngleich eine solche Auswahl nie vollkommen repräsentativ sein kann, so sind doch die punktuellen Einblicke, die im Folgenden gegeben werden, auf dem Spektrum wissenschaftlicher Arbeiten zur direkten Interaktion mit Displays recht breit gestreut.

3.1 Point&Shoot und Sweep

Die beiden zuerst vorgestellten Techniken aus der Arbeit von Ballagas, Rohs und Sheridan [8, 9] werden auf einem Mobiltelefon vor einem Display verwendet. Wie später zu sehen ist, ergänzen sich beide und sind daher gleichzeitig auf dem Handy zugreifbar. Bei aufgerufener Software wird auf dem Handy-Display das aktuelle Bild der eingebauten Kamera angezeigt, wobei ein Fadenkreuz die Bildmitte markiert.



Abbildung 2. Point&Shoot. Der Anwender visiert das gewünschte Objekt an, bewegt den Joystick nach links und wählt damit das Objekt aus. (Quelle: [9])

Im Falle von Point&Shoot (siehe Abbildung 2) kann man ein auf dem großen Display befindliches Objekt „anklicken“, indem man auf die Handy-Anzeige zum Anvisieren des Zielbereiches verwendet. Dazu wird das auf dem Telefon sichtbare Fadenkreuz auf das gewünschte Objekt ausgerichtet und der Joystick in horizontaler Richtung gedrückt und wieder losgelassen (nach links für normales Klicken oder nach rechts für „Drag&Drop“). Damit die Handy-seitige Software den anvisierten Punkt bestimmen kann, erscheint daraufhin auf dem gesamten Display für einen kurzen Moment ein Netz von optischen Markern (Abbildung 2 Mitte). Die Software erkennt den von der Kamera erfassten Marker. Da dieser seine eigene Position auf dem Display codiert und außerdem ein Koordinatennetz (zu sehen auf Abbildung 3) aufspannt, kann das Programm pixelgenau die Stelle berechnen, auf die geklickt werden soll. Diese Information wird an das Display übertragen, welches in üblicher Weise wie auf einen Mausklick reagieren kann (vgl. Abbildung 4).

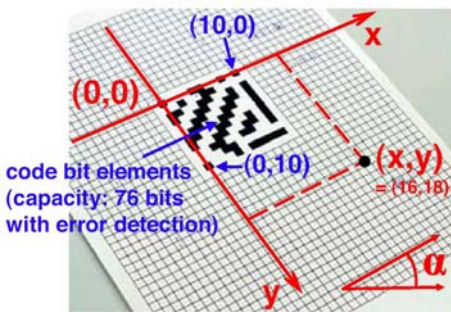


Abbildung 3. Marker mit Koordinatennetz. Punkt (x,y) ist die berechnete Stelle. (Quelle: [9])



Abbildung 4. Rechts: Handy-Bewegung mit gedrücktem Joystick. Links: Bewegung des Mauszeigers am Bildschirm. (Quelle: [9])

Bei der Sweep-Technik ist es egal, wohin das Kamera-Auge gerade zeigt, d. h., es kann irgendwo in die Umwelt zeigen, solange es nicht zugedeckt ist. Die Aktion beginnt, wenn der Nutzer den Joystick in vertikaler Richtung drückt und danach auch gedrückt hält (nach oben zur bloßen Bewegung, nach unten für „Drag&Drop“). Während so der Joystick gedrückt ist, erkennt die Software aus dem laufenden Kamerabild die relative Bewegung des Handys in X- und Y-Richtung und auch eine etwaige Rotationsbewegung um die Z-Achse. Diese Informationen können beispielsweise zur Bewegung eines Mauszeigers verwendet werden. Möglich ist auch die zusätzliche Bewegungserfassung in Z-Richtung.

3.2 Dynamo

Nach den Worten der Autoren von [10] soll das Dynamo-System – ein großflächiges, öffentlich zugängliches, interaktives Display – den Benutzern ermöglichen, Medieninhalte anzuzeigen, gemeinsam zu nutzen und auszutauschen. Das Display kann aus mehreren Anzeigeflächen zusammengesetzt sein (auch eine Anzeige auf dem Tisch ist möglich) und verfügt über sogenannte „Interaction Points“ (wörtlich übersetzt:

„Interaktionspunkte“), das sind Schnittstellen für die Interaktion mit dem Display, die jeweils die Steuerung eines Mauszeigers, die Eingabe von Text und das Einspielen von Medieninhalten ermöglichen.



Abbildung 5. Der Haupt-Bildschirm des Dynamo-Systems. (Quelle: [10])

Es existieren zwei Arten dieser Schnittstellen, die verschiedene Konfigurationen angeschlossener Geräte vorsehen:

- Der „Base Interaction Point“ stellt Tastatur, Maus und USB-Steckplätze zur Verfügung. Über USB können beispielsweise USB-Speichergeräte, MP3-Player oder Digitalkameras (auch Webcams) angeschlossen werden.
- Der „Mobile Interaction Point“ ermöglicht es Benutzern mit Notebook oder PDA, diese als Schnittstellen zu verwenden, da auf ihnen normalerweise Tastatur-, Maus- und Speicherfunktionen vorhanden sind.

Jeder Benutzer verfügt auf dem Bildschirm über einen farbigen Mauszeiger und eine persönliche Palette in übereinstimmender Farbe, auf der seine persönlichen Datenquellen und Speichermedien angezeigt werden. Medieninhalte können per Drag&Drop von der persönlichen Palette auf die allgemeine Anzeigefläche gezogen und dort angezeigt werden, oder umgekehrt durch Ziehen auf die persönliche Palette z. B. in den eigenen USB-Stick kopiert werden. Die Hauptansicht des Displays ist in Abbildung 5 wiedergegeben.

Da das Display von allen angeschlossenen gleichzeitig benutzt wird, man aber evtl. bestimmte Daten nur bestimmten Personen zugänglich machen möchte (oder man die Daten zwar anzeigen, sie aber nicht für andere freigeben möchte), kann man eine Teilfläche des Bildschirms für sich selbst „abstecken“, indem man mit der Maus – wie in Zeichenprogrammen – ein Rechteck aufspannt. In den so markierten Bereich können Medieninhalte kopiert und angeordnet werden. Andere Benutzer haben zunächst keinen Zugriff auf diesen Bereich.



Abbildung 6. Der abgegrenzte Bereich gehört dem „blauen“ Benutzer. Um ihn der „roten“ Benutzerin zugänglich zu machen, wird der Schlüssel auf ihre Palette gezogen. Im sie wieder „auszusperren“ wird ihr Symbol (Foto) entfernt. (Quelle: [10])

Möchte man die Inhalte eines solchen Bereiches anderen Personen zugänglich machen, dann kann man das Schlüsselsymbol, das sich in der rechten oberen Ecke des Bereiches befindet, auf die Palette des betreffenden Benutzers ziehen. Dieser hat dann ebenfalls Zugriff darauf (siehe Abbildung 6).

Sowohl die allgemeine Anzeigefläche als auch die abgegrenzten Bildschirmregionen, die einem bestimmten Benutzer „gehören“, können zum Austausch von Medien verwendet werden. Darüber hinaus können auch die Bildschirme mobiler Endgeräte eingebunden werden, um Dateien vom Geräte-Display auf das Dynamo-Display zu „ziehen“ und umgekehrt. Eine weitere Option ist das Hinterlassen von Nachrichten oder von sogenannten „Media Parcels“ (zu Deutsch: Medienpaketen) für andere Benutzer. In letzteren können neben Dateien auch Schlüssel übergeben werden, die den Zugriff auf abgesteckte Bildschirmbereiche ermöglichen.

Als vertiefende Literatur zu diesem System, seiner Entwicklung und insbesondere zum Einsatz und zur Akzeptanz in einem öffentlichen Mehrbenutzer-Umfeld sei noch die Lektüre der Doktorarbeit von Harry Brignull [11] besonders empfohlen.

3.3 Navigations-Würfel

Als alternatives Eingabegerät zum Umschalten zwischen Fernsehprogrammen – oder allgemeiner: zwischen verschiedenen Videostreams – entstand in der Forschungsgruppe „Embedded Interaction“ ([22], siehe Arbeit von Block et al. [12]) ein Würfel, der eine Interaktion auf spielerische Art ermöglicht: Nimmt man ihn in die Hand und neigt ihn, so sieht man auf dem Display einen virtuellen Würfel im gleichen Neigungswinkel, auf dessen Seiten die verschiedenen Streams in Echtzeit zu sehen sind. Im Falle der Fernsehprogramme erhält man so eine Vorschau mehrerer Programme zur gleichen Zeit und kann gezielt auf den Kanal mit dem gewünschten Inhalt wech-

seln, einfach indem man den Würfel wieder auf den Tisch stellt. Die Seite des virtuellen Würfels, die den Benutzer im Moment des Abstellens „anschaut“, wird dann auf das gesamte Display vergrößert. Abbildung 7 veranschaulicht das Prinzip.



Abbildung 7. Der Würfel in Aktion. (Quelle: [12])

Das Gerät im Inneren ist mit Sensoren auf Basis eines sogenannten Smart-Its [24] ausgestattet. Diese messen die Schwerkraft entlang der drei Raumachsen und senden die Information per Funk an ein weiteres Smart-It, das über die serielle Schnittstelle an den Rechner angeschlossen ist. Die so gewonnenen Rohdaten ergeben zusammen den gedachten Schwerkraft-Vektor durch den Ursprung des Koordinatensystems im Würfel (siehe Abbildung 8). Der Winkel zwischen Y-Achse und diesem Schwerkraft-Vektor gibt schließlich die Lage des Würfels im Raum an.

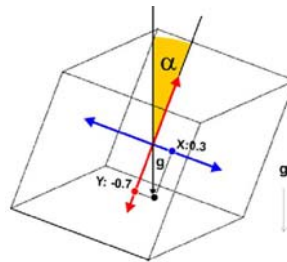


Abbildung 8. Der Vektor a gibt Aufschluss über die Lage des Würfels. (Quelle: [10])

Obwohl bei der Entwicklung dieses Gerätes in erster Linie an den spielerischen Einsatz in der heimischen Unterhaltungselektronik gedacht wurde, sind mögliche andere Anwendungen durchaus augenfällig, wie etwa das Umschalten zwischen verschiedenen Vollbild-Anwendungen (etwa mehreren Chatrooms im Falle von Kommunikationsanwendungen), Desktops (wie unter Windows oder KDE [26] durch Schaltflächen „1“, „2“, „3“, „4“ möglich), Rechnern (zur Verwendung mit nur einer Tastatur; sonst durch sogenannte Keyboard-Video-Mouse- bzw. KVM-Adapter realisiert) oder auch die Steuerung dreidimensionaler Objekte innerhalb einer Anwendung.

3.4 Laserpointer

Bei dem in Stanford entwickelten LumiPoint-System [13] können Anwender durch Laserpointer mit einem großen Display interagieren. (Es ist zwar auch vorgesehen, dass man den Laserpointer bei Bedarf als Stift direkt auf die Oberfläche aufsetzen kann, der Haupt-Fokus liegt laut den Autoren jedoch auf der Arbeit aus der Distanz.)

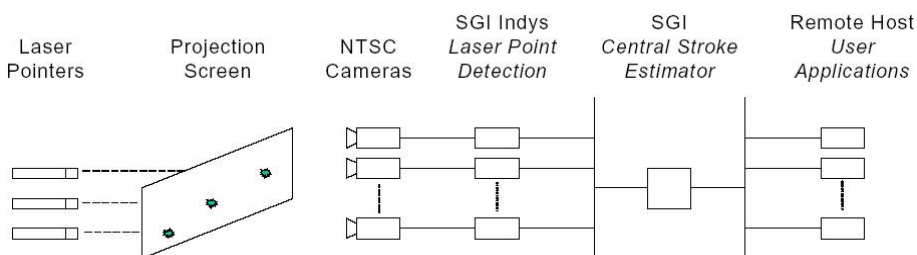


Abbildung 9. Schematischer Aufbau des LumiPoint-Systems. (Quelle: [13])

Der Aufbau ist in Abbildung 9 dargestellt. Dabei wird die verwendete Rückprojektionswand („Interactive Mural“ – deutsch: „Interaktives Wandbild“) von hinten durch Kameras beobachtet, und für jede Kamera erkennt jeweils ein angeschlossener Computer aufgrund der stärkeren Helligkeit die von vorne auftreffenden Laserpunkte. Die Daten der einzelnen Computer werden anschließend von einem zentralen Rechner zusammengeführt, der auch daraus für das gesamte durch die Kameras abgedeckte Bild jeweils Beginn, Fortsetzung und Ende von Laserpunkt-Bewegungen erkennt und Linien, die von verschiedenen Benutzern gezogen werden, jeweils auseinanderhält. Die Display-Software erhält dann – separiert nach zusammengehörigen Bewegungen – die Eingabeereignisse (englisch: „input events“)

- „stroke begin“ (für den Beginn einer Linie),
- „stroke end“ (für deren Ende) und
- „current stroke position“ (für die aktuelle Position des Laserpunktes).

Die Verfasser von [13] weisen besonders darauf hin, dass ihr System in Größe, Auflösung und Nutzerzahl beliebig skalierbar sei (und das angeblich erste mit dieser Eigenschaft). Dies wird mit der Möglichkeit begründet, dem System leicht weitere und/oder höher auflösende Displays, weitere Kameras (zur Erhöhung der überblickba-

ren Display-Fläche oder zur Erhöhung der Erkennungsgenauigkeit) oder weitere Nutzer hinzuzufügen.

Übrigens kommen Maynes-Aminzade et al. ([15], siehe auch Abschnitt 3.5), die in ihren Untersuchungen mit großen Personenzahlen und einer Kino-Leinwand auch mit Laserpointern experimentiert haben, u. a. zu dem Ergebnis, dass trotz technischer Machbarkeit bei einer zeitgleichen Verwendung von 50 Laserpointern die Testpersonen nicht mehr in der Lage sind, ihren eigenen Laserpunkt auf der Leinwand wiederzufinden.

In der eben vorgestellten Arbeit von James Davis und Xing Chen wird eine Hard- und Software-Konfiguration für die Erfassung von Laserpointer-Signalen vorgestellt. Eine Umsetzung der generierten Eingabeereignisse in Befehle, die für ein Betriebssystem oder beliebige Anwendungen darauf unmittelbar verwendbar sind, wird dagegen nicht behandelt. Aus diesem Grunde ist ein Hinweis auf die Arbeit von Olsen et al. [14] angebracht. Deren Systemarchitektur unterscheidet sich zwar in technischen Details vom LumiPoint-System. Ihre Ergebnisse zur Interaktionstechnik sind jedoch unabhängig von der konkreten technischen Umsetzung und werden daher im Folgenden kurz vorgestellt.

Ausgangspunkt bilden fünf Ereignisse, die vom System erkannt werden können:

- LaserOn(X,Y) für das Einschalten des Lasers, d. h. das Erscheinen eines Laserpunktes zu Beginn einer Bewegung
- LaserOff(X,Y) für das Verschwinden des Lasers
- LaserExtendedOff(X,Y) wenn über 2 Sekunden lang kein Laser-Punkt mehr erkannt wurde
- LaserMove(X,Y) für die Fortsetzung einer Bewegung
- LaserDwell(X,Y) für das Verweilen des Laserpunktes an einer bestimmten Stelle (in einem Toleranzbereich)

Als Anwendungsumgebung verwenden die Autoren das XWeb-System (näheres dazu auf der Projekt-Homepage [27]), welches Eingaben von sogenannten interaktiven Clients, also beispielsweise von einer Sprachsteuerung oder von der o. g. Laserpointer-Erkennung, entgegennehmen kann. Die Navigation erfolgt hier durch Manipulation von sogenannten Widgets, wobei je nach Widget und je nach Laserpointer-Event eine Rückmeldung an den Benutzer durch Veränderung des Cursors gemäß Abbildung 10 erfolgt. Natürlich erfolgt auch eine Rückmeldung durch das manipulierte Widget selbst, indem etwa ein angeklicktes Icon rot umrandet wird (vgl. Abbildung 11) oder indem in einer Liste nach oben oder unten gescrollt wird. Auch die Eingabe von Text mittels „Graffiti“ (vgl. [3], in Abschnitt 2.1 bereits erwähnt) ist vorgesehen, siehe auch Abbildung 12.





Tracking		Dwell detected	
Scrolling		Graffiti	

Abbildung 10. Rückmeldung durch verschiedene Cursor-Formen (Quelle: [14])

Wie Olsen und Nielsen feststellen, ist der erste auftretende Laserpunkt in einer Bewegung sowie die danach gezogene Strecke für die Interaktion unbrauchbar, da der

Benutzer nach dem Einschalten des Stifts erst einmal den eigentlich anvisierten Punkt zu treffen sucht. Schon insofern ist das Auftauchen und Verschwinden des Lasers nicht mit dem Aufsetzen und Hochheben einer Maus vergleichbar, deren Zeiger auf dem Bildschirm ja an der zuletzt erreichten Position stehen bleibt und von dort auch wieder weiterbewegt wird. Dies ist auch der Grund für das Einführen des LaserDwell-Events.



Abbildung 11. Angeklicktes Icon.
(Quelle: [14])



Abbildung 12. Texteingabe mit Graffiti.
(Quelle: [14])

3.5 Steuerung durch eine Menschenmenge

Bei ihren Versuchen mit größeren Personengruppen in [15] gehen die Autoren vom 1991 vorgestellten Cinematrix Interactive Entertainment System aus, bei dem die Teilnehmer in einem Kinosaal Schilder hochhalten konnten, wobei jeweils eine von zwei verschiedenfarbigen Seiten nach vorne zeigte. Dies wurde von einem Computer optisch erfasst und ermöglichte der Personengruppe beispielsweise die Steuerung durch ein Labyrinth oder die Teilnahme an einer Abstimmung. Maynes-Aminzade et al. beschäftigen sich wiederum mit einer Gruppe von Personen, die einen ganzen Kinosaal füllen. An dieser Stelle werden zwei Beispiele aus ihrer Arbeit vorgestellt.

Im ersten Beispiel ist eine Kamera am vorderen Ende des Kinosaals aufgestellt und ins Publikum gerichtet. Durch Analyse der Kamerabilder in Echtzeit haben die Teilnehmer so die Möglichkeit, beispielsweise einen Rennwagen oder im bekannten Spiel „Pong“ ein Paddel zu steuern, indem sie sich zur Seite lehnen (zu sehen in Abbildung 13).

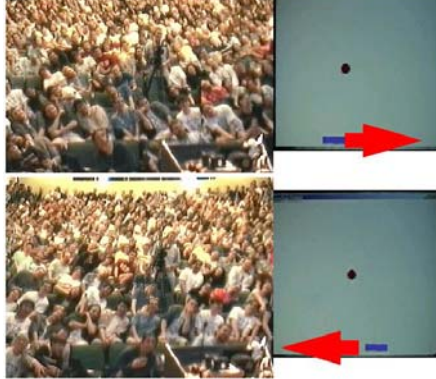
Die dazu notwendige Video-Auswertung auf dem Computer verzichtet auf die Erkennung einzelner Personen. Stattdessen werden Vorlagenbilder, die zuvor aufgenommen worden sind, mit den aktuellen Bildern verglichen und jeweils die Vorlage mit der größten Übereinstimmung ausgewählt. Dazu wird vor Beginn des Spiels jeweils ein Referenzbild vom Publikum aufgenommen, während dieses

- sich nach links lehnt,
- sich nach rechts lehnt und
- ruhig dasitzt.

Die Autoren weisen darauf hin, dass die notwendige, robuste Echtzeit-Verarbeitung mit simplen Bildverarbeitungs-Algorithmen auskommt, die auf handelsüblichen Computern ablaufen können. Zudem können weitere Gesten in unter 30 Sekunden „gelernt“ werden. Da die Referenzbilder zu Beginn der Veranstaltung aufgenommen und dann durchgehend verwendet werden, verschlechtert sich die Erken-

nung bei Veränderung des Publikums, etwa sobald eine ausreichende Anzahl von Personen erst nach Beginn des Spiels Platz nimmt. In diesem Fall kann eine Aufsichtsperson während des Spiels in geeigneten Momenten durch Knopfdruck weitere

Abbildung 13. Das Kino-Publikum spielt Pong, indem sich alle nach rechts bzw. links lehnen. (Quelle: [15])



Referenzbilder aufnehmen, wenn sich das Publikum gerade entsprechend bewegt.

Als Nachteil wird angeführt, dass die Bewegungen der einzelnen Teilnehmer in unterschiedlicher Gewichtung in die Auswertung eingehen. So haben Personen, die ganz vorne sitzen, wegen der größeren Fläche, die sie im Kamerabild einnehmen, größeren Einfluss auf das Ergebnis als andere, die weiter hinten sitzen. Die Autoren machen dagegen geltend, dass das Publikum sich dessen nicht bewusst ist und die Spielfreude nicht darunter leidet.

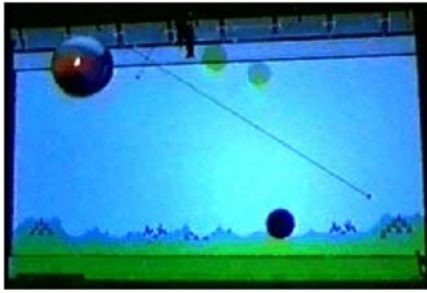


Abbildung 14. In dieser Version des Spiels Missile Command muss der Schatten des Balls die Raketen treffen. (Quelle: [15])

Beim zweiten Beispiel ist die Kamera auf die Leinwand gerichtet. Ins Publikum wird ein Ball (eine Art Wasserball) gegeben, der einen Schatten auf die Leinwand wirft. Durch gegenseitiges Zuspielden können die Teilnehmer den Ball als Cursor verwenden, beispielsweise um Raketen im Spiel „Missile Command“ abzuschießen (siehe Abbildung 14).

Statt des Balles wurde dessen Schatten zur Steuerung herangezogen, da die Erkennung des Balles selbst durch einfallende Reflexionen und Schatten ungleich schwerer ist, zudem ist eine Kalibrierung (die notwendig ist, um Ball bzw. Schatten und den Cursor auf dem Leinwand-Bild zur Deckung zu bringen) im Falle des Schattens einfacher. Der Computer erkennt den Schatten im Kamerabild, indem er den größten dunklen Fleck herausucht. Dabei kommen auch Standard-Techniken aus der sogenannten Blob-Analyse (findet zusammenhängende Regionen in Bildern) zum Einsatz. Durch das Suchen nur nach solchen Schatten, die in der Nähe eines evtl. zuvor erkannten Schattens liegen, wird Zeit gewonnen. Der Versuch, zwei Bälle gleichzeitig in Echtzeit zu erkennen, gelang so ebenfalls.

3.6 Interaktion durch Gesten

Die Erkennung von menschlichen Gesten ist besonders aufwendig, da sich der Benutzer nicht wie andere bewegliche Steuergeräte (3D-Maus, Joystick) an den Computer anschließen lässt, und überdies ist sie sehr beliebt, weil sie im Verdacht steht, besonders intuitive und ergonomische Techniken zur Interaktion zu ermöglichen – und vielleicht auch, weil die Vorstellung, dass ein Computer Befehlen aus einer Quelle gehorcht, die nicht an ihn angeschlossen ist, schon immer etwas Zaubenhaftes hatte, so wie die uralte, hartnäckige Vorstellung, der Computer nütze dem Menschen.

In der Tat ist auf dem Gebiet der Gesten-Erkennung viel gearbeitet und publiziert worden, daher genügt auch hier die Präsentation einer einzigen Referenzarbeit nicht. Stattdessen zunächst zwei Projekte, die reflektierende Marker einsetzen.

In [18] berichten die koreanischen Autoren u. a. über ihre Arbeit zur Erkennung von Handgesten, wobei sie zwei Infrarot-reflektierende Marker in Form zweier Fingerhüte (für Daumen und Zeigefinger) einsetzen. Diese unterscheiden sich in ihrer Länge, so dass der Computer, der die Bilder der eingesetzten Infrarot-Kamera auswertet, die beiden Finger voneinander unterscheiden kann. Abbildung 15 zeigt die von den Autoren vorgeschlagenen Gesten für verschiedenen Bewegungen und zum Auswählen.

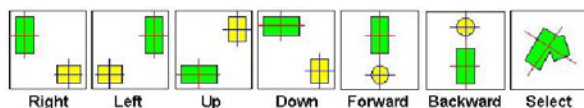


Abbildung 15. Gesten mit Daumen (grün) und Zeigefinger (gelb) (Quelle: [18])

Vogel und Balakrishnan (siehe [17]) verwenden in ihrer Arbeit recht große Marker, die zu einem kommerziellen System zur Verfolgung von Bewegungen gehören. Schwerpunkt ihrer Arbeit ist die Untersuchung von geeigneten Handgesten zur Interaktion mit einem Display. Den Nachteil der Notwendigkeit von Markern nehmen sie im Augenblick in Kauf, bis ausreichend schnelle Technologien für die Erkennung von Gesten ohne Marker zur Verfügung stehen. In der Zwischenzeit liefert das kommerzielle Marker-System schnelle, präzise Daten, mit deren Hilfe an Gesten geforscht werden kann, die später nützlich sein werden, wenn die Technik reif ist.

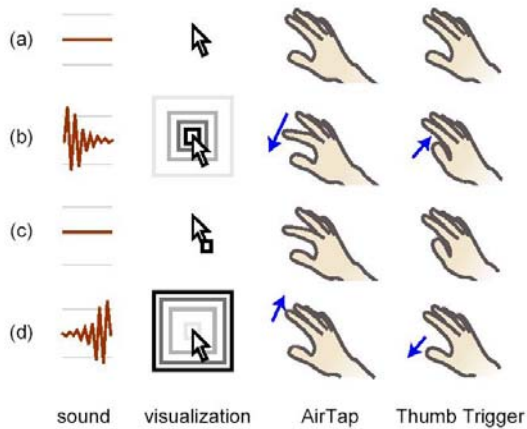


Abbildung 16. Handgesten inklusive Rückmeldungen in Ton und Bild. (Quelle: [17])

Per Hand wird ein Mauszeiger gesteuert. Abbildung 16 zeigt in einer Übersicht die vorgeschlagenen Gesten zur Kodierung von Mausclicks: kein Klicken (a), Klicken (b), geklickt halten (c), den Klick beenden / loslassen (d). Die beiden linken Spalten stehen dabei für die akustischen und optischen Rückmeldungen an den Benutzer, die beiden rechten Spalten für die entsprechenden Handbewegungen. Diese können entweder mit dem Zeigefinger (AirTap) oder mit dem Daumen (Thumb Trigger) ausgeführt werden. Kann die Handstellung nicht eindeutig erkannt werden, so erscheint der Mauscursor auf dem Display rot eingekreist. Für die Bewegung bzw. das Zeigen auf den Bildschirm schlagen die Autoren drei Techniken vor, die auch in Abbildung 17 dargestellt sind:

- Absolute Positionierung, „RayCasting“ (Abb. 17 links): Der Mauszeiger wird an der Stelle positioniert, an der eine gedachte Linie (als Verlängerung des Zeigefingers) auf dem Display auftreffen würde.
- Relative Positionierung mit „Clutching“ – deutsch: „Einkuppeln“ (Abb. 17 Mitte): Mit der offenen Hand wird der Zeiger relativ zum Ausgangspunkt weiterbewegt, entsprechend der Bewegung mit einer Computermaus. Bei geschlossener Hand / Faust kann die Hand neu positioniert werden. Der Cursor bewegt sich währenddessen nicht sondern geht optisch in einen Ruhezustand über.
- Hybride Variante, „RayToRelative“ (Abb. 17 rechts): Mit der geöffneten Hand erfolgt die relative Bewegung des Mauszeigers. Das Neupositionieren der Hand erfolgt mit absoluter Positionierung per Zeigefinger.

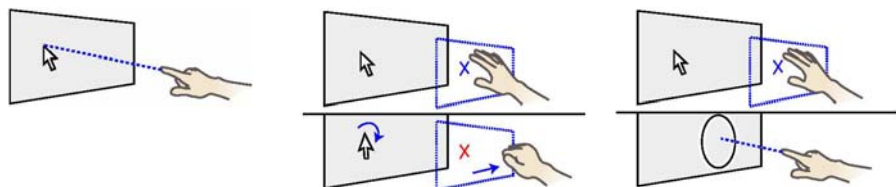


Abbildung 17. Absolute Positionierung (links), relative Positionierung (Mitte) und hybride Variante (rechts). (Quelle: [17])

Die Arbeiten [17, 18] liefern Beiträge für die Erkennung von Handgesten. Das im Folgenden vorgestellte Projekt von Sparacino et al. [16] zielt darauf ab, den Kopf und beide Hände sowie den ganzen Körper für Gesten und deren Erkennung durch Computer heranzuziehen. Vor allem aber werden hier bereits Technologien eingesetzt, die bei der Erkennung von menschlichen Körperteilen und deren Bewegungen keine Marker mehr benötigen. Auf die Verwendung von elektromagnetischen oder Ultraschall-Sensoren zur Positionierung, Handschuhe, spezielle Kleidung, eine besondere Kulisserie, Abstandsmesser oder elektronische Bodenplatten wurde bewusst verzichtet. Der Nutzer sollte ein möglichst natürliches Umfeld betreten und dann sofort mit der Interaktion beginnen können.

Der Arbeitsplatz besteht aus einem großen Display mit einem Tisch davor. Ein Kamera-Paar nimmt Stereo-Bilder auf, aus denen der Computer ein dreidimensionales, Blob-basiertes Modell des Kopfes und der Hände des Benutzers erzeugt. Die Blobs (zusammengehörige Regionen im Bild) „Kopf“, „rechte Hand“ und „linke Hand“ werden in ihren Bewegungen verfolgt, und ggf. Gesten erkannt. Die Gesten starten aus dem Ruhezustand, in dem beide Hände auf dem Tisch liegen.

Getestet wurde das System zusammen mit einem 3D-Browser. Diesem liegt die Einschätzung zugrunde, dass das Navigieren im Netz anhand der Buch-Metapher (mit jeweils einer einzigen angezeigten Seite) überholt ist und beim Nutzer kognitive Dissonanz erzeugt, da im Netz verschiedenste Dienste genutzt werden, die kreuz und quer miteinander verlinkt sind. Beim 3D-Browser startet man auf einem flachen (leeren) Stadtplan. Jede neu aufgerufene Internetseite erscheint darauf als neues Haus, wobei die Elemente der Seite an den Hauswänden angebracht sind. Die vom System erkannten Befehle ergeben sich aus Tabelle 1.

Aktion/Befehl	Geste/Bewegung
dem Link folgen	auf den entsprechenden Link am Bildschirm deuten
zurück zu vorherigem Ort	nach links zeigen
weiter zu nächstem Ort	nach rechts zeigen
aufwärts navigieren	eine Hand nach oben bewegen
abwärts navigieren	Hände zum Körper bewegen
in die Vogelperspektive gehen	beide Hände nach oben bewegen

Tabelle 1. Gesten zur Navigation im 3D-Browser. (Übersetzt aus [16])

Ein in der Testumgebung interagierender Nutzer ist in Abbildung 18 zu sehen. Unklar ist, ob alle Häuser einstürzen, wenn der Benutzer niest, und ob dies aufgrund der Geste oder wegen der Erschütterung passiert.



Move one hand up → navigate up



Point right → go to next location



Move hand towards body → navigate down



Move both hands above head → show aerial view

Abbildung 18. Eine Testperson interagiert über Gesten mit dem 3D-Browser.
(Quelle: [16])

4 Ausblick

In Kapitel 2 wurden einige wesentliche Technologien vorgestellt, die heute bereits standardmäßig in Endgeräten implementiert und somit für eine breite Anwenderschicht verfügbar sind. Unterstellt man, dass Funktechnologien für mittlere Distanzen zwischen 10 und 100 Metern (also für ein sogenanntes „Personal Area Network“, PAN) auch für die indirekte Kommunikation mit Displays auch in Zukunft zweckmäßig und wertvoll sein werden, so lohnt es sich, hier noch einmal auf Bluetooth zu sprechen zu kommen.

In [20] wird offiziell angekündigt, dass die nächste sogenannte „High-Speed“-Bluetooth-Generation auf einer Plattform basieren wird, die neben Standard-Bluetooth gleiche mehrere Standards für die drahtlose Breitband-Übertragung (Wire-

less USB, Wireless FireWire, Wireless IP) unterstützt. Die gleichzeitige Verwendung verschiedener drahtloser Standards in einem Personal Area Network soll problemlos funktionieren. Zusammen mit der extrem erhöhten Datenrate von bis zu 480 Mbps könnte die neue Bluetooth-Generation so ihren Anwendungsbereich erweitern und die Akzeptanz erhöhen. Ein genaues Einführungsdatum ist noch nicht bekannt.

Die Weiterentwicklung mobiler Endgeräte ist zumindest teilweise vorgegeben: Immer bessere Kameras (mit besserer Optik, Autofokus und echtem Zoom, höherer Auflösung) werden in Handys integriert. Diese und PDAs erleben weitere Schübe bei der Rechenleistung und – mit der Integration von High-Speed-Übertragungstechnik wie WLAN und der nächsten Bluetooth-Generation – bei der Bandbreite zur drahtlosen Übertragung, so dass Berechnungen, wie sie zum Beispiel für Point&Shoot (siehe Kapitel 3.1) nötig sind, schneller vonstatten gehen, eine noch genauere Positionierung erfolgen kann oder Anwendungen möglich werden, die wir heute nicht einmal erahnen.

Der in Abschnitt 3.3 vorgestellte Navigations-Würfel als neuartige Eingabeschnittstelle, die auf der physischen Interaktion des Benutzers mit einem instrumentierten Gegenstand beruht, ist eines von zahlreichen Produkten, die das Forschungsgebiet „Tangible Interaction“ (im Deutschen sinngemäß etwa „Interaktion durch Anfassen“) hervorgebracht hat. Grundgedanke ist die Abbildung eines physikalischen Gegenstandes (Würfel) auf ein virtuelles Objekt (Würfeldarstellung auf dem Bildschirm) und umgekehrt. Zu aktuellen und zukünftigen Entwicklungen auf diesem Themengebiet sei auf eine der vorausgegangenen Arbeiten im Rahmen dieses Hauptseminars verwiesen, in der sich Johannes Jüngst ausschließlich mit der Zusammenarbeit per Tangible Interaction befasst.

Die in Abschnitt 3.6 vorgestellten Projekte zeigen, wie präzise die Erkennung von Handgesten schon heute möglich ist. Das natürliche und auch (z. B. in [17]) erklärte Ziel ist hier, Gesten, die alleine mit der Hand durchgeführt werden, ohne die Hilfe von Markern hinreichend deutlich zu erkennen. Im Projekt von Sparacino et al. [16], wo bislang eine eher grobe Erkennung von Kopf und Händen erfolgt, ist daher schon eine zusätzliche Kamera vorgesehen, die auf Orte von besonderem Interesse gerichtet werden kann. Mit dieser könnte dann etwa eine Hand anvisiert und genauer beobachtet werden. Weiterführende Literatur zu Forschungsthemen und Entwicklungsrichtungen im Rahmen der visuellen Gestenerkennung wird in der zusammenfassenden Arbeit von Ying Wu und Thomas S. Huang [21] genannt.

Viele Projekte (wie das Dynamo Project [10] oder andere, die hier schon aus Platzgründen nicht behandelt werden können) betrachten verstärkt die sozialen Aspekte neuer Interaktionstechniken und erproben diese in verschiedenen sozialen Kontexten. Während solcher Feldversuche ergeben sich oft neue Formen der Interaktion (so auch in [11] berichtet), die der Technologie überhaupt erst einen höheren Stellenwert im alltäglichen Leben verleihen. Man darf gespannt sein, welche Interaktionstechniken sich dabei als wirklich nützlich erweisen, denn häufig macht nicht die technisch ausgefeilteste oder für einen Informatiker scheinbar „schönste“ Lösung das Rennen, sondern das Konzept, das der Anwender versteht, annimmt und gerne benutzt.

Literatur

1. Jessie Scanlon. If Walls Could Talk, Streets Might Join In. In: The New York Times, 18.09.2003.
<http://tech2.nytimes.com/mem/technology/techreview.html?res=9B03E7D6153AF93BA2575AC0A9659C8B63> (19.06.2006)
2. Keith Cheverst, Alan Dix, Daniel Fitton, Chris Kray, Mark Rouncefield, George Saslis-Lagoudakis, Jennifer G. Sheridan. Exploring Mobile Phone Interaction with Situated Displays. 1st International Workshop on Pervasive Mobile Interaction Devices (PERMID 2005). München, 2005.
<http://www.caside.lancs.ac.uk/publications/PERMID.pdf> (19.06.2006)
3. Wikipedia, the free encyclopedia. Graffiti (Palm OS).
http://en.wikipedia.org/wiki/Graffiti_%28Palm_OS%29 (19.06.2006)
4. Wikipedia, the free encyclopedia. MIDlet. <http://en.wikipedia.org/wiki/Midlet> (09.07.2006)
5. Wikipedia, the free encyclopedia. OBEX. <http://en.wikipedia.org/wiki/OBEX> (09.07.2006)
6. Jürgen Scheible, Timo Ojala. MobiLenin – Combining A Multi-Track Music Video, Personal Mobile Phones and A Public Display into Multi-User Interactive Entertainment. In: Proceedings of the 13th annual ACM international conference on Multimedia, S. 199-208. ACM Press, New York, USA, 2005.
http://www.leninsgodson.com/mobilenin/MobiLenin_acm_Multimedia2005.pdf (10.07.2006)
7. Unbekannter Autor. CeBIT: Mobilfunk. In: c't, 26:06, 2004, S. 34.
8. Rafael Ballagas, Michael Rohs, Jennifer G. Sheridan. Mobile Phones as Pointing Devices. 1st International Workshop on Pervasive Mobile Interaction Devices (PERMID 2005). München, 2005.
<http://media.informatik.rwth-aachen.de/materials/publications/ballagas2005b.pdf> (19.06.2006)
9. Rafael Ballagas, Michael Rohs, Jennifer G. Sheridan. Sweep and Point & Shoot: Phone-cam-Based Interactions for Large Public Displays. In: CHI '05: Extended abstracts of the 2005 conference on Human factors and computing systems. ACM Press, New York, USA, 2005.
<http://www.vs.inf.ethz.ch/res/papers/Sweep-PointShoot-CHI2005.pdf> (19.06.2006)
10. Shahram Izadi, Harry Brignull, Tom Rodden, Yvonne Rogers, Mia Underwood. Dynamo: A public interactive surface supporting the cooperative sharing and exchange of media. In: Proceedings of the 16th annual ACM symposium on User Interface Software and Technology (UIST), S. 159-168. ACM Press, New York, USA, 2003.
<http://www.cs.nott.ac.uk/~sxi/papers/uist03.pdf> (19.06.2006)
11. Harry Brignull. Understanding and Designing for the Voluntary Adoption of Community Displays. Doktorarbeit, University of Sussex, 2005.
12. Florian Block, Albrecht Schmidt, Nicolas Villar, Hans W. Gellersen. Towards a Playful User Interface for Home Entertainment Systems. In: EUSAI 2004 Proceedings, Springer LNCS 3295, S. 207-217. Springer, Berlin, 2004.
http://www.hcilab.org/documents/Block_TowardsaPlayfulUserInterfaceforHomeof_EUSAI2004.pdf (19.06.2006)
13. James Davis, Xing Chen. LumiPoint: Multi-User Laser-Based Interaction on Large Tiled Displays. In: Displays, 23:05, 2002, S. 205-211.
<http://graphics.stanford.edu/papers/multiuser/lumipoint.pdf> (10.07.2006)
14. Dan R. Olsen Jr., Travis Nielsen. Laser Pointer Interaction. In: Proceedings of the SIGCHI conference on Human factors in computing systems, S. 17-22. ACM Press, New York, USA, März 2001.
<http://icie.cs.byu.edu/Papers/LaserPointer.pdf> (10.07.2006)

15. Dan Maynes-Aminzade, Randy Pausch, Steve Seitz. Techniques for Interactive Audience Participation. In: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces, S. 15-20. MIT, Cambridge, USA, 2002.
<http://www-ui.is.s.u-tokyo.ac.jp/~takeo/course/2005/media/papers/audience.pdf>
(10.07.2006)
16. Flavia Sparacino, Christopher Wren, Ali Azarbayejani, Alex Pentland. Browsing 3-D spaces with 3-D vision: body-driven navigation through the Internet city. In: Proceedings of the First International Symposium on 3D Data Processing, Visualization, and Transmission, S. 224-231. MIT Media Lab, Cambridge, USA, 2002.
<http://alumni.media.mit.edu/~flavia/Papers/3dpvtWeb.pdf> (10.07.2006)
17. Daniel Vogel, Ravin Balakrishnan. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In: Proceedings of the 18th annual ACM symposium on User interface software and technology, S. 33-42. ACM Press, New York, USA, 2005.
http://www.dgp.toronto.edu/~ravin/papers/uist2005_distantpointing.pdf (10.07.2006)
18. Sang Chul Ahn, Tae-Seong Lee, Ig-Jae Kim, Yong-Moo Kwon, Hyoung-Gon Kim. Large Display Interaction using Video Avatar and Hand Gesture Recognition. In: ICIAR 2004 Proceedings, Springer LNCS 3211, S. 261-268. Springer, Berlin, 2004.
<http://web.imrc.kist.re.kr/~asc/papers/paper432-videoAvatar.pdf> (10.07.2006)
19. Rowanne Fleck. How the move to physical user interfaces can make human computer interaction a more enjoyable experience. Physical Interaction Workshop on Real World User Interfaces at the Mobile HCI Conference 2003. Udine, Italien, 2003.
<http://www.medien.informatik.uni-muenchen.de/en/events/pi03/papers/fleck.pdf>
(10.07.2006)
20. WiMedia Alliance. Next Generation High-Speed Bluetooth Wireless Technology To Be Based On WiMedia Ultra-Wideband Platform.
http://www.wimedia.org/imwp/idms/popups/pop_download.asp?contentID=8008
(09.07.2006)
21. Ying Wu, Thomas S. Huang. Vision-Based Gesture Recognition: A Review. In: Proceedings of the International Gesture Workshop '99, Springer LNCS 1739, S. 103. Springer, Berlin, 1999.
<http://wwwhome.cs.utwente.nl/~kosterj/hmi/vision-based-gesture-recognition.pdf>
(10.07.2006)
22. Homepage der Research Group Embedded Interaction. <http://www.hcilab.org/>
23. Homepage des CASIDE Projekts. <http://www.caside.lancs.ac.uk/index.php>
24. Homepage zum Smart-Its Project. <http://smart-its.teco.edu/>
25. Nokia Deutschland. Modellübersicht Mobiltelefone.
<http://www.nokia.de/de/mobiltelefone/638.html> (19.06.2006)
26. Homepage des K Desktop Environment (KDE).
27. Homepage des Xweb-Projekts. <http://icie.cs.byu.edu/ICE/XWeb/>

Die Geschichte der E-Mail und des Instant Messaging

Lucie Drasch

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
lucie.drasch@t-online.de

Zusammenfassung In den 1960er Jahren wurden elektronische Nachrichten innerhalb eines Time-Sharing-Systems ausgetauscht. Mit der Einführung des ARPANET wurde es möglich, Nachrichten über ein Netzwerk zwischen geographisch entfernten Computern auszutauschen. Ray Tomlinson entwickelte 1971 die Software SENDMSG/README, die den Weg dazu bereitete. In den 1980er Jahren entstanden erste Programme, die Instant Messaging ermöglichen, wie TALK und der Internet Relay Chat. Letzterer ermöglichte es mehreren Nutzern sich in Echtzeit zu unterhalten. Das erste Instant Messaging Programm, wie wir es heute kennen, ICQ, wurde 1996 von Mirabilis veröffentlicht. Instant Messaging Programme von AOL, Microsoft und Yahoo folgten. In den 1990er Jahren wurde auch das mobile versenden von elektronischen Nachrichten möglich. Der Erfolg der SMS zog verbesserte Nachfolgeversionen wie EMS und MMS nach sich. Auch E-Mails konnten schließlich über das Handy versendet werden.

1 Einleitung

Im Laufe der Geschichte der elektronischen Nachrichten, haben sich zwei Grundversionen herauskristallisiert. In der einen Version wird die Nachricht an eine Mailbox gesendet und wartet dort bis der Empfänger sie öffnet und liest. Sozusagen eine elektronische Variante der traditionellen Briefpost, wie sie die E-Mail darstellt. Die andere Version ermöglicht eine direkte und zeitgleiche Kommunikation zwischen zwei oder mehreren Personen, ähnlich wie ein Gespräch oder eine Diskussion: das Instant Messaging.

Bemerkenswert bei der Betrachtung der Geschichte der elektronischen Nachrichten ist die Tatsache, dass diejenigen Personen, die die technischen Grundlagen für das Versenden der elektronischen Nachrichten entwickelten gar nicht an die Notwendigkeit oder den Erfolg von solchen Nachrichten glaubten. So war beispielsweise das ARPANET ursprünglich nicht zur Übertragung von Nachrichten gedacht. Erst zwei Jahre nach der Einführung des Netzes kam man auf die Idee einen solchen Dienst zu entwickeln und bereits weitere zwei Jahre später stellte er den größten Anteil der Netzlast dar [4]. Ebenso verhält es sich mit der SMS, die bei der Einführung des GSM-Standards von den Verantwortlichen eher als „Abfallprodukt“ angesehen wurde. Sie entwickelte sich jedoch schnell zur beliebtesten mobilen Kommunikationsform [12].

Die Gründe für die Beliebtheit elektronischer Nachrichten sind vielfältig. Die elektronische Post besitzt viele Vorteile gegenüber der traditionellen Briefpost oder dem Telefon. Sie ermöglicht eine zeitnahe Kommunikation, setzt aber nicht unbedingt voraus, dass Sender und Empfänger zur gleichen Zeit erreichbar sind [4]. Ihren Anfang nahm die Geschichte der elektronischen Nachrichten in den 1960er Jahren.

2 1960er: Erster Nachrichtenaustausch über Computer

2.1 Elektronische Botschaften innerhalb eines Time-Sharing-Systems

Am Anfang der Geschichte der modernen Computer steht die Mainframe-Ära, die in dem Zeitraum zwischen 1950 und 1975 anzusiedeln ist. Typisch für diese Zeit waren, nach heutigen Maßstäben, große und teure Computer. Abbildung 1 zeigt einen PDP-10 Rechner, der beispielhaft für die damals verwendeten Maschinen ist und vor allem an Universitäten zum Einsatz kam. Time-Sharing-Systeme erlaubten mehreren Nutzern die gleichzeitige Nutzung eines solchen Computersystems. Das erste Time-Sharing-System, das so genannte Compatible Time-Sharing-System (CTSS), wurde 1961 am Massachusetts Institute of Technology (MIT) entwickelt. Jeder Nutzer konnte an einem eigenen Terminal arbeiten und dort mehrere Anwendungen laufen lassen. Die Time-Sharing-Systeme arbeiteten nach dem Round-Robin-Prinzip, d.h. der Prozessor teilte jedem Nutzer ein bestimmtes Zeitintervall zu, in dem er seine Prozesse abarbeiten konnte.

Mit dem Programm MAILBOX war es möglich Nachrichten innerhalb eines solchen Time Sharing Systems auszutauschen. Dazu bekam jeder Nutzer eine persönliche Datei/Mailbox auf dem Zentralrechner zugeteilt, auf die nur er zugreifen konnte. Jeder Nutzer konnte Nachrichten an jede Mailbox schicken [1]. Vor allem in großen Unternehmen mit weitläufigen Gebäuden konnte dies eine enorme Arbeitserleichterung darstellen. Die Kollegen konnten auf diese Weise Nachrichten austauschen ohne sich von Schreibtisch erheben zu müssen. Vor allem auch bei unterschiedlichen Arbeitszeiten diverser Angestellter stellte MAILBOX einen Vorteil dar.

Andererseits waren die Möglichkeiten sehr begrenzt, denn die Nachrichten konnten nur innerhalb eines Computersystems bzw. eines Unternehmens ausgetauscht werden. Zudem war die Akzeptanz unter den Mitarbeitern nicht hundertprozent. Viele zogen weiterhin ein persönliches Gespräch der elektronischen Nachricht vor.

2.2 Die Entstehung des ARPANET

Im Jahr 1957 schickte die UdSSR den ersten SPUNIK-Satelliten in die Erdumlaufbahn. Die Reaktionen der USA ließen nicht lange auf sich warten. Ein Jahr später wurde die DARPA (Defense Advanced Research Projects Agency) gegründet, eine Agentur des US-Verteidigungsministeriums. Ihre Aufgabe bestand darin, neue Technologien im Bereich Kommunikation und Datenübertragung zu erschließen [1].

Dazu vergab sie auch Forschungsaufträge an diverse amerikanische Universitäten und private Forschungseinrichtungen. Eines der größten Erfolge dieser Forschungstätigkeiten war das ARPANET. Ein dezentrales Netzwerk, das erstmals 1969 vier amerikanische Universitäten miteinander verband, indem die vor Ort benutzten Rechenmaschinen miteinander vernetzt wurden. Ziel war vor allem die gemeinsame Benutzung von Ressourcen [1]. Im Laufe der Zeit nahm die Anzahl der angeschlossenen Universitäten und Einrichtungen immer mehr zu. Das ARPANET stellt damit den Vorläufer des heutigen Internet dar.



Abbildung 1: PDP-10 Rechner:
Eine Großrechenanlage wie sie in den 1960er und 70er üblich war

3 1970er: Die Geburtsstunde der E-Mail

3.1 Das erste E-Mail-Programm

Wie bereits erwähnt, war das ARPANET vor allem zur gemeinsamen Nutzung von Ressourcen gedacht. Der Austausch von Nachrichten, oder persönliche Kommunikation war von den Entwicklern des Netzes anfangs weder geplant noch als notwendig erachtet.

Diesen Umstand änderte Ray Tomlinson, Ingenieur bei BBN, einem privaten Forschungsunternehmen. Er entwickelte 1971 ein Programm zum Nachrichtenaustausch, das zunächst zur Anwendung in Time-Sharing-Systemen gedacht war [1]. Dieses Programm bestand aus zwei Teilen: zum einen SENDMSG, um Nachrichten zu senden und READMAIL, um Nachrichten zu empfangen und zu lesen [3]. Um zu testen, ob eine Übertragung auch zwischen Computern möglich ist, entwickelte Tomlinson das Dateiübertragungsprotokoll CYPNET. In einem erfolgreichen Versuch

gelang es Tomlinson 1972 eine Nachricht zwischen zwei PDP-10 Rechner auszutauschen, die sich im Labor von BBN befanden. Dies war ein entscheidender Schritt. Ein Austausch von Nachrichten zwischen zwei identischen Maschinen in einem Gebäude war somit möglich. Die nächste Herausforderung bestand darin Nachrichtenübertragungen zwischen heterogenen Rechnern zu ermöglichen, deren Standort hunderte Kilometer voneinander entfernt lag [4]. Von mehreren Seiten kam der Vorschlag, das im ARPANET verwendete Datentransferprotokoll FTP auch für die Übertragung von Nachrichten zu verwenden. Ein entsprechender Versuch war erfolgreich und so konnten schließlich E-Mails über das ARPANET verschickt werden.

Der damalige Leiter der DARPA Stephen Lukasik, gehörte selbst zu den ersten, die die E-Mail für die tägliche Kommunikation nutzten. Mit ihrer Hilfe stand er mit Kollegen in Kontakt und ließ sich über die neuesten Entwicklungen der jeweiligen Forschungen informieren. Er ermunterte jeden im ARPANET diese neue Form der Kommunikation zu nutzen [1].

Schon bald konnte eine rege Zunahme des E-Mail-Verkehrs verzeichnet werden. Bereits 1973 zeigte eine Studie der ARPA, dass 75% des ARPANET-Datenverkehrs von der elektronischen Post eingenommen wurde [5]. Dabei war die E-Mail zum alltäglichen Kommunikationsmittel im ARPANET avanciert und wurde auch zunehmen für profanere Zwecke und private Kommunikation eingesetzt [4].

Larry Roberts entwickelte 1972 das Programm RD (sprich: „read“), das es ermöglichte Nachrichten abzulegen und zu löschen, sowie Listen von Nachrichten aufzurufen. RD wurde ein großer Erfolg und zog eine Menge von Programmen nach sich, die auf ähnlichen Konzepten beruhten. Das wohl beliebteste und am meisten verwandte war MSG, das 1975 von John Vittal entwickelt wurde. Das revolutionäre an diesem Programm waren die Befehle „Weiterleiten“ und vor allem „Antworten“, die es von nun ab unnötig machten ganze Nachrichten nochmals von neuen einzutippen, wenn man darauf Bezug nehmen, oder sie anderen zukommen lassen wollte [1].

3.2 Das @-Zeichen

Ray Tomlinson wurde vor allem bekannt durch die Entwicklung von SENDMSG/READMAIL und CPYNET, doch eine weitere seiner Entscheidungen hat bis heute Bestand und wurde zum „Symbol der vernetzten Welt schlechthin“ [1]. Dabei suchte Tomlinson zunächst einfach nur nach einem Weg, um in der E-Mail-Adresse den Namen des Nutzers und den Namen seiner Host-Maschine zu trennen. Dies sollte ein Zeichen sein, das im Namen des Nutzers keinesfalls vorkommen konnte. Dazu besah er sich die Tastatur seines damals gebräuchlichen Fernschreibers Teletype 33 (siehe Abbildung 2) und wurde fündig: das @-Zeichen, der „Klammeraffe“, welches die Bedeutung „at“ („bei“) hatte. Doch gerade dieses Zeichen gab zu Beginn Anlass zu Streitereien in der Computergemeinschaft [1].

In den 1970ern gab es zwei vorherrschende Betriebssysteme das dominante Tenex und das weniger verwendete Multics. Das Problem war nun, dass das @-Zeichen in Multics der Befehl zur Zeilenlöschung war. Wollte also nun ein Multics-User eine E-Mail verschicken, las der Rechner die Adresse bis zum @-Zeichen und löschte daraufhin den Rest der Zeile. Diesen Umstand hatte der Tenex-User Tomlinson wohl bei seiner Wahl nicht bedacht [1].



Abbildung 2: Fernschreiber wie der „Teletype 33“ wurden in der Mainframe-Ära zur Datenein- und -ausgabe verwendet

3.3 Die MSGGroup und der Header-Krieg

Die Heterogenität der Betriebssysteme und Mailsysteme brachte noch weitere Probleme mit sich. Nachrichtenaustausch zwischen verschiedenen Mailprogrammen vor allem auf verschiedenen Betriebssystemen war mitunter ein schwieriges Unterfangen. Es gab keinen allgemein gültigen Standard. Vor allem die Header stellten dabei ein großes Problem dar. Zu diesem Zeitpunkt bestanden die Header aus Bitcode, der nur von Maschinen lesbar war [1]. Minimale Abweichungen in der Syntax konnten bei der empfangenden Maschine zu unvorhersehbaren Ergebnissen führen. Man konnte also nicht einfach eine E-Mail an jemanden schicken, vorher musste man sicher gehen, dass der sie auch empfangen konnte [1].

Um der Standardisierung von E-Mail Anwendungen Vorschub zu leisten und neue Entwicklungen auf eine Linie zu bringen, wurde 1975 die MsgGroup (Message Services Group) gegründet. Eine Netzwerk-Mailingliste in der sich alle austauschten, die an der Implementierung von Mail-Systemen beteiligt waren [1].

Eines der ersten Themen das auf dieser Plattform diskutiert wurde, war die Header-Frage. Ziel war es einen standardisierten Header zu entwickeln, mit dem alle Mail-Programme zurechtkommen sollten und der leicht auch vom Nutzer gelesen werden konnte. Bereits nach kurzer Zeit entwickelten sich verschiedene Lager, die jeweils ihren Standpunkt durchsetzen wollten. Dies führte mitunter zu heftigen Diskussionen innerhalb der Gruppe. Dabei war nicht nur die Kompatibilität zwischen den Betriebsprogrammen ein zu lösendes Problem. Auch der eigentliche Inhalt der Header stellte einen Streitpunkt dar. Viele Mitglieder vertraten den Standpunkt, dass der Header möglichst viele Angaben und Informationen beinhalten sollte. Dort sollten neben den Absender noch diverse Schlüsselbegriffe, ID und Zeichenzählungen stehen [3]. Die Minimalisten plädierten dagegen für Einfachheit, sie waren der Meinung

Absender und Datum würden genügen, andernfalls wäre der Header mitunter länger als die Nachricht selbst. Einige sprachen sich für den Kompromiss aus, jeder Nutzer solle Gestalt und Umfang des Headers selbst bestimmen können [1].

Eine Spezifikation dessen was zwischen den Hosts ausgetauscht werden sollte wurde schließlich 1977 im Request for Comments (RFC) 724 beschrieben [1]. Das Ergebnis war jedoch nicht zur Zufriedenheit aller, da es inkompatibel zum beliebten Programm MSG war, obwohl dessen Entwickler, John Vittal, selbst bei der Erstellung des RFC 724 mitgewirkt hatte.

Das Wirken der MsgGroup dauerte etwa zehn Jahre, bevor Sie sich Anfang der 80er Jahre langsam auflöste.

3.4 Emotionen im elektronischen Nachrichtenaustausch

Ein genereller Nachteil von reinen Textnachrichten ist es, dass Emotionen schwer zu vermitteln sind. Gestik, Tonfall und Mimik des Senders werden nicht übertragen, was Missverständnissen Vorschub leisten kann. So schlug 1979 ein Mitglied der MsgGroup, mit Namen Kevin MacKenzie, vor neue Zeichen zu verwenden die symbolhaft für Gemütsbewegungen oder Mimik stehen sollten. Die Emoticons waren geboren [1].

Mitunter nahmen die hitzigen Diskussionen in der MsgGroup polemische und beleidigende Formen an. Diese so genannten „Flames“ trieben manche Mitglieder der Gruppe zum Ausstieg und stellen auch heute noch ein Problem in Foren und Chatrooms dar.

Einer der schlimmsten Flames in der Geschichte der MsgGroup wurde durch die FINGER-Kontroverse ausgelöst. Der Stein des Anstoßes war ein Programm, entwickelt von der Carnegie-Mellon-Universität, das es ermöglichte Einblick in die Online-Gewohnheiten anderer Nutzer zu nehmen. Unter anderem konnte festgestellt werden wann sich jemand zuletzt eingeloggt hatte um seine e-Mails zu lesen. Daraus entflammte eine Debatte über die Wahrung der Privatsphäre im Netz. Schließlich wurde FINGER derart abgeändert, dass jeder Nutzer die Informationen blockieren konnte, die er nicht preisgeben wollte [1].

Um Flames zu verhindern wurde im Laufe der Zeit die so genannte Netiquette entwickelt, die Richtlinien zur höflichen und fairen Kommunikation im Netz definiert [6]. Die verschiedenen Chatrooms und Newsgroups haben heute meist eigene Netiquetten die sich im Detail unterscheiden können, jedoch im Grundtenor übereinstimmen.

3.5 PLATO Notes

Plato Notes, ein Time-Sharing-System, wurde 1973 an der Universität von Illinois entwickelt. Es ermöglichte Nutzern öffentliche Berichte über Probleme des Computersystems zu posten [2]. Diese Berichte konnten von anderen Benutzern kommentiert und beantwortet werden. Das Programm zeigte den ursprünglichen Bericht, gefolgt von den Antworten der anderen User, so dass jeder Nutzer den Verlauf des Online-Gesprächs folgen konnte. Das Programm erfreute sich bald solcher Popularität, dass ein Konferenzsystem integriert wurde, das es mehreren

Nutzern ermöglicht sich in Echtzeit online über Hobbies, Religion, Sport Music etc. auszutauschen. Der erste Chat Room war geboren [2]. Bald wurde auch ein E-Mail Programm eingeführt.

Mitte der 70er wurde PLATO kommerzialisiert und in immer mehr Time-Sharing-Systemen eingesetzt.

4 1980er: Erste Schritte Richtung Instant Messaging

4.1 Die Reformierung und Kommerzialisierung der E-Mail

Seit ihrer Einführung wurde die E-Mail mit Hilfe des Dateitransferprotokolls FTP versendet. Mit der wachsenden Komplexität und Verbreitung der elektronischen Post entpuppte sich dies immer mehr als ungenügend. So nutzte man 1982 die Einführung von TCP/IP im Netz um auch bei der E-Mail eine Neuerung durchzuführen. Die E-Mail bekam ihren eigenen Übertragungsmechanismus das „Simple Mail Transfer Protocol“ (SMTP), das einige neue Steuerungsfunktionen beinhaltete und nicht benötigte Funktionen wegließ [1].

In das Jahr 1983 fällt die Abspaltung des rein militärischen MILNET vom ARPANET. In dieser Zeit wurde auch der Begriff INTERNET für den öffentlichen Teil des Netzes geprägt [5].

Durch die wachsende Ausbreitung des Netzes und die rasante Zunahme der Hosts wurde die Identifikation einzelner Hosts zunehmend schwieriger [7]. Zwar war jeder Host durch seinen Namen zu identifizieren, doch konnten mehrere den gleichen Namen wählen. So gab es beliebte Modenamen, die immer wieder verwendet wurden und sich kaum auseinander halten ließen [1].

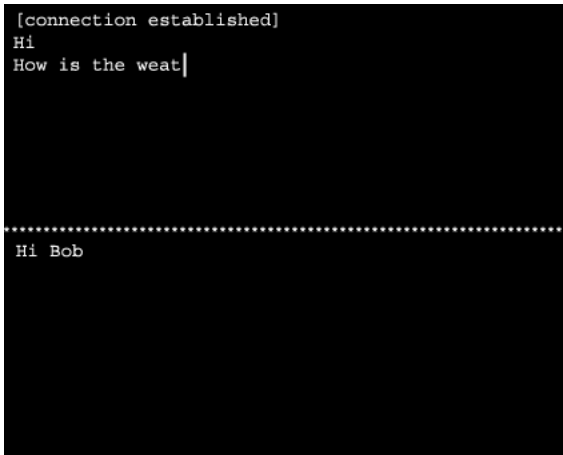
Die Lösung versprach das Domain-Namensystem (DNS) das 1983 entwickelt wurde. Ein neues Adressierungsverfahren, das auf „baumartig strukturierten Hierarchien basiert“ [1]. Dazu wurden sieben so genannte „Topleveldomains“ definiert: *com* (US-Firmen), *edu* (Bildungseinrichtungen), *gov* (Regierungseinrichtungen der USA), *int* (internationale Regierungseinrichtungen), *mil* (militärische Einrichtungen der USA), *net* (Einrichtungen zur Netzverwaltung), *org* (nichtkommerzielle Organisationen). Nun durfte es sieben Computer mit gleichen Namen geben, jeweils einen mit der Endung *.com*, *.edu* und so weiter. Nach einigen Diskussionen wurde das DNS-System 1984 eingeführt.

Anfang der achtziger Jahre kamen auch die ersten kommerziellen E-Mail-Service-Anbieter wie MCI Mail, CompuServe und Telemail auf. Dies waren größtenteils akademische und in sich geschlossene Systeme, in die man sich kostenpflichtig einloggen musste um E-Mails empfangen und senden zu können [16]. Die Gebühren wurden meist pro Nachricht erhoben.

Ab dem Jahr 1989 begannen die kommerziellen Systeme ihre Dienste im öffentlichen Internet anzubieten. Dies brachte eine große Wende im E-Mail-Verkehr, denn die Kosten, bezogen sich nun nur noch auf die Zeit, die man mit dem Internet verbunden war. Für den Versand von E-Mails fielen keine weiteren Kosten an [16].

4.2 TALK und IRC

Ein Vorläufer der heutigen Instant Messaging Programme war TALK das in den 80er und 90er Jahren in Unix Systemen verwendet wurde. Zwei Nutzer konnten sich über das Netz austauschen. Wie Abbildung 3 zeigt war der Bildschirm dabei in der Mitte geteilt und jeder Nutzer schrieb in sein Segment. Das jeweilige Gegenüber konnte Zeichen für Zeichen verfolgen wie die Nachricht eingetippt wurde. Das Nachvollziehen des zeitlichen Verlaufs war dabei jedoch schwierig [9].



```
[connection established]
Hi
How is the weat|

-----
Hi Bob
```

Abbildung 3: Das Unix Programm talk war ein Vorläufer der heutigen Instant Messaging Programme

Mit den Grundgedanken von TALK entwickelte Jarkko Oikarinen 1988 an der Universität von Oulu in Finnland den Internet Relay Chat (IRC) (siehe Abbildung 4). Oikarinen konzipierte den IRC als Client-Server-Programm. Dabei läuft ein Client-Programm auf den lokalen Rechner des Nutzers, das sich via Internet mit dem Server-Programm verbindet.

Der IRC ermöglichte es erstmals mehreren Personen gleichzeitig und in Echtzeit miteinander über das Netz zu kommunizieren. Bald schon erfreute sich der IRC in Finnland größter Beliebtheit und auch ausländische Institutionen zeigten sich interessiert. Da Finnland jedoch zu dieser Zeit noch nicht ans Internet angeschlossen war, konnte keine Verbindung aufgebaut werden [14]. Als schließlich die Internetverbindung zwischen USA und Finnland eingerichtet wurde, breitete sich die Nutzergemeinde weiter aus und bald konnten auch weitere Länder zugreifen. Es entstand ein Netzwerk mit Servern in mehreren Ländern. Bereits 1994 gab es den IRC in 27 Ländern [14]. Ein Server der an das Netzwerk angeschlossen werden sollte, musste gewisse Anforderungen erfüllen. Uneinigkeiten bezüglich dieser Anforderungen führte Anfang der neunziger Jahre zu einer Aufteilung des Netzwerkes in zwei Teilnetze: das Udernet IRC und das EFNet, die beide bis heute in Betrieb sind [15].



Abbildung 5: Der Pager diente als Wegbereiter der SMS

Durch das Aufkommen von GSM Anfang der 90er Jahre und damit der SMS wurden die Pager vollkommen verdrängt. Dabei war die SMS von Experten anfangs nur als unbedeutender Nebendienst der mobilen Telefonie gesehen worden. Die SMS ist eine reine Textnachricht mit einer Kapazität von 160 Zeichen, die anfangs meist nur von den Mobilfunkbetreibern genutzt wurde um Service-Nachrichten an ihre Kunden zu schicken [8]. Ähnlich wie bei der E-Mail hatte niemand die Notwendigkeit eines solchen Dienstes gesehen, geschweige denn mit einem solchen Erfolg gerechnet. Abbildung 6 zeigt den rasanten Anstieg des SMS-Verkehrs in den letzten Jahren.

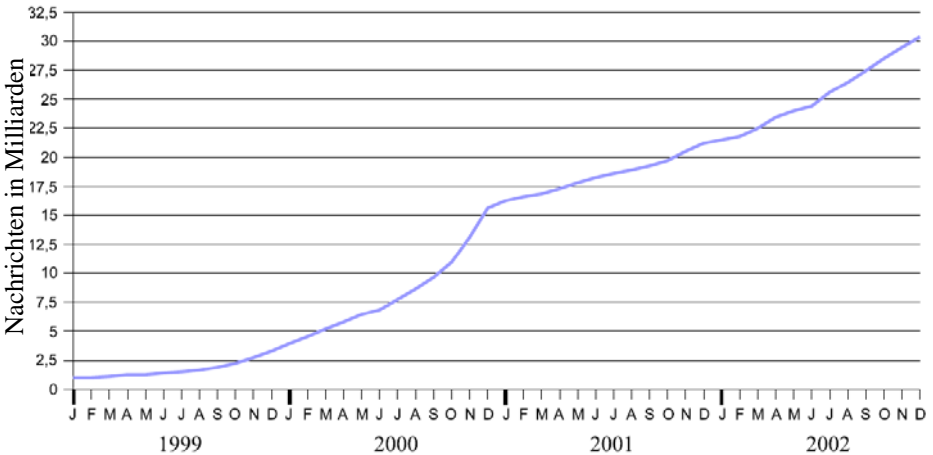


Abbildung 6: Weltweiter SMS-Verkehr pro Monat [8]

Mit der EMS (Enhanced Messaging Service) konnte ab 2000 bereits mehr als nur Text versendet werden. Es war möglich formatierten Text, Animationen, polyphonen Sound und Bilder zu verschicken [8]. Außerdem war es konnten bis zu 255 SMS als eine EMS versenden werden, so dass weit mehr als 140 Zeichen zur Verfügung standen [8].

Als nächster evolutionärer Schritt kann die MMS (Multimedia Messaging Service) angesehen werden. Mit diesem Dienst konnten nun Formate wie JPEG, GIF, MPEG und MIDI versendet werden. Zudem war es möglich verschiedene Formate mittels

einer Präsentationssprache wie SMIL in einer Nachricht zu kombinieren [8]. Während die EMS auf SMS basiert, stellt die MMS eine vollkommen neue Technologie dar. Sie bedient sich unter anderen existierender Protokolle (WAP, SMTP, HTTP) und Formate (SMIL, MIME) [8]. Abbildung 7 zeigt nochmals die jeweiligen Merkmale der mobilen Nachrichtendienste.

Um die Kompatibilität und uneingeschränkten Nachrichtenaustausch zwischen den verschiedenen Endgeräten zu gewährleisten schlossen sich 2001 die Herstellerfirmen Ericsson, Motorola und Nokia zu der Initiative Open Mobile Alliance (OMA) zusammen [8].

Feature	SMS	EMS	MMS
Media supported	Text only	Formatted text, simple media formats, e.g. pictures, animations, sounds	Multiple rich media formats, e.g. video, audio, text
Delivery mechanism	Signaling channel	Signaling channel	Data channel
Store-and-Forward	yes	yes	yes
Confirmation of message delivery	yes	yes	yes
Protocols	SMS specific, e.g. SMPP	SMS specific	WAP and general Internet e.g. MIME, HTTP, SMTP
Platform	SMS Center	SMS Center	MMS Server, MMS Relay, MMS Message Store, MMS User Agent, MMS User Databases

Abbildung 7: Vergleich mobiler Nachrichtendienste [8]

Mit der Einführung von WAP 1999 wurde es möglich E-Mails über mobile Endgeräte zu verschicken und zu empfangen [8].

Als weiterer mobiler Instant Messaging Dienst kann Push-to-Talk gesehen werden. Dieser Dienst ist nur in wenigen Mobilfunknetzen verfügbar und ermöglicht es kurze Textnachrichten an einzelne Nutzer oder Gruppen zu versenden. Erstmals wurde dies 1998 us-amerikanischen Mobilnetz Nextel angeboten.

5.2 E-Mail wird alltäglich

Seit der Einführung des World Wide Web im Jahr 1994, nahm die Anzahl der Internetnutzer rasant zu. Wie die ARD/ZDF-Online-Studie [22] zeigt, nutzen im Jahr 1997 6,5% der deutschen Bevölkerung das Internet. Im Jahr 2005 waren es bereits 57,9%. Dabei blieb die E-Mail der meist genutzteste Dienst im Internet [22]. Das versenden von E-Mails wurde dabei mehr und mehr zum alltäglichen Kommunikationsmittel. Die demographische Strukturierung der E-Mail-Nutzer gleicht sich mehr und mehr der der Gesamtbevölkerung an [22]. Die Angebote web-basierter E-Mail-

Services, macht es den Nutzern zudem möglich, die eigenen E-Mails von jedem Computer auf der Welt mit Onlinezugang zu lesen und zu versenden [16].

5.3 ICQ & Co.

Das erste Instant Messaging Programm mit grafischer Benutzeroberfläche wie wir es heute kennen ist ICQ. Mirabilis, eine kleine israelische Firma, veröffentlichte ICQ (I seek you) (siehe Abbildung 9) 1996. ICQ beinhaltete neben dem Versenden von Instant Messages, Features wie Buddy Listen, sowie die Anzeige des jeweiligen Online-Status. Die Freeware ICQ erfreute sich schnell größter Beliebtheit, bereits sechs Monate nach Erscheinen wurden 850.000 registrierte Nutzer gezählt [8]. Im Mai 1997 veröffentlicht AOL sein Instant Messaging Programm, den AOL Instant Messenger (AIM), dessen Beliebtheit vor allen in den USA schnell zunahm. Ein Jahr später sichert sich AOL die Vormacht im Instant Messaging Bereich, durch den Kauf von Mirabilis und somit ICQ. ICQ konnte zu dieser Zeit bereits 12 Millionen registrierte Nutzer, vor allen in Europa und Asien verzeichnen [8]. Schließlich entwickelten auch Yahoo und Microsoft ihre Instant Messaging Versionen den Yahoo Messenger und MSN Messenger (siehe Abbildung 9) die 1999 auf den Markt kamen. Im Jahr 2003 war der AIM Marktführer mit 100 Millionen registrierten Nutzern, gefolgt von ICQ mit 68 Millionen, MSN mit 66 Millionen und schließlich Yahoo mit 36 Millionen [9].

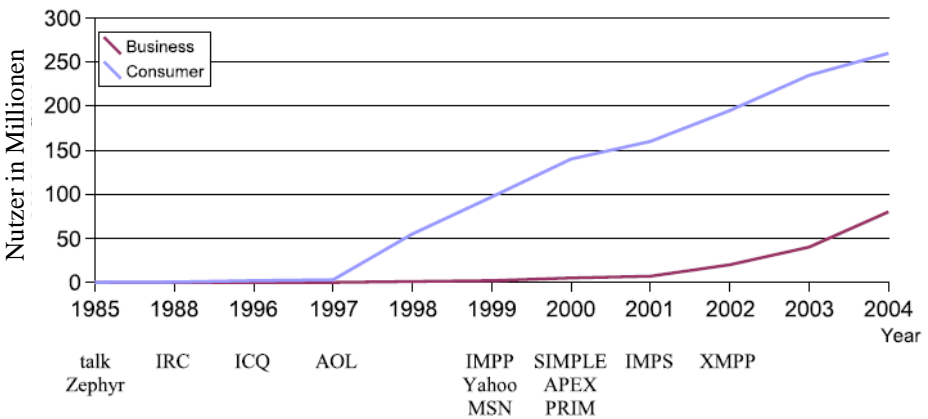


Abbildung 8: Entwicklung der Nutzerzahlen von Instant Messaging Programmen [8]

Für Windows XP wurde von Microsoft der Windows Messenger entwickelt. Dieser ist vor allem auf den Businessbereich ausgerichtet und unterstützt das Protokoll SIMPLE. Microsoft besitzt nun also zwei Instant Messaging Clients, die jedoch miteinander kompatibel sind, da sie das gleiche Netzwerk benutzen.

Bereits Mirabilis meldete für ICQ das US-Patent an, das AOL im Jahr 2002 schließlich erhielt. Dies beinhaltet die direkte Kommunikation zweier Personen in

einem Netzwerk und die Möglichkeit anzuzeigen ob der Kommunikationspartner gerade online oder offline ist. Ein Jahr später sicherte sich Microsoft das Patent auf die Möglichkeit die Tätigkeiten des Kommunikationspartners während einer Instant Messaging Sitzung anzuzeigen.

Wie Abbildung 8 zeigt wurden Instant Messaging Anwendungen schnell populär und erfreuten sich eines wachsenden Nutzerkreises. Jede dieser Anwendungen hatte ihr eigenes Protokoll und eigene Server. Das heißt, dass Nutzer verschiedener Instant Messaging Dienste nicht miteinander kommunizieren konnten [9]. Sogar die beiden Systeme von AOL sind weitestgehend inkompatibel.

Eine Lösung bestand darin Instant Messaging Clients zu implementieren, die auf mehrere Netzwerke zugreifen konnten. So genannte Multi-Protokoll-Clients, wie iChat, Trillian, Gaim, Fire, Proteus, Miranda IM, Adium und viele andere [9]. AOL, Microsoft und Yahoo konfigurieren jedoch mitunter ihre Netzwerke, um Zugriffe solcher Clients zu verhindern oder zu erschweren [9].

Eine andere Möglichkeit bestand darin ein standardisiertes für alle gültiges Protokoll zu definieren, wie es die Internet Engineering Task Force anstrebte [9]. Die Instant Messaging and Presence Protocol (IMPP) Working Group wurde 1999 von der Internet Engineering Task Force (IETF) ins Leben gerufen. Aufgabe der Gruppe sollte es sein Anforderungen und Frameworks für Instant Messaging Programme zu spezifizieren und eventuell einen Standard für ein Instant Messaging Protokoll zu definieren [8]. Die Gruppe konnte sich jedoch nicht auf ein Protokoll einigen, weshalb man entschied den Markt entscheiden zu lassen. So wurden drei neue Gruppen gebildet, die jeweils ein Protokoll definierten: Appliance Exchange (APEX), Presence and Instant Extensions (PRIM) sowie Session Initiation Protocol for Instant Messaging and Presence Leveraging Extensions (SIMPLE) [8].

APEX stieß jedoch auf wenig Interesse und die Spezifikation von PRIM kam nie zu einem Abschluss. SIMPLE stellte sich als der erfolgsversprechendster Kandidat heraus.

Im Jahr 2002 gründete IETF eine weitere Working Group, die das auf XML basierende Instant Messaging and Presence Services (IMPS) entwickelte [8].

Ein Open-Source-Projekt das an der Entwicklung eines Protokolls und eines offenen Instant Messaging Netzwerkes arbeitete ist Jabber. Das Projekt wurde 1998 von Jeremie Miller ins Leben gerufen, um eine Alternative zu den Instant Messaging Netzwerken von AOL, Yahoo und Microsoft zu schaffen [9]. Zwei Jahre später wurde eine erste Version verbreitet. Jabber bediente sich eines dezentralen Konzepts und eines XML-basierten Protokolls. Eine besondere Funktion von Jabber war die Möglichkeit mit Nutzern der Netzwerke von AOL, Yahoo und Microsoft zu kommunizieren [9]. Dabei wurden nicht wie, bei den bereits erwähnten Multi-Protokoll-Clients, die verschiene benötigten Protokolle auf der Clientseite unterstützt, sondern die Nachrichten wurden von den Jabber-Servern weitergeleitet [9].

Im Jahr 2001 wurde die Jabber Software Foundation gegründet, die bestrebt war, die Jabber-Spezifikation durch einen Offenen-Standard-Prozess zu bestätigen [10]. Ein Jahr später arbeitete die XMPP Working Group daran das Jabber Protokoll an die Anforderungen der IETF anzupassen. Diese erkannte schließlich das Jabber Protokoll unter den Namen XMPP (Extensible Messaging and Presence Protocol) als Standard an [10].

Eine Anwendung, die das XMPP verwendet ist Google Talk, ein Programm von Google, das Instant Messaging und Voice-over-IP miteinander verbindet und 2005 veröffentlicht wurde. Google Talk unterstützt dabei nicht nur den eigenen Client, sondern auch alle anderen die mit dem XMPP arbeiten, wie beispielsweise GAIM oder Trillian Pro [11].

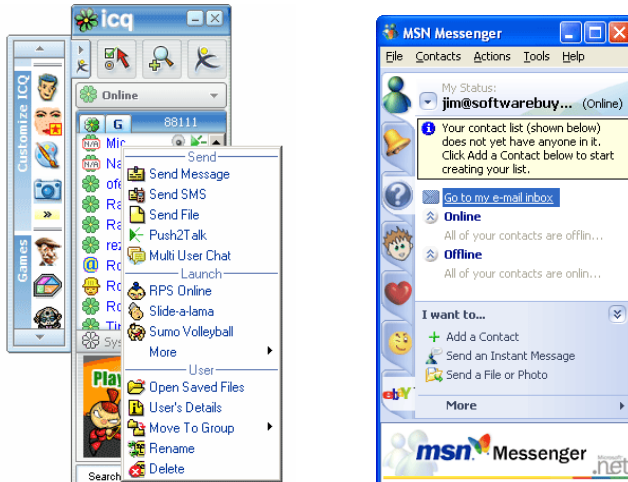


Abbildung 9: Links: ICQ, rechts: MSN Messenger [8]

5.4 Spam und Spim

Unter einer Spam-Mail versteht man eine Nachricht, die vornehmlich zu Werbezwecken an mehrere Empfänger geschickt wird, die diese aber weder angefordert noch gewollt haben [17]. Typisch ist auch, dass der Sender seine Identität verschleiert. Das Pendant im Instant Messaging Bereich wird Spim genannt. Zu vergleichen ist dies mit den Werbesendungen die man täglich im Briefkasten vorfindet. Anders als Postwurfsendungen stellen Spam-Mails allerdings für den Sender einen viel geringeren finanziellen und logistischen Aufwand dar [16].

Obwohl Spam erst in jüngerer Vergangenheit zu einem großen Problem wurde, ist es doch kein neues Phänomen. Tatsächlich wurde bereits in Jahr 1978 die erste Spam-Mail verschickt. Ein Computerhändler schickte damals eine E-Mail an alle im ARPANET, an der amerikanischen Westküste, um sie über einen Tag der offenen Tür zu informieren, an dem das neue Modell des DEC-20 Computers präsentiert werden sollte [18].

In den 80er und Anfang der 90er Jahre wurden vor allem Newsgroups von Spam-Beiträgen heimgesucht. Ein bekanntes Beispiel aus dem Jahr 1994, das das Internet buchstäblich überschwemmte, ist das Anwaltsbüro Canter & Siegel aus den USA, das Services für die Verlosung von Greencards anbot [5].

Mit dem Boom des Internet Ende der 90er Jahre stieg auch die Anzahl der Spam-Mails ins Unermessliche. Heute umfassen Spam-Mails 39% des gesamten E-Mail-Verkehrs [19]. Um dem entgegenzuwirken, wurden verschiedene Abwehrkonzepte entwickelt. Der Spam-Filter gehört heute bereits zur Standardausrüstung eines jeden E-Mail-Programms. Verschiedene Länder wie die USA und die Europäische Union führten Gesetze ein, die den Versand von Spam-Mails einschränken oder verhindern sollen. In den USA verbietet beispielsweise der CAN-SPAM Act, eine Mail in der Absicht zu manipulieren die Identität des Senders zu verschleiern [16].

Auch im mobilen Bereich wurde Spam in den vergangenen Jahren zu einem großen Problem. Spam in Form von SMS und MMS sind keine Seltenheit mehr, was in diesem Jahr sogar die Mobilfunkbetreiber selbst auf den Plan rief. „Auf Initiative des Branchenverbandes GSM Association (GSMA) haben fünfzehn Netzbetreiber eine gemeinsame Initiative gegen die Verbreitung von Spam über Mobilfunknetze gestartet und einen ‚Code of Practice‘ veröffentlicht“ [20]. In Deutschland zählen alle Mobilfunkbetreiber (außer E-Plus) zu den Teilnehmern. Auch die internationalen Firmen Cingular Wireless (USA), der Hutchison-Konzern, Turkcell und Bharti (Indien) haben sich der Initiative angeschlossen. Im „Code of Practice“, sind verschiedene Wege zur Spam-Bekämpfung definiert, an denen sich alle Unterzeichner beteiligen. Dazu gehören beispielsweise Anti-Spam-Konditionen in Verträgen mit neuen Lieferanten, oder die Verwendung von Software der GSM, die zur Aufspürung von mobilen Spam dient [21]. Der volle Inhalt des „Code of Practice“ ist nur den Mitgliedern zugänglich.

6 Ausblick

Die Zukunft der elektronischen Nachrichten liegt wohl in der mobilen Kommunikation. Die Instant Messaging Netzwerke von AOL, Yahoo und Microsoft sind vor allem für Desktop Systeme geeignet. Mobiles Instant Messaging stellt spezielle Anforderungen an Instant Messaging Systeme. Parviainen und Parnes [24] beschreiben ein System das diese Anforderungen erfüllt. Dabei werden Kontaktliste und Nachrichtenverlauf nicht lokal sondern auf den Server gespeichert. So kann sich der Nutzer auf mehreren Endgeräten anmelden und eine aktuelle Sitzung beispielsweise vom Handy auf den PC verlagern ohne die Sitzung unterbrechen zu müssen oder den Online-Status zu verändern.

Der Online-Status wird in aktuellen Instant Messaging Anwendungen vom Nutzer manuell eingestellt oder verändert sich nach einer gewissen Zeit, die vom Nutzer eingestellt wird. Die momentan möglichen, vordefinierten Zustände sind üblicherweise „verfügbar“, „nicht verfügbar“, „abwesend“ oder „offline“. An der Ludwig-Maximilian-Universität München wurde von der Resarch Group Embedded Interaction ein Tangible (greifbares) User Interface entwickelt, mit dessen Hilfe sich Online-Zustände schnell und differenziert einstellen lassen. So kann es beispielsweise einen Zustand „Any suggestions what to do?“ [23] geben. Das User Interface ist eine kleine Box (6,5 cm x 4,5 cm x 4,0 cm) die mit dem Instant Messaging System verbunden ist. Je nach Ausrichtung der Box ändert sich der Online-Status. Der aktuelle Zustand wird auf einem kleinen Display angezeigt.

Literatur

1. Hafner, K., Lyon, M. „Arpa Kadabra oder die Geschichte des Internet“, Heidelberg, dpunkt.verlag, 2000.
2. Moschovitis, C. J., “History of the Internet: A Chronology. 1843 to the Present”, Santa Barbara, Calif. [u.a.], ABC-CLIO, 1999.
3. Naughton, J., „A Brief History of the Future“, London, Phoenix, 2000.
4. Abbate, J., “Inventing the Internet”, Cambridge, London, MIT Press, 2000.
5. Zakon, R.H., „Hobbes’ Internet Timeline v8.1“, 2005, <http://www.zakon.org/robert/internet/timeline/#1960s>, 09.06.2006.
6. Runkehl, J., Schlobinski, P., Siever, T., „Sprache und Kommunikation im Internet“, Oplade, Wiesbade, Westdeutscher Verlag, 1998.
7. Solymar, L., „Getting the Message“, New York, Oxford University Press, 1999.
8. Salin, P., “Mobile Instant Messaging Systems - A Comparative Study and Implementation”, Master’s Thesis, HELSINKI UNIVERSITY OF TECHNOLOGY, Department of Computer Science and Engineering, 2004.
9. Wigtil, H., “Interactive User Modelling and Information Filtering with Hidden Markov Models”, Thesis, Norwegian University of Science and Technology, 2003
10. Finch, T., “The Jabber Protocol in Cambridge”, University of Cambridge Computing Service, 2006, <http://www-uxsup.csx.cam.ac.uk/~fanf2/hermes/doc/talks/2006-01-itstc/paper.html>, 09.06.2006.
11. Weber, V., “Google Talk geht an den Start“, Heise Online, 2005, <http://www.heise.de/newsticker/meldung/63152>, 10.06.2006.
12. Hinderberger, R., del Carmen Martinez, M., “Erfolgswunder SMS”, Fraunhofer Institut FIT, 2005, <http://www.fit-fuer-usability.de/news/praxis/maerz05/sms.html>, 11.06.2006.
13. Reid, E., “Electropolis: communication and community on Internet Relay Chat”, Thesis, Dept. of History, University of Melbourne, 1991,
14. Seidler, K. "Computerfreaks like 2 party". Wissenschaftszentrum Berlin für Sozialforschung FS II 94-104, Berlin, 1994, <http://oswaldism.de/chat/rps/>, 11.06.2006.
15. Karrels, D.R., “Internet Relay Chat Services: GNUWorld”, Thesis, University of Florida, 2003.
16. Cerf, V.G., “Spam, Spim and Spit”, Communications of the ACM, 48(4): 39–43, April 2005.
17. Khong, W.K., “Spam Law for the Internet”, Journal of Information, Law & Technologie (JILT), 3/2001.
18. Riedel, J., Brown, Z., “The Evolution of Spam and SpamAssasin”, Thesis, Worcester Polytechnic Institute, 2004.
19. “More than a Third of E-Mail is Spam”, Business Communications Review, Feb 2006: 6.
20. „Führende Mobilfuncker gründen Initiative gegen mobilen Spam“, Heise Online, 2006, <http://www.heise.de/newsticker/result.xhtml?url=/newsticker/meldung/69669&words=Instant%20Messaging>, 10.06.2006.
21. „Mobile Spam Code of Practice“, GSM Assocoation, 2006, http://www.gsmworld.com/documents/public_policy/digital_divide/mobile_spam.pdf, 14.06.2006.
22. van Eimeren, B., Frees, B., “ARD/ZDF-Online-Studie 2005 - Nach dem Boom: Größter Zuwachs in internetfernen Gruppen”, Media Perspektiven, 8/2005: 362-379.
23. Kranz, M., Holleis, P, Schmidt, A., „Ubiquitous Presence Systems“, Research Group Embedded Interaction, Ludwi-Maximillian-Universität, München, 2006
24. Parviainen, R., Parnes, P., “Mobile Instant Messaging”, Proceedings of ICT 2003, Tahiti, February 23 - March 1, 2003.

Informationsvisualisierung in E-Mail Anwendungen

Ana Monika Ticaloiu

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
`monika.ticaloiu@cip.ifi.lmu.de`

Zusammenfassung In diesem Paper werden verschiedene Informationsvisualisierungstechniken für E-Mail erörtert. Hierfür wird zuerst der Begriff der Informationsvisualisierung näher betrachtet und Einsatzgebiete im Bereich von E-Mail identifiziert. Anschließend werden anhand von Beispielen Informationsvisualisierungstechniken für E-Mail Threads, E-Mail Archive und soziale Netzwerke vorgestellt. Dabei werden die Kriterien Visualisierung und Interaktion näher betrachtet.

1 Einleitung

In den letzten Jahren gewann E-Mail sehr an Popularität und ist heute für fast jeden unverzichtbar geworden. Wurden 1999 laut [7] noch 5 Milliarden E-Mails täglich versendet, werden heute ein Zifaches davon, nämlich 171 Milliarden (vgl. [11]), täglich versendet. Bei 1,1 Milliarden E-Mail Benutzern (vgl. [11]) erhält bzw. versendet jeder durchschnittlich 155 E-Mails täglich. Diese immense Datenflut ist aber nicht nur durch unsere Kontaktfreudigkeit begründet, denn 71% dieser E-Mails sind Spam-Mails, also unerwünschte E-Mails (vgl. [11]). Die Vielzahl der eingehenden und schon vorhandenen E-Mails erschwert es dem Benutzer den Überblick zu bewahren und zeigt die Grenzen herkömmlicher E-Mail Clients auf.

E-Mail wurde ursprünglich für die asynchrone - die zeitlich versetzte - Kommunikation entwickelt. Mit der Zeit kamen aber auch die so genannten PIM (Personal Information Management) Funktionen hinzu (vgl. [1]). Dazu zählen hauptsächlich Personal Archiving, Task Management und Contact Management.

Personal Archiving beschäftigt sich mit der Archivierung von Informationen, die per E-Mail gesendet wurden. Mit Contact Management ist die Organisation und Speicherung von Kontaktdaten gemeint und zu Task Management gehören z.B. Abfrage des aktuellen Aufgabenstatus, Aufbewahrung von relevanten Informationen und Erinnerung an Aufgaben, die erledigt werden müssen.

Herkömmliche E-Mail Clients zeigen bei der Handhabung dieser Funktionen jedoch noch erhebliche Schwächen. Betrachtet man Task Management können wichtige, aufgabenbezogene E-Mails im Posteingang aufgrund der Menge an anderen E-Mails übersehen werden. Beim Personal Archiving treten Probleme dadurch auf, dass der Benutzer bei der Namensgebung von Ordnern oft inkonsistent ist und somit Daten nicht wieder gefunden werden oder diese redundant gespeichert sind.

Eine Schwäche im Bereich Contact Management ist, dass viele E-Mail Clients zwar E-Mail Adressen aus empfangenen E-Mails automatisch extrahieren und speichern können. Dahingegen müssen Telfonnummern, physikalische Adressen und andere Kontaktdaten von dem Benutzer eigenhändig in den E-Mails gesucht und dem richtigen Kontakt zugeordnet werden, was über längere Sicht gesehen viel Zeit in Anspruch nimmt.

Wie eben aufgeführt gibt es eine Vielzahl an Gründen, um E-Mail Clients zu verbessern. Eine Hilfestellung hierbei sind Informationsvisualisierungstechniken. Mögliche Anwendungsgebiete im Bereich E-Mail und schon realisierte Techniken werden in den nächsten Abschnitten beschrieben.

2 Informationsvisualisierung

Die Idee der Informationsvisualisierung ist es, Daten computergestützt auf 2D oder 3D Objekte abzubilden, um diese besser verstehen und analysieren zu können (vgl. [10]).

Vor allem sollen große Datenmengen dargestellt werden, die schwierig zu entschlüsseln sind. Man erreicht somit eine qualitative Darstellung der Daten, die es erleichtert Strukturen, Muster, Trends, Anomalien und Beziehungen zu identifizieren.

Informationsvisualisierung kann für die visuelle Darstellung von Text, Statistiken, Finanzdaten, Software und anderen Arten von Daten eingesetzt werden. Abbildung 1 zeigt eine mögliche Informationsvisualisierung einer Produktpalette für Tees.

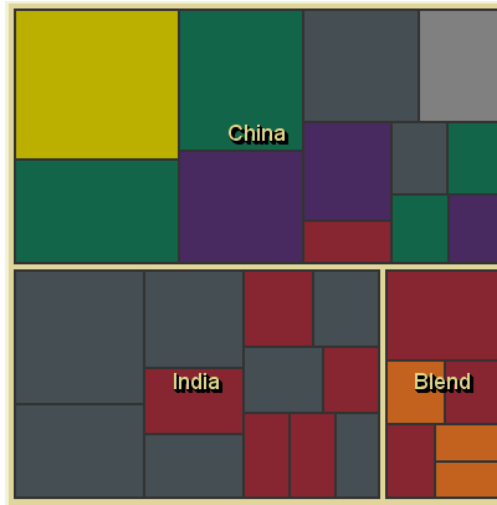


Abbildung 1. Informationsvisualisierung in Form einer Treemap einer Produktpalette für Tees (vgl. [6]).

Eine gute Informationsvisualisierung sollte auch einige Interaktions-/ Manipulationsmöglichkeiten bieten. Es sollte vor allem eine Suchfunktion und eine Zoomfunktion implementieren, um eine Analyse von Datensätzen - um

Gemeinsamkeiten, Unterschiede und Eigenarten festzustellen - zu ermöglichen. Aufgrund der Eigenschaften und der Möglichkeiten, die diese Informationsvisualisierung bietet, scheint sie sich für die Darstellung von E-Mails und deren Attribute sehr gut zu eignen.

Visualisierungstechniken, die im Bereich von E-Mail Archiven eingesetzt werden, können nach [13] in vier Gebiete eingeteilt werden. Die thread-basierte Visualisierung, die Visualisierung von sozialen Netzwerken, die zeit-basierte Visualisierung und die kontakt-basierte Visualisierung. Anhand dieser Einteilung sollen die in den nächsten Kapiteln vorgestellten Techniken kategorisiert werden.

Für die Visualisierung können verschiedene E-Mail Daten herangezogen werden. Laut [4] kann E-Mail Information in die Kategorien Metainformationen, Inhalt, Inner-E-Mail und Inter-E-Mail eingeteilt werden. Im E-Mail Header ist die Metainformation, wie Absender, Empfänger, Zeitpunkt des Sendens und des Empfangens, der Pfad den die E-Mail zurückgelegt hat, Betreff und ob die E-Mail Anhänge enthält, gespeichert. Diese Art der Daten kann vor allem für die Suche, Sortierung und der Identifikation von Trends verwendet werden. Der Inhalt einer E-Mail kann in verschiedenen Dateitypen, z.B. Html, Text oder PDF, vorhanden sein. Die Menge an gebotenen Inhalt kann sehr schwanken, und sehr große E-Mails passen vielleicht nicht auf einen Bildschirm und müssen gescrollt werden. Dieses stellt eine Herausforderung für den E-Mail Client dar und könnte durch die Verwendung von Fokus und kontextbezogenen Techniken eine Verbesserung bringen. Inner-E-Mail Information bietet die Grundlage für statistische Berechnungen, wie die Länge einer E-Mail, Anzahl der Wörter, benötigter Speicherplatz und die Häufigkeit eines bestimmten Wortes. Die Analyse einer Sammlung von E-Mails wird als Inter-E-Mail Analyse bezeichnet. Dieses wird oft bei der Darstellung von Threads benötigt (vgl. [4]).

Im nächsten Schritt werden einige Informationsvisualisierungstechniken vorgestellt, die Ihren Einsatz in unterschiedlichen Bereichen eines E-Mail Clients finden.

3 Visualisierung von E-Mail Threads

Wie schon erwähnt stellen Benutzerschnittstellen heutiger E-Mail Clients nicht die Bedürfnisse der Benutzer zufrieden. Eine Verbesserung könnte laut Venolia und Neustaedter [9] durch die Darstellung von E-Mails in „conversations“ und nicht mehr als einzelne Nachrichten sein.

Eine solche „conversation“ ist vergleichbar mit einem E-Mail Thread, der laut Kerr [3] eine Sammlung einzelner Nachrichten ist, die durch die „reply-to“ Funktion miteinander in Beziehung stehen. Die erste gesendete Nachricht wird als „root“ bezeichnet, jede beantwortete Nachricht wird als „parent“, deren Antwort und alle Antworten auf eine Nachricht wird als deren „children“ bezeichnet. Im Gegensatz zu Threads in Foren, die oft ziemlich lang sind, sind E-Mail Threads meist kurz und bestehen durchschnittlich aus 1-4 Nachrichten.

Die Darstellung in Threads bietet laut [9] gewisse Vorteile. Zum einen kann eine Nachricht in einem besseren lokalen Kontext dargestellt werden und erleichtert es dem Benutzer den Inhalt der E-Mail besser zu verstehen. Auch ein besserer globaler

Kontext ist gegeben, denn durch die kompakte Darstellung können mehr Nachrichten auf einmal angezeigt werden. Außerdem könnten Operationen integriert werden, die Auswirkung auf den gesamten Thread haben. z.B. erhält der Benutzer eine Nachricht, mit einem uninteressanten Thema, kann er den gesamten Thread löschen und somit verhindern, dass er noch weitere Nachrichten zum gleichen Thema erhält.

Um es dem Benutzer zu ermöglichen E-Mail Threads besser nutzen zu können, sollen nun thread-basierte Visualisierungstechniken betrachtet werden. Zuerst werden schon existierende Techniken erläutert, um dann anschließend näher auf Thread Arcs (vgl. [3]) einzugehen.

3.1 Existierende Techniken

Eine weit verbreitete Technik, um Threads darzustellen sind „Tree Diagrams“ (vgl. [3]). Jede Ebene stellt hierbei eine Generation von Nachrichten dar. Abbildung 2 zeigt ein „Tree Diagram“ für einen E-Mail Thread mit 6 Nachrichten. Die Zahlen (nur zu Anschauungszwecken in dieses „Tree Diagram“ eingefügt) bei den jeweiligen Knoten stellen die zeitliche Reihenfolge der Nachrichten dar.

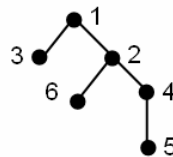


Abbildung 2. Beispiel für ein „Tree Diagram“ mit sechs Nachrichten.

Eine andere Möglichkeit der Visualisierung ist die „TreeTable“ (vgl. [5]). Hierbei wird ein Thread in einer 2D Tabelle dargestellt. Jede Reihe dieser Tabelle entspricht einer Ebene im „Tree Diagram“. Jede Nachricht belegt so viele Zellen, wie es Antworten auf diese Nachricht gab (vgl. Abbildung 3).

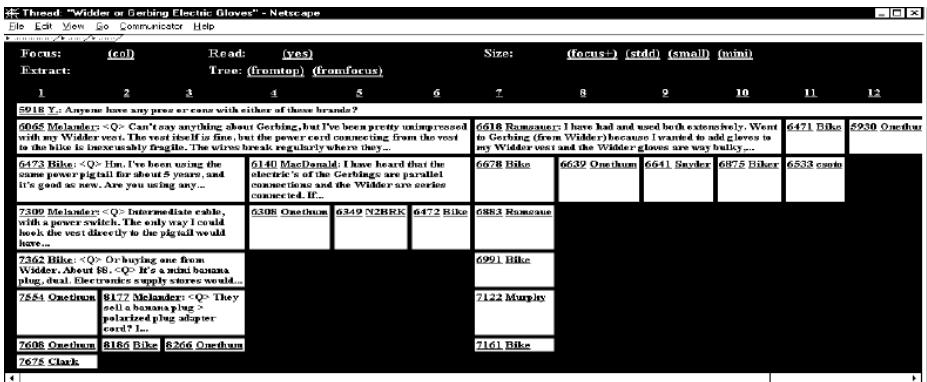


Abbildung 3. Beispiel für eine Tree Table (vgl. [5]).

Eine weitere Möglichkeit ist eine Darstellung entsprechend der „Conversation map“ (vgl. [9]). Ein Thread wird hier als Graph dargestellt, dessen Knoten Personen

entsprechen die an diesem Thread beteiligt waren. Die Kanten stellen entweder Zitate aus anderen Beiträgen oder Antworten dar. (vgl. Abbildung 4).

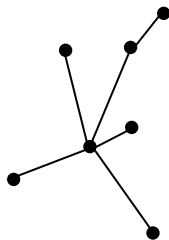


Abbildung 4. Beispiel für eine „Conversation map“ (vgl. [9]).

Betrachtet man nun alle vorgestellten Techniken, stellt man fest, dass sie alle einen E-Mail-Thread als Baum darstellen, und somit die hierarchische Entstehung, also die Beziehungen zwischen E-Mails eines Threads, in den Vordergrund stellen. Mit Hilfe solche Visualisierungstechniken lassen sich Fragen beantworten, wie:

1. Welche Frage wurde als erste gestellt?
2. Auf welche Frage ist diese Nachricht eine Antwort?
3. Gibt es Antworten auf diese Frage?

Nach [12] ist aber auch eine andere Art der Visualisierung denkbar, und zwar die sequentielle Entstehung eines Threads. Diese Art der Visualisierung ist in jedem E-Mail-Client, über die Sortierung der Nachrichten nach Datum, verfügbar. Damit lassen sich Fragen beantworten wie:

1. Welche Nachricht (von zweien) wurde als erste versendet?
2. Welche Nachricht wurde vor einer bestimmten Nachricht versendet?
3. Welche Nachricht ist die aktuellste?

Im Bereich von E-Mail sind beide Arten der Visualisierung, die hierarchische und die sequentielle, von großer Bedeutung. Thread Arcs vereinen beide Arten und realisieren noch weit mehr.

3.2 Thread Arcs

In [3] werden einigen Anforderungen definiert, die für eine optimale Visualisierung von E-Mail-Threads von Bedeutung sind und wodurch sich Thread Arcs von anderen Techniken unterscheiden. Die ersten zwei davon wurden schon in Kapitel 3.1 genannt. Zum eines ist das die „Chronology“, d.h. die E-Mails eines Threads sollten chronologisch korrekt dargestellt werden, z.B. sollte für den Benutzer klar ersichtlich sein, welche E-Mail als erste und welche als letzte angekommen ist. Darüber hinaus ist die Darstellung der „Relationships“ wichtig, also sollten alle „reply-to“ Beziehungen visualisiert sein. Weitere wichtige Kriterien sind „Stability“ und „Compactness“. „Stability“ bedeutet, dass E-Mails unabhängig von der Größe des Threads immer am gleichen Ort dargestellt werden, so dass der Benutzer eine E-Mail immer am gleichen Ort findet. Die „Compactness“ verlangt eine kompakte Darstellung, damit er leicht in einen E-Mail Client eingebunden werden kann.

Betrachtet man unter diesen Gesichtspunkten z.B. das in Kap. 3.1 vorgestellte „Tree Diagramm“ stellt man fest, dass es die Kriterien „Chronology“, „Stability“ und

„Compactness“ nicht erfüllt. „Chronology“ nicht, weil es die zeitliche Darstellung überhaupt nicht berücksichtigt. „Stability“ nicht, denn kommt eine neue Nachricht an, wird teilweise eine andere verschoben (vgl. Abbildung 5). „Compactness“ nicht, da ein Tree Diagram sowohl in die Breite als auch in die Höhe wachsen kann und dadurch sehr große E-Mail Threads nicht kompakt dargestellt werden können. Thread Arcs dagegen erfüllen alle genannten Kriterien.

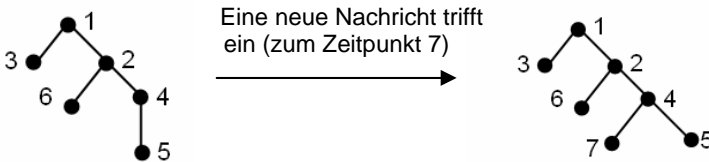


Abbildung 5. Veränderung eines „Tree Diagram“ beim Eintreffen einer neuen Nachricht.

3.2.1 Visualisierung

Thread Arcs bestehen aus linear angeordneten Nachrichtenknoten, die über Beziehungsbögen miteinander verbunden sind.

Die einzelnen Nachrichten werden horizontal, in gleichen Abständen, angeordnet. Hierbei wird links die älteste Nachricht und rechts aktuellste Nachricht dargestellt (vgl. Abbildung 6). Durch diese Reihenfolge erfüllt die Darstellung das Kriterium der „Chronology“. Kommt eine neue Nachricht hinzu, wird diese einfach rechts angehängt, wodurch auch die Anforderung der „Stability“ realisiert wird.

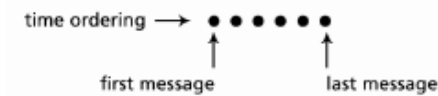


Abbildung 6. Zeitliche Anordnung von Nachrichten in Thread Arcs (vgl. [3]).

Ein E-Mail Thread mit n Nachrichten wird durch einen Thread Arc mit n Knoten und den dazugehörigen „reply-to“ Beziehungen visualisiert. Durch diese Darstellung wird das Kriterium der „Relationships“ und der „Compactness“ verwirklicht, denn es werden alle Nachrichten mit den dazugehörigen Beziehungen dargestellt (vgl. Abbildung 7).

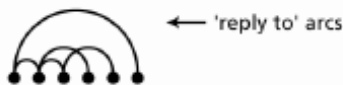


Abbildung 7. Nachrichten werden in Thread Arcs über „reply-to“ Beziehungsbögen miteinander verbunden (vgl. [3]).

Aus Abbildung 7 können nun schon einige Informationen ausgelesen werden, wie die Anzahl der Nachrichten eines E-Mail Threads oder die Anzahl der Antworten auf eine bestimmte Nachricht. Es kann aber auch die Art der Unterhaltung (Anforderung „Sense/ Scanability“) erkannt werden. Handelt es sich zum Beispiel um einen Thread, in dem eine Frage mehrere Antworten erhält. Dieses wird als „Bushy“ bezeichnet (vgl. Abbildung 8). Erhält jede Frage nur eine Antwort (vgl. Abbildung 8), wird der Thread als „Narrow“ bezeichnet.



Abbildung 8. Visualisierung typischer Unterhaltungstypen in Thread Arcs (vgl. [3]).

3.2.2 Interaktion

Integriert man Thread Arcs in einen E-Mail-Client ist es sehr nützlich auch einige Interaktionsmöglichkeiten bereitzustellen, um dem Benutzer das Auslesen von wichtigen Daten zu erleichtern.

„Attribute Highlighting“ ist eine davon. Hierbei ist es z.B. möglich einzelne E-Mails zu selektieren. Geschieht dies, werden alle Nachrichten und deren Beziehungsbögen, die mit dieser Nachricht nicht in Beziehung stehen, ausgeblendet. Alle anderen Nachrichten und deren Beziehungsbögen werden Blau dargestellt. Hierbei wird die Beziehungsart auch berücksichtigt, d.h. „children“ dieser Nachricht erscheinen in einem dunkleren Blau und „parents“ dieser Nachricht in einem helleren Blau (vgl. Abbildung 9).

Abbildung 9. „Attribute Highlighting“ für die Selektion einer Nachricht in Thread Arcs (vgl. [3]).

Aber auch andere Arten des „Attribute Highlighting“ sind möglich. Der Benutzer kann bestimmten „wichtigen“ Kontakten Farbtöne zuweisen, dies wird als „People Highlighting“ bezeichnet. Erhält oder schreibt der Benutzer einem dieser Kontakte eine Nachricht, werden die entsprechenden Nachrichtenknotten in der zugewiesenen



Farbe eingefärbt. Eine weitere Art ist das „Attribute shading“. Hierbei kann entweder die zeitliche Reihenfolge („Time shading“), die Generationstiefe („Generational shading“), die unterschiedlichen Personen in dem Thread („Contributer shading“) oder das eigene Mitwirken („Personal highlighting“) farblich dargestellt werden (vgl. Abbildung 10).

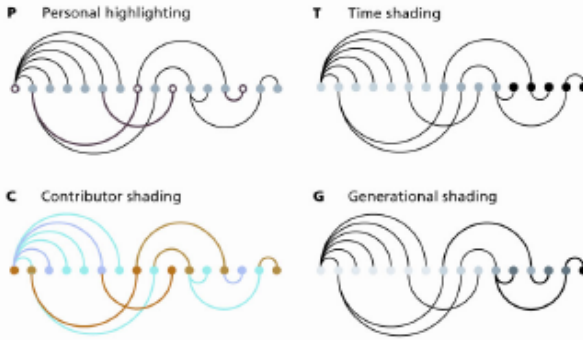


Abbildung 10. Unterschiedliche Typen von "Attribute Shading" in Thread Arcs (vgl. [3]).

Um die gefundenen Kriterien, den Nutzen und die Benutzerfreundlichkeit von Thread Arcs zu validieren, wurden im Rahmen von [3] in einer Fallstudie verschiedene Techniken gegenübergestellt. Die Fallstudie soll im nächsten Schritt kurz erläutert werden.

3.2.3 Fallstudie

Für die Durchführung der Fallstudie wurden drei Visualisierungstechniken, „Tree Diagramm“, „Tree Table“ und Thread Arcs, in einen E-Mail Client integriert. Der Benutzer konnte zwischen den verschiedenen Visualisierungstechniken wechseln. Abbildung 11 zeigt das User Interface des E-Mail Clients.

Acht Personen nutzten diesen E-Mail Client über einen gewissen Zeitraum für ihre Korrespondenz und wurden anschließend zu ihren Erfahrungen interviewt. Es sollen nun die Ergebnisse zusammengefasst werden.



Abbildung 11. ReMail, ein E-Mail-Client der die Visualisierungstechnik Thread Arcs integriert (vgl. [8]).

Laut [3] stellte sich heraus, dass Thread Arcs insgesamt am besten abschnitten. Aber vor allem wurden sie in Bezug auf „Chronology“, „Stability“ und „Compactness“ wesentlich besser bewertet als die restlichen Visualisierungstechniken. Wie vermutet erwies sich „Chronology“ als das wichtigste Kriterium für E-Mail Threads, sowohl als allgemeine Anforderung als auch in Bezug auf „Attribute Highlighting“. In Bezug auf „Relationships“ wurden Thread Arcs im Vergleich zu den anderen Techniken am schlechtesten bewertet. Gründe dafür sind, dass Thread Arcs große Threads sehr kompliziert dargestellt. Darüber hinaus war den Benutzern die Darstellung in Baumdarstellung teilweise bekannt. Im Bereich des „Attribute Highlighting“ kamen „People Highlighting“ und „Contributors Shading“ gleich nach „Time Shading“. Somit erweisen sich Thread Arcs als eine viel versprechende Technik, die weiterentwickelt werden sollte und hoffentlich bald auch in herkömmliche E-Mail Clients eingebunden wird.

4 Visualisierung von E-Mail Archiven

Um die im vorhergehenden Kapitel vorgestellte Informationsvisualisierungstechnik sinnvoll nutzen zu können, muss zuerst die gewünschte E-Mail im Postfach gefunden werden. Da aber viele Leute die Angewohnheit haben ihre E-Mails aufzubewahren und nicht zu löschen, befinden sich teilweise die E-Mails der letzten Jahre im Postfach, was eine Suche erschwert. Um dieses Problem zu umgehen, kategorisieren einige Benutzer ihre E-Mails in dafür angelegten Ordnern.

Obwohl herkömmliche E-Mail Clients einige Filter- und Suchmöglichkeiten anbieten, die das Auffinden einer bestimmten E-Mail erleichtern, können sie keinen Aufschluss über Strukturen, Anomalien und Trends der gespeicherten E-Mails geben. Es treten darüber hinaus auch Probleme bei der Kategorisierung auf, z.B. bei der Wahl eines passenden Ordnernamens oder bei einer konsistenten Namensgebung.

Eine Visualisierung dieser Daten könnte zur Lösung dieser Probleme beitragen. Im nächsten Abschnitt werden zwei Visualisierungstechniken für archivierte E-Mails vorgestellt.

4.1 „ThE-Mail“

Bei „ThE-Mail“ (vgl. [13]) handelt sich um eine Visualisierungstechnik, bei der der Fokus vor allem auf einer zeit-basierten Darstellung liegt, wobei aber auch soziale Aspekte einbezogen werden. Der große Unterschied zu den meisten anderen Techniken ist, dass die Informationen, die für die Darstellung verwendet werden, aus dem Inhalt (vgl. Kapitel 2) und nicht aus dem Header einer E-Mail stammen.

Ziel ist es zum einen Informationen über das Thema zu liefern, das in einer Diskussion mit einem Kommunikationspartner behandelt wird und zum anderen Unterschiede in der Sprache/ Wortwahl bei der Kommunikation mit unterschiedlichen Partnern zu identifizieren.

4.1.1 Visualisierung

„The-Mail“ ist eine typografische Visualisierung des E-Mail-Inhalts eines Postfachs entlang der Zeitachse. Für die Darstellung sind entlang der horizontalen Zeitleiste mehrere Spalten mit Schlüsselwörtern angeordnet. Die Farbe der Schlüsselwörter ist abhängig von der Häufigkeit des Auftretens und deren Aussagekraft. Angezeigt wird immer der E-Mail-Verkehr zu einem bestimmten Kontakt.

Die Schlüsselwörter sind in verschiedenen Informationsebenen organisiert. Im Hintergrund, etwas verblasst und in grau, sind die „yearly words, die Wörter die im letzten Jahr am häufigsten verwendet wurden dargestellt. Im Vordergrund sind die „monthly words in gelb abgebildet. Die „monthly words“ sind die Wörter, die im letzten Monat am häufigsten verwendet wurden und die größte Aussagekraft, für eine bestimmte Beziehung, besitzen. Kommt z.B. ein bestimmtes Wort in einer Kommunikation sehr oft vor und in anderen kein einziges Mal, hat es für diese Beziehung eine hohe Aussagekraft. Je häufiger ein Wort vorkommt und je aussagekräftiger es ist, desto größer wird es dargestellt. E-Mails werden als farbige Kreise dargestellt, die entsprechend der Länge der E-Mail größer oder kleiner sind. Empfangene E-Mails und versendete E-Mails sind in unterschiedlichen Farben dargestellt (vgl. Abbildung 12).



Abbildung 12. „The-Mail“ visualisiert den Inhalt eines E-Mail Postfachs (vgl. [13]).

4.1.2 Interaktion

Abbildung 12 zeigt das gesamte User Interface von „The-Mail“. Auf der rechten Seite wird zunächst der Kontakt selektiert zu dem man den Kommunikationsverlauf betrachten möchte. Fährt der Benutzer in dieser Ansicht mit der Maus über ein Wort, wird es vergrößert und in weißer Schriftfarbe angezeigt (vgl. Abbildung 12 für das Wort „Performance“). So ist es auch möglich sehr kleine Wörter zu lesen. Klickt der Benutzer ein bestimmtes Wort an, gilt es als selektiert. Es erscheinen an der linken Seite alle E-Mails, die dieses Wort enthalten. Innerhalb dieser E-Mails ist das selektierte Wort auch wieder hervorgehoben. Somit kann der Benutzer den Kontext betrachten, in dem das Wort benutzt wird.

Auch eine Suche innerhalb der Schlüsselwörter ist möglich. Nach dem Eintippen des gesuchten Begriffs werden alle Suchergebnisse durch Farbe und Größe hervorgehoben dargestellt (vgl. Abbildung 13 für das Wort „ex“).

Da es vorkommen kann, dass es Monate mit wenig und Monate mit sehr viel E-Mail-Kontakt gibt, bietet „The-Mail“ zwei unterschiedliche Ansichten an. In der erweiterten Ansicht werden auch Monate ohne E-Mail Kontakt, als freier Platz, dargestellt und in der verkleinerten Ansicht werden nur Monate mit E-Mail Kontakt angezeigt.

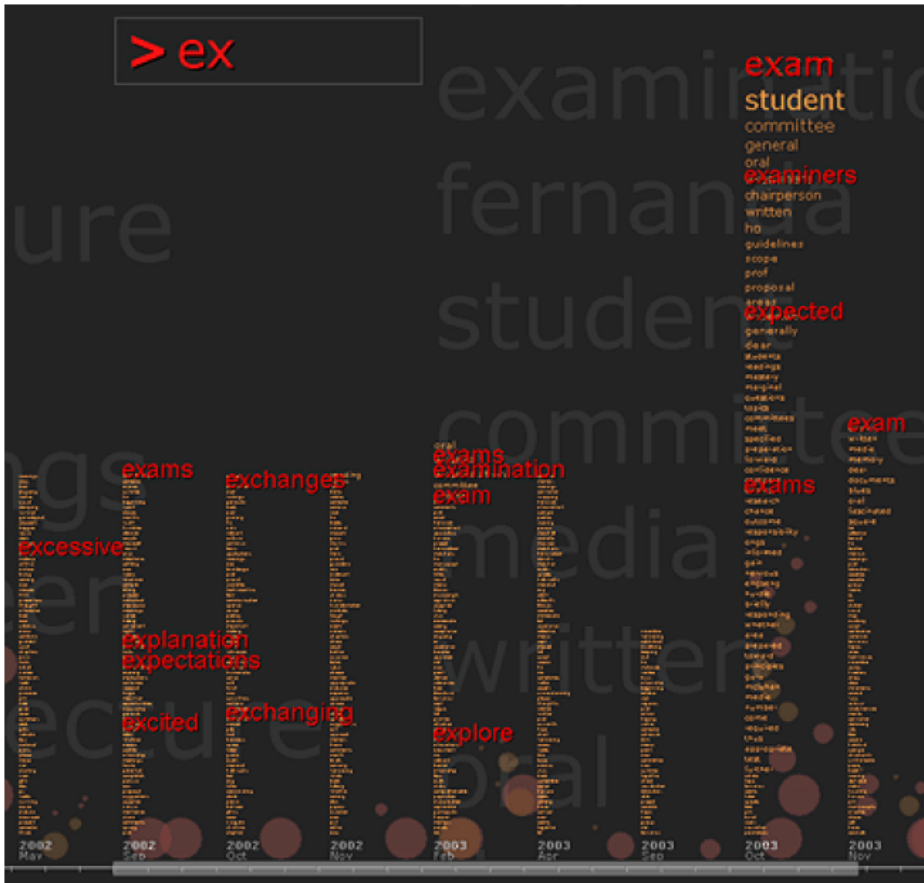


Abbildung 13. Suche nach dem Begriff „ex“ in „The-Mail“ (vgl. [13]).

Mit diesen Möglichkeiten gibt es verschiedene Anwendungsgebiete, für die man „The-Mail“ einsetzen kann. Zum einen kann man sich ein Gesamtbild über die Beziehung zu einem bestimmten Kontakt machen oder/und kann hierbei auch neue Muster und Strukturen identifizieren, über die man sich nicht bewusst war. Man kann aber auch, anhand der Veränderung der Schlüsselwörter, die Entwicklung der Beziehung, z.B. von Arbeitskollegen zu Freunden hin, nachvollziehen.

Dennoch zeigt diese Technik einige große Schwächen. Ein Beispiel ist, dass durch die Gleichbehandlung aller E-Mails, oft verwendete Betreffeinträge, wie es bei

weitergeleiteten E-Mails vorkommt, einen zu hohen Stellenwert erhalten. Auch kommt es öfter vor, dass Schlüsselwörter nicht repräsentativ sind, da es sich z.B. um Wörter der Signatur handelt. Insgesamt beeindruckt diese Technik durch die Vielzahl an Möglichkeiten und die Idee, die Informationen aus dem Inhalt der E-Mails zu gewinnen.

4.2 „Ivyfield“

Auch die Technik „Ivyfield“ (vgl. [14]) legt ihren Fokus vor allem auf eine zeitbasierte Darstellung, wobei aber auch soziale Aspekte berücksichtigt werden. Das besondere dieser Technik ist, dass sie eine Garten-Metapher verwirklicht. Die Informationen, die sie für die Darstellung benötigt, extrahiert sie aus dem Header der E-Mails. Diese Visualisierung soll eine Vielzahl an Möglichkeiten des Managements für E-Mail-Archive zur Verfügung stellen und die Vielfalt an sozialen Informationen, die in solchen Archiven enthalten sind, offen legen.

4.2.1 Visualisierung

In „Ivyfield“ wird jede E-Mail durch ein Efeublatt dargestellt. Jeder Kommunikationspartner, mit dem mehr als eine E-Mail ausgetauscht wurde, wird als Efeuranke visualisiert. Der Stiel dieser Ranke wird als Lebenslinie bezeichnet, da er den Zeitraum der Kommunikation darstellt. Je länger die Lebenslinie, desto länger der E-Mail-Austausch. Empfangene E-Mails werden als grüne Blätter links vom Stiel und gesendete E-Mails als blaue Blätter rechts vom Stiel visualisiert (vgl. Abbildung 14). Für besondere E-Mails werden auch andere Farben verwendet, z.B. werden noch nicht gelesene E-Mails durch ein rotes Blatt und besonders wichtige durch ein lila Blatt angezeigt.



Abbildung 14. E-Mail Kommunikation mit einem Kontakt, in „Ivyfield“ durch eine Efeuranke dargestellt (vgl. [14]).

Die horizontale Achse stellt alle Kontakte anhand der Anfangsbuchstaben ihrer E-Mail-Adressen und die vertikale Achse die Zeit dar. Somit wird an einer Ranke die älteste ausgetauschte E-Mail weiter unten und die aktuellste weiter oben dargestellt. Da ältere E-Mails eher an Wichtigkeit verlieren, ist die Zeitachse nicht linear. Ein aktuelleres Datum, vor allem das der aktuellen Woche, wird eine höhere Priorität zugesprochen, so dass dies mehr Platz in der Zeitleiste einnimmt.

4.2.2 Interaktion

Nach [15] bietet Ivyfield eine Vielzahl von Interaktionsmöglichkeiten an. Um eine bessere Skalierbarkeit zu gewährleisten, werden zwei Ansichten angeboten. Existieren zu viele E-Mail-Kontakte mit denen eine Kommunikation stattfindet, kann die horizontale Achse auf zwei Fenster zur halben Größe aufgeteilt werden. Diese Ansicht geht auf Kosten der vertikalen Achse, da dieser Bereich halbiert wird (vgl. Abbildung 15). Es existiert auch ein Zoom Funktion, mit der die aktuelle Darstellung um 10% vergrößert oder verkleinert werden kann, um Details besser identifizieren zu können.

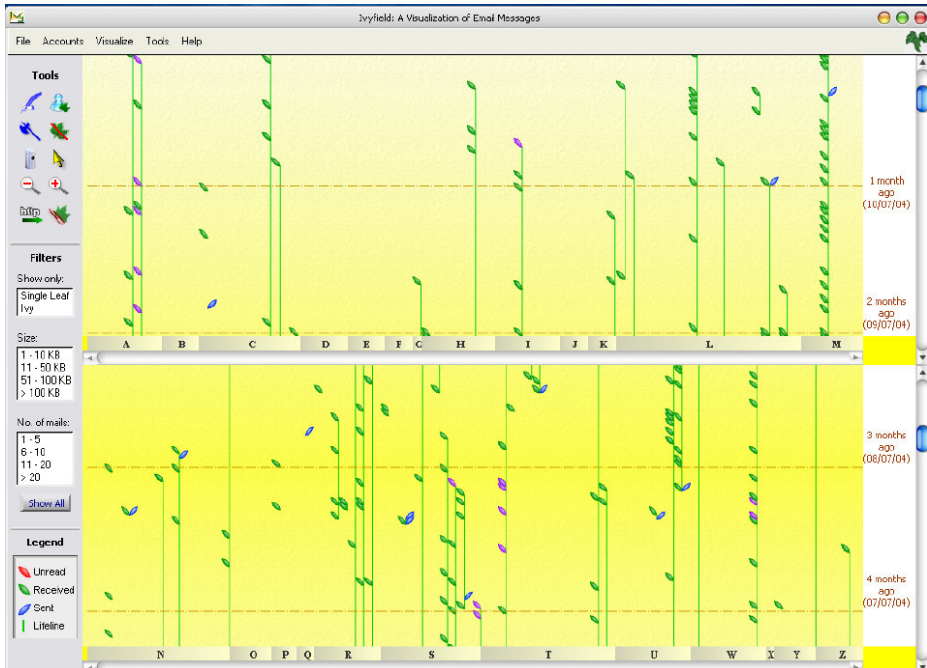


Abbildung 15. Um in Ivyfield eine detailliertere Ansicht zu erhalten wird das Fenster auf zwei aufgeteilt (vgl. [14]).

In einer Visualisierung werden textuelle Informationen, wie z.B. der Absender oder der „Betreff“ einer E-Mail, vernachlässigt. Um dennoch zu gewährleisten, dass der Benutzer alle benötigten Informationen erhält, erscheint beim Fahren mit der Maus über einen Stiel oder ein Blatt ein Fenster, in dem der Absender und/oder Betreff erscheint. Führt er einen Doppelklick auf einem Blatt aus, wird die gesamte Nachricht angezeigt. Führt er dagegen einen Doppelklick auf eine Ranke aus, wird eine neue Ansicht geöffnet, in der nur dieser Thread, mit Preview auf alle enthaltenen E-Mails, dargestellt wird.

Dem Benutzer stehen auch einige Filterfunktionen, wie Größe, Datum und Anzahl der E-Mails, zur Verfügung. So können z.B. nur E-Mails ab einer Größe von 100 KB angezeigt werden, was die Darstellung irrelevanter E-Mails reduziert. Eine weitere Funktion ist die Anzeige noch ungelesener E-Mails in einer Preview Ansicht am linken Seitenrand, ähnlich zu einem Thread.

Vergleicht man beide Techniken, „ThE-Mail“ und „Ivyfield“, scheinen sie sehr unterschiedlich zu sein, obwohl sie den gleichen Fokus haben. Ein Grund dafür ist auf jeden Fall, dass sie verschiedene Typen von Informationen verarbeiten. „ThE-Mail“ extrahiert seine Informationen aus dem Inhalt einer E-Mail und hat somit hauptsächlich mit Schwächen in Bezug auf das Parsing, die Wahl der richtigen Schlüsselwörter, zu kämpfen. Dahingegen verarbeitet „Ivyfield“ die im Header einer E-Mail enthaltenen Metadaten. Diesbezüglich kann es sich kaum von anderen Visualisierungstechniken für E-Mail Archive absetzen. Das besondere an „Ivyfield“ ist, dass es für die Visualisierung eine Garten-Metapher heranzieht.

Im nächsten Abschnitt sollen nun Visualisierungstechniken erläutert werden, deren Fokus auf der Darstellung sozialer Netzwerke liegt.

5 Visualisierung sozialer Netzwerke

Obwohl unsere Lebensweise und Erfahrungen immer weiter digitalisiert werden, sehnen wir uns dennoch nach sozialen Einflüssen. Z.B. nutzen manche Benutzer wenn sie online sind „Web Traffic Monitore“, um einen Eindruck von ihrem Verhalten zu erhalten. Mit Blogs und Friendster haben sich Technologien entwickelt, die es Benutzern ermöglichen, sich sozial zu engagieren und sogar soziale Netzwerke aufzubauen. Dieser Aspekt wurde im Bereich E-Mail noch kaum berücksichtigt, obwohl E-Mail zu den wichtigsten, meist benutzten, sozialen Aktivitäten in der digitalen Welt zählt.

Da weit mehr als die Hälfte der empfangenen E-Mails entweder Spam-Mails oder unwichtige Mails sind, erkennt der Benutzer in einem herkömmlichen E-Mail-Client, in dem E-Mails unabhängig vom Empfänger gleich angezeigt werden, für ihn sozial wichtige E-Mails kaum. Um die soziale Kompetenz von E-Mail besser nutzen zu können, können Visualisierungstechniken eingesetzt werden, die die soziale Wichtigkeit eines Kontaktes in der E-Mail-Kommunikation berücksichtigt.

5.1 „faMailiar“

Im Bereich von E-Mail spielt neben dem sozialen Aspekt auch die Aktualität eine sehr wichtige Rolle. „faMailiar“ ist eine Visualisierungstechnik für die E-Mail Kommunikation, bei der der Fokus vor allem auf der Darstellung von sozialen Netzwerken liegt, wobei aber auch der zeitliche Aspekt mit einfließt (vgl. [4]).

Das Ziel dieser Technik ist es vor allem, Muster und Strukturen innerhalb der E-Mail-Kommunikation mit anderen Benutzern zu erkennen und dadurch zu realisieren, welche soziale Beziehung mehr Zeit benötigt oder mit welcher Beziehung nicht so viel Zeit verschwendet werden sollte. Somit soll sie den Benutzer vor allem beim Zeitmanagement unterstützen.

5.1.1 Visualisierung

Um soziale Wichtigkeit überhaupt darstellen zu können, müssen Parameter eingeführt werden, um diese messen zu können. Dafür verwendet „faMailiar“ eine Kombination

aus zwei Parametern: soziale Kategorie des Absenders und soziale Relevanz der Nachricht.

Die soziale Kategorie des Absenders wird vom Benutzer im Adressbuch editiert. Es reflektiert die soziale Beziehung, den Grad der Intimität, zu diesem Kontakt. Hierfür gibt es vier Kategorien, von sehr intim bis zu überhaupt nicht intim. Alle Kontakte, die der Benutzer in keine Kategorie einordnet, werden automatisch in eine zusätzliche „Default“ Kategorie eingeordnet.

Die soziale Relevanz einer Nachricht, wird automatisch, anhand der in der Nachricht enthaltenen Daten, ermittelt. Dieser Wert schwankt zwischen 0 und 1, je nachdem wie viele „intime“ und wie viele „nicht intime“ Wörter in der Nachricht enthalten sind. Z.B. werden weitergeleitete Nachrichten als „nicht besonders intim“ eingestuft.

Die soziale Kategorie des Absenders wird durch Icons unterschiedlicher Farbe und Form visualisiert. Hierfür werden für die „intimsten“ Kontakte ein warmes gelb verwendet und werden dann linear zum „am wenigsten intimen“ Kontakt, rot, grün und blau. Die Kontakte der „Default“-Kategorie werden in grau dargestellt.

Zusätzlich dazu werden unterschiedliche Kategorien durch Icons unterschiedlicher Form dargestellt. Absteigend zu dem „am wenigsten intimen“ Kontakt, erhalten die Formen immer mehr Ecken. Somit wird ein Kontakt der Kategorie „am intimsten“ durch ein gleichschenkliges Dreieck und ein Kontakt der Kategorie „am wenigsten intim“ durch ein Sechseck dargestellt. Kontakte der „Default“ Kategorie werden als Kreise dargestellt. Versendete E-Mail werden in den entsprechenden Farben als Stern dargestellt (vgl. Abbildung 16). Die soziale Relevanz der Nachricht wird durch Helligkeit dargestellt. Je „intimer“ eine Nachricht, desto heller wird sie dargestellt.

Das User Interface von „faMailiar“ stellt mit Hilfe dieser Icons den gesamten E-Mail Verkehr entlang einer Zeitachse dar (vgl. Abbildung 17).



Abbildung 16. Icons in „faMailiar“ um die soziale Kategorie von Kontakten darzustellen (vgl. [4]).

5.1.2 Interaktion

Der Benutzer hat zwei Ansichten zur Verfügung. Die Tagesansicht stellt horizontal den jeweiligen Tag dar und vertikal die Uhrzeit. Jede Nachricht ist durch ein Icon visualisiert. In der Wochenansicht werden horizontal die Wochen und vertikal die Tage angezeigt (vgl. Abbildung 17). Jedes Icon visualisiert alle E-Mails innerhalb eines Tages, die an Kontakte derselben sozialen Kategorie gesendet oder von diesen gesendet wurden. Dadurch ist es dem Benutzer möglich, entweder ins Detail zu gehen oder Strukturen über längere Zeit zu beobachten.

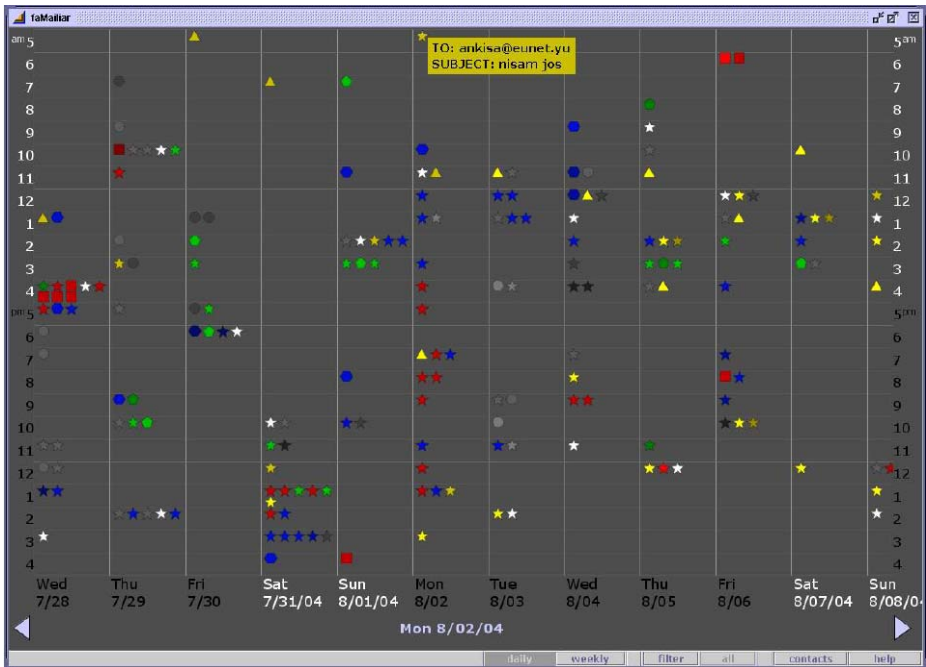


Abbildung 17. Wochenansicht eines E-Mail Postfachs in “faMailiar” (vgl. [4]).

In beiden dieser Ansichten erhält der Benutzer, wenn er mit der Maus über einen Icon fährt, Informationen zu diesem angezeigt. Bei der Tagesansicht kann das nur eine Kurzinformation oder auch die gesamte E-Mail sein, bei der Wochenansicht werden Informationen zu allen E-Mails, die für dieses Icon relevant waren, angezeigt.

Über eine Zoomfunktion kann der Benutzer den Zeitraum, für den die E-Mail-Kommunikation angezeigt wird, festlegen. Er kann auch die Nachrichten anhand bestimmter Attribute filtern, um nur eine Untermenge der Nachrichten darzustellen. Diese Attribute können entweder der Absender, der Empfänger, der Betreff oder ein bestimmter E-Mail-Thread sein. Hierfür werden in der Darstellung die Untermenge normal und die Restlichen verblasst angezeigt.

Zusammenfassend kann man sagen, dass dieses Tool, durch die Nutzung aussagekräftiger Techniken, Benutzer dabei unterstützen könnte festzustellen, wie sich ihr Freundeskreis im letzten Jahr verändert hat.

Im nächsten Abschnitt soll eine weitere Visualisierungstechnik für soziale Netzwerke betrachtet werden, die die Darstellung von Kontakten unterstützt.

5.2 Personal Map

In den meisten Programmen werden Adressbücher immer noch in Listenform, alphabetisch angeordnet, organisiert. Bei der immensen Anzahl der Kontaktdaten in einem E-Mail-Client, bietet dieses System aber keine guten Dienste mehr. Denn Benutzern fällt es immer schwerer, ihre Adressbücher auf dem neusten Stand zu halten und sie in einer sinnvollen Art und Weise zu organisieren. Eine andere Form

der Darstellung könnte hierbei bessere Dienste leisten. In „Personal Map“ (vgl. [2]) werden alle Kontakte automatisch in soziale Kategorien eingeordnet und in einem Tortendiagramm dargestellt (vgl. Abbildung 18).

Das hauptsächliche Ziel von „Personal Map“ ist es, dem Benutzer dabei zu unterstützen einfach und schnell mit anderen Benutzer oder Gruppen online zu kommunizieren. Ein weiteres Ziel ist es, so wenig Benutzerinput wie möglich zu verlangen.

5.2.1 Visualisierung

Bei „Personal Map“ ist der Eigentümer des Adressbuches in der Mitte des Tortendiagramms abgebildet. Alle Kontakte werden dann nach Wichtigkeit für den Eigentümer automatisch in das Diagramm eingeordnet. Hierfür werden die Informationen der Header aller E-Mails, die im letzten Monat empfangen oder versendet wurden, ausgewertet. Je wichtiger ein Kontakt für den Benutzer, desto näher an der Mitte wird er angeordnet. Die Wichtigkeit eines Kontakts hängt von der Häufigkeit der E-Mail Kommunikation ab. Die räumliche Anordnung der Kontakte hängt davon ab, wie häufig bestimmte Kontakte gemeinsam kontaktiert wurden. Je häufiger zwei Kontakte gemeinsam angeschrieben wurden, desto näher zu einander sind sie abgebildet.

Im normalen Leben werden Personen bestimmten Gruppen, z.B. Schulfreunde, zugeordnet. Deswegen werden Kontakte in „Personal Map“ auch in Gruppen geordnet angezeigt. Jede einzelne Gruppe stellt ein Stück im Tortendiagramm dar und wird farblich anders abgebildet. Die Größe der Tortenstücke ist abhängig davon, wie häufig der Benutzer mit den Teilnehmern einer Gruppe E-Mail-Kontakt hatte.

„Personal Map“ ist ein dynamisches Tool, somit kann sich die allgemeine Einordnung eines Kontaktes ändern, oder die Zugehörigkeit zu einer bestimmten Gruppe, aber auch die Größe der Tortenstücke.

5.2.2 Interaktion

„Personal Map“ soll vor allem ein Visualisierungstool für Kontaktdaten sein, in dem der Benutzer sein soziales Netzwerk betrachten kann. Somit existieren nicht besonders viele Interaktionsmöglichkeiten.

Dennoch hat es eine Zoomfunktion, mit der unabhängig voneinander, eingestellt werden kann wie viele Kontakte insgesamt und wie viele Gruppen angezeigt werden. Dieses ist vor allem für die Skalierbarkeit wichtig, denn somit ist die Funktionalität für jede Postfachgröße garantiert. Außerdem bietet es zwei verschiedene Ansichten. Zum einen die Tortendiagramm Ansicht (vgl. Abbildung 18), die im vorherigen Abschnitt schon beschrieben wurde, zum anderen eine Listenansicht.

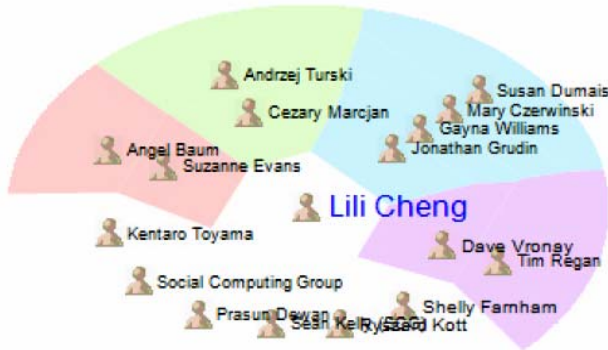


Abbildung 18. Tortendiagramm Darstellung in “Personal Map” (vgl. [2]).

„Personal Map“ bietet ein innovatives Adressbuch, das sehr einfach zu organisieren ist, da die Einordnung der Kontakte automatisch geschieht. Es ist auch denkbar dieses Tool im Alltag einzusetzen, dafür müsste es jedoch in einen E-Mail Client eingebunden werden.

Vergleicht man beide Visualisierungstechniken kann man kaum Gemeinsamkeiten feststellen. „faMailiar“ verfügt zum einen über eine ziemlich ausgeklügelte Technik zur Messung der sozialen Verbundenheit. Andererseits ist die Art der Darstellung, entlang einer Zeitachse, in „faMailiar“ nicht revolutionär. „Personal Map“ dagegen zieht für die Messung der sozialen Wichtigkeit nur die Häufigkeit des E-Mail Kontakts heran, ist aber bei der Darstellung (in Tortenform) und die Visualisierung von dynamischen Gruppen innovativer.

6 Ausblick

Für verschiedene Aspekte der E-Mail Kommunikation wurden Visualisierungstechniken vorgestellt. Zum einen wurde erklärt, warum eine thread-basierte Darstellung für E-Mails sinnvoll ist und einige Techniken dazu erläutert. Dabei wurden vor allem vier Kriterien, „Chronology“, „Relationships“, „Compactness“ und „Stability“, identifiziert, die bei dieser Darstellung eine große Rolle spielen. Auch für die Visualisierung von E-Mail Archiven wurden zwei sehr innovative Beispieltechniken vorgestellt, wobei es sich als sehr schwer erweist so große Datenmengen, wie sie in einem E-Mail Archiv zu finden sind, kompakt und sinnvoll darzustellen. Soziale Netzwerke spielen in der heutigen Zeit eine sehr große Rolle und daher wurden auch dazu Visualisierungstechniken erläutert. Das Problem dabei ist die richtige Datenbasis und Bewertung für den sozialen Status eines Kontaktes zu wählen, so dass die Visualisierung aussagekräftig ist.

Alle vorgestellten Techniken befanden sich noch im Entwicklungsstadium und wurden noch nicht in kommerziellen E-Mail-Clients eingebunden. Nach weiteren Verbesserungen wäre es aber durchaus vorstellbar, dass sich manche der Techniken durchsetzen und von den Benutzern akzeptiert und verwendet werden. Denn wächst

die E-Mail Kommunikation so wie bisher, wären einige Hundert Adressbucheinträge und einige Tausend archivierte E-Mails keine Seltenheit und ohne entsprechende Visualisierung nahezu nutzlos.

Literatur

1. Ducheneaut N., Bellotti V. E-mail as habitat: an exploration of embedded personal information management. In: *interactions*, 8: 5, 2001, S. 30-38.
2. Farnham S., Portnoy W., Turski A., Cheng L., Vronay D. Personal Map: Automatically Modeling the User's Online Social Network. In: *Human-Computer Interaction - INTERACT'03*. Zürich, Schweiz, September 1-5, 2003, IOS Press, S. 567-574.
3. Kerr B., IBM Research. THREAD ARCS: An E-Mail Thread Visualization. In: *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*. Seattle, WA, USA, Oktober 20-21, IEEE Computer Society, S. 211- 218.
4. Mandic M., Kerne A. Using intimacy, chronology and zooming to visualize rhythms in E-Mail experience; In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. Portland, OR, USA, April 02-07, 2005, CHI '05. ACM Press, New York, S. 1617-1620.
5. Newman P. S. Exploring discussion lists: steps and directions. In: *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*. Portland, Oregon, USA, July 14-18, 2002. JCDL '02. ACM Press, New York, 126-134.
6. Peet's Coffee. Abgerufen am 18.05.2006. http://www.peets.com/selector_tea/tea_selector.asp.
7. Regents of the University of California. How much information? 2003 - Internet. Abgerufen am: 18.05.2006. <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/internet.htm>.
8. Rohall S. L., Gruen D., Moody P., Wattenberg M., Stern M., Kerr B., Stachel B., Dave K., Armes R., Wilcox E. RE-Mail: a reinvented E-Mail prototype. In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. Wien, Österreich, April 24-29, 2004, CHI '04. ACM Press, New York, S. 791-792.
9. Sack W. Conversation map: a content-based Usenet newsgroup browser. In: *Proceedings of the 5th international Conference on intelligent User interfaces New Orleans, Louisiana, United States, January 09-12, 2000. IUI '00*. ACM Press, New York, S. 233-240.
10. Schmidt A., Butz A. Vorlesung Mensch-Maschine-Interaktion 2, Sommersemester 2005, *Information Visualization, Basic Visualization Techniques*. Abgerufen am 18.05.2006. <http://www.medien.informatik.uni-muenchen.de/lehre/ss2005/mmi2/vorlesung.html>.
11. Software & Support Verlag GmbH. 1.1 Billion E-Mail Users Worldwide: Study. Abgerufen am 18.05.2006. <http://www.entwickler.de/zonen/portale/psecom.id.99.nodeid.,lang.India,-sda.8678.,html>.
12. Venolia G. D., Neustaedter C. Understanding sequence and reply relationships within E-Mail conversations: a mixed-model visualization. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Ft. Lauderdale, Florida, USA, April 05-10, 2003. CHI '03. ACM Press, New York, S. 361-368.
13. Viégas F. B., Golder S., Donath J. Visualizing E-Mail content: portraying relationships from conversational histories. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Montréal, Québec, Canada, April 22-27, 2006. CHI '06. ACM Press, New York, S. 979-988.
14. Zhicheng L. Ivyfield: Visualizing E-Mail Archives. Undergraduate Research Opportunity Program (UROP) Project Report, Department of Computer Science, School of Computing, National University of Singapore, 2003/2004.

Instant Messaging und Anwendungen

Claudia Ruch

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
ruchc@ifi.lmu.de

Zusammenfassung Instant Messaging (IM) gewinnt sowohl im privaten als auch im geschäftlichen Bereich immer mehr an Bedeutung. Im Folgenden wird eine Einführung in IM gegeben. Dabei werden wichtige Begriffe und Funktionalitäten sowie technische Aspekte vorgestellt. Neben dem Nutzerverhalten werden Einsatzgebiete von IM dargestellt, wobei ein großer Schwerpunkt auf die geschäftliche Kommunikation gelegt wird. Abschließend werden aktuelle Forschungsthemen betrachtet.

1 Einleitung

In den 1960er Jahren waren die Instant Messages (sofortige Nachrichten) noch eine Zukunftsvision, die in Science-Fiction Geschichten des amerikanischen Psychologen und Schriftsteller Paul A. Linebarger beschrieben wurden und Nachrichten verkörpern, die „mit Überlichtgeschwindigkeit über interstellare Distanzen hinweg verschickt werden konnten“ [12]. Diese Nachrichten galten damals noch als unerschwinglich [12]. Mittlerweile hat sich dies sehr gewandelt. 80% der Internetuser im Alter von 14 bis 19 Jahren kommunizieren über Instant Messaging (IM) [14]. Auch in der Geschäftswelt wird diese Technologie immer mehr eingesetzt und verzeichnet großes Wachstum. Aufgrund der immer stärkeren Verbreitung von IM stellt diese Arbeit die wichtigsten Begriffe im Zusammenhang mit IM dar. Neben dem Verhalten des Nutzers werden Einsatzgebiete aufgezeigt und relevante Forschungsarbeiten erläutert.

Diese Arbeit ist in vier Punkte gegliedert. Zu Beginn werden allgemeine Grundlagen zu IM vorgestellt. Dazu gehören eine Erklärung des Begriffes, sowie eine kurze Abgrenzung zu ähnlichen Technologien. Es wird darauf aufbauend beschrieben, welche Funktionalitäten IM Anwendungen besitzen und was es neben Anwendungen auf dem Desktop PC noch für Arten von IM gibt. Nach einem kurzen Einblick in die Problematik, die sich aus proprietären Übertragungsprotokollen ergibt wird ein Überblick über gängige Standardprotokolle gegeben. Zum Abschluss der Grundlagen über IM wird ein abstraktes Modell vorgestellt, das den IM- und Anwesenheitsdienst beschreibt.

Im zweiten Teil der Arbeit wird das Verhalten von Nutzern und verschiedene Einsatzgebiete von IM dargestellt. Dazu werden anfangs verschiedene Anwendergruppen vorgestellt. Darauf aufbauend wird anhand von unterschiedlichen Nutzergruppen vorgestellt, wie sich Benutzer in IM Systemen verhalten und welche

Merkmale an der Art der Kommunikation bei diesen Nutzergruppen auffallen. Es wird anhand dieser Gruppen auch gezeigt für welche Zwecke IM verwendet wird. Zum Abschluss des Punktes wird ein Beispiel für die Einführung eines IM System in einem Unternehmen vorgestellt.

Anschließend werden Arbeiten aus der Forschung vorgestellt. Zum einen handelt es sich um ein Beispiel aus dem Bereich des mobilen IM, das besondere Anforderungen an die Technologie stellt und somit ein zentrales Thema der Forschung ist. Zum anderen wird die Entwicklung eines greifbaren Bedienelements für IM Anwendungen vorgestellt, das in der realen Welt platziert ist und den Benutzer dabei unterstützt, Einstellungen an seinem IM Client vorzunehmen.

Im abschließenden Teil der Arbeit werden die wichtigsten Punkte zusammengefasst. Außerdem wird die Verwendung von IM bei den verschiedenen Nutzergruppen verglichen und ein abschließendes Fazit gegeben.

2 Instant Messaging

Beginnend werden theoretische Grundlagen zu IM aufgezeigt, die als Grundlage für die restliche Arbeit dienen. Es geht dabei um Begriffsdefinitionen, sowie die funktionalen und technischen Aspekte von IM.

2.1 Definition von IM

Da es für IM, zu Deutsch sofortiger Nachrichtentransfer, in der Literatur keine eindeutige Definition gibt, werden zunächst einige Definitionen vorgestellt. Darauf aufbauend wird eine Arbeitsdefinition festgelegt, welche als Grundlage dieser Arbeit dienen soll.

Das IEC (International Engineering Consortium) beschreibt den Begriff folgendermaßen: “Instant messaging (IM) is an Internet protocol (IP)–based application that provides convenient communication between people using a variety of different device types” [3].

Eine weitere Definition nach Debbabi et al. [1] lautet: “Instant Messaging allows sending and receiving messages in real-time independent of locale.“

Mannan et al. erklärt die Bedeutung von IM wie folgt: “Instant Messaging (IM) is a type of communications service over the Internet that enables individuals to exchange text messages and track availability of a list of users in real-time” [6].

Während die erste Definition den Begriff sehr allgemein hält und er somit auf viele Arten von elektronischen Nachrichten zutrifft, kommen bei den anderen Definitionen zwei wichtige Eigenschaften zur Geltung. Einerseits läuft die Kommunikation beim IM in beinahe Echtzeit ab und andererseits gibt es die Möglichkeit zu erkennen, welche Benutzer gerade verfügbar sind. Diese Verfügbarkeitsanzeige ist sehr stark mit IM verbunden, wird in nahezu allen Quellen genannt und als Presence Service (Anwesenheitsdienst) bezeichnet.

Da der Begriff der Anwesenheit eine wichtige Rolle im IM spielt, wird er ebenfalls kurz definiert. Mannan et al. beschreibt Anwesenheit in seiner Arbeit [6] folgendermaßen: “Presence information reveals whether a user is logged in to an IM server or

not.“ Der Anwesenheitsdienst liefert also die Information, ob ein Anwender mit dem IM Service verbunden ist und wird von den meisten IM Systemen unterstützt.

Daraus ergibt sich folgende Definition für diese Arbeit. Unter IM versteht man den beinahe in Echtzeit stattfindenden Austausch von Nachrichten zwischen Anwendern, die denselben IM Dienst nutzen. Die meisten IM Dienste unterstützen auch die Anzeige von Anwesenheitsinformationen, die anzeigt ob ein User mit dem Server verbunden ist oder nicht.

Zur Email kann man den Begriff IM laut Debbabi et al. [6] wie folgt abgrenzen: “The primary difference between instant messaging and email is that instant messaging is intended for immediate end-user delivery.“ Der Unterschied liegt also darin, dass die gesendete Nachricht beim IM sofort auf dem Bildschirm des Empfängers erscheint, wenn der Nutzer angemeldet ist. Bei der Email kann der Abruf der Nachricht auch länger dauern.

2.2 Funktionalitäten von IM Diensten

Nachdem der Begriff des IM vorgestellt wurde, werden Funktionalitäten von IM Diensten beschrieben. Die Basisfunktionalitäten, die IM Anwendungen heutzutage bieten, hat Vogiazou [9] in ihrer Arbeit in vier Bereiche eingeteilt.

Zunächst gibt es die Basisfunktionalität der Kommunikation zwischen zwei Benutzern, die durch das senden und empfangen von Nachrichten realisiert wird. Für jeden Dialog gibt es ein eigenes Nachrichtenfenster, in dem gesendete und empfangene Nachrichten angezeigt werden. Normalerweise können Nachrichten zur späteren Kenntnisnahme auch an Personen geschickt werden, die gerade nicht verfügbar sind. Die meisten Anwendungen haben auch die Möglichkeit Nachrichten in einem Nachrichtenverlauf zu speichern, andere zeigen vorhergehende Nachrichten in dem Fenster an, in dem auch die Kommunikation stattfindet. Neben der eins-zu-eins Kommunikation unterstützen die meisten IM Dienste auch Gruppensitzungen, wobei man mit mehreren Teilnehmern gleichzeitig in einem Fenster kommunizieren kann [9].

Als zweite Funktionalität ist das Kontaktlistenmanagement zu nennen. Die Kontaktliste wird auch häufig als Buddy-Liste bezeichnet und man kann damit Profile von Freunden und Kollegen anlegen und diese in Gruppen organisieren. Es ist auch möglich eine Liste anzulegen, die Benutzer enthält, die blockiert werden sollen und somit keinen Kontakt mit der eigenen Person aufnehmen können. Die meisten IM Dienste haben die Möglichkeit durch Audiotöne zu signalisieren, wann Benutzer der Kontaktliste online oder offline gehen.

Eine weitere wichtige Funktionalität ist das Management der Verfügbarkeit eines Benutzers. Jeder Anwender hat einen Anwesenheitsstatus, der widerspiegelt, ob er gerade verfügbar ist. Dieser Status wird in der Kontaktliste der anderen Benutzer angezeigt. Die bekanntesten Zustände sind „online“ und „offline“. Den „online“ Status differenzieren unterschiedliche IM Dienste noch genauer. Abbildung 1 zeigt Auswahlmenüs zur Einstellung von Verfügbarkeitsinformationen von verschiedenen IM Anbietern. Die Zustandseinstellungen können entweder vom Benutzer manuell durchgeführt oder automatisch durch das System veranlasst werden. Der IM Client stellt den Zustand automatisch auf „offline“, falls der Benutzer längere Zeit keine Mausbewegungen durchgeführt hat.

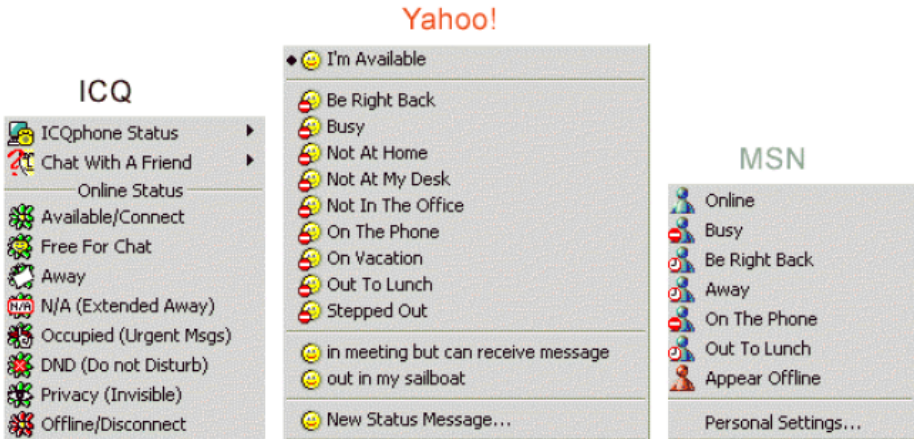


Abbildung 1. Abbildung von Auswahlmensüs zur Anwesenheitskontrolle von unterschiedlichen Instant Messenger Anbietern [9]

Als vierte Basisfunktionalität definiert Vogiazou [9] die Datenbanksuche. Benutzer können darin anhand von Interessen, Benutzernamen und anderen Eigenschaften andere Anwender finden und mit ihnen in Kontakt treten. Allerdings ist es den Usern freigestellt diese Funktionalität zu deaktivieren um nicht mit ungewünschten Gesprächspartnern kommunizieren zu müssen.

Neben diesen Basisfunktionalitäten bieten viele Anbieter zusätzliche Optionen an. Einige IM Dienste bieten die Möglichkeit andere Nachrichtenarten, wie SMS, Push-to-talk oder Voice-Over-IP zu integrieren. Außerdem können mit vielen Anwendungen auch Dateien übertragen und der Text formatiert werden.

2.3 Arten von IM

In dieser Arbeit sollen zwei verschiedene Arten von IM vorgestellt werden. Dabei handelt es sich um rein stationäres und mobiles IM.

Es gibt einerseits Anwendungen von IM Systemen, die auf einem Desktop PC oder Laptop laufen. Die bekanntesten proprietären Anwendungen sind ICQ, der Yahoo! Messenger, der AOL Instant Messenger und der MSN Messenger. Da jeder Anbieter sein eigenes Protokoll hat, sind die Dienste nicht miteinander kompatibel. Es gibt aber mittlerweile Anwendungen, die mehrere IM Systeme verwalten können, d.h. man kann mithilfe von einem Programm alle Profile, die man bei unterschiedlichen Anbietern hat, benutzen. Die bekanntesten Anwendungen sind Trillian und Miranda.

Andererseits gibt es von einigen oben genannten Herstellern IM Systeme, die auf mobilen Endgeräten laufen. Sobald an einer Konversation ein mobiles Endgerät beteiligt ist, spricht man von mobilem IM. Ein mobiles Endgerät stellt laut Parviainen [7] andere Anforderungen an IM, bietet aber auch neue Möglichkeiten. Mobiles IM muss mit Ressourceneinschränkungen zurechtkommen, da ein mobiles Endgerät eine kleinere Displaygröße, niedrigere CPU Leistung und Speichermöglichkeiten bietet. Außerdem muss berücksichtigt werden, dass die Verbindungsstärke je nach

Netzabdeckung stark schwanken kann. Es sollte auch möglich sein mehrere Geräte, wie z. B. PC und Handy gleichzeitig zu nutzen oder das Gerät mit dem man arbeitet einfach zu wechseln. Es ist bei mobilem IM auch denkbar, dass Kontextinformationen des Anwenders automatisch mit in den Status eingearbeitet wird. Die heutigen IM Anwendungen bieten diese Möglichkeiten noch nicht an.

2.4 Standardprotokolle

Standardisierungskomitees wie die IETF (Internet Engineering Task Force) oder die IMPS (Instant Messaging and Presence Services) Gruppe der OMA (Open Mobile Alliance) arbeiten daran, einheitliche Protokolle und Frameworks für IM und Anwesenheitsdienste zu definieren [1]. Im Folgenden sollen drei Protokolle kurz erläutert werden.

Die IETF hat ein Standardprotokoll entwickelt, das sich SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions) nennt. Es baut auf dem SIP (Session Initiation Protokoll) auf, das für Verbindungsaufbau, -kontrolle, - und -abbau von Kommunikationssitzungen zuständig ist [1]. SIMPLE erweitert das SIP durch Methoden für IM und Anwesenheitsdienste. Das SIP-basierte SIMPLE Protokoll zielt darauf ab Instant Messaging und Anwesenheitsdienste in die traditionelle Telefonkommunikation und in Webkonferenztools zu integrieren. In der Industrie hat dieses Protokoll bereits Zuspruch von zwei der größten Softwarekooperationen, Microsoft und IBM, erhalten [9].

Neben SIMPLE gibt es das ebenfalls von der IETF standardisierte Protokoll XMPP (Extensible Messaging and Presence Protocol), das seinen Ursprung 1998 in einem Projekt von Jeremie Miller hatte. Das Ziel des Projektes war es ein vollständig kompatibles offenes Instant Messaging Protokoll als alternative zu den proprietären Protokollen einzuführen. XMPP ist das Kernprotokoll des XML-basierten Jabber Protokolls [9].

Ein Protokoll, das für mobile Umgebungen optimiert ist, ist das Wireless Village Protokoll, das von Ericsson, Motorola und Nokia initiiert wurde und von der der IMPS Arbeitsgruppe der OMA standardisiert wird. Es beinhaltet Anwesenheitsmanagement, IM, Gruppenmanagement und das Verwalten von geteiltem Inhalt. In den Anwesenheitsdienst werden neben der Verfügbarkeitsanzeige auch Informationen über den Aufenthaltsort und das Gerät mit einbezogen [1].

2.5 Modelle der Umsetzung

Zur Umsetzung der beiden Kernfunktionen des IM, des Anwesenheits- und Nachrichtenübermittlungsdienstes wurde von der IETF IMPP Arbeitsgruppe ein abstraktes Modell [11] eingeführt. Sowohl proprietäre als auch offene Standards greifen auf dieses Modell zurück [8].

Das Modell des Anwesenheitsdienstes besteht aus 2 Typen von Clients, wie in Abbildung 2 zu erkennen ist. Einerseits gibt es einen Client, der sich Anwesenheitsdarsteller nennt und Informationen über die Verfügbarkeit eines Nutzers an einen Anwesenheitsdienst liefert. Die Beobachter dagegen rufen diese Information über die Verfügbarkeit ab. Eine Applikation enthält immer beide Typen. Der Anwesenheits-

darsteller stellt die eigenen Informationen anderen Applikationen zur Verfügung und der Beobachter ruft Informationen zu den bekannten Profilen ab [11].

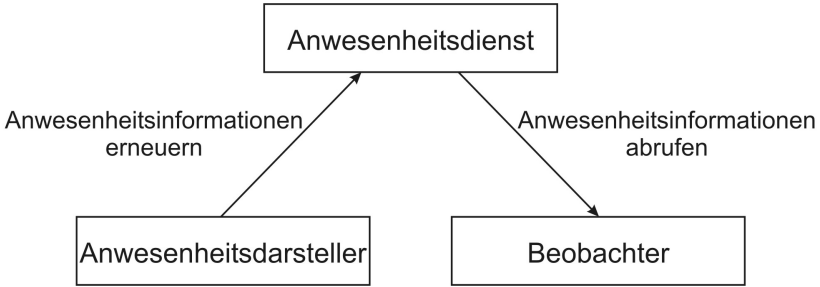


Abbildung 2. Modell des Anwesenheitsdienstes nach [8]

Die Beobachter werden, wie in Abbildung 3 zu sehen ist, noch einmal genauer klassifiziert. Es gibt Abonnenten, Abrufer und Poller. Ein Abonnent ist ein Beobachter, der sich bei einem Anwesenheitsdarsteller angemeldet hat und eine Mitteilung bekommt, sobald sich ein Status geändert hat. Der Abrufer dagegen fragt Gegenwartsinformation beim Anwesenheitsdienst ab. Der Poller ist ein Abrufer, der Abfragen in regelmäßigen Intervallen durchführt [11].

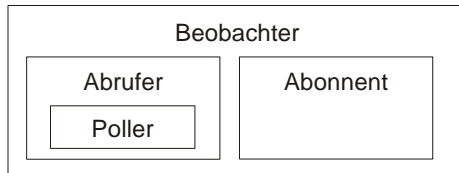


Abbildung 3. Klassifikation von Beobachtern nach [8]

Neben dem Anwesenheitsdienst gibt es auch einen IM Service, der ebenfalls aus zwei Arten von Clients besteht (siehe Abbildung 4). Da wären einerseits die Sender zu nennen, die geschriebene Nachrichten verschicken. Der Posteingang auf der anderen Seite bekommt Nachrichten vom zuständigen IM Service geliefert.

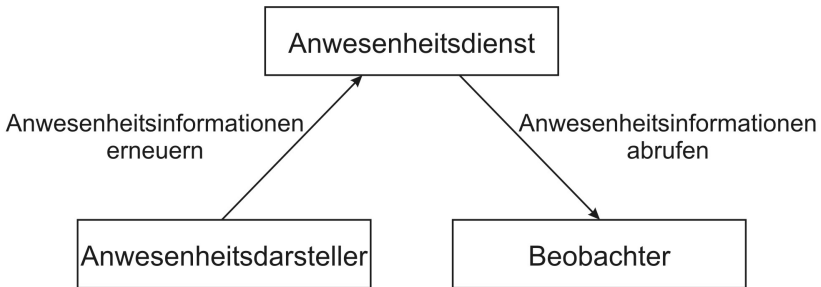


Abbildung 4. Modell des IM Dienstes nach [8]

Betrachtet man den Anwesenheits- und IM Dienst zusammen, dann laufen auf dem IM Server die beiden Dienste, die den IM Clients zur Verfügung gestellt werden. Die Clients liefern durch den Anwesenheitsdarsteller und den Sender Informationen, die der Server an andere Applikationen weitergibt. Außerdem bekommt der IM Client alle an ihn adressierten Nachrichten und Auskünfte zur Verfügbarkeit anderer Nutzer über seinen Posteingang und durch eine Art der angesprochenen Beobachter. Die Kommunikation zwischen unterschiedlichen Clients läuft also über einen Server ab.

3 Nutzerverhalten und Einsatzgebiete

Nachdem im vorherigen Teil der Arbeit die theoretischen Grundlagen für das Thema IM geschaffen wurden, wird nun dargestellt, welche Nutzer IM verwenden. Anschließend wird gezeigt, wie das Nutzerverhalten im privaten und geschäftlichen Bereich zu charakterisieren ist und welche Einsatzgebiete IM hat.

3.1 Nutzer von IM

Bei IM Anwendungen unterscheidet man zwischen Anwendungen im geschäftlichen und im privaten Bereich. Auf nachfolgender Abbildung der IDC, die eine Tochter der International Data Group ist [16], ist zu erkennen, dass sich IM im privaten Bereich schon sehr weit durchgesetzt hat. Im geschäftlichen Bereich ist dagegen weiterhin noch starkes Wachstum zu verzeichnen.

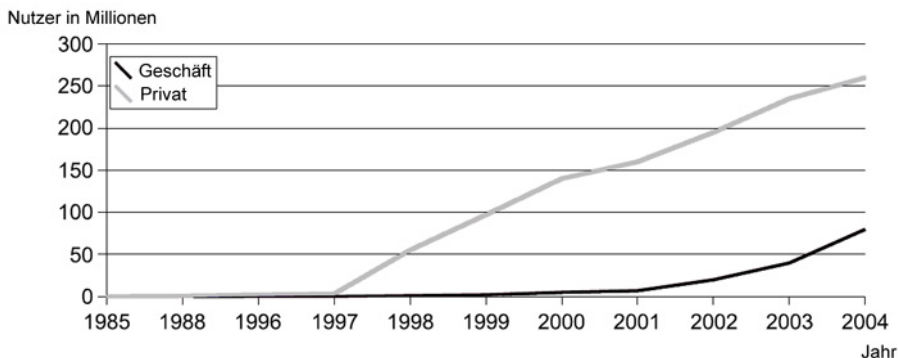


Abbildung 5. Wachstum von IM im privaten und geschäftlichen Bereich nach [8]

Die Benutzer von IM sind also sowohl Privat- als auch Geschäftsleute. Das Alter der User wird in Abbildung 6 näher beschrieben. Man kann erkennen, dass die Nutzungshäufigkeit von IM bei Jugendlichen am höchsten ist und mit zunehmendem Alter sehr stark fällt.

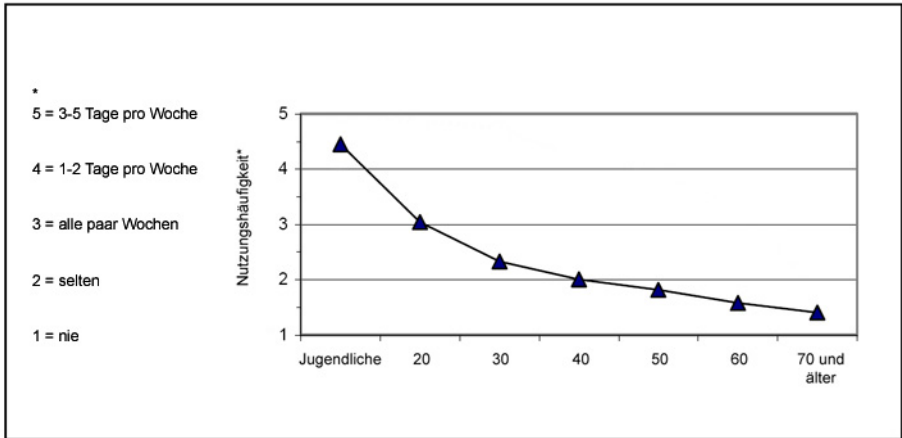


Abbildung 6. Nutzungshäufigkeit von IM in verschiedenen Altersklassen nach [15]

In den nächsten Abschnitten werden diese beiden Usergruppen genauer beschrieben. Es sollen das Nutzerverhalten und die Kommunikationscharakteristika für beide Gruppen dargestellt werden. Außerdem werden mögliche Einsatzgebiete des IM für diese Nutzer vorgestellt.

3.2 Privater Gebrauch von IM

Stellvertretend für die Gruppe der privaten Anwender wird das Nutzungsverhalten von Jugendlichen dargestellt, da sich die Kommunikation mit IM in dieser Gruppe am meisten durchgesetzt hat. Griter et al. [2] haben eine Studie zum Thema IM bei Jugendlichen durchgeführt und untersucht, wie oft IM genutzt wird, welche Technologien verwendet werden und für welche Zwecke es eingesetzt wird. An der Studie haben 16 User im Alter zwischen 15 und 20 Jahren teilgenommen.

Die Teilnehmer der Studie geben an, dass sie einen so hohen und lang anhaltenden IM Gebrauch haben, da für sie ein großes Bedürfnis besteht sich einerseits Freunden anzupassen und andererseits Kontakte zu pflegen. Da IM Systeme gratis sind und die erforderliche Hardware und Verbindung sowieso bereits zur Verfügung steht, ist es zudem eine kostengünstigere Lösung als die SMS. Alle Teilnehmer haben auch E-mailzugänge, die regelmäßig geprüft werden. Aber die Email wird für formale Zwecke, wie z. B. Kommunikationen mit Lehrern verwendet [2].

Griter et al. teilen die Studenten, die an der Studie teilgenommen haben anhand der Gebrauchshäufigkeit und den Verbindungsprofilen in 2 Gruppen ein. Unter einer diskreten Verbindung beschreiben Griter et al. den Verbindungsaufbau über ein Modem oder einen Computer, der mit anderen Personen geteilt wird und es dem Teenager dadurch nicht möglich ist durchgängig online zu sein. Anwender, die diesem Profil zugeordnet sind, beschreiben, dass ihre IM Sitzungen nicht länger als 3 Stunden dauern. Jugendliche, die über eine kontinuierliche Verbindung, d. h. eine DSL oder Ethernet Verbindung, oder einen eigenen Computer verfügen, lassen ihren Instant

Messenger dagegen bis zu 24 Stunden am Tag laufen, auch wenn sie gerade nicht am Computer sind. IM Fenster bleiben teilweise mehrere Tage offen und es werden immer wieder Nachrichten hinzugefügt [2].

Die Kommunikation über IM ist bei der betrachteten Gruppe hauptsächlich auf Freunde beschränkt, die sich in der realen Welt kennengelernt haben. Bei Hochschulstudenten, die nicht mehr zu Hause wohnen ist IM auch eine Möglichkeit mit der Familie in Kontakt zu treten. Der Messenger, der verwendet wird ist abhängig von der Gruppenzugehörigkeit. Es ist zu beobachten, dass eine soziale Gruppe von Freunden einen bestimmten Instant Messaging Client benutzt und neue Gruppenmitglieder werden dazu eingeladen diesen auch zu verwenden. Ein Teilnehmer an der Studie gibt an, dass er mit seinen Hochschulfreunden ein bestimmtes IM System gewählt hat und ein anderes mit seinen Oberschulfreunden verwendet. Auch der zu MSN, Yahoo! und ICQ vollständig kompatible IM Client Jabber wird eingesetzt um das Problem zu lösen, dass Freunde unterschiedlichen IM Systeme haben [2].

Die drei Haupteinsatzgebiete von IM in der betrachteten Gruppe sind die Durchführung von informellen Gesprächen, das Planen von Verabredungen und die Zusammenarbeit bei Hausaufgaben. Der erste Anwendungsbereich dient dabei zur Kontaktpflege. Die informellen Gespräche, welche sich auch als Geplauder des Alltages bezeichnen lassen, bilden die wichtigste Einsatzform. Themen bei diesen Gesprächen sind Ereignisse des Tages, Fernsehserien oder Klamotten, die gerade getragen werden [2].

Zusätzlich wird IM zur Planung eingesetzt, was den zweiten Einsatzbereich darstellt. Es werden Kinoverabredungen oder andere soziale Aktivitäten koordiniert. Der Vorteil ist, dass man auf dem Computer auch für die Verabredung relevante Informationen, wie z. B. den Filmbeginn im WWW recherchieren kann [2].

Als drittes Einsatzgebiet bei dieser Gruppe ist die Zusammenarbeit bei Hausaufgaben zu nennen. IM wird eingesetzt um sich gegenseitig zu unterstützen und miteinander über Hausaufgaben zu diskutieren. Ein Teilnehmer an der Studie gab an, dass er IM als Lehrtool benutzt. Da er als Assistent arbeitet, gibt er für jüngere Kommilitonen Uhrzeiten bekannt, an denen er über IM zur Unterstützung bei Programmieraufgaben erreichbar ist. Alle Teilnehmer der Studie gaben außerdem an, dass sie IM nutzen während sie noch andere Arbeiten auf dem Computer erledigen.

3.3 Geschäftlicher Gebrauch von IM

Im Bereich der geschäftlichen Kommunikation über IM Systeme wird eine Studie über IM am Arbeitsplatz von Nardi et al. [10] vorgestellt. Sie beschreibt wie die Kommunikation abläuft, zu welchen Zwecken IM genutzt wird und was es für weitere Auswirkungen von IM am Arbeitsplatz gibt. Dazu wurde das Nutzerverhalten von 20 Mitarbeitern untersucht.

Folgende Konversation aus der oben genannten Studie zeigt, wie eine typische IM Sitzung zwischen der Sekretärin Melissa und ihrem Chef Alan abläuft. Melissa teilt sich eine Arbeitsnische mit Jackie, die als Sekretärin für Alan's Chef Sam Jones arbeitet.

melissa (8:33:32 AM): *The fire is out???????? [there has been an embarrassing public relations problem]*

Auto response from alan: (8:33:32 AM): *I'm idle...may be asleep. [Alan was there but working on another computer. The message was a personalized automatic response.]*

alan (8:33:45 AM): *not quite...still putting it out*

melissa (8:37:13 AM): *I can send some water. Just talked with Georgina....Marsha is running around with her head cut off!!!!*

alan (8:37:29 AM): *just put Carl on my calendar at 10 am, for half-hour. [Carl was able to help solve the problem.]*

melissa (8:37:45 AM): *You got it!!!!*

melissa (8:38:43 AM): *By the way....I can go to lunch if I can catch a ride with you...Beth has the car for lunch.*

alan (8:38:56 AM): *fine with me!*

alan (8:39:12 AM): *also, do you know when will sam jones be back?*

[Melissa turned to Jackie who kept Sam's calendar and asked her about Sam's schedule.]

melissa (8:40:39 AM): *Sam will be coming in on June 1 as of this moment*

alan (8:40:56 AM): *oh...not here this fri, eh?*

melissa (8:41:11 AM): *NO....He is in Hawaii at the moment.*

alan (8:41:24 AM): *right...for the shareholders meeting.*

melissa (8:42:09 AM): *You got it...Making Gail crazy needing paperwork from Stan's group yesterday at 4pm and they are out on an Offsite....*

alan (8:42:34 AM): *: -)*

Abbildung 7. Gesprächsablauf über ein IM System im Büro [10]

Anhand des Kommunikationsablaufes kann man erkennen, dass innerhalb von knapp 10 Minuten, ein Problemthema, das „fire“ angesprochen wurde, ein Meeting mit Carl und eine Verabredung zum Mittagessen vereinbart wurde. Außerdem wurden Informationen über den Terminplan von Alans Chef ausgetauscht und über die Atmosphäre im Büro gesprochen. Währenddessen gingen beide Gesprächspartner nebenbei anderen Aktivitäten nach. Alan hatte gleichzeitig ein Telefongespräch abgewickelt und seine Emails gelesen [10].

Der allgemeine Tenor der Nachrichten aus obigem Beispiel ist eher locker, informativ und freundlich. Es wurden auch kleine übliche Witze gemacht. Im Gegensatz dazu würde man sich bei geschäftlichen Emails eher förmlich ausdrücken. Alan äußert sich folgendermaßen darüber: „I use email more like the adult thing. IM is more the fun thing“ [10].

Zum Charakter der Kommunikation in diesem Gespräch kann man sagen, dass es sich um sehr kurze, präzise Nachrichten handelt. Isaacs et al. [4] hat in einer anderen Studie, in der 21.213 Konversationen ausgewertet wurden, herausgefunden, dass ein Dialog im Durchschnitt 4 Minuten und 23 Sekunden dauert und dabei 17,2 Nachrichten geschrieben werden. Die einzelnen Mitteilungen bestehen dabei aus durchschnittlich 13,5 Wörtern. Dies zeigt, dass IM für sehr kurze Gespräche verwendet wird. Weitere Charakteristiken von IM Konversationen sind laut Isaacs et al. [4], dass man mehrere Arbeiten zugleich machen kann, was bei einem Gespräch bei dem sich die

Gesprächspartner von Angesicht zu Angesicht gegenüberstehen nicht möglich wäre. Als dritte Eigenschaft ist zu nennen, dass während Gesprächen oft das Medium gewechselt wird. Dies ist häufig der Fall, wenn der Dialog zu komplex wird und man sich persönlich trifft, oder das Telefon mit einbezieht [4].

Nardi et al. [11] hat vier Einsatzgebiete von IM am Arbeitsplatz identifiziert. IM wird für schnelle Fragen und zur Klärung, zur Koordination und Zeitabstimmung, zur Abstimmung von spontanen Meetings und dazu verwendet um mit Freunden und der Familie in Kontakt zu bleiben. Der entscheidende Vorteil von IM gegenüber anderen Medien im ersten Einsatzgebiet, der Klärung von schnellen Fragen, ist seine Effizienz. IM erlaubt einen schnelleren Austausch als Email, umgeht aber den Aufwand einer direkten Kommunikation von Angesicht zu Angesicht [11]. Ein Programmierer der an der Studie von Nardi et al. [11] teilgenommen hat beschreibt, dass er Anfragen für kleine Änderungen an einer Webseite bekommt, die Modifikationen durchführt und über IM schnelles Feedback über den Erfolg bekommt.

Auch im zweiten Anwendungsgebiet, der Koordination und Zeitabstimmung wird IM aufgrund seiner Unmittelbarkeit bevorzugt. Bei der Terminkalenderabstimmung sollte man zur besseren Planung so schnell wie möglich den Kalender der betroffenen Personen kennen. Bei einer Email besteht die Gefahr, dass sie erst mit großer Verzögerung gelesen wird. IM hat dabei außerdem den Vorteil, dass man sich Zeit spart, weil man nicht auf Formalitäten achten muss [11].

Zusätzlich wird IM genutzt um Treffen, die von Angesicht zu Angesicht stattfinden zu arrangieren. Trotz des Drucks in der heutigen Arbeitswelt, in der sich die Pflege sozialer Kontakte als schwierig gestaltet, ist soziale Bindung eine wichtige Komponente. Die Mitarbeiter verabreden sich z. B. zum Essen, wenn sie eine Pause haben, oder zu kurzfristigen Meetings. Die Koordination von spontanen Angelegenheiten wird bevorzugt per IM gemacht, da es schnell abläuft [11].

Der letzte Punkt der Einsatzgebiete von IM am Arbeitsplatz, ist der soziale Kontakt zu Freunden und der Familie. Auch während der Arbeitszeit werden Planungen für den Feierabend vollzogen [11]. Doch obwohl IM die Möglichkeit bietet private Angelegenheiten mit Freunden, der Familie und Arbeitskollegen zu besprechen liegt der Schwerpunkt der Gespräche mit 62% auf Themen, die die Arbeit betreffen [4].

Eine andere wichtige Erkenntnis, die Nardi et al. [11] in ihrer Studie aufgedeckt haben ist, dass Mitarbeiter ein besonderes Gefühl für soziale Verbundenheit zu Arbeitskollegen entwickeln. Viele Teilnehmer erkennen einen Mehrwert darin, dass sie anhand der Kontaktliste sehen können, welcher ihrer Arbeitskollegen gerade online ist. Diese Tatsache erweckt in den Mitarbeitern eine besondere Empfindung. Ein Teilnehmer aus der Studie beschreibt es folgendermaßen: „You feel like you know where other people are, so you feel like you're not the only one working on a weekend.“[11] Instant Messaging hat also die Funktion Menschen auf eine gewisse Weise zu verbinden. So werden auch einfach nur kleine „Guten Morgen“ Nachrichten ausgetauscht, die man nicht in einer Email schicken würde [11].

3.4 Planung und Einführung eines IM Systems am Arbeitsplatz

Im Folgenden wird beispielhaft ein IM System in einem Unternehmen beschrieben. Huang et al. [13] haben ein IM System in einem Unternehmen eingeführt, das die Kommunikation von Mitarbeitern unterstützen soll, die gerade nicht an ihrem Arbeitsplatz sitzen.

Zunächst soll dargestellt werden, warum es sinnvoll ist ein solches System einzuführen. Einerseits zeigt eine Studie nach Whittaker et al. [17], dass Mitarbeiter während ihres Arbeitstages viel Zeit an Aktivitäten verbringen, die nicht direkt an ihrem Schreibtisch, sondern außerhalb ihres persönlichen Arbeitsplatzes stattfinden. Somit wird die Möglichkeiten von IM eingeschränkt. Die Aktivitäten, die Mitarbeiter außerhalb ihres Arbeitsplatzes erledigen, sind formale und informale Besprechungen. Andererseits zeigt die bereits aufgeführte Studie von Nardi et al. [10], dass IM Dienste Mitarbeiter bei ihrer Arbeit unterstützen können. Da sie aber bei Besprechungen die nicht an ihrem Schreibtisch, aber im gleichen Gebäude stattfinden, keinen Zugang zu ihrem Instant Messenger haben, ist es schwierig Kontakt zu Arbeitskollegen aufzunehmen. Die Mitarbeiter müssten erst wieder zurück zu ihrem Schreibtisch um Nachrichten zu schicken. Da von Huang et al. [13] beobachtet wurde, dass Kontaktaufnahmen zu Mitarbeitern oftmals unmittelbar vor Meetings geschehen, ist es sinnvoll eine Möglichkeit der Kontaktaufnahme an einem nahen öffentlichen Ort im Bürogebäude zu haben.

Das eingesetzte System besteht aus zwei Teilen. Einerseits aus einem Event Display, das Informationen zu bevorstehenden Veranstaltungen und Ankündigungen darstellt und andererseits aus einem einfachen Messaging Client, der es Benutzern erlaubt Instant Messages an Mitarbeiter, die in der Kontaktliste aufgeführt sind zu senden. Beim Senden gibt es unterschiedliche Modi. Man kann eine Nachricht an einen bestimmten Nutzer oder durch eine Broadcast Aufruf an alle Benutzer der Kontaktliste verschicken. Es gibt keinen Anmeldeprozess, sondern alle Nutzer verwenden die gleiche IM Here Kennung. Auf nachfolgender Abbildung kann man das System erkennen, das direkt vor dem Gemeinschaftsraum aufgestellt wurde. Dieser Ort wurde gewählt, da er sehr zentral liegt und dort sehr viele Besprechungen stattfinden.



Abbildung 8. Darstellung des IM Here Systems [13]

Während der ersten sechs Wochen nach der Einführung des Systems wurden Evaluationen durchgeführt. Dabei wurde festgestellt, dass die Mitarbeiter das System generell positiv aufgenommen haben. Sie fanden es sehr nützlich, vor allem wegen der Nähe zum Gemeinschaftsraum. Typische Szenarien für den Einsatz dieses Systems waren die Erinnerung an ein Treffen, wenn jemand nicht pünktlich war oder die Klärung kurzer Fragen, die während eines Meetings aufkamen. Obwohl das Tool eigentlich dazu entwickelt wurde, damit Leute Anfragen vom Gemeinschaftsraum aus schicken können, kam auch die Kommunikation in die andere Richtung vor. Es wurden dabei Anfragen von Benutzern aus den Büroräumen an Personen die gerade zufällig vorbeikamen geschickt. Eine solche Anfrage stellt auch folgendes Beispiel dar:

11:11:53AM Dave: Dear passerby -- Can you tell me if [the common room] is scheduled from 1- 2PM? (I'd look at the server myself, but I can't get there from here for some weird reason...)

11:14:51AM IM Here: 1 to 2 PM today looks open. 2 to 2:30 is booked.

11:15:06AM Dave: Thanks! (I'm at 2 - 2:30) -- BTW - who are you?

Abbildung 9. Gesprächsablauf auf dem IM Here System [13]

In diesem Beispiel wird ausgeführt, wie ein Mitarbeiter eine Nachricht an den IM Here Client schickt um zu klären, ob der Gemeinschaftsraum zwischen 13 und 14 Uhr besetzt ist. Diese Anfrage wird dann von einem vorbeilaufenden Mitarbeiter beantwortet. Dadurch ergibt sich eine Zeitersparnis, weil der Mitarbeiter nicht selbst zum Gemeinschaftsraum gehen musste. Außerdem haben Anwender noch eine andere unerwartete Nutzungsmöglichkeit des Im Here Messenger zusammen mit dem Event Display entdeckt. Teilweise werden Events aus dem Eventkalender kopiert und mit Hilfe des Messengers verschickt [13].

4 Forschungsgebiete

In diesem Abschnitt werden zwei Einblicke in die aktuelle Forschung gegeben. Zunächst werden Anforderungen und Möglichkeiten von mobilen IM Systemen dargestellt und eine mögliche Umsetzung eines solchen Systems beschrieben. Anschließend soll in einem zweiten Punkt ein greifbares Bedienelement vorgestellt werden, mit dessen Hilfe man seinen Status verändern kann.

4.1 Mobiles IM

Wie bereits unter 2.3 beschrieben setzten vor allem mobile Endgeräte neue Anforderungen, wie kleinere Displaygröße, verminderte Speicherkapazität und CPU Leistung, sowie Schwankungen der Signalstärke, an das Instant Messaging. Parviainen et al. [7]

beschreiben ein mobiles IM System, das für die gleichzeitige Nutzung von mobilen und stationären Endgeräten ausgelegt ist.

Heutige IM Systeme, wie ICQ, AIM und der MSN Messenger sind primär für Desktop Systeme entwickelt worden und setzen voraus, dass der Anwender nur an einem Gerät, bzw. an einem zu jedem Zeitpunkt angemeldet ist. Im mobilen IM System von Parviainen et al [7] kann ein Benutzer auf mehreren Geräten gleichzeitig angemeldet sein. Daten, wie die Kontaktliste sowie alle Freunde, die darin aufgenommen sind und der Nachrichtenverlauf müssen somit auf dem Server und nicht lokal auf dem Gerät gespeichert werden. Außerdem wird die Kontaktliste zwischen den Geräten synchronisiert, wenn man auf einem Gerät z. B. einen neuen Kontakt hinzufügt. Es muss auch einen einfachen Mechanismus zum Wechseln des Gerätes geben, ohne, dass man seinen Status ständig zwischen Online und Offline wechselt. Eingehende Nachrichten sollen dabei auf alle angemeldeten Clients geschickt werden. Auch der Punkt der Verbindungsstärke wird berücksichtigt, da diese in mobilen Netzen stark schwanken kann. Der Status wird nicht sofort auf nicht verfügbar geändert, wenn nur für kurze Zeit kein Signal zur Verfügung steht. Es ist zusätzlich möglich, dass man bei mobilen Anwendungen zusätzliche Information über den Aufenthaltsort des Anwenders an andere Teilnehmer weitergibt, was aber erst in nachfolgenden Arbeiten mit aufgenommen wird und in dieser Arbeit von Parviainen noch nicht berücksichtigt ist [7].

Die oben beschriebenen Ziele können durch einen Mobilien Instant Messenger Server (MIM Server), der auf einem PC, wie beispielsweise dem normale Desktop PC, läuft, realisiert werden. Verschiedene IM Dienste können über mobile IM (MIM) Plugins angeschlossen werden, die den Server mit Kontaktlisten, Statusinformationen, Funktionen zum Setzen des eigenen Status und zum Senden und Empfangen von Nachrichten, versorgen. Der Server kümmert sich um die gesamte Synchronisation und speichert erforderliche Daten, wie ungelesene Nachrichten usw. Die verschiedenen Clients können sich am Server anmelden und bekommen für sie relevante Informationen zugeschickt. Der MIM Server ist immer zu allen IM Systemen angeschlossen. Wenn längere Zeit kein Client verfügbar ist, kann der Zustand auf vorkonfigurierte Werte gesetzt werden. Falls Clients keine ständige Verbindung mit dem Server halten können, können sie auch in einen Polling Modus laufen, in dem sich das Gerät in bestimmten Zeitabständen mit dem Server verbindet und synchronisiert [7].

Die vollständige Architektur für das IM System von Parviainen et al. [7] kann man auf der folgenden Abbildung erkennen.

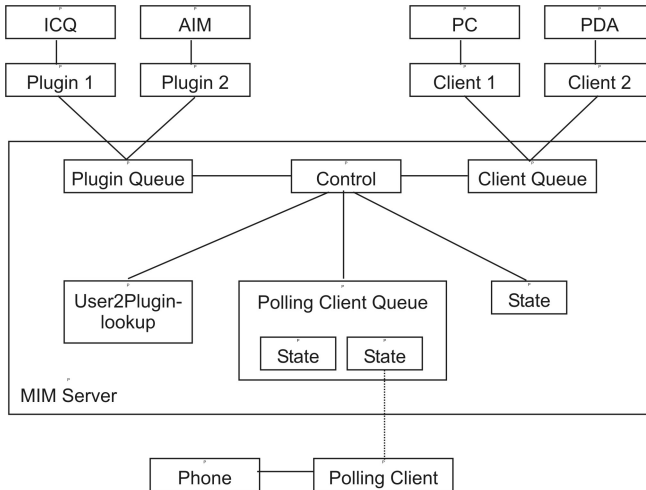


Abbildung 10. Abbildung der Architektur des mobilen IM System nach Parviainen et al. [7]

In Zukunft wird noch weiter an diesem Projekt gearbeitet, damit auch Ortsinformationen über den Anwender mit berücksichtigt werden können.

4.2 Greifbare Bedienelemente im IM

In den meisten IM Systemen kann man seine Verfügbarkeit anhand von vordefinierten Zuständen auswählen. Laut einer Studie der Forschungsgruppe Embedded Interaction [5] an der 16 Personen teilgenommen haben, stellen nur 2 Personen diesen Zustand manuell ein. Somit wird die Verfügbarkeit bei den meisten Leuten automatisch durch das System angepasst, was sich dadurch zeigt, dass der Zustand nur zwischen verfügbar und nicht verfügbar schwankt und nicht differenzierter ausgedrückt wird, in welcher Situation sich der Benutzer befindet. Es wurde außerdem beobachtet, dass Menschen dazu neigen mit physischen Objekten zu spielen, während sie etwas erledigen [5].

Aus diesen beiden Tatsachen ist an der Forschungsgruppe Embedded Interaction die Idee entstanden, ein greifbares User Interface zu entwerfen, mit dessen Hilfe man den Status einer Anwendung einstellen kann. Das daraus entwickelte Bedienelement kann man in folgender Abbildung erkennen.

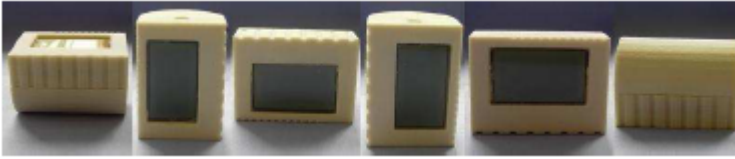


Abbildung 11. Darstellung der verschiedenen Verfügbarkeitszustände mithilfe des greifbaren Bedienelements [5]

Es handelt sich um eine Box, die 6.5 cm x 4.5 cm x 4.0 cm (Breite x Höhe x Tiefe) groß ist und sechs verschiedene Verfügbarkeitszustände darstellen kann. Jeder Zustand ist einer bestimmten Ausrichtung der Box zugeordnet. Zusätzlich hat sie ein Display auf dem Feedbackinformation dargestellt werden kann. Ausgestattet mit Sensoren kann die Box erkennen, wie sie ausgerichtet ist und wann sie gedreht wird, und somit einen neuen Status an die IM Anwendung melden [5].

Das greifbare Bedienelement wurde in oben genannter Studie evaluiert. Obwohl das Gerät den Personen ohne zusätzliche Erklärungen übergeben wurde konnten sie schnell erkennen, wie es verwendet wird. Zusammenfassend kann man sagen, dass das Ergebnis positiv ausgefallen ist, da der Prototyp schnell akzeptiert wurde und die Teilnehmer großes Interesse gezeigt haben. Weiterführende Entwicklungen sollen dazu dienen, den Prototypen zu verbessern und die Funktionalität dadurch zu erweitern, dass neben der Anwesenheitsfunktionalität auch Kontextinformation geliefert wird [5].

5 Schluss

Zusammenfassend lässt lassen sich sagen, dass IM ein wichtiger Bestandteil der täglichen Kommunikation geworden ist. Die sehr hohe Bedeutung von IM bei Jugendlichen wurde anhand der Nutzerstudie in 3.2 dargestellt. Man konnte erkennen, dass diese Gruppe IM zum Führen von informellen Gesprächen, zur Planung von Verabredungen und zur Zusammenarbeit nutzen. Die Wahl der verwendeten Software ist abhängig von der Gruppenzugehörigkeit und es ist eher unwahrscheinlich, dass man mit Menschen kommuniziert, die man nicht aus der realen Welt kennt. Im Gegensatz zur Email, werden in dieser Gruppe über IM weniger formale Konversationen geführt.

Im geschäftlichen Bereich ist die Bedeutung von IM noch nicht so ausgeprägt wie im Privaten, aber der Trend geht hier immer stärker in dieselbe Richtung. Man hat gesehen, welche Vorteile die sehr unkonventionelle und schnelle Kommunikation von IM gegenüber der Email hat. Neben der Vereinfachung der Kommunikation, spielt der soziale Aspekt eine herausragende Rolle. Die Einsatzgebiete im geschäftlichen Bereich sind die Klärung von schnellen Fragen, die Koordination und Zeitabstimmung, die Abstimmung von spontanen Meetings und die Pflege von sozialen Kontakten mit Freunden und der Familie. Der Tenor der Kommunikation lässt sich im Gegensatz zur Email als eher umgangssprachlich und locker bezeichnen. Die Kom-

munikation beim IM zeichnet sich dadurch aus, dass die Nachrichten kurz sind, die Aufmerksamkeit des Users ist nicht allein der IM Kommunikation gewidmet und bei komplizierten Gesprächen ist es üblich, dass das Kommunikationsmedium gewechselt wird.

Das Beispielsystem in 3.4 hat gezeigt, dass IM nicht nur am Arbeitsplatz direkt, sondern auch in Räumen Einzug erhält, in denen mehrere Benutzer ein IM System verwenden. Wie man Beispiel von IM Here erkennen konnte, kann IM auch eine Vereinfachung der Kommunikation abseits des Arbeitsplatzes darstellen.

Nach Betrachtung der beiden Benutzergruppen kann man feststellen, dass das Nutzerverhalten sehr ähnlich ist. Grundsätzlich sind die Einsatzgebiete bei den betrachteten Gruppen sehr ähnlich, nur die Inhalte der Kommunikation sind verschieden. Auch der nicht formale Gesprächsablauf und das Wechseln der Aufmerksamkeit zwischen verschiedenen Anwendungen sind identisch. Es werden von beiden Gruppen Gespräche mit Leuten, die bereits aus der Realität bekannt sind, bevorzugt.

Trotz der vielen Anwendungsmöglichkeiten und Vorteile, die das IM bietet ist die aktuelle Forschung daran interessiert, weitere Verbesserungen zu entwickeln. Die Anpassung herkömmlicher IM Systeme an Voraussetzungen für mobiles IM, sowie die Einführung von Kontextinformationen stellen einen wichtigen weiteren Schritt für das kontinuierliche Wachstum von IM dar. Zusätzliche neue Ideen für die einfachere und erfolgreichere Bedienung von IM Applikationen, wie das beschriebene greifbare Bedienelement zur Kontrolle des Anwesenheitsstatus stellen erfolgsversprechende neue Möglichkeiten dar.

Literatur

1. Debbabi, M., Rahman, M., "The War of Presence and Instant Messaging: Right Protocols and APIs", In: Proceedings of the 1st IEEE Consumer Communications and Networking Conference, Las Vegas, NV, USA, January 2004, S. 341-346.
2. Grinter, R. E., Palen, L., „Instant messaging in teen life”, In: Proceedings of the 2002 ACM conference on Computer supported cooperative work, New Orleans, Louisiana, USA, November 16-20, 2002.
3. International Engineering Consortium, "Instant Messaging Definition und Overview", http://www.iec.org/online/tutorials/instant_msg/index.html, abgerufen am 06.06.2006.
4. Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D. J., Kamm, C., "The character, functions, and styles of instant messaging in the workplace", In: Proceedings of the 2002 ACM conference on Computer supported cooperative work, New Orleans, Louisiana, USA, November 16-20, 2002.
5. Kranz, M., Holleis, P., Schmidt, A., „Ubiquitous Presence Systems“, In: Proceedings of the 2006 ACM Symposium on Applied Computing (ACM SAC), Dijon, France, April 2006.
6. Mannan, M., van Oorschot, P. C., „Secure public Instant Messaging: A Survey” In: Proceedings of the 2nd Annual Conference on Privacy, Security and Trust (PST'04), Fredericton, NB, Canada, Oct. 2004, pages 69–77.
7. Parviainen, R., Parnes, P., "Mobile Instant Messaging", In: Proceedings of ICT 2003, Tahiti, February 23 - March 1, 2003.
8. Salin, P., "Mobile Instant Messaging Systems - A Comparative Study and Implementation", Master's Thesis, Helsinki University of Technology, Helsinki, 2004.
9. Vogiazou, Y, The Open University, "Wireless Presence and Instant Messaging", November 2002, http://www.jisc.ac.uk/uploaded_documents/tsw_02-07.pdf, abgerufen am 06.06.2006.
10. Nardi, B. A., Whittaker, S., Bradner, E.; "Interaction and outeration: instant messaging in action", In: Proceedings of the 2000 ACM conference on Computer supported cooperative work, Philadelphia, Pennsylvania, United States, December 2000, p.79-88.
11. Day, M., Rosenberg, J., Sugano, H., The Internet Society, "RFC 2778 / RFC2778 A Model for Presence and Instant Messaging", <http://www.rfc-archive.org/getrfc.php?rfc=2778>, February 2000.
12. wikipedia, "Instant Messaging", http://de.wikipedia.org/wiki/Instant_Messaging, abgerufen am 06.06.2006.
13. Huang, E. M., Russell, D. M., Sue, A. E., "IM Here: Public Instant Messaging on Large, Shared Displays for Workgroup Interactions", In: Proceedings of CHI 2004, Vienna, Austria, 2004.
14. ECIN, "Neue Kommunikation boomt", 19.01.2006, <http://www.ecin.de/news/2006/01/19/09101/index.html>, abgerufen am 06.06.2006.
15. Boneva, B. S., Quinn, A., Kraut, R. E., Kiesler, S., Shklovski, I. "Teenage communication in the Instant Messaging Era" In: R. Kraut, M. Brynin, and S. Kiesler (Eds.), Domesticating Information Technology Oxford University Press.
16. IDC, "About IDC", <http://www.idc.com/about/about.jsp>, abgerufen am 20.06.2006.
17. Whitaker, S., Frolich, D., Daly-Jones, O., "Informal Workplace Communication: What is it Like and How Might We Support it?" In Proceedings of CHI 1994, Boston, Massachusetts, USA, 1994.

„Generation Blogging“: Weblogs erobern das Internet

Ina Zumbruch

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
zumbruch@cip.ifi.lmu.de

Zusammenfassung Seit einigen Jahren gibt es einen Trend im Internet mit Namen Blogging. Normale Bürger machen Teile ihres Lebens öffentlich oder verstehen sich als Journalisten, die neue Nachrichten entdecken und ihren Mitbürgern zur Verfügung stellen. Auch Unternehmen haben das Bloggen als neue Art der Unternehmenspräsentation entdeckt. Es stellt sich also die Frage, was das Bloggen so attraktiv macht? Die folgende Arbeit gibt einen Überblick über die Erfolgsgeschichte des Bloggings und die verschiedenen Arten von Blogs. Dabei wird auf gesellschaftliche, kommunikationswissenschaftliche und technische Aspekte eingegangen.

1 Einleitung

Wenn man Blog oder Weblog hört, dann denkt man – sofern einem der Begriff geläufig ist – wahrscheinlich als erstes entweder an eine Linksammlung oder an ein öffentliches Tagebuch im Internet. Die Tätigkeit, ein Blog zu aktualisieren, nennt man „blogging“ und jemand, der ein Blog betreibt und Artikel dafür verfasst, ist ein „Blogger“ [9]. Doch was genau macht eine Internetseite zu einem Blog oder Weblog? Gill [9] identifiziert sieben grundlegende Charakteristika:

1. *Formatierung*: In umgekehrter chronologischer Reihenfolge
2. *Aktualität*: Datumsvermerk zu jedem Eintrag
3. *Zuordnung*: Links zu Artikeln, Dokumenten und anderen Blogs, die mit dem aktuellen Artikel in Bezug stehen
4. *Zugriff auch auf alte Inhalte*: Einträge werden archiviert
5. *Blogrolling*: Verwandte Blogs werden verlinkt
6. *Leichte Syndikatsbildung*: RSS oder XML
7. *Stimme*: Leidenschaft etwas mitzuteilen

Diese Punkte werden in der vorliegenden Arbeit näher erläutert. Zuerst erfolgt ein kurzer Überblick über die noch junge Geschichte der Weblogs, anschließend werden die vielfältigen Gründe untersucht, die einen Internetnutzer zum Blogger werden lassen. Da technische Begriffe auch beim Bloggen eine Rolle spielen, werden die wichtigsten Begriffe in einem kleinen Blogging-ABC erläutert. Zum Abschluss erfolgt ein Ausblick, in welche Richtung sich das Bloggen entwickeln könnte.

2 Die Erfolgsgeschichte der Weblogs

Der Begriff Weblog wurde erstmals 1997 von Jorn Barger verwendet, der das Blog „Robot Wisdom“¹ betreibt [18]. Doch worin unterscheiden sich Blogs von „normalen“ Webseiten? Barger bezeichnete als Blogs Webseiten, die in umkehrter chronologischer Reihenfolge auf andere Webseiten verweisen, die der Weblogger interessant findet [18] [1]. Weitere Merkmale von Blogs: Sie werden regelmäßig modifiziert und die Leser können Kommentare zu den einzelnen Einträgen hinterlassen [1].

1997 war auch die Geburtsstunde von Slashdot². Der College-Student Rob Malda aus den USA entwickelte das Portal aus freier Software, das sich anfangs kaum von anderen Weblogs unterschied. Er veröffentlicht darauf „Neuigkeiten für Nerds“, also Wissenswertes für Computerfreaks. Slashdot erfreut sich großer Beliebtheit und hat vier Millionen Besuche bzw. 80 Millionen Pageviews pro Monat (Stand Mai 2004) [18]. Ein typischer Slashdot-Artikel ist etwa einen Absatz lang und verweist auf andere interessante Webseiten (Abbildung 1). Zu jedem Artikel gehört auch ein Diskussionsforum, in dem sich jeder über den Artikel äußern kann. Ferner können Besucher von Slashdot eigene Artikel einsenden. Malda und seine Kollegen entscheiden dann, ob ein Artikel veröffentlicht wird. Seit August 2001 kann jeder auch ein eigenes persönliches Journal führen. Damit ist die Webseite zu einem Weblog-Hosting-Service geworden. [18]

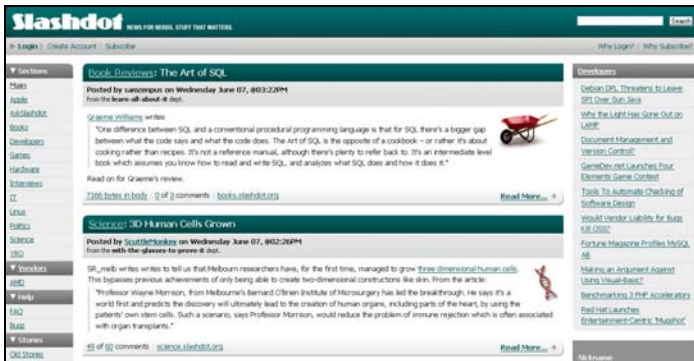


Abbildung 1: Das Slashdot-Blog bietet in kurzen Artikeln Neuigkeiten für Computerinteressierte.

Im April 1999 beschloss Peter Merholz den Begriff „Weblog“ als „Wee-Blog“ auszusprechen oder es in der Kurzform „Blog“ zu nennen. Er vermerkte diesen Gedankengang auf seiner Homepage und der Begriff hat sich seitdem durchgesetzt. [16]

Doch es gibt einen Streitpunkt darüber, wie genau Blogs auszusehen haben. Wie bereits oben beschrieben, waren es zunächst Linksammlungen, die mit Kommentaren versehen werden konnten. Doch im Laufe der Zeit veränderte sich die „Blogging-Welt“ (auch Blogosphäre genannt). Immer mehr Menschen nutzten ihr Blog als eine

¹ www.robotwisdom.com

² www.slashdot.org

Art Online-Tagebuch [18] [1]. Diese Entwicklung wurde vor allem dadurch vorange-
 trieben, dass im Jahr 1999 verstärkt Blog-Hoster eine einfache Möglichkeit boten,
 Texte im Internet zu veröffentlichen. „Von Hand“ gebastelte Seiten, auf dem eigenen
 Server oder gemietetem Webspace, waren somit nicht mehr nötig. LiveJournal³
 (Abbildung 2), Slashdot (Abbildung 1) und Blogger.com⁴ (Abbildung 3) sind die
 Pioniere dieser Szene [18] [1].



Abbildung 2: LiveJournal bietet jedem Internetnutzer die Möglichkeit sein eigenes Blog zu betreiben. Motto: „Teile dich vielen, wenigen oder einzelnen mit“.



Abbildung 3: Drei einfache Schritte zum Blogger.com-Blog: Anlegen, Namen vergeben, Vorlage auswählen.

Weltweit gibt es mittlerweile rund 200 Millionen Blogs (Stand: Februar 2006) [28]. Sieben Monate zuvor, im Juli 2005, waren es „nur“ 70 Millionen [29]. Daran sieht man, dass sich die Blogosphäre großer Beliebtheit erfreut. Ein Grund ist, dass in

³ www.livejournal.com

⁴ www.blogger.com

Blogs oft auf andere Blogs verlinkt wird [26] und die Blogger so von immer neuen Weblogs erfahren und auch angespornt sind ihrerseits eigene Blogs zu betreiben.

Die deutsche Blogosphäre umfasste Ende 2005 geschätzte 200.000 Weblogs [27]. Im Februar 2006 waren es schon rund 300.000 [26]. Auch hier ist ein starkes Wachstum zu verzeichnen. Von den gezählten Blogs ist jedoch eine nicht genau zu bestimmende Anzahl inaktiv [1]. Sie wurden angelegt, um einen bestimmten Zweck zu erfüllen (z.B. Dokumentation eines Urlaubs) oder um das Bloggen einmal auszuprobieren und dann nicht wieder benutzt.

Die Vielfalt der Blogs ist groß. Deshalb im Folgenden ein Überblick über die verschiedenen Arten. In dieser Arbeit sollen die Begriffe Weblog und Blog als Synonyme verstanden werden.

3 Weblog ist nicht gleich Weblog

Herring et al. [11] haben drei grundlegende Arten von Blogs identifiziert:

1. Persönliche Journale (meist in Form von Online-Tagebüchern),
2. "Filter"-Blogs, die Infos von anderen Webseiten sammeln und kommentieren (also sehr eng angelehnt an die ursprüngliche Form des Weblogs) und
3. Wissens-Logs, bei denen Forscher oder Projektgruppen ihre Untersuchungen verbreiten

Darüber hinaus gibt es noch Mischformen, die sich aus den oben genannten drei Blog-Arten ergeben.

Ferner haben Herring et al. [11] in ihrer Analyse festgestellt, dass Weblogs zwischen Standard-Webseiten und asynchroner computervermittelter Kommunikation (wie z.B. Diskussionsforen) liegen. Abbildung 4 zeigt, dass Standard-Webseiten im Normalfall eher selten aktualisiert werden, aber oft multimediale Inhalte enthalten. Dagegen wird in Foren meist sehr rege in Form von vielen reinen Texteinträgen diskutiert. Weblogs werden in der Regel zwar auch regelmäßig, aber nicht ganz so oft mit neuen Einträgen versehen, können dafür aber auch Multimediainhalte enthalten. Herausgegriffen wurden zwei spezielle Weblog-Arten: Online-Journale (LiveJournal) liegen dabei näher an Standard-Webseiten, während Community-Blogs (Slashdot) sich mehr an Foren orientieren. Die zuvor erwähnten drei grundlegenden Weblog-Arten kann man wie folgt einordnen: Persönliche Journale und Online-Tagebücher können in die Nähe von Online-Journalen und Filter-Blogs bei Community-Blogs eingeordnet werden. Wissens-Logs konnten Herring et al. nicht auf dieser Skala einordnen, dazu seien weitere Dimensionen notwendig, schreiben sie in ihrer Studie.

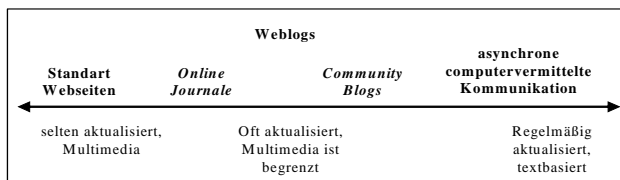


Abbildung 4. Weblogs auf einer Skala zwischen Standard-Webseiten und asynchroner computervermittelter Kommunikation [11]

Weiterhin kann man nach Nardi et al. [22] Weblogs in private (Passwort geschützte) und öffentliche Blogs unterteilen. Die Mehrheit der Blogger sieht den Schutz ihrer Privatsphäre jedoch als nicht so wichtig an, denn nur ein kleiner Teil der Blogs ist tatsächlich nur mit einem Passwort zugänglich. Wenn man sich aber die Gründe anschaut, die einen Internetnutzer zum Blogger machen, ist diese Tatsache eventuell weniger überraschend. Dies soll im Folgenden näher untersucht werden.

4 Die „Generation Blogging“ und ihre Weblogs

So, wie es nicht nur die eine Art von Blog gibt, sind auch die Gründe, der „Generation Blogging“⁵ beizutreten, vielfältig. Laut Nardi et al. [22] haben viele erst mit dem Bloggen angefangen, nachdem andere sie explizit dazu aufgefordert haben. Die meisten waren anfangs skeptisch, dann aber begeistert, als sie herausgefunden haben, dass andere ihre Blogs lesen und „sie eine Stimme haben, die gehört wird“. Daraus kann man schließen, dass die soziale Aktivität beim Bloggen eine große Rolle spielt.

Blogging-Hosts wie Livejournal und Blogger haben, wie zuvor bereits erwähnt, zum großen Erfolg der Weblogs beigetragen. Im Kapitel „Technik“ wird näher darauf eingegangen. Im Folgenden werden zunächst einige Gründe beschrieben, die einen Blogger zum Betreiben eines Weblogs motivieren. Die Selbstdarstellung ist dabei sicherlich für viele ein Hauptgrund.

4.1 Ich blogge, also bin ich: Blogging als soziale Aktivität & Selbstdarstellung

Blogs beschäftigen sich mit allen Bereichen des Lebens: Alltägliches (Tagebuch), Urlaub, Familien-Ereignisse, Dokumentation eines Krankheitsverlaufs, Kommentare zu politischen Ereignissen oder gerichtlichen Entscheidungen, Veröffentlichung von Gedichten... [26] [23].

Man kann sagen, dass das Verfassen eines Blogs (insbesondere, wenn es sich um persönliche Journale handelt) einen hohen Stellenwert im Leben des Bloggers einnimmt. Descartes philosophischen Schluss „Ich denke, also bin ich“, könnte man somit für die „Generation Blogging“ in „Ich blogge, also bin ich“ abwandeln. Die Blogger identifizieren sich über ihr Blog und verbreiten ihre persönlichen Gedanken darüber. Sie wollen andere an ihrem Leben teilhaben lassen - auch wenn in den meisten Fällen das Publikum eher der Blogger selber oder vielleicht ein paar gelegentliche Leser sind. Es ist somit eine Art persönliche Aufzeichnung, die das eigene Leben dokumentiert [26] [23].

Als einen weiteren Beweggrund identifizierten Schiano et al. [26] und Nardi et al. [22] [23] den Wunsch, in Kontakt mit Freunden und Verwandten zu bleiben, falls man räumlich getrennt ist (Leben und arbeiten in einer anderen Stadt oder eine Urlaubsreise). Diese Art des Blogs ist dann eine Art Newsletter, über den die Angehörigen über Neuigkeiten aus dem Leben des Bloggers erfahren können – und das ganze sozusagen in „Echtzeit“. Daraus haben sich dann auch Blogs in Form von Fotoalben (Photo-

⁵ Markus Chr. Koch und Astrid Haarland haben ein Buch mit dem Titel „Generation Blogger“ geschrieben (Mitp-Verlag, 2003), das neuen Bloggern den Einstieg erleichtern soll.

Blogs = Phlogs) gebildet. Die große Verbreitung von Digitalkameras hat ebenfalls dazu beigetragen, dass sich diese Art des Weblogs großer Beliebtheit erfreut. Einige nutzen ihr Blog auch als eine Art Fortschrittbericht. Beispielsweise, um die Krankheits- bzw. Genesungsgeschichte eines Angehörigen zu dokumentieren.

Dies alles könnte natürlich auch per E-Mail geschehen, aber Bloggen ist einfacher, da man sich nicht an bestimmte Adressen erinnern muss und auch keinen ungefragt belästigt. Blogs kann man also als eine Art „Pull-Service“ bezeichnen, bei der der Adressat aktiv werden muss, während E-Mail eher als eine Art „Push-Service“ angesehen werden kann, bei dem der Empfänger Informationen erhält, ob er nun will oder nicht. [26] [23]

Doch der Wunsch nach Selbstdarstellung ist nicht der einzige Grund für Blogger ein Weblog zu betreiben und es regelmäßig zu aktualisieren. Auch der Wunsch mit anderen in Kontakt zu treten und eigene Ideen auszutesten spielt für die Community-Blogger eine Rolle.

4.2 Community Blogging: Die neue Art der anonymen Kritik

Community-Blogger nutzen ihr Blog, um Dinge und Ereignisse zu kommentieren, die sie beschäftigen. Das können technische Ratschläge sein, aber auch kulturelle und gesellschaftliche Kritik spielt in vielen Blogs eine Rolle. So haben sich viele Blogger mit dem Irak-Krieg beschäftigt. Andere prangern die Todesstrafe in den USA an. Das Weblog ist also auch zu einer Forum der Demonstrationsbewegung geworden. [23]

The image shows a screenshot of a blog page. The main content is a poem titled "Gedichtwelt" by M. Matze. The poem is about hope and the shadow of fear. Below the poem, there are two comments from users. The first comment is from "Teddykneger" and the second is from "M. Matze" (the author). The page layout includes a header, a sidebar with "Über mich" and "Freunde", and a main content area with the poem and comments.

Abbildung 5: Ein Beispiel für ein Gedichts-Blog (links das veröffentlichte Gedicht und rechts die zugehörige Diskussion)⁶

Nardi et al. [22] [23] haben auch herausgefunden, dass für viele das Bloggen „thinking by writing“ ist, sie also ihre Ideen in ihrem Blog vor einem kleinen Publikum testen können. Einige hoffen, dass ihre Beiträge in der Zukunft vielleicht einmal in einem Magazin oder einer Zeitung veröffentlicht werden. Bis es soweit ist, üben sie in ihrem Weblog, den direkten „Draht“ zu ihrem Publikum haben sie ebenfalls: Die Kommentar-Funktion hilft ihnen, ihren Schreibstil zu verbessern oder Ideen für neue

⁶ http://gedichtwelt.blog.de/2006/03/25/neu_und_neu~674591#comments

Artikel zu finden. Aber nicht nur Diskussionen um das aktuelle politische Geschehen haben ihren Platz in der Blogging-Welt gefunden, auch Hobby-Dichter schätzen das Weblog, um ihre Gedichte zu veröffentlichen. Auch hier feilen die Autoren oft an ihren Texten in Zusammenarbeit mit ihren Lesern (Abbildung 5). Das frühe Veröffentlichliche hilft ihnen also beim Schreiben.

Doch diese Form des Bloggings kann man auch kritisch sehen, da sich normalerweise die Anzahl der Kommentare in Grenzen hält [11]. Stellt sich die Frage, welchen Stellenwert diese Funktion dann noch einnimmt. Blogger können sich laut Nardi et al. [22] hinter ihren Blogs verstecken. Viele Community-Blogger schätzen die anonyme Kritik und halten ihre Kritiker sozusagen „eine Armlänge auf Abstand“. Sich mit den Kommentaren zu einem Artikel auseinanderzusetzen ist weniger persönlich und intensiv, als wenn sie ihre Gedanken oder Gedichte direkt mit jemandem diskutieren würden [23] [22]. Ein weiterer Vorteil ist laut Nardi et al. [22], dass das Feedback nicht sofort erfolgt, sondern dass die Autoren erst in Ruhe ihre Artikel zu Ende schreiben können. Die Diskussion findet mit einem größeren zeitlichen Abstand statt und oft auch mittels anderer Medien wie beispielsweise E-Mail, Instant Messaging oder Telefon [6].

Doch welchen Stellenwert hat diese Art der Informationsverbreitung? Kann sich nun jeder Journalist oder Dichter nennen, da er oder sie bereits Artikel oder Gedichte veröffentlicht hat? Dieser Frage soll im Folgenden nachgegangen werden.

4.3 Nachrichten-Blogs: Amateur- oder ernstzunehmender Journalismus?

Es stellt sich die Frage, was überhaupt einen Journalisten ausmacht. Laut dem Deutschen Journalisten-Verband (djuv)⁷ ist Journalist/in, wer

- hauptberuflich
- produktiv oder dispositiv
- Informationen sammelt, auswertet und/oder prüft und Nachrichten unterhaltend, analysierend und/oder kommentierend aufbereitet,
- sie in Wort, Bild und/oder Ton
- über ein Medium
- an die Öffentlichkeit vermittelt
- oder den publizistischen Medien zu dieser Übermittlung bereitstellt.

Hauptberuflich betreibt wohl kaum ein Blogger sein Weblog, weil die Verdienstmöglichkeiten eher gering sind. Aber alle anderen Punkte kann wohl jeder Blogger, der es möchte, auch abdecken. Viele Blogger sammeln Informationen, werten diese aus, prüfen sie und bereiten damit Nachrichten auf, die unterhaltend, analysierend oder kommentierend sein können. Das Medium, über das sie dies an die Öffentlichkeit vermitteln ist das Weblog. Mitchell [17], Regan [25] und Lasica [13] formulieren es so, dass mit Hilfe der immer einfacher gewordenen Blogging-Technik jeder Leser auch zum Redakteur (Produzent, usw.) werden kann. Die einfache Formel für Journalismus: Finde die richtigen Leute, stelle die richtigen Fragen und schreibe es auf [13].

Es stellt sich die Frage, ob Blogger damit Amateur-Journalisten sind, die dem Berufsbild schaden, da sie journalistische Prinzipien unter Umständen nicht beachten,

⁷ <http://www.djuv.de/journalist/berufsbild/ueberblick.shtml>

oder eine wertvolle weitere Informationsquelle im Internet darstellen. In den Augen vieler Journalisten sind Blogs von Amateuren gemacht, schlecht geschrieben, von egozentrischer Natur und enthalten eine zu große Meinungsäußerung [25]. Auf der anderen Seite sehen Blogger die Arbeit von Journalisten oft als stagnierend und zu elitär an. Außerdem bemängeln sie, dass Journalisten zu viel Zeit darauf verwenden für ihre eigene Berufsgruppe zu schreiben, anstatt für die große Öffentlichkeit [25].

Mitchell [17] schreibt, dass Weblogs eine Gefahr hinsichtlich der journalistischen Qualität darstellen und in den Bereich der reinen Meinungsäußerung abgleiten können, wenn sie von unerfahrenen Journalisten betrieben werden, die Informationen ungeprüft veröffentlichen. In einer Redaktion durchläuft ein Artikel meist mehrere Stationen und wird von mindestens einer weiteren Person redigiert. Das ermöglicht einerseits eine gewisse Qualitätssicherung [17], andererseits gehen auch viele Ideen dabei verloren. Ein Blogger dagegen kann frei von solchen Restriktionen Nachrichten veröffentlichen. Als Editoren wirken dabei die Internetgemeinde, die Kommentare zum Artikel abgeben kann [13].

Ein Problem des traditionellen Journalismus kann darin bestehen, dass die Anzahl der Nachrichtenquellen bei einem lokalen Ereignis nicht ausreicht, wenn es in einem Ort nur eine begrenzte Anzahl von verfügbaren Medien gibt [25]. Blogs stellen dagegen oft eine größere Meinungsvielfalt zur Verfügung.

Nach Lasica [13] kann somit der traditionelle Journalismus durchaus vom Blogging profitieren. Es sei

- eine neue Form des Journalismus,
- die Einflussnahme auf Nachrichten-Auswahl bietet,
- das Vertrauen des Lesers erhöht,
- die Nachrichten personalisiert und
- Interaktivität ermöglicht.

In der Tat gibt es hauptberufliche Journalisten, die die Möglichkeiten des Bloggens zu schätzen wissen. Sie veröffentlichen auf ihrem Blog Hintergrundmaterial oder führen Nachrichten weiter, die in den Medien, für die sie hauptberuflich schreiben, keinen Platz finden. Laut Mitchell [17] ist Dan Gillmore⁸ (Kolumnist der San Jose Mercury News) der Pionier der bloggenden Journalisten. Ein weiteres sehr bekanntes politisches US-Weblog ist das des Journalisten Andrew Sullivan⁹.

⁸ <http://bayosphere.com/blog/dangillmor/>

⁹ http://time.blogs.com/daily_dish/



Abbildung 6: Das BILDblog kontrolliert die BILD-Zeitung und informiert hier die Leser über einen inkorrekten WM-Plan.

Im deutschsprachigen Raum kommentiert das BILDblog¹⁰ Artikel der Bild-Zeitung (Abbildung 6). Die Blogger haben es sich zur Aufgabe gemacht, die Macher der großen Boulevard-Zeitung zu kontrollieren und Fehler oder Ungenauigkeiten klarzustellen und zu korrigieren. Journalistisch aktive Blogger können somit auch die Funktion eines „Medienwächters“ einnehmen.

Neben der kreativen Freiheit, Schnelligkeit, Interaktivität kann auch das nicht Vorhandensein von Marketing-Einschränkungen ebenfalls ein großer Vorteil von Blogs [14] darstellen. Die Blogger müssen – im Gegensatz zu den meisten traditionellen Medien – nicht auf Anzeigenkunden oder Leser Rücksicht nehmen. Denn Blogs werden nicht speziell für bestimmte Personengruppen geschrieben, sondern von ihnen gefunden [14].

Doch es gibt auch kritische Anmerkungen. Graswurzel-Journalismus (auch partizipativer Journalismus genannt) bezeichnet die journalistische Aktivität eines Bürgers oder einer Gruppe von Bürgern¹¹. In der Kommunikationswissenschaft ist in dem Zusammenhang auch der Begriff Agenda setting¹² bekannt. Dies besagt, dass Medien zwar nicht bestimmen was wir denken, aber worüber wir nachdenken. Und wenn man außerdem noch davon ausgeht, dass nur ein gewisser Anteil der Bevölkerung politisch und sozial aktiv, gebildet und interessiert ist und auch nur diese Menschen die neue Form der Meinungsäußerung und –bildung über das Internet wahrnehmen, kommt man zum Begriff des „Opinionleaders“. Die Wissenskluft zwischen gebildeten und ungebildeten Menschen wird somit immer größer. [12]

4.4 Lern-Blogs: Vom Lern-Tagebuch zum Wissens-Log

In Learning Communities schließen sich lernwillige Menschen zusammen, um sich gegenseitig beim Lernen zu unterstützen und ihre Fortschritte in einer Art „Lern-Tagebuch“ zu dokumentieren. Wagner [31] und Downs [4] haben diese Form des

¹⁰ <http://www.bildblog.de/>

¹¹ <http://de.wikipedia.org/wiki/Graswurzel-Journalismus>

¹² http://de.wikipedia.org/wiki/Agenda_Setting

Bloggens näher untersucht. Bei der Verwendung von Lern-Logs müssen die Studenten ihre Ergebnisse regelmäßig an den Lehrer senden, der diese dann auf der Kurs-Webseite veröffentlicht [31]. Dies dauert meistens sehr lange und ist für den Lehrer aufwendig. Bei der Verwendung von Blogs kann jeder Kursteilnehmer mit einfachen Mitteln seine Ergebnisse selber veröffentlichen. Der Aufbau von Wissens-Logs (Knowledge-Logs) wird somit einfacher [31]. Die Kursteilnehmer können schneller auf die Ergebnisse der Kommilitonen zugreifen und diese untereinander diskutieren und kommentieren. Die Dokumentation des Lernprozesses übernimmt dabei die Blogging-Software [31].

Downs [4] sieht sogar noch vielfältigere Möglichkeiten. Er zeigt die Möglichkeit auf, die Blogs in drei Kategorien aufzuteilen: Ein Blog für den gesamten Kurs (für Mitteilungen, die den gesamten Kurs betreffen), ein persönliches, aber öffentliches (um den Studenten die Möglichkeit zur Diskussion untereinander zu bieten) und ein privates (zum Notieren von Ideen und um Hilfestellung vom Lehrer zu erhalten). Viele Studenten schätzen die Möglichkeit der direkten Kommunikation untereinander, da sie so motiviert seien, mehr als sonst zu schreiben [4]. Aber auch die Lehrer und Organisatoren haben Weblogs als Arbeitserleichterung entdeckt: Die Organisation sei so viel einfacher [4].

4.5 Corporate Blogging: Wettbewerbsfähiger durchs Bloggen

Auch Unternehmen haben Weblogs als Kommunikationsmittel entdeckt. Wacka [30] nennt die folgenden betriebswirtschaftlichen Schlagwörter als Gründe ein Corporate Blog zu betreiben:

- Business Leader: Das eigene Unternehmen als Experten positionieren.
- Customer Relationships: In einem Forum, in dem es nicht ums Verkaufen geht, ist die persönliche Beziehung zu den Kunden oft besser. Kunden können Tipps austauschen oder Produkte diskutieren.
- Media Relations: Die Beziehungen zu den Medien intensivieren und pflegen. Mit den Medien in einen Dialog treten, anstatt nur Pressemitteilungen zu versenden.
- Internal Collaboration: Blogs zum Gedankenaustausch zwischen den Mitgliedern von Projektteams nutzen.
- Knowledge Management: Informationen können einfach gefunden werden und Bloggen ist außerdem eine Art "Learning by doing".
- Recruitment: Zum attraktiven Arbeitgeber werden, da durch die transparente Firmenpolitik potenzielle Arbeitnehmer auf das Unternehmen aufmerksam werden.
- Testen von Ideen und Produkten: Blogs sind informell. Kommentare von Benutzern helfen Ideen oder Produkte zu testen.
- Gutes Ranking in Suchmaschinen: Links zu anderen Seiten, und dass viele Seiten auf die eigene Seite verlinken, ist für viele Suchmaschinen fürs Ranking wichtig.

Beispielsweise haben Hapag-Lloyd Express, SAP und Microsoft das Corporate Blogging für sich entdeckt. Hapag-Lloyd Express¹³ bietet seinen Kunden die Möglichkeit ihre Urlaubserlebnisse zu bloggen und auch Mitarbeiter erzählen über ihre Ferienreisen. SAP¹⁴ und Microsoft¹⁵ informieren über neue Ideen und Produkte und bieten die Möglichkeit sich über „Best Practices“, also gute Ideen aus der Praxis, zu informieren.

4.6 Moblogs: Weblogs werden mobil

Das Bloggen von Fotos ist bei vielen Bloggern sehr beliebt. Wie bereits oben beschrieben, werden persönliche Erlebnisse dokumentiert (z.B. Urlaubsreisen), so dass Freunde, aber auch Fremde am eigenen Leben teilhaben können [32]. Bekannte Foto-Blogs sind Flickr.com¹⁶ und Photoblogs.com¹⁷. Video-Blogs sind ebenfalls beliebt. FrameDrop¹⁸ und Vidblogs¹⁹ sind bekannte Vertreter dieser Weblog-Art.

Doch Fotos und Videos werden meist unterwegs gemacht. Neue Handys verfügen über Foto- und Video-Möglichkeiten, was den Wunsch nahe legt, das mobile Veröffentlichen von Videos und Fotos zu ermöglichen: Das Mobile Blogging (Moblogging), auch „Mail-to-Web“ genannt, war geboren [32].

Gross und Kleppe [10] haben FrameDrops entwickelt. Diese Architektur ermöglicht mobiles Bloggen – auf einfache Weise. Denn der Entstehungsort der Fotos und Videos ist oft wichtig, muss aber meist per Hand in die Metadaten eingetragen werden. FrameDrops ermöglicht das automatische Positionieren von Einträgen mit Hilfe von GPS. Dabei werden Videos in einen geografischen sowie chronologischen Kontext gebracht. Der mobile Blogger sendet seine Videos per E-Mail an den FrameDrops-Server. Dieser konvertiert die Videos ins Flash Video Format (FLV) und ein Vorschaubild (JPG). Die zum Video oder Bild gehörenden Metadaten wie beispielsweise Datum, Zeit, Abmessungen, Auflösung und Ortsinformationen enthalten wertvolle Informationen und werden deshalb extrahiert. Video und Foto werden dann mittels XML-RPC²⁰ an den Antville Webserver²¹ gesendet, der Macro-Code zurückschickt (z.B. Code, der Attribute und Werte enthält, das den Objekt-Identifikator und Layout beschreibt). [10]

Noch hat nicht jeder ein GPS-fähiges Mobiltelefon, um Fotos ohne Aufwand mit geografischem Kontext versehen zu können. Momentan müssen solche Informationen noch per Hand in die Metadaten eingefügt werden. Diese Möglichkeit bietet Mappr²². Fotos, die in Flickr geblotgt werden, können mit Mappr mit geografischem Kontext versehen werden. Abbildung 7 zeigt die Möglichkeit, sich anschließend anhand eines

¹³ <http://blog.hlx.com/>

¹⁴ <http://www.sap.com/community/pub/blogs.epx>

¹⁵ <http://www.microsoft.com/Communities/blogs/PortalHome.mspx>

¹⁶ <http://www.flickr.com/>

¹⁷ <http://www.photoblogs.org/>

¹⁸ <http://www.framedrops.net/>

¹⁹ <http://www.vidblogs.com/>

²⁰ <http://www.xmlrpc.com/>

²¹ <http://antville.helma.org/project/about/features/>

²² <http://www.mappr.com/>

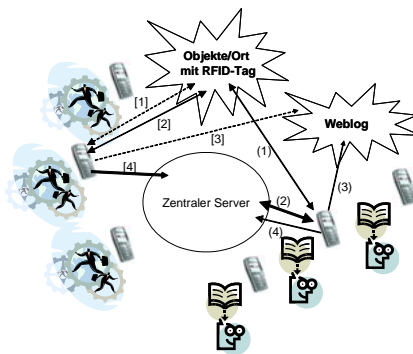
Schlagwortes (hier Route 66) die darunter gespeicherten Fotos anzeigen zu lassen. Die Fotos erscheinen auf einer Landkarte, an den Orten, an denen sie gemacht wurden. Dies gibt einen Ausblick auf das Potenzial des mobilen Bloggens.



Abbildung 7: Fotos, die auf Mappr.com unter dem Schlagwort „Route 66“ gespeichert sind, zeigen den Verlauf der Route 66.

Eine weitere Möglichkeit ist die Ortserfassung mittels RFID. Cheng et al. [3] beschreiben ein Beispielszenario wie folgt (siehe Abbildung 8): A hat gerade in einem Restaurant gegessen und möchte seine Erfahrungen und Einschätzung mit anderen Teilen. Mit seinem Handy bewertet er das Essen und die Atmosphäre im Restaurant. Anschließend schwenkt er das Handy über einen RFID-Tag auf den Tisch. Seine Kritik wird dann auf seiner Weblog-Seite veröffentlicht. Der zentrale Server führt einen Index, auf dem die Tag-ID dem Weblog-Eintrag zugeordnet wird. B steht nun vor dem Restaurant und möchte wissen, wie das Essen dort ist. Er schwenkt sein Handy über den RFID-Tag, der auf der Speisekarte vor dem Restaurant angebracht ist und bekommt eine Liste von Einträgen angezeigt, die zu dieser Tag-ID gehören. B wählt den Eintrag von A aus, durch Kommentare können A und B nun miteinander kommunizieren.

- [1] Erfahrungen über Objekt (mit RFID-Tag ausgestattet) erfassen.
- [2] RFID-Tag erfassen.
- [3] Eintrag auf Weblog-Seite veröffentlichen.
- [4] Aktualisierung des Index der RFID-Tag-ID und des Weblog-Eintrags.



- (1) RFID-Tag erfassen.
- (2) Kommunikation mit dem Server herstellen, um Liste der relevanten Einträge zu bekommen.
- (3) Direkte Kommunikation mit Web-Präsenz (Herunterladen des Weblog-Eintrags).
- (4) Hochladen der Bewertung/ des Kommentars zum Eintrag.

Abbildung 8: System-Architektur für Blogging mit RFID. [3]

5 Die Technik: Was man als Blogger wissen sollte

Nach Downs [4] kann man die Blogging-Software in zwei grundlegende Varianten unterteilen: Hosting-Services und installierte Applikationen. Es gibt zwar auch einige wenige Blogger (meist der ersten Stunde), die ihre Weblogs noch von Hand programmieren, aber solche Blogger, wie beispielsweise Rebecca Blood²³, werden immer weniger [5], da es immer schwieriger wird, bei den technischen Neuerungen, wie z.B. Trackback (siehe unten), mitzuhalten [1].

Hosting Services ist eine Webseite, die einem alles zur Verfügung stellt, was man zum Erstellen eines Blogs benötigt[4]. Es bietet ein Formular zum Eintragen des Inhalts, einige Werkzeuge zum Erstellen eines Templates und zugriff auf diverses Zubehör [4]. Das Blog wird gehostet auf dem Zentralcomputer des Hosting-Service (die URL spiegelt dabei den Namen wieder) [4]. Diese Blog-Art ist ähnlich zu Services, die das Hosten von eigenen Webseiten anbieten (z.B. Yahoo!GeoCities²⁴) oder eigenen Diskussionsgruppen (z.B. Yahoo!Groups²⁵) [4]. In diese Kategorie fallen beispielsweise die bereits oben erwähnten Seiten Blogger.com und LiveJournal [4]. Diese web-basierten Blogs sind in diversen Programmier- und Skriptsprachen, wie beispielsweise Java, Perl, PHP und ASP, geschrieben [15].

Installierte Applikationen stellen die andere Blog-Variante dar [4]. Dabei wird ein Software-Stück, das vom Provider bereitgestellt wird, auf dem eigenen Webespace installiert [4]. Diese Kategorie ist ähnlich wie web-basierte Applikationen wie beispielsweise ColdFusion²⁶ und Hypermail²⁷ [4]. Die Anzahl derer, die diese Blog-Art verwenden ist sehr viel geringer, aber ihr Wissenstand über Blogs und deren technische Realisierung ist dafür umso größer [4]. MoveableType²⁸ von Six Apart's ist der bekannteste Vertreter bei den installierten Applikationen [4]. Es bietet dem Blogger viele Optionen, die das Anpassen an individuelle Bedürfnisse ermöglichen [4]. Weitere Vertreter dieser Kategorie sind Radio Userland²⁹, WordPress³⁰ und Drupal³¹ [5].

5.1 Das kleine Blogging-ABC: Permalinks, Trackbacks & Co.

Blogger haben eine „eigene Sprache“. Da ist von Permalinks, Trackbacks, Blogroll und Pingback die Rede. Doch was bedeutet dies? Im Folgenden nun deshalb ein kleines Blogging-Wörterbuch.

Permalinks: Dadurch hat jeder Blog-Eintrag eine permanente Ortsangabe (URL), auf die referenziert werden kann. Vorher konnte man durch Weblog-Archive nur durchs Browsing (Durchstöbern) navigieren. Jetzt können Blogger auf einzelne Weblog-Einträge verweisen, indem sie diese referenzieren. [1] [5]

²³ <http://www.rebeccablood.net>

²⁴ <http://geocities.yahoo.com>

²⁵ <http://groups.yahoo.com/>

²⁶ <http://www.adobe.com/de/products/coldfusion/productinfo/overview/>

²⁷ <http://www.hypermail.org/>

²⁸ <http://www.moveabletypes.org>

²⁹ <http://www.userland.com>

³⁰ <http://wordpress.org>

³¹ <http://drupal.org>

Crossblog-Talk: Diese Kommentarfunktion ermöglicht das weblog-übergreifende Diskutieren von Einträgen. Ein Blogger kann damit einen Eintrag in einem anderen Blog vornehmen. Dies ist die Kommentarfunktion, ohne die Blogs heute undenkbar sind. [1]

Trackback: Die automatisierte Version von Crossblog-Talk. Eingeführt wurde es 2001 von MovableType. Über entgegengesetzte Hyperlinks kann das referenzierte Weblog verfolgt werden. Unter jedem Blog-Eintrag befindet sich eine Spezial-URL (vom Blogger angegeben oder automatisch im Hintergrund erstellte Trackback-URLs), die von einem anderen Blog zur Verlinkung verwendet werden kann („Ich habe dein Blog hier mit diesem Text zitiert“). Zuvor unsichtbare inhaltliche Verbindungen werden so sichtbar. Das zitierte Blog aktualisiert sich nach dem Aufruf der Trackback-URL selbst und fügt den entsprechenden Auszug aus dem zitierenden Blog seinem Trackback-Bereich hinzu. [1] [5] [18]

Pingback: Ähnlich wie Trackback. Sobald ein Blog-Eintrag veröffentlicht wird, werden alle darin enthaltenen Links darüber informiert („angepingt“). [5]

Blogroll: Einige Weblog-Services bieten die Möglichkeit oft verwendete Links auf der Seitenleiste (sidebar) automatisch anzuzeigen. Es zeigt somit die Quellen, mit denen der Blogger arbeitet und stellt sie den Lesern ebenfalls zur Verfügung. [5]

Syndication: Damit wird die Integration von Inhalten eines Blogs in ein anderes bezeichnet. In der Regel beschränkt es sich auf Schlagzeilen. [18]

Aggregation: Die Zusammenführung von Inhalten zu einem Thema. Möller [18] schreibt dazu: „In der Praxis gibt es zahlreiche so genannte Feed-Reader, die Schlagzeilen von Blogs und News-Websites herunterladen und diese zusammenführen, um dem Leser das mühsame Klicken von Seite zu Seite auf der Suche nach Neuigkeiten zu ersparen. Die letzten Schlagzeilen der Websites werden als XML-Dateien angeboten. Für die enthaltenen Beiträge werden jeweils URL, Kurzbeschreibung, Publikationsdatum und Kommentar-Link angeboten.“

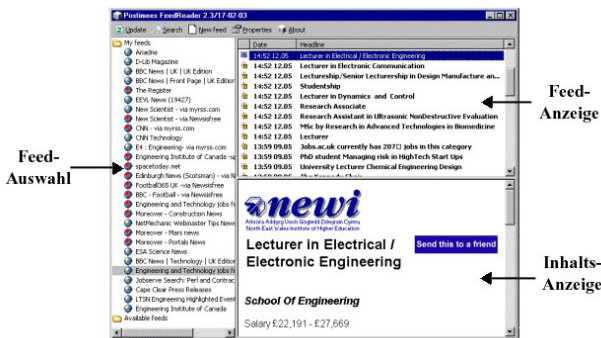


Abbildung 9: FeedReader ermöglicht eine Auswahl diverser Feeds (links), die Anzeige eines gewählten Feeds (rechts oben) und den Inhalt eines ausgewählten Eintrags (recht unten) [21].

*RDF*³²/*RSS*³³: RSS steht für Really Simple Syndication, Rich Site Summary oder RDF Site Summary - je nachdem wen man fragt oder von welcher Version man spricht [21]. RSS ist RDF (Resource Description Framework)-basiert und somit auch XML-basiert [24]. RDF bietet die Möglichkeit Inhalte mit Metadaten zu versehen, um zu verstehen, worum es sich inhaltlich handelt [24] (dies ist auch Grundlage für das sog. Semantic Blogging, siehe Kapitel 5.3). Die zuvor erklärten Begriffe Syndication und Aggregation werden oft im Zusammenhang mit RSS verwendet. Um diese Arten der Inhaltsverwendung zu ermöglichen, gibt es diverse *RSS-Reader*, die entweder web-basiert (z.B. NewsIsFree³⁴) oder als eigenständige Programme (z.B: FeedReader³⁵, siehe Abbildung 9) arbeiten [21]. So können Blogger ihre Inhalte automatisch der Blogging-Gemeinschaft zur Verfügung stellen [15]. Ihre Leser müssen lediglich den sogenannten *RSS-Feed* abonnieren. Abbildung 10 [21] stellt die Möglichkeiten von RSS grafisch am Beispiel der BBC dar. Es werden so Inhalte für andere Webseiten, Desktop-Reader, Suchmaschinen und Online-Zusammenfassungen angeboten. Die BBC profitiert davon, dass man, um die Originalnachricht lesen zu können, in allen Fällen ihre Webseite besuchen muss und darauf entsprechend verlinkt wird.

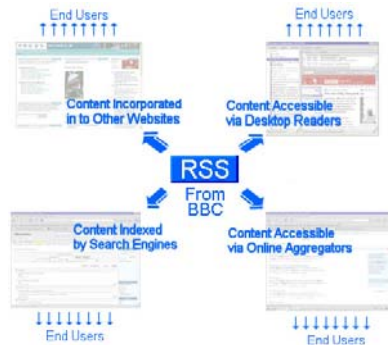


Abbildung 10: RSS-Feed der BBC stellt Inhalt für verschiedene Anwendungsbe-
reiche (andere Webseiten, Desktop-Reader, Suchmaschinen, Online-
zusammenfassungen) zur Verfügung. [21]

Eine RSS-Datei besteht aus einem Channel und mehreren Einträgen (Items). Der Aufbau könnte wie folgt aussehen [21]:

```
<channel>
<title>BBC News | UK | UK Edition</title>
<link>http://news.bbc.co.uk/hi/uk/default.stm</link>
<description>BBC UK News updated every minute of every day</description>
</channel>
```

³² <http://www.w3.org/RDF/>

³³ <http://www.rss-specifications.com/rss-specifications.htm>

³⁴ <http://www.newsisfree.com>

³⁵ <http://www.feedreader.com/>

```

<item>
<title>Clare Short quits post over Iraq</title>
<description>Clare Short quits the cabinet, accusing Tony Blair of breaking his
promises over the UN's role in rebuilding Iraq.</description>
<link>http://news.bbc.co.uk/hi/uk_politics/3019871.stm</link>
</item>

```

Zuerst wird der Channel mittels Titel, Link und Kurztext beschrieben und anschließend die dazugehörigen Inhalte (Items). Diese Items besitzen wiederum eine Beschreibung, die aus einem Titel, dem Kurztext und dem Link besteht.

Durch diese vielfältigen Verlinkungsmöglichkeiten (Permalinks, Trackback, Pingback, Blogroll, RSS), sind die Blogs untereinander nicht nur logisch, sondern auch physikalisch verbunden [6]. Dies ermöglicht ein großes Publikum zu erreichen.

5.2 Blog-Promotion: Wie Weblogs populär werden

RSS ermöglicht es, Blog-Inhalte effizient mit Software auszuwerten. Laut Möller [18] machen Seiten wie Daypop.com, Blogdex.net, Blogsnow.com und Technorati.com sich dies zunutze, um beispielsweise die populärsten von Bloggern verlinkten News-Stories und Blog-Beiträge aufzulisten. Eine Liste der Blogger mit den meisten eingehenden Links (Trackbacks, Pingbacks) ist ebenfalls möglich. Da nun die Suchmaschine Google beispielsweise Seiten mit vielen eingehenden Links für wertvoller hält als solche ohne, haben die Blogger einen nicht unerheblichen Einfluss auf die Platzierung von Suchergebnissen. Blogger nutzen dies gelegentlich gezielt für so genannte „Google-Bomben“. [18]

Wer beispielsweise nach „Miserable Failure“ („jämmerlicher Versager“) sucht, landet je nachdem welche Seite mehr Erfolg hatte bei George Bush oder Michael Moore. Dies ist zwar nur eine Spielerei, aber die Auswirkungen auf aktuelle Suchergebnisse ist schwer zu quantifizieren. Auch spezielle Blog-Suchmaschinen wie beispielsweise Feedster³⁶ oder Technorati³⁷ tragen zur Verbreitung von Weblogs bei. [18] Blogs können ja nach Popularität in Kategorien eingeteilt werden. Sog. A-Blogs sind sehr beliebt und sozusagen die Meinungsmacher der Blogging-Welt. [5] Der bereits oben eingeführt Begriff des „Agenda setting“ kommt hier also wieder zum Tragen.

5.3 Semantic Weblogs: Komfortable Suche und individuelle Kategorisierung

Die Vielzahl der Blogs macht es schwierig, sich einen Überblick zu verschaffen. Cayzer [2] zeigt auf, dass semantische Anfragen wie beispielsweise „finde alle Blog-Einträge über ein Paper, das Autor X geschrieben hat“ oder „finde alle Blog-Begriffe über meine Freunde“ noch nicht selbstverständlich sind. Auch die Strukturierung von beliebigen Blog-Inhalten nach den eigenen Wünschen (Kategorisierung) ist noch ein

³⁶ <http://www.feedster.com/>

³⁷ <http://www.technorati.com/>

Traum vieler Blogger. Als dritten Punkt führt Cayzer den Wunsch nach einer semantischen Navigation auf, also das effiziente Auffinden von Blogs (z.B. „stimmt zu“ oder „Teil von“). Cayzer [2] hat einen „Semantic Blogging Demonstrator“³⁸ gebaut, der dies realisiert. Um die notwendigen Metadaten zu erfassen benutzt er RDF³⁹. Das Hauptelement von RDF ist ein Subjekt-Prädikat-Objekt-Konstrukt. Damit kann eine graphenartige Struktur kreiert werden mit Subjekten (oder Objekten) als Links und Prädikaten als Eigenschaften. Damit ist es möglich, Blogs semantisch aufzuwerten. Durch Abfrage der RDF-Graphen (z.B. mit der RDF-Anfragesprache SPARQL⁴⁰) ist es möglich Anfragen wie oben gefordert zu stellen.

Noch schöner wäre es jedoch, wenn die Meta-Informationen nicht extra eingegeben werden müssten, sondern man bereits vorhandene Informationen nutzt. Möller et al. [19][20] nutzen dazu die Informationen, die bereits auf dem Desktop vorhandenen sind. Abbildung 11 zeigt die semiBlog-Architektur, die wie folgt funktioniert: Der Benutzer hat diverse Anwendungen auf seinem Computer, die darin enthaltenen Objekt-Informationen werden jeweils mit einem sog. Wrapper zu RDF Metadaten umgewandelt. Zusammen mit den text-basierten Einträgen des Benutzers wird diese Metainformation zu einer XML-Repräsentation kombiniert. Um den Blog-Eintrag zu generieren muss die XML-Datei noch umgewandelt werden (momentan die nicht-semantischen Informationen in HTML und die semantischen Informationen in RSS 1.0, das zum HTML-Dokument verlinkt). Nun kann der Blog-Eintrag, angereichert mit den Desktop-Informationen, publiziert werden. Abbildung 12 zeigt einen solchen Blog-Eintrag: Informationen zur Person „Eyal“ sowie die URL zur semiBlog-Webseite wurden eingefügt.

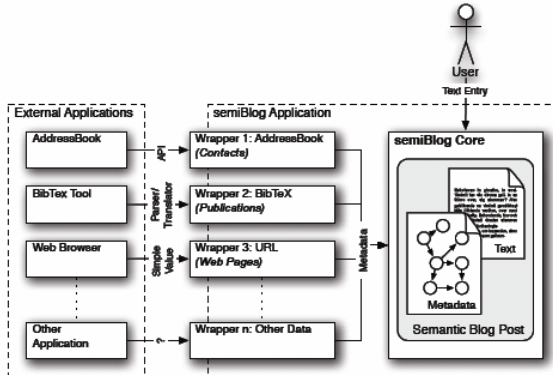


Abbildung 11: Mit der semiBlog-Architektur Inhalte aus Desktop-Metadaten erstellen. [20]

³⁸ <http://www.semanticblogging.org>

³⁹ <http://www.w3.org/RDF/>

⁴⁰ <http://www.w3.org/TR/rdf-sparql-query/>



Abbildung 12: Screenshot eines Blog-Eintrags mit angefügten Metadaten. [20]

6 Ausblick: Das ganze Leben ist ein Blog (bzw. im Blog)?

Mit Hilfe eines Weblogs kann man sich also auf einfache Weise der Internet-Welt mitteilen und je nach Belieben sein Leben vor der Internetgemeinschaft ausbreiten, Bürger-Journalismus betreiben, Diskussionen führen, sich über andere Blogger informieren, ... – die Möglichkeiten sind vielfältig, wie wir oben gesehen haben. Und für jeden Bereich gibt es eine spezielle Art von Weblog: normales Textblog, Fotoblog, Videoblog, Audioblog (Podcast). Die Blogger sind auch im technischen Bereich nicht auf sich gestellt. Web-basierte Blogs ermöglichen das einfache Einrichten und „Befüllen“ eines Blogs. Wer gerne eigene Funktionen einbindet und nicht vom Hosting-Service abhängig sein möchte, der kann eine eigenständige Blogging-Software auf seinem eigenen Webspace installieren.

Erst in letzter Zeit rücken auch die Weblog-Leser stärker in den Fokus der Forschung. Physikalische Verknüpfungen zwischen einzelnen, logisch zusammengehörigen Blog-Einträgen waren der erste Schritt (Trackback, Pingback, Blogroll, ...). Mit Semantic Blogs gibt es erste Ansätze, um die große Anzahl und Vielfalt der Weblogs noch besser in den Griff zu bekommen – denn das schönste und informativste Blog nützt wenig, wenn es niemand findet bzw. man innerhalb des Blogs nichts findet.

Die Blogging-Welt wird immer größer und auch zunehmend mobiler. Fast jeder hat heute ein Fotohandy, mit dem man schnell ein paar Fotos schießen oder sogar kurze Video-Filme aufnehmen kann. Wenn sich UMTS durchsetzt, ist es auch nicht auszuschließen, dass mehr Handy-Nutzer die Web-Funktionalität ihres Mobiltelefons nutzen. So könnten diese mobil gewonnenen Informationen noch besser in das persönliche Blog übertragen werden. Dann sind auch Ansätze wie von Gemmell et al. [8] [7] denkbar. Sie haben einen Prototypen entworfen, der das passive Aufzeichnen des eigenen Lebens ermöglicht. Über eine Kamera, die man um den Hals trägt, wird die Umgebung aufgezeichnet und die gewonnenen Daten mittels GPS mit Positionsdaten versehen. Gemmell et al. [8] [7] nutzen dann MyLifeBits, einen persönlichen, auf Lebenszeit ausgerichteten, digitalen Speicher, um die gewonnenen Daten zu sammeln. Diese Daten kann man nun auch durchsuchen – sofern man der Technik vertraut, ist es eine weitaus bessere Alternative als Fotos, Briefe u.a. in einem Schuhkarton aufzubewahren. Blogger könnten diese Möglichkeit ebenfalls nutzen und einen Teil der so gewonnenen Daten auf einfache Weise veröffentlichen. Die Möglichkeiten des Bloggens sind in jedem Fall noch nicht ausgeschöpft.

Literatur

1. Blood, Rebecca: "How blogging software reshapes the online community". In: *Communication of the ACM* 47 (12), S. 53-55, 2004.
2. Cayzer, Steve: "Semantic blogging and decentralized knowledge management". In: *Communications of the ACM* 47 (12), S. 47-52, 2004.
3. Cheng, Y., Yu, W., and Chou, T.: "Life is sharable: blogging life experience with RFID embedded mobile phones". In: *Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Amp, Salzburg, Austria, 2005 & MobileHCI*, vol. 111. ACM Press, New York, NY, S. 295-298, 2005.
4. Downes, Stephen: "Educational Blogging". In: *EDUCAUSE Review*, 39 (5), S. 14-26, 2004. <http://www.educause.edu/ir/library/pdf/erm0450.pdf>
5. Du, H. and Wagner, C.: "Weblog Success: Exploring the Role of Technology", 2006-02-02. http://wagnernet.com/tiki/tiki-download_file.php?fileId=27 bzw. <http://www.pacis-net.org/file/2005/282.pdf>. Abgerufen am 2006-06-15.
6. Efimova, L., de Moor, A.: "Beyond personal webpublishing: An exploratory study of conversational blogging practices". *Proc. HICSS'05*, 2005.
7. Gemmell, J., Williams, L., Wood, K., Lueder, R. and Bell, G.: "Passive Capture and Ensuing Issues for a Personal Lifetime Store." In: *Proceedings of CARPE '04*, S. 48-55, 2004. <http://research.microsoft.com/~jgemmell/pubs/CARPE2004.pdf>
8. Gemmell, Jim, Bell, Gordon, Lueder, Roger, Drucker, Steven and Wong, Curtis: "MyLifeBits: Fulfilling the Memex Vision". In: *ACM Multimedia*, S. 235--238, 2002. <http://research.microsoft.com/~jgemmell/pubs/MyLifeBitsMM02.pdf>
9. Gill, K.: "How can we measure the influence of the blogosphere?" Presented at the Workshop on the Weblogging Ecosystem at the 13th International World Wide Web Conference, New York, 2004. http://faculty.washington.edu/kegill/pub/www2004_blogosphere_gill.pdf
10. Gross, T., Kleppe, M. "FrameDrops: A Mobile VideoBlog for Workgroups and Virtual Communities"
11. Herring, Susan C., Scheidt, Lois Ann, Bonus, Sabrina, Wright, Elijah: "Bridging the gap: A genre analysis of weblogs". In: *Proceedings 37th Annual HICSS Conference, Big Island, Hawaii, 2004*. <http://www.ics.uci.edu/~jpd/classes/ics234cw04/herring.pdf>
12. Kavanaugh, A., Zin, T.T., Carroll, J.M, Schmitz, J., Pérez-Quinones, M, Isenhour, P.: "When Opinion Leaders Blog: New forms of citizen interaction". In: *7th Annual International Conference on Digital Government Research*. San Diego, 2006. <http://java.cs.vt.edu/public/projects/digitalgov/papers/DG06.KavanaughEtAl.3.3.06.pdf>
13. Lasica, J.D., "Blogs and Journalism Need Each Other ". *Nieman Reports*, (2003). <http://www.nieman.harvard.edu/reports/03-3NRfall/70-74V57N3.pdf>
14. Lasica, J.D., „Blogging as a Form of Journalism“. *Online Journalism Review*, veröffentlicht am 2001-05-24, modifiziert am 2002-04-29, <http://www.ojr.org/ojr/workplace/1017958873.php>. Abgerufen am 2006-06-06.
15. Lindahl, Charlie and Elise Blount: "Weblogs: Simplifying Web Publishing". In: *Computer* 36, S. 114-116, 2003.
16. Mescellany, Peter: "Play with your words". Mai 2002. <http://www.peterme.com/archives/0000205.html>. Abgerufen am 2006-06-01.
17. Mitchell, Bill: "Weblogs, A Road Back to Basics". *Nieman Reports*, 2003. <http://www.nieman.harvard.edu/reports/03-3NRfall/65-68V57N3.pdf>
18. Möller, Erik: "Die heimliche Medienrevolution - Wie Weblogs, Wikis und freie Software die Welt verändern". 1. Auflage, Hannover, Heise Zeitschriftenverlag, S. 115-158, 2005.
19. Möller, K., Breslin, J. G., und Decker, S.: "semiBlog – Semantic Publishing of Desktop Data". In: *14th Conference on Information Systems Development (ISD2005)*, Sweden, 2005. <http://sw.deri.org/~knud/papers/semiBlogISD2005Word/semiBlogISD2005print.pdf>

20. Möller, K., Bojars, U. und Breslin, J. G.: "Using Semantics to Enhance the Blogging Experience". To appear in 3rd European Semantic Web Conference ([ESWC2006](#)), Budva, Montenegro, June 2006.
21. Moffat, M.: „RSS – A Primer for Publishers & Content Providers“. EEVL , Heriot-Watt University Library, Großbritannien, 2003-08-20. http://www.techxtra.ac.uk/rss_primer/ Abgerufen am 2006-06-15.
22. Nardi, B. A., Schiano, D. J. und Gumbrecht, M.: "Blogging as social activity, or, would you let 900 million people read your diary?" In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, Chicago, Illinois, USA, November 2004. CSCW '04. ACM Press, New York, S. 222-231. <http://home.comcast.net/~diane.schiano/CSCW04.Blog.pdf>
23. Nardi, B. A., Schiano, D. J., Gumbrecht, M., and Swartz, L.: "Why we blog". Communication of the ACM 47 (12), S. 41-46, 2004. <http://doi.acm.org/10.1145/1035134.1035163>
24. Petrovic, M., Liu, H., and Jacobsen, H.: "CMS-TopSS: efficient dissemination of RSS documents". In: Proceedings of the 31st international Conference on Very Large Data Bases, Norway, 2005. VLDB Endowment, S. 1279-1282. <http://www.vldb2005.org/program/paper/demo/p1279-petrovic.pdf>
25. Regan, Tom: "Weblogs Threaten and Inform Traditional Journalism". Nieman Reports, 2003. <http://www.nieman.harvard.edu/reports/03-3NRfall/68-70V57N3.pdf>
26. Schiano, D. J., Nardi, B. A., Gumbrecht, M., and Swartz, L.: "Blogging by the rest of us". In: CHI '04, Extended Abstracts on Human Factors in Computing Systems, Austria, 2004. CHI '04. ACM Press, New York, S. 1143-1146. <http://home.comcast.net/~diane.schiano/CHI04.Blog.pdf>
27. Schmidt, Jan, Wilbers, Martin: „Wie blogge ich? Die Weblog-Umfrage 2005“. Bericht der Forschungsstelle „Neue Kommunikationsmedien“, Working Paper Nr. 06-01, Universität Bamberg, 2006. <http://www.fonk-bamberg.de/pdf/fonkbericht0601.pdf>
28. The Blog Herald: „The Blog Herald Count February 2006: 200 million blogs in existence“, 2006. <http://www.blogherald.com/2006/02/02/the-blog-herald-blog-count-february-2006-200-million-blogs-in-existence/>. Abgerufen am 2006-06-01.
29. The Blog Herald: "Blog Count for July: 70 million blogs", 2005. <http://www.blogherald.com/2005/07/19/blog-count-for-july-70-million-blogs>. Abgerufen am 2006-06-01.
30. Wacka, F.: „Why blog for business?“ Malmö, Schweden. <http://www.corporateblogging.info/basics/why/default.asp>. Abgerufen am 2006-06-11.
31. Wagner, Christian: "Put another (B)Log on the Wire: Publishing Learning Logs as Weblogs". In: Journal of Information Systems Education, 14 (2), S. 131-132, 2003.
32. Wang, H.-C., Deng, Y.-S., Chiu, S., „Beyond Photoblogging – New Directions of Mobile Communication“. In: Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Salzburg, Austria. September 19 - 22, 2005. MobileHCI '05, vol. 111. ACM Press, New York, S. 341-342.

Web-based collaboration systems

Thomas L. Lederer

LFE Medieninformatik

Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
thomas.lederer@stud.ifi.lmu.de

Zusammenfassung Collaboration (im Sinne der Zusammenarbeit über räumliche und zeitliche Distanz) ist im modernen Kommunikationszeitalter unentbehrlich geworden. Diese Arbeit befasst sich mit der Thematik. Nach einer kurzen Betrachtung der Gründe für die Benutzung web-basierter Collaboration Systemen werden einige solcher Systeme vorgestellt, untersucht und verglichen. Der Vergleich beinhaltet Technik, Anwendung, Stärken und Schwächen der jeweiligen Systeme.

1 Einleitung

1.1 Überblick

Kapitel 1 gibt einen kurzen Einblick in die Thematik, warum Collaboration (und speziell web-basierte) heutzutage unverzichtbar sind.

Im folgenden Kapitel 2 werden vier Collaborationsoftware Systeme vorgestellt.

Diese werden dann in Kapitel 3 miteinander verglichen, und es werden ihre Besonderheiten erläutert.

Eine Zusammenfassung und ein Fazit findet sich in Kapitel 4.

1.2 Warum Collaboration?

In der modernen Zeit gewinnt gemeinsames und verteiltes Arbeiten immer mehr an Bedeutung. Für dieses Vorgehen sprechen viele Gründe: um Team-Synergien zu nutzen, entferntes Arbeiten (z.B. Home-Office oder Büros in verschiedenen Städten) aber auch um Zeit zu sparen, oder effektiv zu nutzen. Kurzum: mehr als eine Person soll im Zuge der Collaboration, eventuell zeitgleich, auf den selben Daten, Dokumenten usw. arbeiten können.

Die meisten großen Firmen sind nicht mehr nur an einem Ort angesiedelt, so dass sie Prozeduren und Werkzeuge benötigen, damit ihre

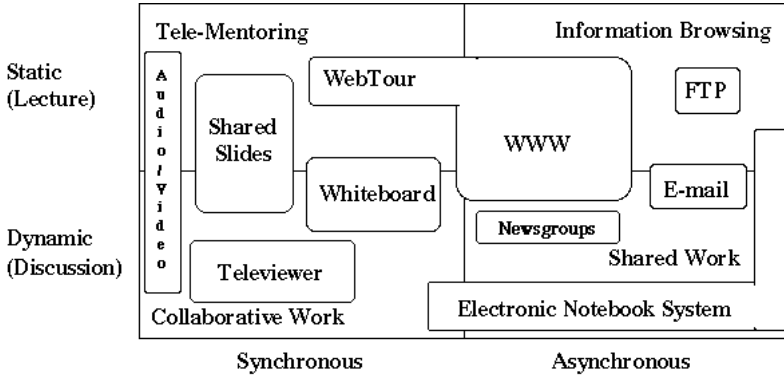


Abbildung 1. Kategorisierung von Tools für Zusammenarbeit. [1]

Mitarbeiter effizient zusammen arbeiten können. Dieses Arbeiten über räumliche (und meist auch zeitliche) Distanz nennt man Collaboration.

Diese Zusammenarbeit wird immer wichtiger, zum einen um den Mitarbeitern die Freiheit zu gewähren wo und wann sie ihr Arbeitspensum erledigen, zum anderen aber auch um die Fertigkeiten aller Mitarbeiter (die eventuell weltweit verstreut sind) im gleichen Projekt nutzen zu können. Die Standortfrage ist hier auch nicht völlig ohne Belang, da weltweit agierende Firmen mit Vorliebe dort so genannte Kompetenzzentren einrichten, wo es am wirtschaftlichsten für sie ist.

Während jeder Zusammenarbeit werden verschiedene Kommunikationskanäle benutzt, je nach dem welches Medium am zweckmäßigsten erscheint. Während einer Videokonferenz kann man die Mimik und eventuell die Gestik des Gegenüber sehen, was oftmals die Kommunikation erleichtert. Oder man bezieht sich auf ein Bild, das aber dann auch jeder sehen können muss. Ein Konzept, das durch die Sprache (Telefon) relativ schwierig zu erklären ist, wird so (Video, Grafiken) sehr einfach klargemacht werden können.

Um echte Collaboration zu realisieren muss eine Software also mehrere verschiedene Medien beherrschen, und auch den Wechsel zwischen diesen unterstützen. In der Abbildung 1 [1] werden verschiedene Methoden der Zusammenarbeit nach zwei Achsen aufgetragen: Synchronität und Interaktivität. Darin gibt es verschiedene Szenarien: Tele-Mentoring, Informationsbrowsing, Collaborative Work und Shared Work.

Und während manche dieser Anwendungen sich auch überlappen, so gibt es keine Applikation, die alle Bedürfnisse erfüllt, und an dieser Stelle

setzt Collaborationsoftware ein. Solche Systeme versuchen alle nötigen Anwendungen unter einem Dach zu vereinen.

Um die Bedeutung und den potenziellen Nutzen solcher Systeme herauszuarbeiten werden im Folgenden vier Systeme vorgestellt, miteinander verglichen und ihre Arbeitsweise, technischen Hintergründe, Stärken und Schwächen gegenüber gestellt.

1.3 Warum web-based Collaboration?

Das Internet hat sich zu einem allgegenwärtigen Medium entwickelt, das von fast jedermann zu jeder Zeit an jedem Ort benutzt werden kann. Fast alle Rechner sind an das Netz angeschlossen und viele andere Geräte bekommen mittlerweile auch eine Schnittstelle zum Internet - z.B. Mobiltelefone, PDAs, Fernsehgeräte etc.

Falls Collaborationsoftware, beziehungsweise Collaboration ganz allgemein, alltäglich werden soll, muss sie eigentlich mit dem Netz interagieren können und somit Teil des Netzes werden. Ein anderes wichtiges Argument dafür, an das Netz angeschlossen zu sein, ist die Fülle der Informationen, die über das Netz zugänglich sind, und bei web-basierten Lösungen nahtlos integriert werden können. [2]

Bedient man sich darüber hinaus sogar der Technologie des Internets (JAVA, html, CGI-Skripte, AJAX), hat man in der Regel gute Chancen, dass die Software auf vielen Systemen eingesetzt werden kann.

2 Vorstellung verschiedener Systeme

Im folgenden Kapitel werden vier verschiedene Web-basierte Collaboration Systeme und ihre Eigenheiten vorgestellt.

2.1 Stanford NRE Collaboratory System

Das Collaboratory System [3] des Stanford National Research Enterprise (NRE) ist ein einfaches Client - Server System, um für Mitarbeiter gemeinsam Anwendungen zu starten. Die Benutzer starten hierbei eine Verbindung mit einem Kollegen, und falls diese zustande kommt werden vordefinierte Anwendungen auf beiden Seiten gestartet. Serverseitig werden hier Benutzerdaten und Programmpfade gespeichert, so dass auf den Clientrechnern nur wenige Zeilen Code nötig sind.

Das Konzept sieht hierbei vor, dass jeder Benutzer einen Teil der Software installieren muss, und dann das System benutzen kann. Außerdem

**Pacific Northwest National Laboratory
Environmental Molecular Sciences Laboratory**

**Collaborative Research Environment
(CORE)**

Session Member Name:

CREATE A SESSION

(your session will be listed in the active sessions)

New Session Name

Choose Tools

Audio/Video White Board
 WebTour Data Notebook
 Televiewer On Line Mass Spectrometer

Create New Session

ACTIVE SESSIONS

Click on a session name to join.
Click on your member name to leave the session.

<i>Sessions</i>	<i>Tools</i>	<i>Member Name</i>
Electronic Notebook	Televiewer	Jim Myers
	White Board	Paul Keller
	Data Notebook	Debbie Payne
Mass Spectroscopy	Audio/Video	Jim Callis
	WebTour	John Price
	Televiewer	

Abbildung 2. CORE Benutzerschnittstelle im Browser. [1]

besitzt jeder Benutzer ein sog. Rolodex (Karteikartensystem), in dem er hinterlegen kann, mit welchem anderen Benutzer er welche Programme gemeinsam aufrufen möchte.

Bei Bedarf wird dann aus dem Rolodex heraus eine Verbindung initiiert, und das System startet dann die jeweils richtigen Programme um den Teilnehmern das gemeinsame Arbeiten zu ermöglichen. Das System läuft in der Urform gänzlich über die Kommandozeile, eine Implementierung für das X Window System (X11) ist angedacht.

Das System von Stanford steuert ausschließlich Unix Tools an, die dann zum zusammen arbeiten genutzt werden können. Zum letzten Stand der Entwicklungen konnten folgende Programme benutzt werden:

- NV - Network Video
- VAT - Visual Audio Tool
- wb - WhiteBoard
- Ytalk - talk client
- xmx - shared X server

2.2 EMSL CORE

Das CORE [1] System (Collaborative Research Environment) des EMSL (Environmental Molecular Sciences Laboratory) stellt aus der Benutzersicht eine Erweiterung des WWW dar, die in einem Browser (siehe Abbildung 2) gesteuert wird. Hierbei wird durch das CGI-Skript und die gespeicherten MIME-Types Programme gestartet, die die Zusammenarbeit ermöglichen.

Diese sind allerdings, ähnlich wie bei dem System der Stanford Universität, externe Programme. CORE bietet allerdings die schon erwähnte Oberfläche, die im Browser erzeugt wird und somit als Bestandteil des WWW erscheint. Dies soll die Bedienung vereinfachen. Darüber hinaus bietet CORE eine Sitzungsverwaltung an und man kann in eine bestehende Sitzung einsteigen.

Wie in der Abbildung 2 zu sehen ist, kann man dabei sehen, welche Benutzer sich mit welchen Tools in der Session aufhalten. Von CORE wurden zum letzten Stand folgende Werkzeuge angeboten:

- Audio/Video - Konferenzsoftware
- WebTour - Synchronisierung von Browsern
- Filesharing
- Chatbox
- Televiewer - Remote Display eines Rechners
- White Board
- EMSL Data Notebook

Die Absicht hinter CORE ist, eine zentrale plattformunabhängige Möglichkeit zu schaffen und verschiedene elektronische Collaborationtools anzusteuern.

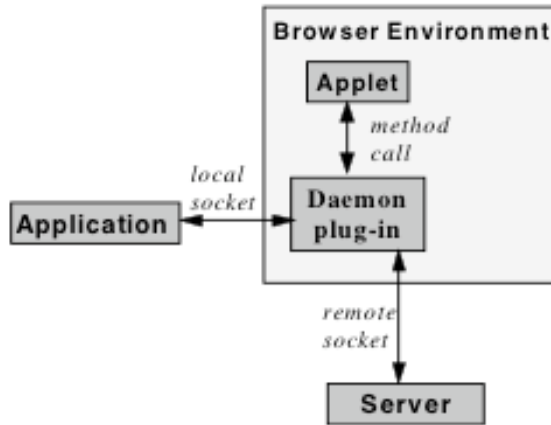


Abbildung 3. TANGO Kommunikationsschema. [5]

2.3 NPAC TANGO

TANGO [4], entwickelt vom NPAC (Northeast Parallel Architectures Center) versucht die Collaboration über JAVA Applets zu realisieren, die quasi nahtlos in ein Browserfenster integriert werden können und damit sehr viel näher am WWW und dessen Informationen sind, als reine JAVA-Applikationen. Hier werden primär die Collaboration Funktionen in Applets ausgeführt. Aber auch lokale vorhandene Applikationen können (dann beidseitig) integriert werden. Eine Steueranwendung koordiniert die Abläufe. Einen ersten Überblick über die Systematik gibt Abbildung 3.

Benutzer finden sich hier in Sessions zusammen und eine Datenbank regelt das zeitgleiche Zugreifen auf die selben Daten. Das Verwenden von JAVA Applets hat darüber hinaus den Vorteil der Modularität, d.h. Komponenten werden erst geladen, wenn sie benötigt werden.

TANGO hat eine wesentlich komplexere Struktur als die beiden vorhergehenden Systeme. Abbildung 4 zeigt die etwas komplexeren Beziehungen zwischen den Elementen des Systems.

Das Konzept von TANGO sieht eine explizite Erweiterbarkeit über ein API vor, so dass relativ problemlos jedwede Funktion angebunden werden könnte. Beispielsweise sieht man in Abbildung 5 [5] ein Whiteboard und in Abbildung 6 [5] ein Applet mit einer Chatanwendung.

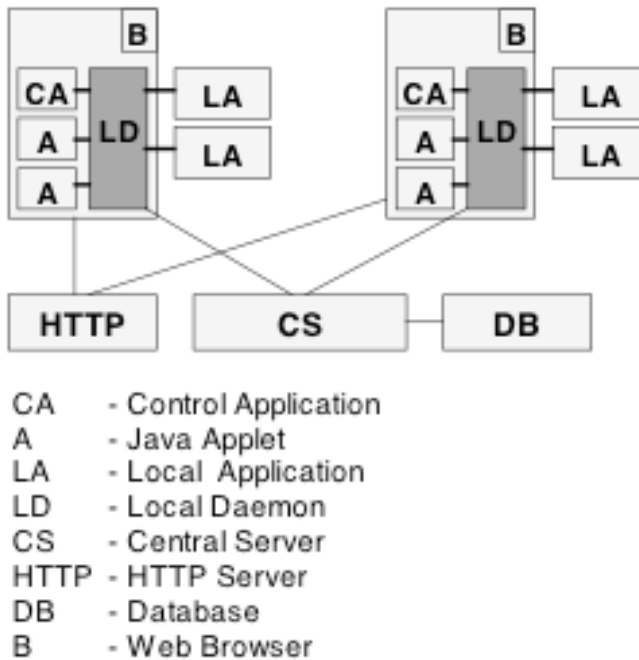


Abbildung 4. Globale Architektur des TANGO Systems. [5]

2.4 NCSA Habanero Project

Das Habanero-Projekt [6] des NCSA (National Center for Supercomputing Applications) geht einen anderen Weg.

Habanero ist in JAVA realisiert und ist damit auch im Wesentlichen plattformunabhängig. Alle Bestandteile der Collaborationssoftware befinden sich in einer geschlossenen Oberfläche, die sich dem Benutzer kompakt darstellt und daher relativ einfach zu bedienen ist. Alle Teilnehmer einer Session (siehe Abbildung 7) sehen das Selbe und können über die Oberfläche miteinander kommunizieren und interagieren. Solche Sessions können auch aufgezeichnet und zu einem späteren Zeitpunkt abgespielt werden. Einen Überblick über die angebotenen Funktionen zeigt das Hauptmenü in Abbildung 7. [7]

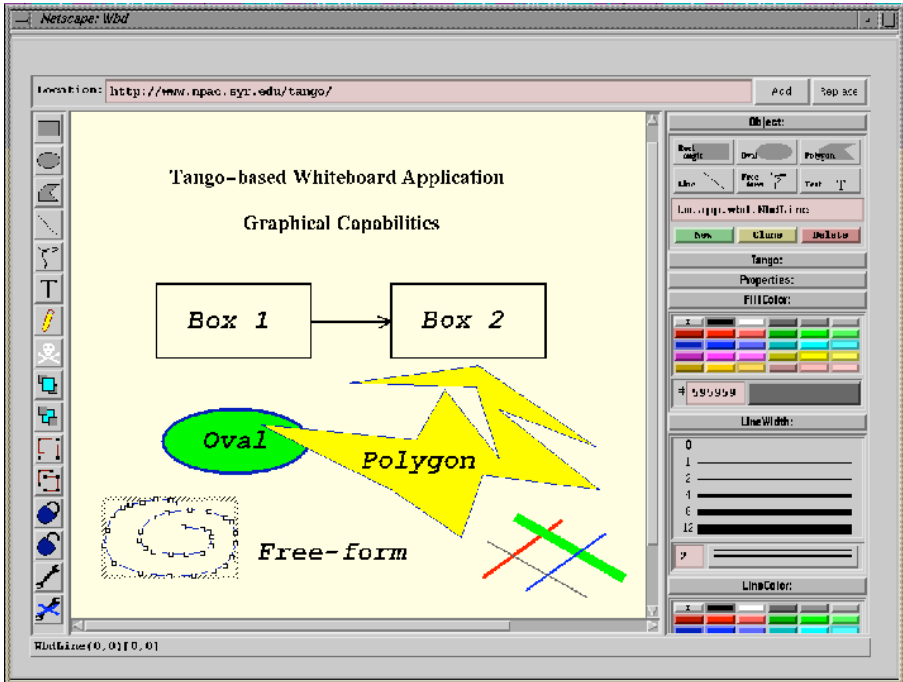


Abbildung 5. TANGO Whiteboard Applet. [5]

3 Vergleich der Systeme

3.1 Technik

In diesem Abschnitt werden die Techniken der Systeme näher betrachtet, um anschließend einen Vergleich durchführen zu können.

Allgemein: Zur technischen Realisierung ist natürlich auch der Entstehungszeitraum der jeweiligen Systeme zu beachten. Alle hier vorgestellten Systeme entstanden am Ende des zwanzigsten Jahrhunderts:

- Stanford NRE: 1995
- EMSL CORE: 1996
- NPAC's TANGO: 1997
- NCSA Habanero 1999

Stanford NRE Collaboratory System: Das System von Stanford bietet eine sehr eingeschränkte grafische Interaktionsmöglichkeit, die durch CGI-Skripte erzeugt wird, die auf dem Server laufen. Das System wurde nur



Abbildung 6. TANGO Chat Applet. [5]

für UNIX-Systeme konzipiert und erfordert die Installation von Skripten auf den jeweiligen Client Rechnern (durch `make`). Außerdem müssen die MIME-Types des Systems dem Betriebssystem bekannt gemacht werden, so dass dann Kommunikationsanfragen vom serverseitig laufenden CGI-Skript gestartet werden können. Auch denkbar ist die Versendung einer Einladung als Datei per Email.

Das Netz wird hierbei direkt genutzt, um die Verbindung zwischen zwei Benutzern herzustellen, die Nutzer aufzulisten und deren Kommunikationsdaten zu speichern. Nach Zustandekommen einer Verbindung wird über das Netz nur noch durch die Programme kommuniziert, in denen zusammen gearbeitet wird. Die Daten, mit denen die Benutzer arbeiten, passieren das Serversystem nicht. Darüber hinaus ist eine Benutzung jeweils auf zwei Nutzer beschränkt.

EMSL CORE: Die grafische Oberfläche bei CORE ist ähnlich der des Stanford Systems, da auch diese aus einem (im Browser laufenden) CGI-Skript erzeugt wird. Bei CORE stehen aber ein Sessionmanager (auf dem Server) und Desktop-Steuer-Applikationen (=Clients) zur Verfügung, doch die wirkliche produktive Arbeit wird (wieder ähnlich dem

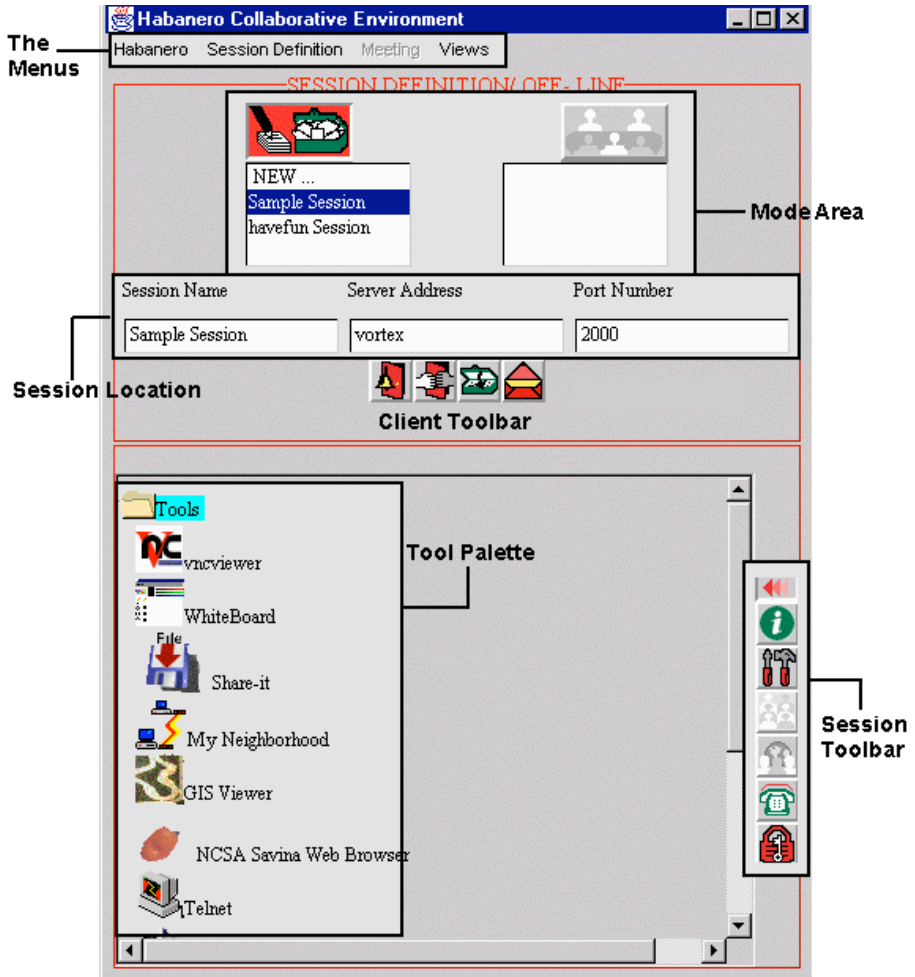


Abbildung 7. Habanero Hauptmenü. [6]

Collaboratory System) in den aufgerufenen Anwendungen geleistet. Dabei werden den externen Applikationen Parameter derart übergeben, so dass diese sich in die Collaboration Umgebung einpassen, das heißt dem Server bekannt gemacht werden.

Ebenso wie beim Stanford System kommunizieren die aufgerufenen Programme direkt miteinander und nicht mehr über den Sessionmanager. Dennoch halten alle Clients Kontakt mit dem Server, um die Informationen über die Sitzungen zu aktualisieren, was auch den wesentlichen

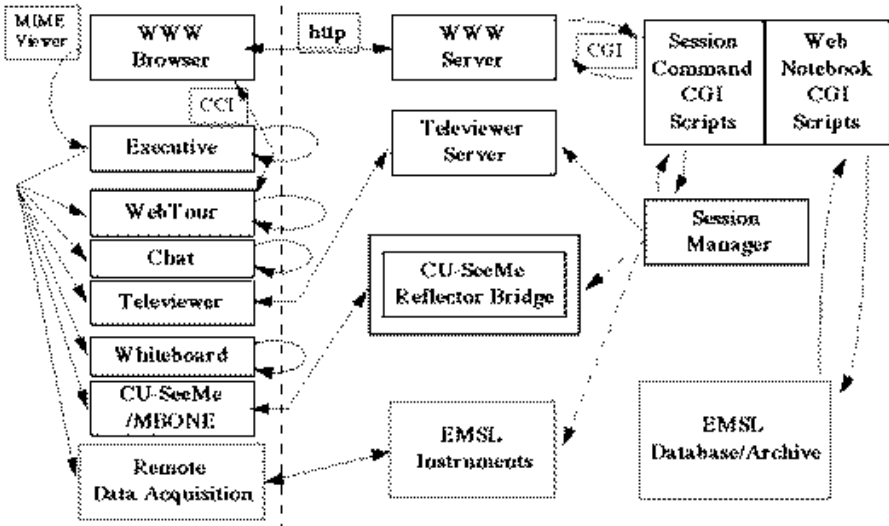


Abbildung 8. Kommunikationswege bei CORE. [1]

Unterschied zum vorhergehenden System darstellt. Außer dieser Kommunikation wird auf das Netz nicht zugegriffen.

Die Kommunikationswege bei des CORE Systems sieht man in Abbildung 8.

NPAC TANGO: Das TANGO System hingegen besitzt zwar auch eine Steuerzentrale, diese ist jedoch identisch mit der Zugriffsoberfläche (Web-Interface, siehe Abbildung 9 [8]).

Wie in Abbildung 3 zu sehen ist existiert zwar die Möglichkeit, externe Applikationen aufzurufen und ähnlich wie bei CORE in die Umgebung einzupassen, gedacht ist das System jedoch so, dass verschiedenste Applets durch ein API an das TANGO System angebunden werden können. Hierbei sind APIs in verschiedenen Sprachen (C, JAVA und JavaScript) verfügbar, um verschiedene Anwendungen anbinden zu können.

Ein weiterer Vorteil hierbei ist das modulare Applet System, das TANGO zu Grunde liegt. Dadurch werden nur die Applets geladen, die auch in der momentanen Sitzung benötigt werden. Darüber hinaus macht sich das System das WWW explizit zu Nutze, da die Informationen aus dem Netz in Applets verarbeitet werden können und sollen. Auch das Einbinden von Applets in Webseiten wird hierbei genutzt.

NCSA Habanero: Wie schon in der Einführung des Habanero-Systems erwähnt bietet dieses System eine komplett in sich abgeschlossene Ar-

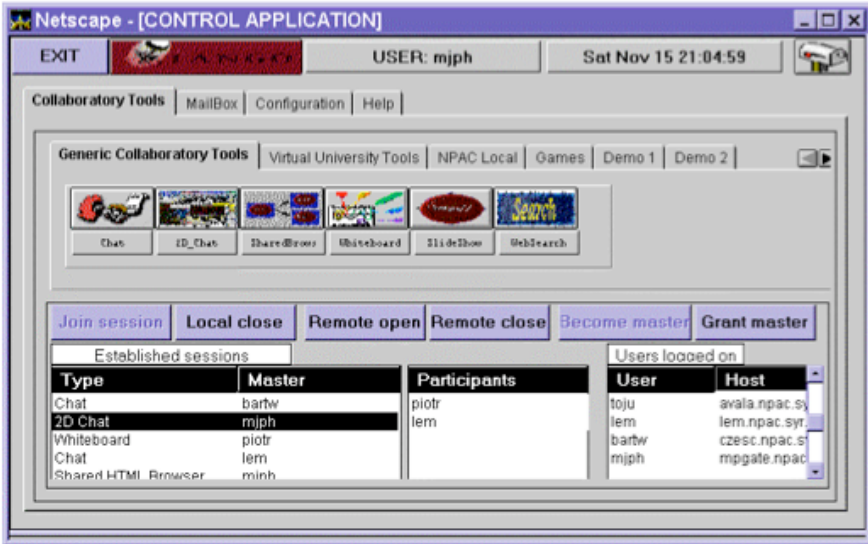


Abbildung 9. TANGO Kontrollapplikation. [5]

beitsumgebung. Alle Komponenten, mit denen die Benutzer zusammen arbeiten sollen, müssen in das System integriert sein.

Als besondere Fähigkeit von Habanero ist sicherlich das Aufzeichnen und Wiedergeben von Sitzungen zu nennen. Dies läßt sich mit externen Applikationen nur schwerlich realisieren. Andererseits ist der Funktionsumfang aber auch durch die im System eingebundenen Module begrenzt.

Alles in Allem ist Habanero eine JAVA-Applikation die auf Collaboration ausgelegt wurde und ausdrücklich als web-basierte Applikation verstanden werden soll, da auch hier nicht nur die Kommunikation über die Verbindungen durch das Netz transportiert wird, sondern auch auf Informationen aus dem WWW, die dann weiterverarbeitet werden können, über das Netz zugegriffen werden kann.

3.2 Konzept & Anwendung

Nun sollen Konzepte und Anwendungsmöglichkeiten der Systeme einander gegenübergestellt werden.

Während alle Systeme natürlich versuchen, Collaboration so einfach wie möglich zu machen, gehen sie doch von völlig verschiedenen Ausgangspunkten an das Thema heran.

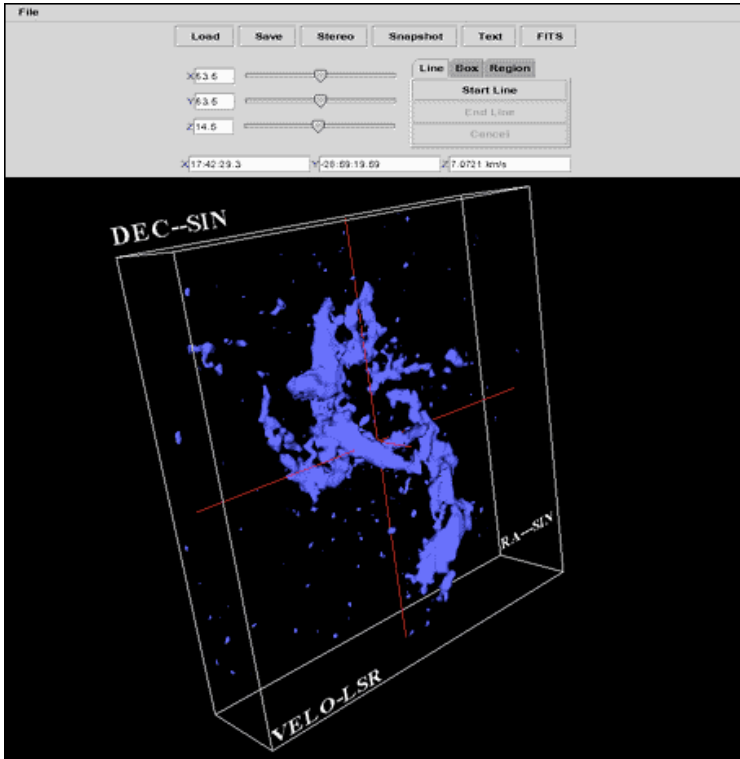


Abbildung 10. Habanero Image Viewer. [6]

Das Habanero-Projekt beispielsweise wurde als unabhängige Applikation erstellt, wobei die üblichen Anwendungen implementiert wurden: Chat, Messaging, Image Viewer (siehe Abbildung 10), Whiteboard. Eine funktionale Erweiterung ist zwar möglich, aber nicht durch den Nutzer vorgesehen.

Falls man also Collaborationsoftware benötigt, die die Grundfunktionen zusammenfasst und keine speziellen Anwendungen benutzt oder braucht, so käme das Habanero-System in Betracht.

Sollten spezielle Programme benötigen werden, dann sind diese sicherlich in den anderen Systemen leichter unterzubringen. TANGO hat diverse APIs, mit deren Hilfe man die Kommunikation zwischen den Elementen selbst implementieren kann, so dass man nicht darauf angewiesen ist, welche Funktionen bereits implementiert sind. Im CORE und dem Stanford System sind Aufrufe an neue Anwendungen relativ leicht einzubinden. Dafür muss man eventuell ohne Sessionverwaltung auskommen, beziehungsweise eine Beschränkung auf UNIX-Systeme hinnehmen.

Die beiden letztgenannten waren auch prinzipiell nur für den hausinternen Gebrauch entwickelt worden, wobei man sich der bereits (einzeln) benutzten Programme bedienen wollte, andererseits aber auch wußte, dass diese Programme vielleicht durch andere irgendwann ersetzt werden sollten. Eine Entwicklung oder Implementierung von Collaborationtools direkt in die Software selbst war also nicht vorgesehen.

Auch in der Anwendung unterscheiden sich die Systeme grundlegend. Während man bei den Stanford und CORE Systemen hauptsächlich mit bereits bekannten (Teil-)Programmen agiert, erfordert das Benutzen von TANGO und Habanero das Erlernen einer neuen Oberfläche. Allerdings gestalten sich die GUIs meist so einfach, dass dies kaum eine Rolle spielt, und es in der Regel (da fast alle Programme ein GUI besitzen) dem Benutzer, der noch nicht mit der Umgebung vertraut ist, einfach fällt, mit einer grafischen Benutzeroberfläche zurecht zu kommen. Der Vorteil besteht dann aber immer noch darin, dass sich der Benutzer nicht die Anwendung von verschiedenen Programmen aneignen muss, sondern die ähnliche Funktionsweisen von einem Teil des Programmen auf einen anderen Teil übertragen kann.

3.3 Stärken & Schwächen

Je nachdem, von welchem Standpunkt man aus die jeweiligen Konzepte betrachtet, haben alle vier vorgestellten Systeme ihre Stärken und Schwächen. Diese sollen hier nochmals zusammengefaßt werden, um auch zu erklären, unter welchen Voraussetzungen man welche Ergebnisse erwarten darf.

Stanford NRE Collaboratory System: Eine große Stärke des Stanford Systems ist sicherlich der sparsame Umgang mit Ressourcen, da kein eigenes Programm im Hintergrund lief, sondern die Aufrufe direkt an das System gegeben wurden. Für den gedachten Verwendungszweck war das Programm ausreichend.

Ein Schwachpunkt ist, dass aus heutiger Sicht das System nicht mehr die Anforderungen der Nutzer an Programme erfüllt. Darüber hinaus lief das System lediglich auf Unix Rechnern und ermöglichte es den Nutzern nicht, im Collaboration System andere Mitarbeiter zu finden. Außerdem sind die angebotenen Programme mittlerweile relativ veraltet und kaum noch im Einsatz. Da das System aber von der Stanford Universität für ihren hauseigenen Gebrauch entwickelt wurde, erfüllte es alle Bedingungen, aber eben auch nur für die Stanford Universität.

EMSL CORE: Die Weboberfläche (siehe Abbildung 2), die das CORE System dank seines CGI-Skriptes hat, erleichtert es dem Nutzer ein Bild davon zu bekommen, welche Sitzungen gerade laufen und welche Möglichkeiten er hat, um eine neue Sitzung zu eröffnen. Hiermit hat CORE einen Vorteil gegenüber dem Stanford System, trotzdem bietet es in keiner Weise die Möglichkeiten eines TANGO oder des Habanero-Systems an.

Da CORE keine eigenen Anwendungen besitzt kann man die Programme benutzen, die bereits zur Verfügung stehen, muss aber evtl. die Anbindung selbst vornehmen. Ansonsten geht das CORE auch relativ sparsam mit den Ressourcen des Systems um. Die Technologie hat sich aber mittlerweile derart weiterentwickelt, dass auch die Systematik hinter CORE als veraltet anzusehen ist. Auch CORE wurde vom EMSL für den Eigenbedarf entwickelt und band hauptsächlich bereits verwendete Applikationen ein.

NPAC TANGO: TANGO läuft in einer Browserumgebung und ist durch den Einsatz von JAVA Applets relativ plattformunabhängig. Sollen externe Programme zum Einsatz kommen, muss man zwar (wie bei CORE) selbst eine Anbindung schaffen, dies ist hier aber durch die APIs besonders einfach, so dass prinzipiell jede Art von Anwendung mit TANGO laufen können sollte.

Die Oberfläche bei TANGO läuft bereits vollkommen grafisch, so dass auch unerfahrenere Benutzer gut damit zurecht kommen sollten. Besonders zu erwähnen ist ausserdem der spezielle Mechanismus, der Module erst lädt, wenn sie benötigt werden, was wiederum eine Schonung der Ressourcen (in diesem Fall der Netzleitungen) entspricht, dies wird speziell deutlich daran, dass theoretisch alle Bestandteile der Software über einen Browser ausgeführt werden können (als Applets).

NCSA Habanero: Der größte Vorteil von Habanero ist vermutlich auch seine größte Schwäche. Da Habanero ein abgeschlossenes Paket ist, fehlt ihm jede Erweiterbarkeit und jede Anbindung nach aussen.

Andererseits entstehen genau dadurch die großen Vorteile des Systems: Einheitliche Oberfläche, leichte Bedienbarkeit, kompaktes System. Dennoch muss hier eine Anwendung installiert werden und die gesamte Anwendung muss gestartet werden, bevor man sich am System anmelden kann.

Eine besonders nützliche Funktion von Habanero ist die Möglichkeit des Aufzeichnens von gesamten Sitzungen und deren Wiedergabe, was zum Beispiel einem Mitarbeiter, der an einem Termin nicht teilnehmen konnte, hilft, den Schaffensprozeß nachzuvollziehen.

4 Fazit & Ausblick

Wenn man sich die unterschiedlichen Systeme ansieht wird klar, dass bevor entschieden werden kann welche Methodik zum Einsatz kommen soll, eine Bedarfsanalyse zentrales Element der Planung sein muss. Hierbei müssen Fragen gestellt werden wie:

- Welche Betriebssysteme sollen genutzt werden?
- Existieren kritische Anwendungen, die zugänglich gemacht werden müssen?
- Wie statisch ist das System? Wie oft sollen neue Anwendungen hinzugefügt werden?
- Welchen technischen Kenntnisstand haben die Mitarbeiter, die das System nutzen sollen?

Außerdem ist heutzutage jeder, der Collaboration betreibt daran interessiert eine nahtlose Integration seiner übrigen Werkzeuge vorzufinden: Email, WWW, FTP, Kalender, etc.

Schließlich bleibt bei der Beobachtung der Entwicklung auf diesem Gebiet und der Erfahrungen mit Collaboration Systemen in der Vergangenheit festzuhalten, dass auf die Benutzung von Webtechnologien auch bei Collaboration Systemen in Zukunft nicht mehr verzichtet werden darf, da das Internet zu einer omnipräsenten Größe heranwächst, an der niemand mehr vorbeikommt.

Literatur

1. Payne, D. A., Myers, J. D.: *The EMSL Collaborative Research Environment (CORE) - Collaboration via the World Wide Web.*
<http://collaboratory.emsl.pnl.gov/presentations/papers/core.wetice96.html> (Zuletzt besucht: 17.06.2006)
2. Bajaj, C., Cutchin, S., Morgia, C., Paoluzzi, A., Pascucci, V.: *Web Based Collaborative CAAD* (pages 326-327)
<http://citeseer.ist.psu.edu/22568.html>
3. *Stanford NRE Collaboratory System*
<http://www-flash.stanford.edu/collab/collab.html> (Zuletzt besucht: 17.06.2006)
4. Beca, L., Cheng, L., Fox, G. C., Jurga, T., Olszewski, K., Podgorny, M., Sokolowski, P., Stachowiak, T., and Walczak, K.: *Tango - a collaborative environment for the world-wide web* Technical report, Northeast Parallel Architectures Center, Syracuse University
<http://citeseer.ist.psu.edu/132092.html>
5. Bernholdt, D. E., Fox, G. C., Markowski, R., McCracken, N. J., Podgorny, M., Scavo, T. R.: *Synchronous Learning at a Distance: Experiences with Tango*
<http://grids.ucs.indiana.edu/ptliupages/publications/disted/npacexperiencessc98.html> (Zuletzt besucht: 17.06.2006)

6. *NCSA Habanero*
<http://access.ncsa.uiuc.edu/Stories/Habanero/Habanero.html> (Zuletzt besucht: 17.06.2006)
7. *The Habanero Client User Interface*
<http://www.isrl.uiuc.edu/isaac/Habanero/Docs/client/> (Zuletzt besucht: 17.06.2006)
8. Trzaska, R., Klasky, S., Major, T., Podgorny, M., Fox, G. C., Bernholdt, D. E.: *TANGO Interactive for Remote Consulting and Group Software Development*
<http://www.nat.org/2003/december/tango.pdf> (Zuletzt besucht: 17.06.2006)

Social Networking for Shared Music Collections

Gert Kauntz

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
Clandestino69@gmx.de

Zusammenfassung Die Musiksuche beschränkt sich längst nicht mehr auf den Musikladen in der realen Welt, sondern wird immer häufiger mittels Internet betrieben. Dazu bauen Recommender-Systeme auf zwei verschiedene Ansätze, den kollaborativen und den Inhaltsbasierten Ansatz. Ersterer bedient sich virtuellen Communities und macht sich das Feedback der User zu Nutze. Hier werden Methoden unterschieden, die sich entweder implizitem oder explizitem Rating bedienen. Der inhaltsbasierte Ansatz filtert hingegen Ähnlichkeiten der Musik heraus und ermittelt anhand deren die Empfehlungen. Dabei unterscheidet man die Filterung durch Metadaten und jene durch so genannte MIR Systeme. Beide Ansätze bringen jedoch diverse Probleme mit sich. Die Kombination soll zur Behebung der Schwächen beitragen und so eine Musikempfehlung via Internet ermöglichen, die den Musikgeschmack des Users am besten trifft.

1 Einführung

Ob eMule, Musicload, Amazon oder iTunes, alle verbinden zwei Gemeinsamkeiten: Musik und Internet. So wird das virtuelle Netz zunehmend zum Ersatz vom Musikladen um die Ecke.

Die Vorteile sind offensichtlich: es ist bequem, spart Zeit und die Auswahl ist unbegrenzt. Dies zieht jedoch einen wesentlichen Nachteil nach sich. Wer blickt noch durch, bei dieser Fülle von Angeboten und Informationen?

Beim Musikladen in der realen Welt, kann man meist auf die Fachkundigkeit des Musikfachverkäufers vertrauen. Er informiert über Neuheiten und gibt Tipps die den persönlichen Musikgeschmack treffen.

In der virtuellen Welt fehlt solch ein Ansprechpartner. Wie kann dieses Problem gelöst werden. Wie kann eine Applikation dabei behilflich sein neue Musikgruppen und Lieder, die gefallen, zu entdecken?

Dazu wurden zwei verschiedene Ansätze entwickelt: Der Content-Basierte und der Kollaborations-Basierte Ansatz. Die Vorgehensweise ist bei beiden Ansätzen ähnlich. Der User interessiert sich für eine bestimmte Band, bzw. einen bestimmten Song. Das System ermittelt daraufhin eine Reihe von Lieder und/oder Bands die dem User ebenfalls gefallen könnten.

Die Methode mit der die Applikation die passenden Bands und Songs filtert ist jedoch grundverschieden.

Die Architektur von Recommender Systemen wird in Abbildung 1 grafisch dargestellt. Sabine Kuzdas verwendet diese Grafik zur Veranschaulichung in Ihrem Artikel über Schwachstellen von Recommender Systemen[3]. Die Abbildung vereint Content-Basierung und Kollaborations-Basierung. Nach einer Musikanfrage des Benutzers an den Server ermittelt das System anhand einer Korrelations-Basierten Datenbank oder einer auf Ratings basierenden Datenbank Empfehlungen, die an den User weitergeleitet werden. Im kollaborativen Fall werden zusätzlich Ratings des Users an das System gesendet.

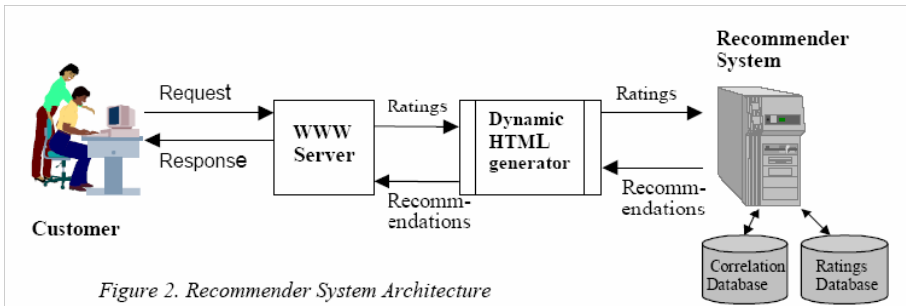


Figure 2. Recommender System Architecture

Abbildung 1. Architektur von Recommender Systemen. Eine Anfrage des Users an den Server wird an das Recommender System weitergeleitet. Daraufhin wird im kollaborativen Fall die auf Ratings basierte Datenbank, und im inhaltsbasierten Fall die Korrelations-Basierte Datenbank durchsucht und entsprechende Empfehlungen an den User zurückgegeben.

2 Content-Basierte Filterung in Music-Recommender Systemen

Balabanovic definiert Content-Basierte Filterung folgendermaßen: „We consider a pure content-based recommendation system to be one in which recommendations are made for a user based solely on a profile built up by analyzing the content of items[1].“ Content-Basierte Empfehlungen werden also lediglich anhand der Analyse des Inhalts eines Items abgegeben.

Ebenfalls aus dem Artikel von Sabine Kuzdas[3] stammt Abbildung 2. Hier wird die grundsätzliche Funktionsweise einer Content-Basierten Filtermaschine dargestellt. Aus der Eingabe des Benutzers wird von der Content-Basierten Maschine ein spezifisches Benutzerprofil erstellt, welches in einer Datenbank gespeichert wird und bei weiteren Anfragen ergänzt werden kann. Anhand dieses Profils werden Items aus einer Datenbank eruiert und dem User empfohlen.

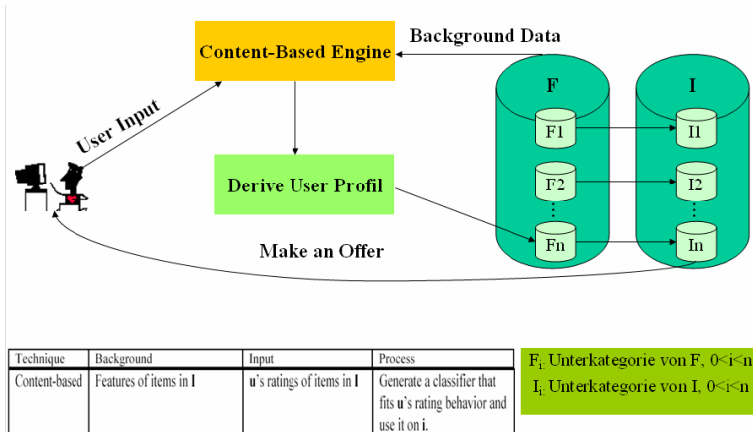


Abbildung 2. Darstellung einer Content-Basierten Filtermaschine. Eine Anfrage des Users wird an die Content-Basierte Maschine gesendet. Jene leitet aus den Daten ein spezifisches User-Profil ab, welches in einer Datenbank gespeichert wird. Daraus werden aus einer Datenbank, in der die Items gespeichert sind, jene dem Benutzer empfohlen die die höchste Übereinstimmung mit dem Benutzerprofil haben.

Die Analyse des Inhalts kann jedoch auf Basis von Metadaten, als auch auf Basis einer Analyse diverser Kriterien wie Rhythmus, Tempo, etc. stattfinden. Dieser Methode bedienen sich die so genannten MIR-Systeme. Eine Kombination aus Metadaten- und MIR-Systemen ist ebenfalls möglich.

2.1 Content-Basierte Filterung anhand von Metadaten

Die einfachste, aber auch uneffektivste Methode des Music Recommending ist die Filterung anhand von Metadaten.

Jeder Song hat bestimmte Attribute wie Herkunftsland, Sprache, Genre, Titel, Interpret, Erscheinungsdatum, Verkaufsrang, etc. Anhand dieser einfach zu bestimmenden, statischen Eigenschaften werden Lieder ermittelt, die dem Benutzer am besten gefallen könnten.

Die Filterung kann entweder stattfinden, indem nur Lieder gefunden werden, die ähnliche Metadaten wie die Anfrage besitzen, oder aber, es werden die Metadaten mit dem Userprofil in Relation gesetzt. Dabei könnte beispielsweise im Profil gespeichert sein, welche Sprache der User bevorzugt, oder ob ihm Mainstream Musik oder eher unbekannte Songs zusagen. Dies muss nicht explizit vom User spezifiziert werden, sondern kann aus den Suchanfragen abgeleitet werden.

Die Filterung anhand von Metadaten wird heutzutage fast nur noch als zusätzliches Mittel eingesetzt. Ein System, welches ausschließlich auf diese Methode baut, wäre

nicht mehr konkurrenzfähig, da es die ungenaueste Methode ist Empfehlungen zu geben.

2.2 Content-Basierte Filterung mittels „Music Information Retrieval (MIR)“

Wikipedia definiert MIR folgendermaßen: „Music information retrieval or MIR is the interdisciplinary science of retrieving information from music.“[21] Es handelt sich also um die interdisziplinäre Wissenschaft der Informationsgewinnung aus Musik. Diese Informationen können Rhythmus, Tempo, Frequenz, etc. sein. Es ist beispielsweise möglich bestimmte wiederkehrende Beatfrequenzen zu messen und daraus ein Schema zu kreieren. Dieses Schema kann mit dem anderer Songs verglichen werden und so diejenigen filtern, die vom Stil her am ähnlichsten sind[2]. Sollte sich der Benutzer beispielsweise für „Morcheeba“ interessieren, deren Stil meist langsame Beats sind, könnte das System „Massive Attack“ ermitteln, die diesem Schema ebenfalls entsprechen.

Hinter Country, Heavy Metal, Brit-Pop, etc. stecken meist ähnliche Rythmen und Klangfarben. So können die meisten Lieder anhand von diesen Attributen einem Genre zugeordnet werden, und beliebig weiter spezialisiert werden. Der Kreis der gewünschten Lieder wird so anhand eines „Lieblingsliedes“ immer weiter eingeschränkt, abhängig davon wie sehr die Attribute spezialisiert werden.

Die Analyse der Songs erfolgt dynamisch. Die Anfrage wird analysiert und mit zuvor analysierten Songs verglichen. Daraus erstellt das System anschließend seine Empfehlungen.

2.2.1 Methoden und Techniken von MIR-Systemen

MIR-Systeme verwenden unterschiedliche Methoden und mathematische Algorithmen zur Informationsgewinnung. Die wichtigsten sind hier beschrieben. Eine Kombination davon wird ebenfalls häufig eingesetzt.

- **Fast-Fourier Transformation (FFT)**
Ein akustisches Signal besteht aus vielen übereinander liegenden Wellen mit verschiedenen Frequenzen und Amplituden[5]. Zum Analysezeitpunkt befindet sich dieses Signal in einer auf Zeit basierenden Form. Um Informationen über das Spektrum ermitteln zu können, muss es in eine auf Frequenz basierende Form (Frequenzteil, Amplitude) umgewandelt werden. Dafür dient die Fast-Fourier Transformation[6].
- **Wavelet Transformation (WT)**
Bei der Wavelet Transformation wird mittels zeitlich lokalisierten Wellen (Wavelets) analysiert, im Gegensatz zur FFT, die mittels unendlichen Sinus- oder Kosinusschwingungen scannt[5].

- Short-Time Fourier Transformation (STFT)
Ähnlich wie bei den Wavelets wird hier ebenfalls nur auf einem zeitlich begrenzten Ausschnitt gearbeitet[5].
- Diskrete Kosinus Transformation (DCT)
Auch hier wird eine Transformation von Zeit- in Frequenz basierte Darstellung vorgenommen. Es werden aber nicht alle Frequenzen gleichmäßig betrachtet was zu einer Glättung des Spektrums führt. Hochfrequente, jedoch kleine Änderungen zwischen den Werten werden entfernt. Die DCT führt dazu, dass die berechneten Koeffizienten unkorreliert sind und dadurch jeder eine eigene Aussagekraft hat[5].
- Hauptkomponentenanalyse (PCA)
Bei einem Vektor mit sehr vielen Komponenten, tragen nicht alle zu einer Verbesserung des Ergebnisses bei. Um Überspezialisierung zu vermeiden und den Vektor auf die wichtigen Komponenten zu reduzieren, wird die Hauptkomponentenanalyse eingesetzt. Hierbei wird für die Ausgangsdaten ein geeigneteres Koordinatensystem berechnet, dass die Richtung der Hauptkomponenten berücksichtigt. Dabei werden neue Koordinatenachsen gefunden entlang derer die größte Streuung in den Daten besteht. Eine genauere Beschreibung findet sich in der Ausarbeitung von Julian Hartmann von der Universität Münster[5].

2.2.2 Beispiele für MIR Systeme

- Eine Auflistung diverser MIR-Systeme findet sich unter der URL: <http://mirsystems.info/>[25]. Zu jedem System wird zusätzlich eine detaillierte Beschreibung geboten.
- Eines dieser Systeme ist das kommerzielle Music Recommender System MusicSurfer [8]. Es bedient sich ausschließlich der Methode des „Music Information Retrieval“ und verzichtet dabei auf die zusätzliche Spezifizierung durch Metadaten. Es handelt sich somit um ein reines MIR-System. Nach der Eingabe eines Künstlers werden die Lieder dieses Interpreten angezeigt. Anhand eines Anfrageliedes kann anschließend eine „Similarity Map“ erstellt werden. Die Nähe der Kugeln beschreibt die Nähe zur Anfrage, und die Farbe das Genre. Die Karte kann nach diversen Kriterien geordnet werden. Beim Rollover über eine Kugel wird das entsprechende Lied mit Interpret und diversen Zusatzinformationen angezeigt. Die Map wird anhand von Rhythmus, Harmonie und Instrumente erstellt. Abbildung 3 zeigt das den Zustand des Systems nach einer Anfrage und nach der Anforderung einer „Similarity Map“.

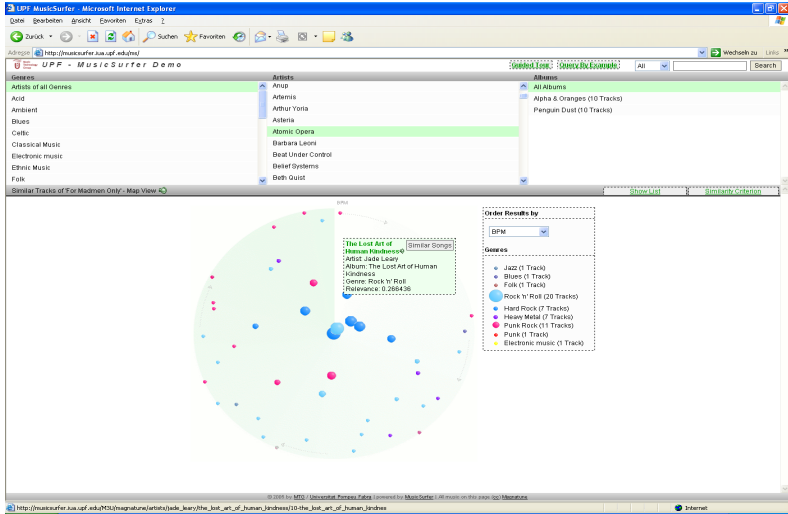


Abbildung 3. Das Music Recommender System MusicSurfer Zustand nach der Anforderung einer „Similarity Map“ für ein Anfragelied

2.3 Probleme und Nachteile Inhaltsbasierter Filterung

Der Content-Basierte Ansatz zieht jedoch diverse Nachteile und Probleme nach sich.

- Die Überspezialisierung ist ein generelles Problem von inhaltsbezogenen Recommender Systemen. Das System liefert dem Benutzer nur eine sehr eingeschränkte Bandbreite an Items, da evt. die Attribute zu genau spezialisiert wurden und somit gewisse Items nicht gefunden werden, obwohl sie von Interesse wären.[3]
- Ein weiteres Problem dieser Art von Filterung ist die Tatsache, dass davon ausgegangen wird, dass der User ähnliche Songs und Interpreten mag, wie seine Suchanfrage. Die Empfehlungen decken nur ein geringes Spektrum ab, und zwar jene Songs, die von Rhythmus, Klangfarbe, etc. ähnlich sind. So entgehen dem Benutzer Lieder die nicht gefunden werden können, da sie keinen inhaltlichen Bezug zum Anfragesong haben, jedoch ebenfalls von Interesse wären.
- Ebenfalls zu erwähnen ist, dass durch diese Technik kein Feedback durch die Benutzer der Plattform erlangt wird.

- Die Suchanfrage ist statisch, solange keine neuen Objekte hinzugefügt werden. Die Empfehlungen sind dadurch stark limitiert. Eine Suchanfrage wird immer die gleichen Ergebnisse liefern, solange keine neuen Songs und/oder Interpreten in die Datenbank eingepflegt wurden.

3 Kollaborative Filterung in Music Recommender Systemen

Musik beinhaltet von Natur aus eine starke soziale Komponente. Oft werden neue Songs und Interpreten durch Freunde entdeckt, indem man gemeinsam Musik hört und sich gegenseitig neue Lieder empfiehlt[7]. Angenommen zwei Freunde kaufen sich eine neue CD, von ersterem ist bekannt, dass er einen ähnlichen Geschmack wie man selbst hat, wobei der zweite Freund eher zu anderer Musik tendiert. Mit größerer Wahrscheinlichkeit wird man sich die gleiche CD wie der erste Freund kaufen.

Dieses Prinzip versucht die kollaborative Filterung aufzugreifen und zu erweitern, indem die Empfehlung nicht an einer oder wenigen Personen orientiert sind, sondern an vielen, oft hunderten Personen, mit einem ähnlichen Musikgeschmack. Dabei machen sich die Systeme soziale Netzwerke zum Nutzen.

Van Schyndel definiert kollaborative Filterung folgendermaßen: Kollaborative Filterung macht sich das Feedback der Benutzer zum Nutzen, um die Qualität des Materials, welches dem Benutzer präsentiert wird zu verbessern[14].

Abbildung 4 stellt die grundsätzliche Funktionsweise einer Kollaborations-Basierten Filtermaschine dar[3].

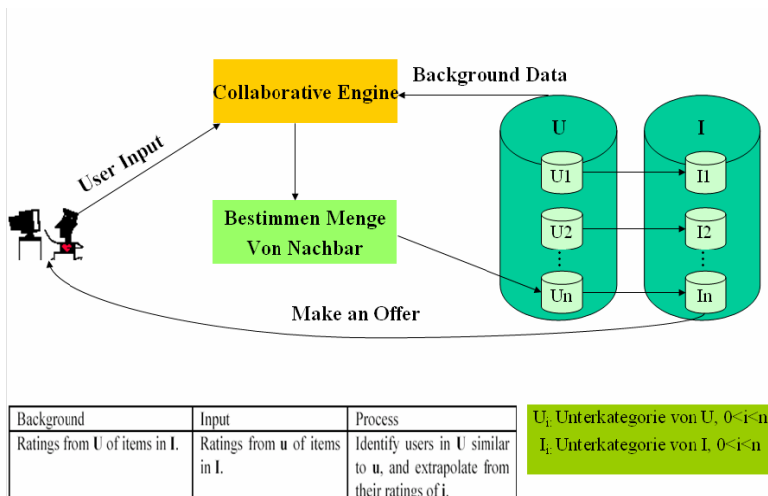


Abbildung 4: Darstellung einer kollaborativen Filtermaschine, Die Eingabe des Benutzers wird an das System weitergeleitet. Dieses bestimmt aus einer User Datenbank und der Eingabe die Menge von nächsten Nachbarn und ermittelt jene Items, die von diesen Usern hoch bewertet wurden. Jene werden dann letztendlich empfohlen.

3.1 Geschichte kollaborativer Filterung

„I am he as you are he as you are me and we are all together“ lautete eine Textzeile aus John Lennons und Paul McCartneys Song „I am the Walrus“. Somit kann man sagen, dass die Beatles bereits 1967 den Grundgedanken des „Collaborative Filtering“ geprägt haben.

Erstmals wurde der Begriff der „kollaborativen Filterung“ jedoch erst 1992 von Goldberg et. Al erwähnt[10]. Zur Reduzierung der Flut von eMails am Xerox Palo Alto Research Center wurde das System „Tapestry“ entwickelt. Jeder Mitarbeiter konnte nach Lesen einer Nachricht diese bewerten. Andere Mitarbeiter konnten so einen Filter definieren, indem sie nur solche eMails zuließen die von bekannten Mitarbeitern als relevant eingestuft wurden.

Das erste Music Recommender System das „kollaborative Filterung“ einsetzte hieß Ringo[11] und wurde 1994 entwickelt. Die User beschreiben Ihre Hörvorlieben indem sie einige Künstler bewerten. Daraus entwickelt das System ein Benutzerprofil. Mit der Zeit ändert sich das persönliche Profil. Ringo verwendet diese Profile und den Vorstoß der kollaborativen Filterung um personalisierte Ratschläge zu geben.

3.2 Rating-Systeme

Um personalisierte Empfehlungen geben zu können, muss die Applikation über ein Rating-System verfügen, anhand dessen die gewünschten Items ermittelt werden.

Dabei werden zwei Methoden unterschieden denen sich solch ein Rating-System bedient, das explizite und das implizite Rating. Sie schließen sich jedoch nicht gegenseitig aus können und auch parallel zueinander verwendet werden.

3.2.1 Das explizite Rating System

Explizites Rating ist allgegenwärtig in unserer Welt, ob ihm Restaurant, wo man sich zwischen ein bis drei Sternen entscheiden muss, bei Hotels die ebenfalls in Bewertungskategorien aufgeteilt sind oder Filme, die von den Filmkritikern bewertet werden. Da liegt die Übertragung auf Recommender Systeme nahe.

In expliziten Rating Systemen wird der User aufgefordert ein bestimmtes Item zu bewerten. Die grundsätzliche Vorgehensweise ist folgende: Der User untersucht ein Item und weist ihm dann einen Wert auf einer Bewertungsskala zu[12]. Daraus wird ein Profil des Raters erstellt. Anschließend können durch den kollaborativen Vergleich mit anderen Profilen bessere Empfehlungen erstellt werden.

Der wesentliche Nachteil dieser Methode ist jedoch, dass damit dem User die „Last“ der Bewertung auferlegt wird[10], was oft als unangenehm empfunden wird. Jeder Bewertungsvorgang ist mit gewissen zeitlichen Kosten verknüpft, welche wiederum im Gegensatz zu dem Nutzen stehen, den der User aus den Bewertungen zieht. Dies führt auch dazu, dass viele Benutzer überhaupt nicht werten, da ihnen der Nutzen im Vergleich der Kosten zu gering ist. Damit spiegeln die expliziten Ratings nur eine Minderheit wieder, was wiederum die Qualität der Empfehlungen beeinträchtigt.

3.2.2 Das implizite Rating System

Abhilfe schafft hier das implizite Rating, welches darauf beruht, dass der User seine Bewertungen „en passe“ abgibt. Potentiell produziert jeder Benutzer der mit einem System interagiert implizite Daten[12]. Indikatoren die ein explizites Rating ersetzen sind beispielsweise die Häufigkeit, die ein User ein bestimmtes Item aufruft, die Zeit die auf einer Seite verbracht wird, auf der das Item angezeigt wird, die Anzahl der Klicks auf ein Item, oder die Kombination dieser Kriterien.

Die Genauigkeit die durch implizites Rating erreicht wird, wird von Konstan et Al. folgendermaßen beschrieben: „Our initial studies show that we can obtain substantially more ratings by using implicit ratings and that predictions based on time spent reading are nearly as accurate as predictions based on explicit numerical ratings.”[22] Die Genauigkeit lässt sich folglich mit der eines expliziten Rating Systems vergleichen, es bietet jedoch zusätzlich den Vorteil der Abgabe von deutlich mehr Bewertungen.

Hauptproblem bei diesem Vorgehen ist der Datenschutz. Selbst wenn die Anonymität des Users gewährleistet wird, wird es doch von vielen Benutzern nicht gewünscht vom System analysiert zu werden.

3.3 Methoden und Techniken Kollaborations-Basierter Filterung

Es gibt diverse Techniken mit denen kollaboratives Filtern realisiert werden kann.

Dabei ist die Laufzeit und die Präzision, also die Schnelligkeit und die Genauigkeit mit der Items aus der Datenbank ermittelt werden können von entscheidender Bedeutung.

3.3.1 Die „nächste Nachbarn“ Methode

Marko Balabanovic setzt in seinem „Fab System“ die Filterung mittels nächster Nachbarn ein[1]. Der Gedanke dahinter ist folgender: Es sollen Items empfohlen werden, die von Usern mit möglichst ähnlichem Geschmack bereits hoch gewertet wurden. Das Problem ist die entsprechenden User effizient aus der Datenbank zu ermitteln. Das „Fab System“ und diverse andere kollaborative Systeme bedienen sich dafür des „nearest neighbour“ Algorithmus. Für jeden Benutzer wird eine Anzahl von „nächsten Nachbar Benutzer“ ermittelt, deren in der Vergangenheit getätigten Ratings am meisten mit dem Profil des Benutzers korrelieren.

Technisch gesehen ist der „nearest neighbour“ Algorithmus ein auf Lernen basierter Algorithmus, der den Abstand zwischen Item-Vektoren innerhalb einer Vektorraumrepräsentation misst. Dieser Abstand beschreibt die Ähnlichkeit von Items (nächste Nachbarn). Hier wird oft die Kosinus Ähnlichkeit verwendet[23]. Bei einer Datenbank von n Items benötigt der „nächste Nachbarn“ Algorithmus lineare Laufzeit $O(n)$. Die Präzision dieser Methode ist äußerst hoch.

3.3.2 Die „Eigentaste“ Methode

Ein weiterer Algorithmus nennt sich „Eigentaste“ und ist ein kollaborativ-basierter Filterungsalgorithmus in konstanter Zeit[17]. Er verwendet „universal queries“ um User ratings auf einem bestimmten Satz von Items zu eruieren und wendet dann eine Komponenten Analyse (PCA, siehe Punkt 2.2.1) an, um die Fülle der Ergebnismatrix zu spezialisieren. Dies ermöglicht eine schnelle Verarbeitung der Daten und Erstellung einer Empfehlung. Bei einer Datenbank mit n Usern braucht ein herkömmlicher Algorithmus, der auf die nearest-neighbour Technik setzt eine Laufzeit von $O(n)$, wohingegen Eigentaste bei gleicher Präzision konstante Laufzeit $O(1)$ erreicht.

3.4 Beispiele für kollaborative Filterung

- Die wohl bekannteste Firma die kollaborative Filterung einsetzt ist Amazon [15]. Es handelt sich hier nicht um ein spezielles Music Recommender System, sondern die Plattform bietet vor allem auch die Möglichkeit Literatur, Musik und diverse andere Items zu erwerben. Um dem Benutzer einen Kauf „schmackhaft“ zu machen und den User auf andere Items, die ihm ebenfalls gefallen könnten aufmerksam zu machen bedient sie sich zusätzlich zum kommerziellen Aspekt eines Recommender Systems, welches die Kaufhandlung fördern soll. Bei einer Anfrage erscheinen sofort diverse andere Produkte die ebenfalls gefallen könnten. Bei Eingabe von der irischen Gruppe „Coldplay“ erscheint neben den Alben auch die Gruppe Keane mit ihrem aktuellen Album. Sie steht unter der Rubrik: „Das könnte Ihnen gefallen“. Amazon bedient sich sowohl einer impliziten als auch einer expliziten kollaborativen Filterung. Dabei wird verfolgt welcher Kunde, welche Art von Musik kaufte. Die Empfehlung wird dann nach dem Prinzip: „Kunden die diesen Artikel gekauft haben, haben auch jenen Artikel gekauft.“ geliefert. Neben dieser impliziten Form verwendet Amazon ein Bewertungssystem, dass die Möglichkeit bietet Produkte nach dem Kauf explizit zu bewerten. Potentielle Kunden, die sich ebenfalls für dieses Produkt interessieren, können dann anhand der Ratings sich ein Urteil über das eventuell zu kaufende Produkt bilden. Abbildung 5 zeigt die resultierende Seite einer Suchanfrage, auf der neben den gesuchten Produkten auch Empfehlungen gemacht werden. Im speziellen Fall wurde nach Produkten der Gruppe Keane gesucht. Neben den Alben dieser Gruppe wird in der linken Spalte auch das aktuelle Album der Band „The Chieftains“ empfohlen.

Alle Ergebnisse für: Keane

Schnellsuche

Das könnte Ihnen gefallen



[Tears of Stone](#)
~ The Chieftains

Verwandte Produkte

Klassik

- [Brigitte Wellness - Schön entspannt](#)
- [Celtic Minstrel / New Frontsleev](#)
- [Brigitte Wellness - Power und Energie](#)

Alle 176 Treffer für die Kategorie Kategorien und "Keane"

Ordnen nach:

- 

1. [Under the Iron Sea](#) von Keane
Island (Universal) (9. Juni 2006)
Audio CD
Durchschnittliche Kundenbewertung: ★★★★★
Gewöhnlich versandfertig bei Amazon in 24 Stunden.

Amazon-Preis: EUR 14,95

Alle Angebote ab EUR 9,99
- 

2. [Hopes and Fears](#) von Keane
Island (Universal) (10. Mai 2004)
Audio CD
Durchschnittliche Kundenbewertung: ★★★★★
Gewöhnlich versandfertig bei Amazon in 24 Stunden.

Statt: EUR 9,97

Abbildung 5. Ergebnis der Suchanfrage "Keane" bei dem System Amazon[15]. Neben den Alben dieser Gruppe wird auch eine Empfehlung für ein Album der Gruppe „The Chieftains“ (linke Spalte) geliefert.

- Eine Idee von Hewlett Packard war es ein kollaboratives System zu erstellen, das Musik und Freundschaft kombiniert, der „Music Buddy“[13]. Die Applikation beruht auf der Annahme, dass bei Leuten die zum Teil gleiche Musik auf ihrem Computer gespeichert haben die Wahrscheinlichkeit hoch ist neue Musik zu entdecken, die ebenfalls interessiert. Anfangs wird die Musik Sammlung vom User auf einen zentralen Server hochgeladen und angezeigt. Bei einem Klick auf ein Lied werden jene User angezeigt die ebenfalls dieses Lied in ihrer Sammlung haben. Es besteht nun die Möglichkeit in den Sammlungen dieser User zu browsen und so neue Musik zu entdecken.

3.5 Probleme kollaborativer Filterung

- Eines der Hauptprobleme dieser Filtertechnik ist das „Kaltstartproblem“[3]. Da dieser Ansatz besser funktioniert, je mehr User ihre Empfehlungen abgeben haben, bzw. je mehr Leute bereits die entsprechende Plattform genutzt haben, ist es für eine treffsichere Suche essentiell ein gewisses Maß von Benutzern zählen zu können. Wird aber eine Recommender Plattform neu freigeschaltet, die diese Technik verwendet, führt dies häufig zu ungenauen oder schlimmer noch gar keinen Ergebnissen bei der Suche, da noch niemand oder sehr wenige

das System verwendet haben. Die Konsequenz ist oft Frustration und Abwanderung der Benutzer.

- Ein spezielles Problem von Systemen, die auf explizites Rating bauen ist die Spärlichkeit der Rezensionen. Jene Systeme haben oft mit dem Problem zu kämpfen, dass sehr wenige Kunden das Produkt bewerten. Nur etwa 1% der Benutzer tätigen im Nachhinein Bewertungen[4].
- Verwand mit dem „Kaltstartproblem“ ist das „New Item Problem“. Wenn ein neues Lied in die Datenbank hinzugefügt wird, wird es anfangs sehr selten oder gar nicht empfohlen, da noch keine Rezensionen über es bestehen, bzw. da es noch nie referenziert wurde[3].
- Bei dem „Black Box Problem“ tritt die Schwierigkeit der Undurchsichtigkeit auf. Die Ausgabe kann nicht gesteuert werden[3]. Es werden manchmal auch Songs und Interpreten gefunden, die rein „zufällig“ mit der Suchanfrage verlinkt sind. Angenommen, man sucht nach einer Band, die sehr neu oder noch relativ unbekannt ist. Wurde diese Band erst von wenigen Usern ermittelt und/oder bewertet, so werden abstruse Verlinkungen mit dieser Band entstehen, die mehr nach Zufall als nach Empfehlung aussehen könnten.
- Eines der größten Probleme des Kollaborations-Basierten Filtern ist die „Mainstreamisierung“. User mit eher außergewöhnlichem Musikgeschmack bleiben bei kollaborativen Systemen oft auf der Strecke, da mit großer Wahrscheinlichkeit wenig oder gar keine anderen User vorhanden sind mit demselben Musikgeschmack. Diese Gruppe der User macht jedoch einen hohen Anteil der Hörer insgesamt aus. Laut einer Musikforschungsstudie von der Speedfacts GmbH[24] haben nur 25,2 Prozent ihre Präferenz im so genannten „Mainstream Pop“. Das Mainstream Spektrum wird zwar oft referenziert, ist aber nicht sehr breit. Hingegen ist das Spektrum selten gekaufter und gehörter Musik deutlich breiter. Kollaborative Filterung deckt also nur einen Teil der Hörer ab, wohingegen ein beachtlich großer Teil auf der Strecke bleibt. Folglich scheitert in diesem Fall das Prinzip der kollaborativen Filterung und es werden keine treffenden Ergebnisse ermittelt.

4 Kombination von Kollaboration und Inhalt

Wie eingangs erwähnt ist es möglich durch eine Kombination von Inhaltsbasierter Suche und Kollaborations-Basierter Suche die Nachteile beider Filterungsverfahren auf ein Minimum zu begrenzen, aber dennoch die Vorteile beider Techniken zu nutzen.

Es werden also sowohl inhaltliche, als auch kollaborative Kriterien auf die Suchanfrage angewendet. Dabei ist das Maß, wie stark die jeweiligen Anteile auf die letztendliche Empfehlung einfließen noch zu bestimmen. Dies stellt eine Schwierigkeit dar, die aber, verglichen mit den Problemen beider Methoden, eine überschaubare Fehlerquelle ist. Ein oft verwendeter Ansatz ist die Kombination von kollaborativer Filterung mit persönlichen Agenten, auch personal Agents genannt.

4.1 Kombination von kollaborativer Filterung und Personal Agents

Der Ansatz ist den kollaborativen Algorithmus um eine Inhaltsbasierte Komponente zu erweitern. Idee dabei ist es, je weniger Rezensionen über ein Suchitem bestehen, desto mehr kommt der Inhaltsbasierte Aspekt zum tragen.

Wir betrachten also die Schwächen von kollaborativen Filterungstechniken und die Möglichkeit diese zu beheben. Eine Möglichkeit ist der Einsatz von Agenten, die bestimmte Kriterien wie Rhythmus, Klangfarbe, etc, aber auch Metadaten wie Land, Sprache, usw. spezifizieren können. Diese fließen dann in das kollaborative Suchergebnis mit ein.

Dies hätte den Effekt der Lösung des Kaltstart Problems, da bereits Empfehlungen gemacht werden können, auch wenn noch keine Userwertungen vorhanden sind.

Auf diese Weise würden auch das New Item Problem, sowie das Problem der Spärlichkeit der Rezension gelöst.

User mit außergewöhnlichem Musikgeschmack würden bei dieser Methode ebenfalls mit Empfehlungen bedient werden, wodurch die implizit selbst wieder Empfehlungen abgeben über jene Musik, und so diese auch kollaborativ referenziert werden kann.

Das Black-Box Problem würde so ebenfalls unterbunden, da eine Willkür bei wenigen Ratings durch Inhaltsbasierung unterbunden wird.

Abbildung 6 zeigt eine Darstellung der verschiedenen Formen von Filterungsmethoden. Während jene links-oben allein auf kollaborative Filterung setzt, und die beiden unten ausschließlich auf Agenten, also auf eine Inhaltsbasierte Lösung, findet man in der Lupe rechts-oben eine Kombination aus Agenten und Kollaboration.

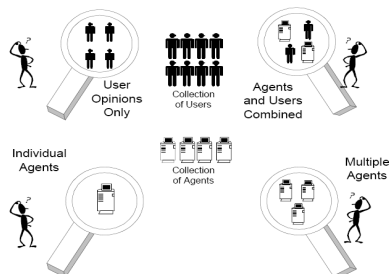


Abbildung 6. Vier Typen von Recommender System. Die Abbildung links-oben verwendet User Empfehlungen, während die Abbildungen unten nur auf Agenten basieren. Die Kombination davon ist die Abbildung rechts-oben.

4.2 Beispiele für inhalts- und kollaborativ kombinierte Systeme

- Liveplasma ist ein Music Recommender System, welches hauptsächlich auf Kollaborations-Basierte Filterung baut, jedoch auch Content-Basierte Filterung verwendet[19]. Der Ablauf ist folgender: Der User gibt in das Suchfeld eine Band ein, die er gerne hört. Darauf hin ermittelt Liveplasma eine Karte mit diversen Bands, die unterschiedlich nahe zu der Suchanfrage angeordnet sind. Dabei ist die Wahrscheinlichkeit bei jenen Bands die näher dem Item sind höher, das sie den Benutzer interessieren, als bei Items die weiter weg von der Suchanfrage liegen. Die Größe der Kugeln sagt aus wie populär die jeweilige Band ist, bzw. wie oft sie referenziert wird. Neben den kollaborativen Kriterien sind die Daten auch nach Metadaten gruppiert, wie beispielsweise Stil, Epoche, etc.

Abbildung 7 zeigt einen Screenshot des Systems, nachdem eine Anfrage getätigt wurde und die entsprechenden Referenzitems ermittelt wurden. Im speziellen Fall wurde nach der britischen Band „The Verve“ referenziert.

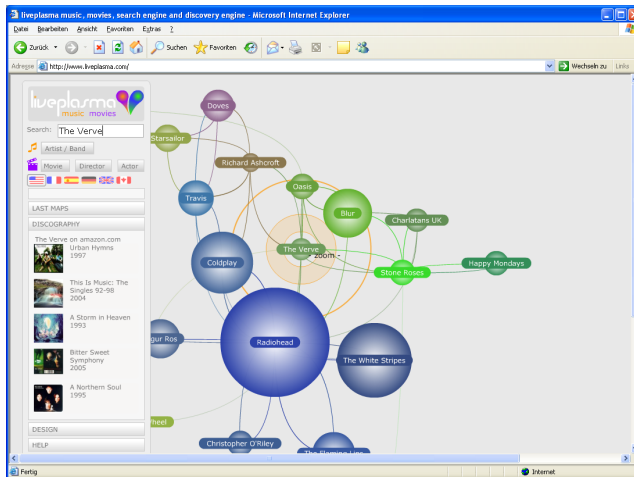


Abbildung 7. Das Music Recommender System LivePlasma, nachdem eine Anfrage getätigt wurde. Je näher ein Item an dem Referenzitem angeordnet ist, desto größer ist die Wahrscheinlichkeit, dass es den Benutzer interessiert. Im speziellen Fall wurde die Band „The Verve“ referenziert und als nächste Nachbarn die Bands Blur, Oasis, Radiohead und Coldplay ermittelt.

- Pandora, ein früher reines kontent-basiertes System hat, hat nun ebenfalls Kollaborations-Basierte Filterung mit ins Konzept aufgenommen. Das Prinzip ist folgendes: Nach Eingabe einer Band oder eines Liedes wird ein persönlicher Radiosender erstellt. Die Idee von Pandora ist Lieder nicht nach Genre, Land oder anderen Metadaten zu empfehlen, sondern rein nach Analyse des Liedes, also (ursprünglich) ein typisches MIR-System. Dabei spielen Attribute wie Melodie, Harmonie, Rhythmus, Instrumente, Arrangement, etc eine Rolle. Dies klingt nach einem rein inhaltsbezogenen Ansatz, der anfangs auch alleine so verwendet wurde. Pandora räumte jedoch ein, dass es eine Breite von Songs gibt, die nicht von musikspezifischen Kriterien erfasst werden können, also nicht durch Rhythmus, Melodie oder Instrumente. Diese Lieder sollen durch die neu eingesetzte kollaborative Filterung ebenfalls gefunden und empfohlen werden. Zusätzlich bietet Pandora jetzt auch die Möglichkeit jeden Song zu bewerten und so zu verhindern, das, im Falle einer negativen Bewertung, der gleiche Song erneut in das „persönliche Radio“ aufgenommen wird. Gleichzeitig wird auch der Schluss gezogen, wenn viele User eine Band oder ein Lied negativ bewerten, wird dieses nicht mehr so oft in „persönlichen Radiostationen“ erscheinen.

Abbildung 8 zeigt einen Screenshot des Systems, dabei stellt jedes Rechteck innerhalb des „Radios“ ein Lied dar. Das gerade abgespielte wird eingblendet, während der Rest ausgeblendet erscheint.

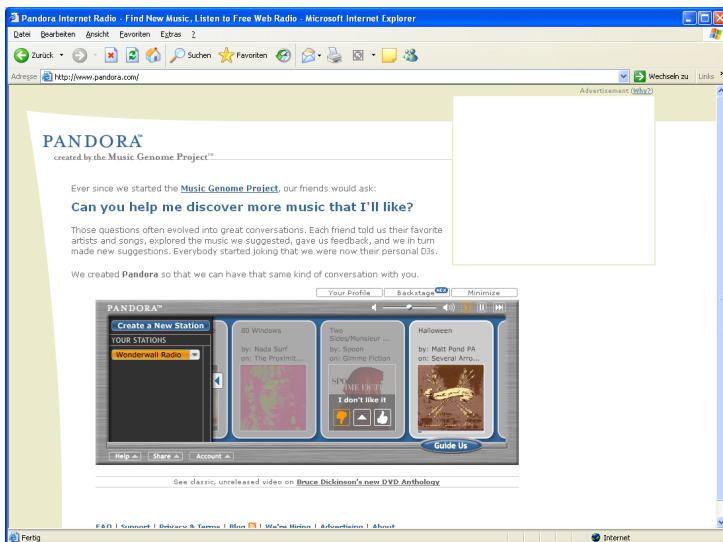


Abbildung 8. Das Music Recommender System Pandora, nachdem ein „persönlicher Radiosender mittels einer Anfrage erstellt wurde.

5 Zusammenfassung

Es wurde Inhaltsbasierte und kollaborative Filterung besprochen. Dabei lässt sich die Inhaltsbasierte Filterung in Systeme aufteilen, die entweder auf Metadaten oder auf „Music Information Retrieval“ basieren. Eine Kombination beider Ansätze kann verwendet werden, da sie sich nicht gegenseitig ausschließen. Hingegen wird bei kollaborativer Filterung zwischen Systemen unterschieden, die explizites Rating oder implizites Rating verwenden. Auch hier ist eine Kombination möglich, wie beispielsweise Amazon sie verwendet.

Technisch gesehen bauen die meisten kollaborativen Systeme auf „nearest-neighbour“ Algorithmen, während ein häufig eingesetzter Algorithmus zur Analyse von Songs in MIR-Systemen die Fast-Fourier Transformation ist.

Bei der Diskussion von Inhaltsbasierter gegenüber kollaborativer Filterung muss abgewägt werden, wer solch ein System verwendet. Während für User, welche zu so genannter „Mainstreammusik“ tendieren eindeutig das kollaborative System die besseren Empfehlungen abgibt, kommen Benutzer mit außergewöhnlichem Musikgeschmack bei solchen Systemen zu kurz und werden durch Inhaltsbasierte Systeme besser bedient. Zusätzlich resultieren zahlreiche Probleme des kollaborativen Filterns daher, dass es keine oder zu wenig Rezensionen über ein Item gibt. Probleme die das Inhaltsbasierte Filtern betreffen sind unter anderem die Überspezialisierung und das Problem, das nicht alle Items gefunden werden, die interessieren, sondern nur jene die von den Attributen dem Anfrageitem, bzw. dem Userprofil entsprechen.

Die Kombination von Inhaltsbasierung und Kollaborationsbasierung kann viele Probleme lösen, schafft jedoch das Problem in welchem Maß welche Filtermethode angewandt wird.

Um diese Kombination zu realisieren werden häufig Agentensysteme eingesetzt, die die kollaborative Filterung ergänzen. Im Fall von Pandora wird jedoch das Inhaltsbasierte Ergebnis um eine kollaborative Komponente ergänzt.

Grundsätzlich ist zu sagen, dass keine von beiden Methoden alleine ausreicht um ein Music Recommender System zu entwickeln, welches alle User ausreichend bedient. Hier muss eine Kombination erfolgen. In diese Richtung wird der Trend in nächster Zeit gehen.

Ebenfalls zu erwähnen ist, dass die Zahl der User, die Music Recommender Systeme in Zukunft verwenden werden steigen wird, was die Bedeutung von kollaborativen Systemen stärken wird.

Literatur

1. Marko Balabanovic, Yoav Shoham. Fab: Content-based, Collaborative Recommendation
2. Otmar Hilliges, Phillip Holzer, Rene Klüber, Andreas Butz. AudioRadar: A metaphorical visualization for the navigation of large music collections
3. Sabine Kuzdas. Schwachstellen von Recommender Systemen
4. Badrul M., Sarwar, Joseph A. Konstan, Al Borchers, Jon Herlocker, Brad Miller and John Riedl. Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System
5. Julian Hartmann. Music Information Retrieval, Automatische Genre-Klassifikation
6. CCRMA Department of Music Stanford University; Mathematics with the Discrete Fourier Transform (DFT) with music and audio applications, 1996-2003 CCRMA, URL: <http://ccrma.stanford.edu/~jos/mdft>
7. Loren Terveen and Will Hill. Beyond Recommender Systems: Helping People Help Each Other
8. Das Music Recommender System MusicSurfer. <http://musicsurfer.iaa.upf.edu/>
9. Marko Balabanović. Recommender systems: Interfaces and Architectures
10. David Goldberg, David Nichols, Brian M. Oki, Douglas Terry. Using collaborative filtering to weave an information tapestry. In: Communications of the ACM archive, Volume 35, Issue 12, S. 61 - 70. 1992
11. Upendra Shardanand. Social Information Filtering For Music Recommendation, Sep 1994
12. David M. Nichols, Lancaster University. Implicit Rating and Filtering
13. Barry Brown¹, Abigail J. Sellen, Erik Geelhoed. Music Sharing as a Computer Supported Collaborative Application
14. Alexandra Uitdenbogerd and Ron van Schyndel. A Review of Factors Affecting Music Recommender Success
15. Deutsche Website von dem amerikanischen Konzern Amazon. <http://www.amazon.de>
16. Lee Becker, David Brown, Mark Claypool and Carolina Ruiz. Research in Recommender Systems at Worcester Polytechnic Institute. <http://www.cs.wpi.edu/Research/refer>
17. Ken Goldberg, Theresa Roeder, Dhruv Gupta, and Chris Perkins. Eigentaste: A Constant Time Collaborative Filtering Algorithm. <http://eigentaste.berkeley.edu/>
18. Nathaniel Good, J. Ben Schafer, Joseph A. Konstan, Al Borchers, Badrul Sarwar, Jon Herlocker, and John Riedl. Combining Collaborative Filtering with Personal Agents for Better Recommendations
19. Das Music Recommender System Liveplasma. <http://www.liveplasma.com>
20. Das Music Recommender System Pandora. <http://www.pandora.com>
21. Die online Enzyklopädie Wikipedia. <http://en.wikipedia.org>
22. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl, J. (1997), Applying collaborative filtering to usenet news, Communications of the ACM
23. Stuart E. Middleton, Intelligence, Agents and Multimedia group (IAM group), Interface Agents, a review of the field. <http://www.ecs.soton.ac.uk/~sem99r/>
24. Speedfacts GmbH, Frankfurt. <http://www.speedfacts.com>
25. Website die diverse MIR-Systeme auflistet und beschreibt. <http://mirsystems.info/>

Mobile communities and friend finders

Benjamin Rabinowitsch

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
ben@rabinowitsch.de

Zusammenfassung Diese Ausarbeitung zeigt eine Reihe von Aspekten, die bei dem Aufbau und der Bereitstellung von mobilen Gemeinschaften und mobilen Freundsuch-Diensten Beachtung finden sollten. Da diese mobilen Applikationen oftmals ortsgebundene Features anbieten sind hierbei technische Hintergründe von Location-based Services genauer zu betrachten. Neben bereits bestehenden Diensten, wie dem Angebot von Nokia namens Sensor und anderen Beispielen, die im Laufe dieses Textes noch genauer betrachtet werden, bietet dieser Bereich eine breites Spektrum an potentiellen kommerziellen und frei verfügbaren Möglichkeiten. Der aktuelle Stand wird also genauso beleuchtet, wie Zukunftsgewandte Modelle. Ebenfalls nicht ungeachtet gelassen werden sollen die sozialen Aspekte, die mobile Gemeinschaften und Freund-Finder innerhalb der heutigen Gesellschaft bieten.

1 Einleitung

Menschen sind soziale Geschöpfe, die ein natürliches Verlangen nach Kommunikation und Interaktion mit anderen Artgenossen haben. Deshalb bilden Menschen bereits in früher Kindheit Gemeinschaften.

Eine Gemeinschaft zeichnet sich dadurch aus, dass in ihr Individuen vertreten sind, die gemeinsame Interessen haben und die gleichen Vorstellungen vertreten. Innerhalb der Gemeinschaft trifft man sich und tauscht nützliche oder interessante Informationen aus.

Damit eine Gemeinschaft funktioniert, ist eine gewisse Anzahl von aktiven Teilnehmern nötig. Nur wenn eine bestimmte Prozentzahl an aktiven Mitgliedern erreicht wird, bekommt die Gemeinschaft ausreichend viele Impulse, um weiterhin interessant für sämtliche Mitglieder zu sein. [1]

In den letzten Jahren haben sich zahlreiche virtuelle Gemeinschaften entwickelt, die im Internet vertreten sind und dort in hohem Maße genutzt werden. Diese Virtualisierung bringt den Vorteil mit sich, dass Gruppenmitglieder nicht weiter an Ort und Zeit gebunden sind, wenn sie untereinander kommunizieren möchten. Die einzelnen Mitglieder können sich jederzeit mit Hilfe des Internets in die Gemeinschaft einloggen und dort neu erstellte Informationen abrufen, bearbeiten und auf sie eingehen oder selbst neue Beiträge erstellen. [2]

Werden diese virtuellen Gemeinschaften nun um den mobilen Aspekt erweitert, bietet sich ihnen die Möglichkeit, weitere Features anzubieten und zu nutzen. Zur

Nutzung dieser erweiterten Möglichkeiten sind lediglich ein mobiles Endgerät und ein Zugang zur Gemeinschaft notwendig. Über mobile Endgeräte verfügt heutzutage nahezu jeder. Die Verbreitung mobiler Telefone lag im Jahr 2005 bei Jugendlichen zwischen 13 und 19 Jahren bei 97%. [3]

In dieser Ausarbeitung sollen neben den mobilen Gemeinschaften auch erweiterte Aspekte wie Friend Finder und je nach Nutzungsart mögliche technische Hintergründe näher beleuchtet werden. Die mobilen Freundsuch-Angebote können entweder bereits eingetragene Freunde verwalten und anzeigen oder Personen nach bestimmten Kriterien als potentiellen Freund ausmachen.

Oftmals wird ein mobiler Dienst, oder eine bestimmte Information bzw. ein Feature innerhalb dieses Dienstes, nur ortsabhängig oder nur zu einer bestimmten Zeit an diesem Ort angeboten. Dann spricht man von Location-based Services, deren Funktionsweisen und Möglichkeiten im späteren Verlauf im Zusammenhang mit mobilen Gemeinschaften näher betrachtet werden.

2 Grundlagen

2.1 Mobile Communities

Es gibt eine Reihe von Aktivitäten, die typisch für Gemeinschaften sind. Diese werden weitestgehend auch von virtuellen und mobilen Zusammenschlüssen übernommen. Im Vordergrund steht auf allen drei Stufen der Gemeinschaften die Kommunikation mit anderen Teilnehmern einer Gruppe.

Die Kommunikation innerhalb einer Gemeinschaft kann auf eine Reihe unterschiedlicher Arten vollzogen werden. Kommunikation kann direkt sein, d.h. man sendet einer bestimmten Person eine Nachricht, wie z.B. per E-Mail, SMS oder innerhalb eines Chats. Sie kann aber auch indirekt von statten gehen. Hierbei wird eine Nachricht hinterlegt und jeder, der diese lesen möchte und auch – falls notwendig – die Erlaubnis dazu hat, kann sie abrufen und bei Bedarf darauf eingehen.

Neben der direkten und indirekten Kommunikation kann man beispielsweise auch zwischen synchroner und asynchroner, sowie zwischen automatisch und manuell angestoßener und vielen weiteren Arten der Kommunikation unterscheiden. [5]

Die Kommunikation innerhalb einer Gemeinschaft dient verschiedenen Zwecken. Sie kann zum einfachen Informationsaustausch genutzt werden, wie auch zur gegenseitigen Hilfe bei Problemen. Eine Gemeinschaft kann aber auch dazu dienen, direkte Interaktion zu Menschen aufzubauen, mit denen wir bisher keinen Kontakt hatten. So werden beispielsweise Kontakte zwischen Personen initiiert, die ein bestimmtes Interessensgebiet teilen oder gleiche Hobbys haben. [4]

Dieser Aufbau einer Verbindung zweier sich bisher unbekannter Menschen ist auch eng mit dem Funktionsmodell des Friend Finder verbunden.

2.2 Friend Finder

Es gibt verschiedene Arten von Freundsuch-Angeboten. Neben der Unterscheidung zwischen mobilen und nicht mobilen Angeboten muss man auch zwischen Angeboten differenzieren, die lediglich der Verwaltung von bestehenden Freundschaften dienen, und solchen, die die Möglichkeit bieten, neue Leute auf Grund verschiedener Aspekte kennen zu lernen.

Mobile Friend Finder können z.B. auf dem mobilen Endgerät anzeigen, wenn ein Bekannter einen bestimmten Radius zu mir betritt. Mit Hilfe dieser Applikation kann man dann mit dieser Person in Kontakt treten und beispielsweise ein reales Treffen ausmachen. Mit einem derartigen Feature dient die Mobilität des Endgeräts dazu, die virtuelle Gemeinschaft in die reale Welt zu verlagern.

Des Weiteren ist es mit einem derartigen Freundsuch-Angebot auch möglich, mir bislang völlig fremde Personen anzeigen zu lassen, wenn die entsprechende Applikation sowohl auf meinem, als auch auf dem Endgerät einer sich in meiner Nähe befindlichen Person vorhanden ist. In diesem Fall errechnet die Applikation, dass ein gewisses Maß an Überschneidungen zwischen meinen Interessen und Hobbys und denen der fremden Person vorhanden sind. [4] Auch dieser Dienst bietet einen direkten Übergang von der virtuellen in die reale Welt.

Sowohl die mobilen Gemeinschaften als auch die mobilen Freundsuch-Applikationen können ortsunabhängig genutzt werden, bieten aber besondere Features, wenn bestimmte Ereignisse innerhalb einer kurzen Entfernung zu meinem mobilen Endgerät geschehen.

Erweiterte Möglichkeiten können angeboten werden, wenn diese Dienste innerhalb eines bestimmten Kontextes benutzt werden. So bieten ortsbasierte Dienste spezielle Aspekte an, die innerhalb eines bestimmten Bereichs zum Tragen kommen.

2.3 Location-based Services

„Standortbezogene Dienste sind über ein Netzwerk erbrachte mobile Dienste, die unter Zuhilfenahme von positions-, zeit- und personenabhängigen Daten dem Endbenutzer selektive Informationen bereitstellen oder Dienste anderer Art erbringen.“[5]

Hierbei unterscheidet man zwischen so genannten reaktiven und proaktiven Diensten. Bei den reaktiven Diensten muss der Nutzer die gewünschte Information explizit anfragen, wohingegen er durch das betreten eines bestimmten Bereichs innerhalb des Location-based Services (LBS) einen proaktiven Dienst auslöst.

Um einen ortsabhängigen Dienst anbieten zu können, benötigt man die Möglichkeit, mobile Endgeräte innerhalb des Zielgebiets möglichst genau lokalisieren zu können.

Unter Lokalisation innerhalb eines LBS versteht man das Messen bestimmter Abstände und Winkel des Endgeräts zu bereits festgelegten Punkten. Das System errechnet dann aus den relativen Daten die absoluten Koordinaten des Nutzers. Findet diese Berechnung unter zu Hilfenahme des mobilen Endgerätes statt, so spricht man von Lokalisation.

Dem gegenüber steht der Begriff Tracking. Dieser besagt, dass sich das mobile Endgerät nicht am Prozess der Standortbestimmung beteiligt. Somit kann das Gerät auch nicht beeinflussen, welcher Aufenthaltsort berechnet wird. Wird ein Tracking

durchgeführt, so wird die Ortsinformation über eines oder mehrere Endgeräte zentral abgespeichert und kann zur Überwachung genutzt werden.

2.3.1 Reichweite von LBS

Je nach Anwendung kann die Reichweite eines LBS angepasst werden. Abhängig davon, wie groß die Reichweite sein soll, verwendet man verschiedene technische Lösungen. Welche Lösungen man je nach Reichweite des Dienstes nutzen kann, wird in Abschnitt 2.3.2 gezeigt.

Möchte man ein sehr kleines Gebiet mit seinem ortsgebundenen Dienst abdecken, so benötigt man eine andere Infrastruktur, als dies für ein wesentlich weitläufigeres Areal von Nöten sein kann.

Ein kleiner LBS kann sich auf lediglich wenige Räume innerhalb eines Gebäudes beschränken, oder er umfasst einen ganzen Campus, wie wir es in einem späteren Beispiel sehen werden. Größere ortsabhängige Dienste können bereits eine komplette Stadt abdecken oder sogar ein ganzes Land, wie es in Deutschland seit Einführung der LKW-Maut auf Autobahnen bekannt ist. Hierfür ist natürlich ein besonders hoher Aufwand erforderlich gewesen, welcher mit großen Kosten verbunden war.

2.3.2 Technische Lösungen

Zur Erstellung eines LBS können eine Reihe von technischen Mitteln benutzt werden. Jede Technik bietet dem Betreiber Vor- und Nachteile. Man muss hierbei zwischen Kosten und Nutzen abwägen. Im Folgenden werden einige Infrastrukturen betrachtet, die für die Errichtung und Aufrechterhaltung eines Netzwerks für ortsbasierte Dienste.

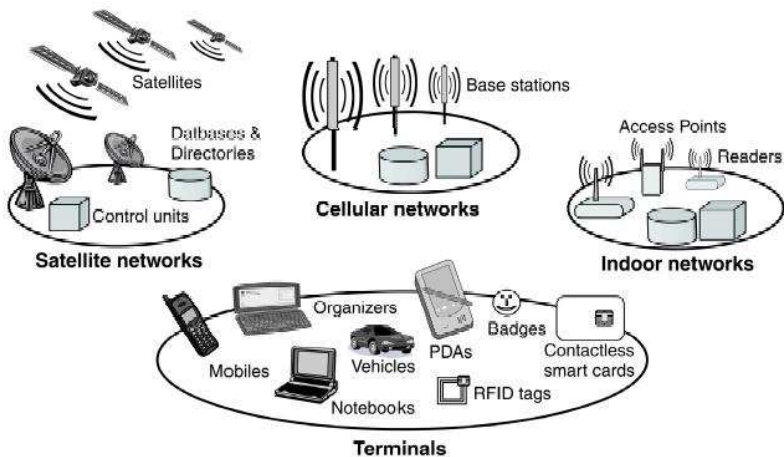


Abbildung 1. Verschiedene Infrastrukturen, die ortsabhängige Dienste ermöglichen [6]

Auf Abbildung 1 sind die drei verschiedenen Arten von Techniken zu erkennen, die zur Nutzung von ortsabhängigen Diensten eingesetzt werden.

Es handelt sich hierbei zum einen, links oben zu sehen, um Netzwerke, die satellitengestützt arbeiten. Diese bieten die höchstmögliche Reichweite an und decken je nach Dienst den gesamten Globus ab, wie dies beim Global Positioning System (GPS) der Fall ist.

Die zweite Gruppe, die in Abbildung 1 mittig oben platziert zu sehen ist, sind die zellularen Netze. Sie arbeiten mit Basisstationen, die eine Reichweite von wenigen hundert bis zu einigen Kilometern haben. Die Reichweite hängt davon ab, wie viele Nutzer innerhalb einer solchen Funkzelle gleichzeitig Daten übertragen wollen. Je mehr Nutzer den Dienst in Anspruch nehmen wollen, desto geringer ist die Reichweite der einzelnen Station. Der Grund hierfür liegt in der gemeinsamen Nutzung der Luftschnittstelle aller User, die über dieselbe Basisstation kommunizieren.

Die dritte Art von Netzwerk ermöglicht die Ortserkennung innerhalb von Gebäuden. Hierbei können z.B. drahtlose Netzwerke genutzt werden, für die man so genannte Access Points braucht. Diese bieten eine Abdeckung von einigen Metern und ermöglichen eine relativ genaue Positionierung innerhalb eines Gebäudes.

Eine Technologie mit der dies möglich ist, wird Fingerprinting genannt. Hierbei wird eine Vielzahl von Messungen der Signalstärke der einzelnen Access Points an verschiedenen Punkten innerhalb des Gebäudes durchgeführt. Jedem Ort wird in einer Datenbank ein bestimmter Fingerabdruck, bestehend aus den Signalstärken der einzelnen Access Points, zugewiesen.

Ebenfalls für die Erkennung innerhalb von Gebäuden ist die Nutzung von RFID-Etiketten und –Lesern sinnvoll. Hierbei sind im gesamten Gebäude Sender oder Empfänger angebracht und eine Person, die innerhalb des Raums geortet werden soll, trägt das jeweilige Gegenstück dazu mit sich.

Durch ihre kurze Reichweite bietet diese Technologie im Vergleich zum WLAN-Fingerprinting, die Ungenauigkeiten von einigen Metern besitzt, eine genauere Lokalisierung, ist dafür aber im Gegensatz zu WLAN-Netzen mit der Ausstattung einer zusätzlichen Infrastruktur verbunden.

Dafür ist der administrative Aufwand bei der Nutzung von RFID-Etiketten sehr gering. Einmal installiert sind diese zuverlässig und bieten über einen sehr langen Zeitraum eine präzise Ortung. WLAN-Fingerprinting hingegen benötigt immer wieder neue Messungen, sobald ein Teil der Infrastruktur ausgewechselt, die Sendeleistung verändert oder ein massiver Gegenstand innerhalb des Gebäudes anders platziert wird.

Die Vielzahl an Empfängern deutet der untere Teil der Abbildung an. [5]

2.4 Überschneidungen

Nicht bei allen Applikationen kann man eine strikte Trennung von mobiler Gemeinschaft, mobilem Freundschaft-Angebot und ortsbasiertem Dienst vornehmen.

Die meisten Applikationen sind Mischformen aus zwei dieser Dienstarten oder sogar aus allen drei (Siehe Abbildung 2). Mobile Gemeinschaften bieten oft die Möglichkeit, bisher fremde Menschen kennen zu lernen. Seien es Menschen, die bisher in keiner Form in der Gemeinschaft aufgetaucht sind, oder welche, die zwar bereits in der Gemeinschaft integriert sind, aber nicht zu seinen eigenen Freunden zählen. Lernt

man diese Leute mit Hilfe der Gemeinschaft kennen und freundet sich an, hat die Gemeinschaft auch als Freundsuch-Dienst agiert.

In dieser Ausarbeitung und in später gezeigten Beispielen möchte ich in erster Linie auf die mobilen Gemeinschaften und Freundsuch-Angebote eingehen. Diese weisen untereinander starke Überschneidungen auf, aber auch mit ortsabhängigen Diensten. So benötigt z.B. der Dienst Active Campus, der eine mobile Gemeinschaft anbietet, eine Infrastruktur, um aktuelle Informationen auf sämtliche in der Gemeinschaft vorhandene Endgeräte schicken zu können. Welche Informationen gesendet werden hängen allerdings vom Standort jeder einzelnen Person ab. Dies macht den LBS-Aspekt innerhalb von Active Campus aus. [8]

Ein Dienst wie Nokia Sensor hingegen benötigt keinerlei Infrastruktur. Bei diesem Angebot überschneiden sich, wie später zu sehen sein wird, mobile Gemeinschaft und Freundsuch-Angebot. [10]

Solche Überschneidungen, wie in Abbildung 2 vereinfacht dargestellt, findet man zwischen LBS, mobiler Gemeinschaft und Freundsuch-Dienst sehr häufig.

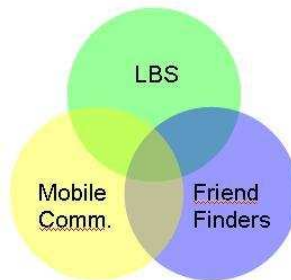


Abbildung 2. Überschneidungen zwischen LBS, mobiler Gemeinschaft und Freundsuch-Angebot [11]

3 Beispiele

3.1 Friend Zone

Ein Beispiel für eine mobile Gemeinschaft, die den Aspekt eines ortsgebundenen Dienstes nutzt, ist FriendZone. Diese Applikation dient als Instant Messenger für mobile Endgeräte und unterstützt zusätzlich das Feature, ortsnahe Personen anzeigen zu können, wenn diese die gleiche Software auf ihrem mobilen Endgerät benutzt und beide gleiche Interessen oder Hobbys haben.

Bezahlt werden musste der Dienst über die normale Rechnung des Mobilfunkanbieters.

Die Möglichkeit, mit einem in der Nähe befindlichen Menschen per Chat zu kommunizieren, dabei aber dennoch anonym zu bleiben, erfreute sich hoher Beliebtheit. Hierbei bestand sowohl die Chance, mit einer Person privat zu kommunizieren, als

auch mit mehreren Teilnehmern zu chatten, die sich allerdings alle in der gleichen Zone befanden.

Bei einer Studie zu dieser Applikation, die im ersten Quartal 2001 in der Schweiz stattfand und an der 600 Personen teilnahmen ergab, dass nahezu 70% der Nutzer besonders die Option attraktiv, mit jemandem per mobilem Endgerät zu chatten, ohne seine Identität preisgeben zu müssen, wie es bei SMS-Nachrichten der Fall ist. [7]

Die Ortsbestimmung innerhalb dieses Dienstes ging ausschließlich vom Netzbetreiber aus, was den Vorteil mit sich brachte, keine neue Infrastruktur errichten zu müssen.

Der Nachteil bestand jedoch darin, dass die Zonen recht weite Flächen umfassten, wodurch eine relativ genaue Lokalisierung unmöglich war.

3.2 Active Campus

Im Gegensatz zur Friend Zone ist Active Campus ein Location-based Service für mobile Gemeinschaften, der sich nicht ausschließlich der Freizeit widmet, sondern das Zusammenleben, aber vor allem auch -studieren vereinfachen soll. [8]

Entwickelt wurde Active Campus, da sich die Universität von San Diego innerhalb kürzester Zeit von einer sehr geringen Zahl studierenden zu einer sehr großen Universität entwickelte. Das schnelle Wachstum führte auch zu einem Ausbau des Campus-Geländes, was die Orientierung für neue Studenten stets verkomplizierte. Zudem war es ein Ziel der Universität, die Studierenden in einem Netzwerk zusammenzufassen, um das Potenzial der einzelnen Studenten erkennen und fördern zu können.

Im Rahmen dieses Projekts wurden 500 WLAN-fähige PDA an studierende verteilt. Dank des Campus-weiten WLAN-Zugangs, sind diese Studierenden in der Lage, wissenschaftliche Fragen zu erforschen und die Lehrgemeinschaft durch die mobilen Endgeräte aufrecht zu erhalten.

Die Applikation, die jeder Student auf seinem PDA hat, ermöglicht es ihm, sich leicht auf dem gesamten Areal zurechtzufinden und Freunde, Kommilitonen, Büros und Forschungseinheiten schnell und einfach ausfindig zu machen und über aktuelle Ereignisse stets informiert zu sein. Wie die Benutzeroberfläche aussieht, ist auf Abbildung 1 zu erkennen.



Abbildung 3. Benutzeroberfläche der Active Campus-Applikation [9]

Auf der linken Seite von Abbildung 3 erkennt man eine schriftliche Übersicht über in der Nähe befindliche Personen und Einrichtungen auf dem Campus-Gelände. Auf der rechten Seite der Abbildung sieht man einen Kartenausschnitt des Geländes auf dem man sich befindet mit allen Gebäuden und relevanten Veranstaltungen. Zu jedem Ereignis findet man einen vertiefenden Text, zu den man per Link gelangt,

Probleme, die mit der Implementierung dieser Applikation einhergegangen sind und die bei solchen Anwendungen immer wieder auftreten, sind das Unterbringen aller relevanten Informationen auf dem kleinen Display des mobilen Endgerätes, sodass sie dennoch leicht erfassbar für den Nutzer bleiben und die geforderte minimale Ressourcennutzung der Anwendung, um dem Anwender möglichst viel Rechenkapazität für seine primären, studiumsrelevanten Aufgaben zu überlassen. Außerdem muss eine derartige Applikation stets Erweiterbar sein, da sich das gesamte Umfeld ebenfalls weiterentwickelt.

Eine solche Anwendung soll dem Studenten helfen, sich schnell an die neue Umgebung zu akklimatisieren und bestehende Potentiale der Studierenden und des Campus selbst zu beschleunigen.

Das Gerät selbst hat die bereits angedeutete Schwäche der geringen Rechenleistung und beschränkten Visualisierungsmöglichkeit. Dem gegenüber steht ihre Mobilität und Unaufdringlichkeit dem Nutzer gegenüber. [8]

Um die Rechenleistung des PDA nicht zu sehr zu strapazieren, wurde eine Client-Server-Lösung implementiert, die es ermöglicht, möglichst viel Rechenleistung auf den Server auszulagern, damit die primären Prozesse des Studenten nicht in Mitleidenschaft gezogen werden.

Das Projekt zeigte einige positive Aspekte auf, die für spätere Applikationen in diesem Bereich gut übernommen werden können. Andere Dinge stellen noch Probleme dar, die ausgeräumt werden müssen. Die Frage ist, wer dafür sorgt, dass die Informationsflut nicht so groß wird, dass sie ein einzelner Endnutzer nicht mehr filtern kann und wenn es eine dritte Person macht, ob man eine mögliche Zensur verhindern kann. Die Technischen Schwierigkeiten waren bei diesem Projekt nur ein sehr geringes Problem. Überhaupt kein Problem stellte die Verbindung der Endgeräte mit dem ortsabhängigen Dienst dar, viel mehr waren es die Geräte selbst, die nur bedingt gute Nutzbarkeit boten. Dennoch war das Projekt ein Erfolg, der in künftigen Diensten weiter ausgebaut werden kann, wenn man die Usability der Endgeräte erhöhen und die Menge an Informationen sinnvoll steuern kann.

3.3 Live Adressbook

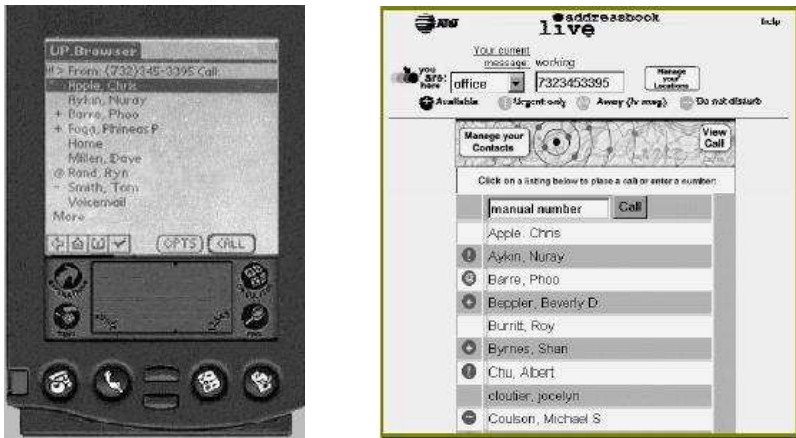


Abbildung 4. Einträge im Live-Adressbuch, links auf einem mobilen Endgerät, rechts in einem Web-Browser [9]

Das Live Adressbuch ist eine Applikation für mobile Endgeräte, die dem Nutzer erweiterte Informationen über seine Kontakte liefern kann, wie in Abbildung 4 zu sehen ist. So können störende Anrufe verhindert werden oder Informationen über den Aufenthaltsort des Anzurufenden eingeholt werden, bevor der Anruf getätigt wird. [8]

Dieses Live Adressbuch ist ein netzwerkbasierendes Adressbuch, das dem Anwender neben statischen auch dynamische Informationen über seine Kontakte bietet. Neben den eigenen Kontakten können in diesem Adressbuch auch Informationen über sämtliche andere Nutzer dieses Dienstes eingesehen werden. Die Voraussetzung hierfür ist allerdings, dass der Nutzer den Namen und die Telefonnummer eines eingetragenen kennt. Ist dies der Fall, bekommt der Anwender z.B. angezeigt, wo sich die andere Person gerade aufhält und welchen Status sie angegeben hat, also ob sie verfügbar ist oder nicht.

Dass man den Namen und die Telefonnummer wissen muss, hat den Vorteil, dass nicht jeder Zugriff auf persönliche Daten hat, was aus Datenschutz-Gründen sehr zu begrüßen ist.

Ein gebotenes Feature durch diese mobile Applikation ist die so genannte „Click-To-Dial“-Funktion. Dank dieser klickt man einfach auf die anzurufende Person und eine Verbindungsbrücke zwischen dem Anrufer und dem Anzurufenden wird aufgebaut. So kann man mit einem anderen Endgerät das Gespräch initiieren, als man es führt. Das Live Adressbuch erkennt anhand des momentanen Aufenthaltsorts, wo sich beide Personen aufhalten. Zudem bietet das Live Adressbuch die Möglichkeit aus einem normalen Gespräch eine Telefonkonferenz zu erzeugen, indem man z.B. einen eingehenden Anruf einfach zu den aktiven Gesprächen hinzufügt.

Als Nutzer des Live Adressbook hat man die Möglichkeit drei verschiedene persönliche Status-Informationen zu editieren. Ähnlich wie bei Online-Messangern lassen sich im Punkt Verfügbarkeit verschiedene Optionen auswählen. Hier gibt der Nutzer an, ob er grundsätzlich oder nur in dringenden Fällen erreichbar ist oder ob er nicht erreichbar ist. Diese eingestellte Information haben alle anderen Live Adressbook-Nutzer dann in ihrer mobilen Anwendung und können dementsprechend einen Anruf starten. Des Weiteren Stellt der User ein, an welchem Ort er sich gerade befindet. So kann beim Click-And-Dial z.B. der Anruf genau an den Standort gestellt werden, den man bereits zu einem früheren Zeitpunkt eingestellt und gerade aktiviert hat. Außerdem bietet es die Möglichkeit geschäftliche Anrufe nicht zuzulassen, wenn man sich gerade zu Hause befindet. Die dritte modifizierbare Variable ist eine Nachricht, die man optional in Form eines Textes hinzufügen kann, in der beispielsweise nähere Erläuterungen zum derzeitigen Status gemacht werden können. [8]

Durch diese Applikation können Adressbuch-Einträge stets aktuell gehalten werden, obwohl man sie nicht zu modifizieren braucht. Dies geschieht durch den Eintrag dessen, bei dem sich etwas ändert, automatisch.

Der Aufenthaltsort einer Person wird jedoch nicht automatisch bestimmt. Neben der technischen Schwierigkeit, die nur mit Hilfe eines LBS gelöst werden könnte, bietet diese Variante zudem eine höhere Privatsphäre. Der angezeigte Standort ändert sich nur, wenn dies der Nutzer explizit wünscht.

Dieses Projekt wurde in einem sechswöchigen Test erprobt und lieferte folgende Ergebnisse: Obwohl das Live Adressbook genutzt wurde, wurden nicht alle Aufenthalts- und Statusänderungen eingetragen. Der vermutete Grund liegt darin, dass es ein relativ hoher Aufwand ist, die Veränderung einzutragen, wenn man den Standort häufig wechselt und dieser Aufwand nicht im Verhältnis zum vermeintlichen Nutzen steht. Dennoch wurde das Live Adressbook regelmäßig genutzt. Die implementierte Instant-Messaging-Funktion kam häufig zum Einsatz.

Insgesamt wurde eine Vielzahl der angebotenen Features recht positiv aufgenommen. Dies verdeutlicht, dass die Erweiterung von Gemeinschaften um eine mobile Applikation oft zusätzliche Möglichkeiten bietet und damit komplementär zu anderen Diensten angeboten werden kann. Reale Welt, bzw. virtuelle Gemeinschaft stehen in diesem Beispiel nicht in Konkurrenz zueinander, sondern ergänzen sich sinnvoll. [9]

3.4 Nokia Sensor

Eines der bekanntesten Beispiele auf dem Sektor der mobilen Gemeinschaften mit integrierter Friendsuch-Funktion ist die von Nokia angebotene Applikation „Sensor“, die kostenlos von der Seite www.nokia.com herunter geladen werden und auf jedem dafür geeigneten Mobiltelefon installiert werden kann.

Nokia Sensor ist eine Applikation, die für spontane, kontaktfreudige Menschen konzipiert worden ist. Mit Hilfe dieser Applikation ist dem Nutzer möglich, seine eigenen personalisierten Seiten zu erstellen. Diese Seite kann von anderen Nokia Sensor-Nutzern betrachtet werden, die sich in Reichweite befinden. Ebenso kann man sich deren Profile und erstellte Seiten ansehen. Die Reichweite beträgt durch die Verwendung von Bluetooth ca. 10 Meter. Befindet man sich innerhalb dieser 10 Meter-Reichweite, wird einem der andere Nutzer angezeigt, sobald die Umgebung

scannt. Da Bluetooth verwendet wird und somit keine Verbindung über einen Netzanbieter aufgebaut werden muss, entstehen den Nutzern dieser Applikation keine Kosten, abgesehen vom Verbrauch des Akkus.

Um Nokia Sensor benutzen zu können, ist es lediglich notwendig, die Software aus dem Internet herunter zu laden und auf dem mobilen Endgerät zu installieren. Zur Übertragung auf das Mobiltelefon kann ebenfalls Bluetooth, aber auch eine Kabelverbindung genutzt werden. Bei neueren Handys ist ein solches Datenkabel standardmäßig beim Kauf enthalten.



Abbildung 5. Das Ausgangsmenü von Nokia Sensor [10]

Sobald man sein eigenes Folio, unter dem in Abbildung 5 zu sehenden zweiten Punkt „My folio“, kreierte hat, kann man mit anderen Nutzern interagieren.

Unter Folio versteht man die eigens gestalteten Seiten, die andere Nutzer betrachten können. Dieses Folio besteht aus 5 Seiten, wobei die erste Seite eine Startseite darstellt, wie in Abbildung 6 zu sehen ist, die eine grobe Übersicht über den Anwender bietet.



Abbildung 6. Die verschiedenen Seiten eines Folios von Nokia Sensor [10]

Die Seiten 2 und 3 des Folios bieten die Möglichkeit der freien Gestaltung. So kann man hier weitere persönliche Informationen einfügen, wobei man die einzelnen Felder selbst beschriften kann. Nicht benutzte Felder oder Seiten werden nicht angezeigt.

Die vierte Seite dient zur der Bereitstellung von Bildern, Videos und anderen medialen Dateien. Dies bietet die Möglichkeit, andere Leute am letzten Urlaub oder an selbst komponierten Soundfiles teilhaben zu lassen. Das senden und empfangen dieser

Dateien wird ebenfalls über die Bluetooth-Schnittstelle abgewickelt, weshalb hierfür ebenfalls keine Kosten anfallen. Die Übertragungsgeschwindigkeit ist allerdings nicht sonderlich hoch, sodass man bei Datenverkehr zwischen seinen Mobiltelefonen darauf achten sollte, sich nicht aus der Reichweite zu bewegen, da die Verbindung sonst abgebrochen werden muss.

Seite 5 bietet eine Gastbuch-Funktion, wie sie bereits aus dem Internet bekannt ist. Hier kann man seinem Kommunikationspartner eine Nachricht hinterlassen, die von allen anderen Nutzern ebenfalls gelesen werden kann. Man kann auf bereits geschriebene Einträge eingehen, oder neue Beiträge hinzufügen.

Die gesamte Applikation ist nur dann für andere sichtbar, wenn sie auf dem mobilen Endgerät auch läuft. Dies bedeutet nicht, dass sie geöffnet sein muss, da sie im Standby-Modus auch im Hintergrund laufen kann, und man sich mit anderen Features seines Handys beschäftigt. Verlässt man das Programm komplett, bricht die Bluetooth-Verbindung automatisch ab.

Läuft die Applikation und startet man den Scan, werden alle Endgeräte mit laufendem Sensor-Dienst angezeigt. Auch andere Bluetooth-Dienste werden angezeigt, diese bieten aber nicht die Sensor-spezifischen Features. Mit den gefundenen Sensor-Nutzern kann man auf verschiedene Weise interagieren.

Beispielsweise kann man Sensor-Nachrichten verschicken. Dieses Nachrichtensystem ist kostenfrei und funktioniert mit allen Sensor-Nutzern innerhalb der Reichweite. Im Gegensatz zu Einträgen im Gästebuch sind die Nachrichten, die man direkt an eine Person schickt nicht öffentlich zugänglich. Diese privaten Nachrichten kann man senden, ohne die Telefonnummer seines Gegenübers zu kennen. Diese Nachricht kann Text und oder Bilder umfassen. Sein Pseudonym und Nutzfoto werden standardmäßig mit gesendet, dies kann jedoch deaktiviert werden. Einfache Textnachrichten können auch an mobile Endgeräte gesendet werden, die nicht Sensor verwenden, aber Bluetooth aktiviert haben. Alle empfangenen Nachrichten werden auf dem Handy lokal gespeichert.

Weitere Kommunikationsmöglichkeiten bieten das Gästebuch, in das man kostenlos Einträge schreiben kann, die Option des Sendens oder Empfangens von Bildern, Videos, Musik und anderer Daten, sowie das Verschicken und Erhalten von Visitenkarten.

Ähnlich zu Instant Messagern, wie AIM, ICQ, Yahoo Messenger oder MSN, die sich im Internet bereits seit langem etabliert haben, gibt es in Nokia Sensor ein Buddy-System. Dies bedeutet, dass man einem Bekannten eine Anfrage auf Freundschaft sendet, die dieser annehmen oder ablehnen kann. Nimmt er sie an, erscheint er in einer Liste, die stets aktuell gehalten wird und in der angezeigt wird, ob sich ein Freund innerhalb des Bluetooth-Bereichs befindet.

Ebenfalls lokal gespeichert wird eine Statistik, in der abzulesen ist, wie oft sein eigenes Folio innerhalb eines bestimmten Zeitraums aufgerufen und betrachtet worden ist.



Abbildung 7. Statistik der Aufrufe des eigenen Folios [10]

Nicht nur Nachrichten werden auf dem Handy gespeichert, sondern auch Verläufe innerhalb der eigenen Zone. So hält Sensor z.B. fest, wie oft das eigene Folio innerhalb eines bestimmten Zeitraums aufgerufen wurde. In Abbildung 7 sieht man eine Übersicht der Aufrufe des eigenen Folios innerhalb eines 7-Tage-Zeitraums. Die Länge des Zeitraums ist skalierbar. Mit Hilfe dieses Features kann man erkennen, an welchen Tagen und zu welchen Zeiten man besonders häufig gescannt wurde. Beim Abgleich mit seiner Tagesgestaltung kann man so Korrelationen zwischen der Häufigkeit von Scans und dem Aufenthalt an bestimmten Orten feststellen. [9]

Dadurch dass Nokia Sensor kostenlos anbietet, erfreut es sich stets wachsender Beliebtheit und wird auf immer mehr Handys installiert und benutzt. Zwar entstehen dem Konzern durch das erstellen der Software Kosten, diese sind aber einmalig und überschaubar und bieten dafür einige Vorteile. Zum einen etabliert sich das Produkt mittelfristig auf dem Markt, was eine mögliche Kommerzialisierung ab einer bestimmten kritischen Masse ermöglicht, in erster Linie aber kann Nokia Sensor die Kaufentscheidung eines unschlüssigen Mobilfunknutzers dahingehend beeinflussen, sich weiterhin für Nokia-Endgeräte zu entscheiden oder auf Nokia umzusteigen.

4 Technische Aspekte

Anhand der Beispiele ist zu erkennen, dass mobile Gemeinschaften oder Freundsuchdienste auf ganz unterschiedliche Arten funktionieren können. Die technische Umsetzung der verschiedenen Angebote liefert jeweils Vor- und Nachteile, die hier aufgezeigt werden sollen.

4.1 Funknetzbetreiber

Das Beispiel FriendZone hat gezeigt, dass eine mobile Applikation mit Hilfe der bestehenden Infrastruktur der Netzbetreiber möglich ist. Nutzer wurden aufgrund der Zelle lokalisiert, innerhalb derer sie sich befanden. Durch diese Lösung musste für diesen Dienst keine neue Infrastruktur errichtet werden, was mit hohen Kosten verbunden gewesen wäre. Gleichzeitig war dieses Freundsuchangebot dennoch nicht preiswert, da der Netzanbieter für jede Interaktion Geld verlangte, weil jedes Mal eine gewisse, wenn auch geringe, Netzlast anfiel. Hätte der Netzbetreiber selbst durch die

mobile Gemeinschaft profitiert, ohne dafür Geld verlangen zu müssen, so hätte diese Option weitaus bessere Chancen, als sie es momentan zu haben scheint.

Zudem die gebotene Leistung keineswegs sonderlich hoch einzuschätzen war, denn innerhalb einer Funkzelle eines Mobilfunknetz-Betreibers sind mobile Endgeräte nur sehr grob zu lokalisieren. Eine solche Funkzelle umfasst oftmals mehrere hundert Meter bis zu wenigen Kilometern. Innerhalb eines solchen Radius können nur grobe, ortsgebundene Dienste angeboten werden, bei denen es auf den exakten Standort kaum ankommt.

Wesentlich unproblematischer ist eine solche Lösung für eine rein mobile Gemeinschaft, die keinerlei Wert auf ortsbasierte Dienste oder Angebote legt.

Der Funknetzbetreiber bietet also für mobile Gemeinschaften ein ausreichendes Maß an Nutzbarkeit, fordert dafür aber relativ hohe Kosten, angesichts der Tatsache, dass beispielsweise eine SMS je nach Anbieter 10 bis 20 Cent kostet, was bei 160 übertragenen Zeichen ein immenser Preis ist. Für ortsbasierte Dienste ist ein Netzbetreiber also eher ungeeignet.

4.2 WLAN

Eine andere Lösung suchte das Projekt Active Campus. Um eine genauere Lokalisierung innerhalb des Campusgeländes vornehmen zu können, wurde ein Drahtloses Netzwerk installiert und den Studenten ein WLAN-fähiges mobiles Endgerät ausgehändigt. Da der WLAN-Standard die lizenzfreie Nutzung des ISM-Bandes im 2,4GHz-Bereich vorsieht, entstehen neben der Infrastruktur und den mobilen Endgeräten keine laufenden Kosten. Da kein hoher Datenverkehr zwischen den mobilen Geräten und dem Netzwerk entstehen, besteht auch kein Problem mit der möglichen Anzahl der Nutzer, die sich allesamt die Luftschnittstelle teilen müssen. Diese technische Lösung per WLAN-Standard funktioniert sehr gut und bietet auch die Möglichkeit der unkomplizierten Erweiterbarkeit.

Ein Nachteil dieser Variante ist jedoch der hohe Stromverbrauch. Dieser führt zu häufigem Aufladen des Endgeräts, was eine Beeinträchtigung der Nutzbarkeit darstellt. Insgesamt bietet die WLAN-Variante allerdings ein solides System, das einmal eingeführt nur noch inhaltlich mit hohem Aufwand gewartet werden muss. Die Hardware bedarf weniger, seltener Modifikationen.

4.3 Bluetooth

Was bei WLAN einen Nachteil darstellte, nämlich der enorme Energieverbrauch, ist ein Pluspunkt bei Bluetooth-Varianten. Im Standby-Modus verbraucht Bluetooth nur einen Bruchteil dessen, was WLAN benötigt. Auch beim senden von Dateien ist der Verbrauch akzeptabel, liefert gegenüber 802.11 aber keinen derart immensen Vorteil wie im Ruhemodus.

Ähnlich zu der Netzwerkvariante, wie sie in Active Campus genutzt vorgestellt wurde, entstehen dem Nutzer von Bluetooth-Gemeinschaften keine Kosten beim interagieren mit anderen Mitgliedern, abgesehen vom Stromverbrauch.

Ein weiterer Vorteil von Bluetooth gegenüber WLAN ist die Möglichkeit, des Ad-hoc- Verbindungsaufbaus. Dies bedeutet, dass jederzeit, unabhängig von einer Infra-

struktur, ein Verbindungsaufbau zwischen Teilnehmern einer Gruppe durchgeführt werden muss. Dieses Fehlen einer Infrastruktur bedeutet niedrige Kosten für den Anbieter des Dienstes und Ortsunabhängigkeit der Nutzer.

Neben den Vorteilen, den die Unabhängigkeit von einer Infrastruktur bietet, sind aber auch Nachteile vorhanden. Z.B. ist die Gemeinschaft während der meisten Zeit nicht in der Lage, untereinander zu interagieren. Dies ist nur möglich, wenn sich innerhalb eines 10-Meter-Radius ein Mitglied befindet.

Nokia Sensor und ähnliche Bluetooth-Anwendungen sind also weniger für mobile Gemeinschaften geeignet, da man innerhalb eines 10-Meter-Radius mit Freunden auch persönlich in Interaktion treten kann. Wesentlich besser geeignet sind derartige mobile Applikationen, um als Freundfinder zu agieren. Will man reale Menschen kennen lernen, gibt es keinen besseren Ort dafür, als die unmittelbare Umgebung. Durch eine kurze Interaktion mit Hilfe des anonymen technischen Geräts, kann die Scheu abgebaut werden und man kann die andere Person real kennen lernen.

5 Soziale Betrachtungsweise

Wie bereits im letzten Abschnitt und auch schon zu einem früheren Zeitpunkt angedeutet, können mobile Applikationen dabei helfen, reale Kontakte aufzufrischen oder neue Kontakte zu knüpfen.

Bei Active Campus kommen Kommilitonen und Kommilitoninnen verschiedener ethnischer und sozialer Hintergründe zusammen, die sich ohne die Unterstützung des elektronischen Hilfsmittels vielleicht nie getroffen hätten. Man wird innerhalb dieses Systems auf Dinge des realen Lebens aufmerksam gemacht, die man anders möglicherweise niemals erfahren hätte. Im Gegensatz zur virtuellen Gemeinschaft bietet die mobile Gemeinschaft einen wesentlich stärkeren Bezug zur realen Umgebung, da die sie auf Ereignisse innerhalb der Umwelt eingehen und auf sie aufmerksam machen kann. Bei virtuellen Gemeinschaften steht der fiktive Raum im Vordergrund, in den man sich bei der Nutzung begibt. Innerhalb dieses Raums tritt man mit anderen Menschen in Interaktion und verbringt dort zusammen einen großen Teil seiner Zeit. Nur selten kommt es zu so genannten Ausbrüchen aus dieser Welt. Unter Ausbruch ist zu verstehen, dass sich Akteure des virtuellen Raums auch in der Realität begegnen.

Durch mobile Gemeinschaften wird die reale Begegnung gefördert. Zum einen dadurch, dass man auf die Nähe eines Freundes oder eines potentiell interessanten Menschen aufmerksam gemacht wird, zum anderen indem man auf Ereignisse innerhalb der Gesellschaft hingewiesen wird, die in der Nähe stattfinden. Nimmt man an einem solchen Event teil, interagiert man zwangsläufig wieder mit Menschen, meistens mit fremden. Aufgrund dieser Aspekte lässt sich sagen, dass mobile Gemeinschaften und Friend Finder unter der sozialen Betrachtungsweise sinnvolle Einrichtungen sind.

6 Zusammenfassung

Diese Ausarbeitung gibt einen Überblick über die verschiedenen Varianten, die es unter den mobilen Gemeinschaften und Friendsuch-Angeboten gibt. Neben der Funk-

tionsweise solcher Dienste, und der technischen Unterschiede, die es je nach Anforderungsprofil eines Angebots gibt, wurden einige Beispiele aufgezeigt, die teilweise nicht über die Testphase hinausgingen. FriendZone z.B. war ein Pilotprojekt, das einige lohnenswerte Aspekte aufgewiesen hat, aber auch einige Schwächen besaß. Ebenfalls Interessante Features bot das Live Adressbook, das die Handhabung von mobilen Endgeräten durch eine zentrale Datenbank erleichtern sollte.

Die mobile Gemeinschaft Nokia Sensor weist viele Stärken auf, weshalb es sich bereits relativ gut etabliert hat. Dass der Dienst kostenlos ist und man auf interessante Art und Weise mit seinen Mitmenschen in Kontakt kommt, macht ihn gerade bei jungen Nutzern beliebt.

Auch Active Campus bietet hilfreiche Features, die den Bereich des universitären Umfelds leichter nutzbar macht und bei relativ geringen Kosten einen hohen Mehrwert bietet.

Das Entwicklungspotential solcher mobilen Applikationen ist allerdings bei weitem noch nicht ausgeschöpft. Sowohl Mehrwertdienste, die zum einen dem Anbieter und zum anderen dem Nutzer Vorteile liefern, ohne dabei etwas kosten zu müssen, als auch kostenpflichtige Dienste, die für den Nutzer große Vorteile bei akzeptablen Preisen bieten, bieten weitere Chancen.

Das Beispiel Nokia Sensor zeigt einen Dienst auf, der für den Betreiber ebenso rentabel, wie für den Anwender sinnvoll ist. Hier kann man von einer Win-Win-Situation sprechen, denn der Bereitsteller der Software, Nokia, profitiert durch Markenbindung an das mobile Endgerät und der Teilnehmer besitzt die Möglichkeit mit seinem Handy andere Dienste in Anspruch zu nehmen, als mit anderen Mobiltelefonen.

Das Angebot solcher mobilen Gemeinschaften und Freundsuch-Diensten wird sich im Laufe der kommenden Jahre stark vermehren, wenn weitere Anbieter erkennen, dass diese Dienste für sie von Vorteil sind. Ob das Angebot dabei kostenpflichtig oder unentgeltlich nutzbar ist, spielt dabei eine untergeordnete Rolle. Dies hängt von der jeweiligen Applikation und der Zielgruppe ab.

Literatur

1. Fremuth, N., Reichwald, R., Ney, M.: Mobile Communities – Erweiterung von Virtuellen Communities mit mobilen Diensten, <http://cosmos.in.tum.de/downloadFiles/FremuthNeyFORMAT.pdf>
2. Koch, M., Groh, G., Hillebrand, C.: Mobile Communities – Extending Communities into the Real World, to appear in: Proc. American Conf. on Information Science (AMCIS2002), Dallas, TX, Aug. 2002, S.1
3. Sabine Feierabend, Thomas Rathgeb, Medienpädagogischer Forschungsverbund Südwest, Basisstudie zum Medienumgang 12- bis 19-jähriger in Deutschland, Stuttgart, November 2005
4. Koch, M., Groh, G., Hillebrand, C.: Mobile Communities – Extending Communities into the Real World, to appear in: Proc. American Conf. on Information Science (AMCIS2002), Dallas, TX, Aug. 2002, S.2
5. Axel Küpper, Location-based Services - Fundamentals and Operation, John Wiley & Sons, 1. Kapitel

6. Lorenz, A., Schmitt, C., Oppermann, R., Eisenhauer, M., Zimmermann, A., Location and Tracking in Mobile Guides. In: Proceedings of 4th Workshop on HCI in Mobile Guides in conjunction with the 6th Workshop on Artificial Intelligence in Mobile Systems (AIMS), 09.09.2005, Salzburg, <http://www.mguides.info/>
7. Conference on Human Factors in Computing Systems, CHI 2003, Ft. Lauderdale, Florida, USA, April 5-10, 2003. ACM 2003, ISBN 1-58113-637-4, Burak, A., Sharon, T., Analyzing Usage of Location Based Services, <http://xenia.media.mit.edu/~taly/publications/chi03-cell.pdf>
8. Griswold, W.G., Boyer, R., Brown, S.W., Truong, T.M., Bhasker, E., Jay, G.R., Shapiro, R.B.: Active Campus – Sustaining Educational Communities through Mobile Technology. Department of Computer Science and Engineering, University of California, San Diego, <http://www-inst.eecs.berkeley.edu/~maratb/readings/ac.pdf>, S. 2-4
9. Milewski, A., Smith, T.M., Providing Presence Cues to Telephone Users, AT&T Labs, Research, 100 Schulz Drive, Red Bank, NJ 07701, 2000, S. 89-92
10. Nokia GmbH, <http://europe.nokia.com/nokia/0,,73655.00.html>
11. Eigene Grafik

Mobile Multimedia Messaging

Christoph Pahre

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
pahre@informatik.uni-muenchen.de

Zusammenfassung Diese Arbeit beschäftigt sich mit dem (Forschungs-) Gebiet „Mobile Multimedia Messaging“. Mit einer Hinführung zum Thema, wobei besonders auf das Paradebeispiel MMS eingegangen wird, beginnt diese Ausarbeitung. Daran schließt dann ein Überblick über die verschiedenen Bereiche, die in diesem Zusammenhang eine Rolle spielen, an. Im zweiten Teil dieser Arbeit werden dann zwei Themengebiete genauer untersucht. Es wird zunächst der Einfluss von „Mobile Multimedia Messaging“ auf soziale Verhaltensweisen dargestellt. Dabei beschränkt sich die Arbeit auf drei ausgewählte Gesichtspunkte. Dazu gehören Verlust von Emotionen bei der Kommunikation über Nachrichten, Aspekte der Privatsphäre und Auswirkungen auf die Interaktion durch die Möglichkeit multimediale Daten einbinden zu können. Mit der Adaption von Inhalten für mobile Endgeräte beschäftigt sich dann der zweite Teil. Abschließend werden dann noch die wichtigsten Punkte zusammengefasst und ein kurzer Ausblick auf zukünftige Entwicklungen gegeben.

1 Einleitung

„Mobile Multimedia Messaging“ lässt sich in mobile multimediale Nachrichtenübertragung übersetzen, genauer trifft es vielleicht aber mobile Nachrichtenübertragung mit multimedialem Inhalt. Nachdem nun versucht wurde den Begriff möglichst treffend ins Deutsche zu übersetzen sollte jeder eine gewisse Vorstellung gewonnen haben, was sich hinter diesem Begriff verbirgt. Um unterschiedliches Verständnis auszuschließen soll zunächst genauer dargestellt werden, was in dieser Arbeit unter diesem Begriff zu verstehen ist. „Mobile Multimedia Messaging“ ist im Zusammenhang mit dieser Ausarbeitung als Forschungsgebiet zu sehen, welches sich mit der Schnittmenge folgender Aspekte beschäftigt:

- „Mobile“: Von Interesse sind mobile Endgeräte, wie bspw. Mobiltelefone und PDAs.
- „Multimedia“: Beim Inhalt der Nachrichten handelt es sich um multimediale Daten, welche von Videos über Klingeltöne bis hin zu Präsentationen alles sein können.

- „Messaging“: Die Übermittlung von Daten findet in Form von Nachrichten (=Datenpaket) statt.

Berührt eine Applikation beispielsweise alle oben genannten Bereiche, fällt sie in das Forschungsgebiet des „Mobile Multimedia Messaging“.

Bevor nun die Materie tiefergehend erörtert wird, sei noch ein kurzer Überblick über diese Arbeit gegeben. Zunächst werden die verschiedenen zum Forschungsbereich des „Mobile Multimedia Messaging“ gehörenden Gebiete identifiziert. Diese werden dann anschließend kurz und bündig dargestellt, um einen Überblick über die ganze Thematik zu geben. Eine Anwendung, die sich dem Bereich des „Mobile Multimedia Messaging“ zuschreiben lässt, schließt das 2. Kapitel dann ab. Es handelt sich dabei um die MMS welche hier beispielhaft genauer beleuchtet wird und das Thema veranschaulichen soll.

Daran schließt sich dann Kapitel 3 an, welches einige der zuvor erwähnten Bereiche tiefergehend behandelt. Diese sind: Die Adaption von Inhalten und die sozialen Aspekte, wobei hier die Aspekte der Privatsphäre, die Auswirkungen von multimedialem Inhalt auf die Interaktion und die Übertragung von Emotionen abgehandelt werden.

Die Arbeit wird dann durch eine Zusammenfassung der Arbeit und einem kurzen Ausblick geschlossen.

2 Bereiche des Mobile Multimedia Messaging

Der Begriff des „Mobile Multimedia Messaging“ umfasst eine ganze Reihe von verschiedenen Themen und Gebieten. Alles von der Übertragungstechnik [7] über die gesellschaftlichen Auswirkungen [3] bis hin zur Adaption von Inhalten [5] spielt in diesem Zusammenhang eine Rolle. In Abbildung 1 sind die wichtigsten Bereiche dargestellt. Dieses Diagramm soll einen ersten Einblick in die gesamte Thematik geben. Im Folgenden wird auf jeden der genannten Bereiche kurz eingegangen und anhand eines Beispiels erörtert.

Geräte

Außer den allgegenwärtigen Mobiltelefonen gibt es noch eine ganze Reihe anderer mobiler Endgeräte. Dazu zählen: Smartphones, PDAs (Personal Digital Assistant), BlackBerry-Handhelds, aber auch Notebooks und Subnotebooks. Da die letzten beiden Typen nur im weitesten Sinne unter die mobilen Endgeräte fallen, werden sie in dieser Arbeit nicht betrachtet [8].

Verbindungsart

Zunächst muss hier zwischen zwei grundlegend verschiedenen Verbindungen unterschieden werden. Auf der einen Seite gibt es Kommunikation welche über den Service Provider bzw. genauer gesagt über dessen Netz führt. Auf der anderen Seite kann eine direkte Verbindung zwischen einem oder mehreren mobilen Endgeräten ohne vermittelnde Station aufgebaut werden [6,7].

Es gibt verschiedene Möglichkeiten eine direkte Verbindung zwischen zwei Geräten herzustellen. Dazu zählen Infrarot, Bluetooth und WLAN (Wireless Local Area Network), als neuester Vertreter dieser Sparte [6,7].

Globale Dienste, wie beispielsweise das Telefonieren, finden bei mobilen Endgeräten über so genannte Service Provider als Vermittler statt. Um dieser Rolle gerecht zu werden wird ein möglichst weitläufiges Funknetz zur Verfügung gestellt, welches der Nutzer bei Bedarf ansprechen kann. Die zu übermittelnden Daten werden zunächst per Funk übertragen und legen dann größere Strecken im Kabelnetz des Service Providers zurück. Je nach dem welches Endgerät angesprochen wird, wird das letzte Stück per Kabel (Telefon / Modem) oder per Funk (Mobiltelefon, PDA) überwunden. Es gibt verschiedene Generationen von Mobilfunknetzen die unterschiedliche Leistungen für den Kunden bereitstellen können (SMS, WAP) [9].

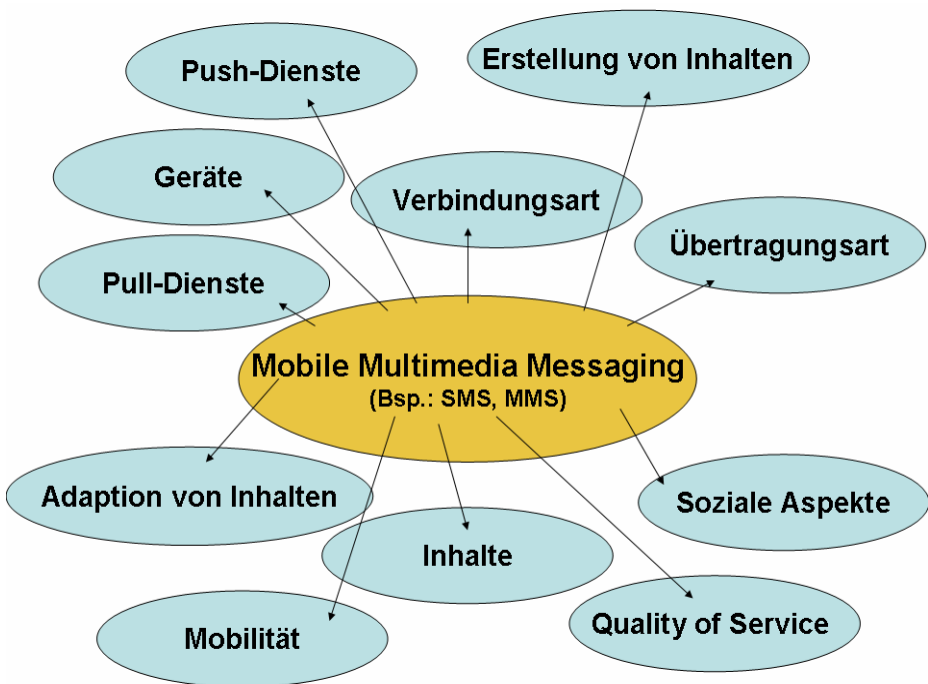


Abbildung 1. Bereiche des "Mobile Multimedia Messaging" [1,2,3,4,5,6,7].

Übertragungsart

Abhängig von den zu übertragenden Inhalten und deren Anforderungen wird die zu nutzende Übertragungs- bzw. Kommunikationsart gewählt. Es kann elementar zwischen nachrichtenorientierter (message-oriented) und datenflussorientierter (stream-oriented) Kommunikation unterschieden werden. Beide Arten haben verschieden Vor- und Nachteile, aber der wesentliche Unterschied besteht in drei Bereichen: Zuverlässigkeit, Geschwindigkeit und Sicherheit [10].

Bei der nachrichtenorientierten Übertragung spielt die Geschwindigkeit eine untergeordnete Rolle, Zuverlässigkeit und Sicherheit haben Vorrang. Als Beispiel kann hier der Download von Software für ein mobiles Endgerät genannt werden [10].

Im Gegensatz zur nachrichtenorientierter Übertragung spielt bei der streamorientierten Kommunikation vor allem die Geschwindigkeit eine Rolle. Die zu übertragenden Daten (Sound, Video) sollen in Echtzeit beim Empfänger ankommen und abgespielt werden. Da allerdings nicht ganz auf ein Minimum an Qualität verzichtet werden kann muss zwischen „Quality of Service“ und der Geschwindigkeit ein vernünftiges Mittelmaß gefunden werden. Ein Beispiel hierfür wäre das Internetradio [10].

Mobilität

Endgeräte, wie Mobiltelefone, sind heutzutage nahezu uneingeschränkt mobil. Dies ist dem so genannten „Mobility Management“ zu verdanken, welches eine der Hauptfunktionen von GSM- und UMTS-Netzen ist. Es ermöglicht einzelne Nutzer im Netz zu „orten“, um dann verschiedene Dienste bereitstellen zu können, wie beispielsweise das Versenden und Empfangen von SMS oder das Telefonieren [11,12].

Heutzutage ist die Mobilität über Landesgrenzen und Netze verschiedener Service Provider hinweg gewährleistet und somit steht dem Nutzer die verschiedenen Dienste nahezu auf der ganzen Welt zur Verfügung [11].

Inhalte

Der Begriff „Mobile Multimedia Messaging“ enthält das Wort Multimedia und weist schon darauf hin, dass verschiedene (multimediale) Daten bei der mobilen Nachrichtenübermittlung eine Rolle spielen. Text, Bild, Video und Audio oder eine beliebige Kombination dieser (z.B. Video mit Untertiteln) sind Inhalte die in diesem Zusammenhang genannt werden müssen [1,2,15].

Zunächst war es nur möglich Text per Nachricht (SMS) zu übertragen, aber mittlerweile kann eine Nachricht Multimediadaten (MMS) enthalten und eignet sich somit auch als Trägermedium für verschiedene Dienste, wie z.B. den Download von Klingeltönen [1].

Quality of Service (QoS)

Der Begriff „Quality of Service“ beschreibt die Qualität der Übertragung. Hierbei spielt insbesondere die Übertragungsgüte (Einfluss haben hier Bandbreite, Jitter, Delay und Fehlertoleranz) eine entscheidende Rolle. Es gibt vier Klassen des QoS im Zusammenhang mit Mobilfunk, die an verschiedene Bereiche der Übertragung mehr oder weniger kritische Anforderungen stellen, da sie für spezielle Aufgaben konzipiert wurden [10,13, 14].

Da wäre zunächst die so genannte „background“-Klasse. Hier wird besonderer Wert auf eine möglichst geringe Fehlerrate gelegt. Unkritisch sind hierbei die Bandbreite, der Jitter und die Verzögerung (Delay) der Übertragung [13].

Die zweite Klasse „interactiv“ ist ähnlich der Ersten, legt jedoch gesteigerten Wert auf ein geringes Delay, da es bei der Nutzung interaktiver Dienste eingesetzt wird und der Nutzer keine langen Wartezeiten haben sollte [13].

Verteilerdienste nutzen die Klasse „streaming“, welche besonderen Wert auf die zur Verfügung stehende Bandbreite legt. Jitter und Bitfehler sind unkritisch einzustufen [13].

Die letzte Klasse „conversational“ ist für die direkte Kommunikation gedacht, sprich das Telefonieren und ähnelt dem „streaming“, legt aber starken Wert auf ein geringes Delay und auch der Jitter muss sich im Rahmen halten [13].

Pull-Dienste

Jeder Mobilfunkanbieter stellt seinen Kunden verschiedene Dienste zu Verfügung. Dazu gehören das Telefonieren, SMS, MMS und vieles mehr. Die Inanspruchnahme manchen Dienste geschieht ohne eine Aktion des Nutzers (Push-Dienste) und manche werden erst durch aktives Handeln angestoßen, hierbei handelt es sich um die Pull-Dienste (Pull = ziehen).

Eine Vielzahl von Leistungen werden erst auf Abruf, Pull, bereitgestellt. Dazu zählen Downloads, Internetinhalte, das Versenden von Nachrichten, der Aufbau eines Telefongesprächs und vieles mehr.

Push-Dienste

Der genaue Gegensatz zu Pull-Diensten sind die Push-Dienste, sie werden ohne Nutzerinteraktion ausgelöst. Als Beispiel sei hier das Zustellen von SMS und MMS genannt. Allerdings sind hier starke Veränderungen im Gang und es wird in verschiedene Richtungen geforscht und entwickelt. Vor allem der Bereich von „Location-Aware-Advertising“ erscheint wirtschaftlich interessant. In [6] wird ein System vorgestellt um ebensolche Werbung realisieren zu können. Allgemein wird nach einer effizienten und flexiblen Lösung gesucht um mobile Nutzer ansprechen zu können [6,16].

Erstellung von Inhalten

Die Erstellung von Inhalten für mobile Endgeräte unterscheidet sich stark von Content für andere Geräte, wie bspw. den PC. Dies liegt an den Einschränkungen, die mobile Geräte aufweisen. Dazu zählen die vergleichbar geringeren Eingabe- (Fehlen einer Mouse, normale Tastatur) und Ausgabemöglichkeiten (Kleines Display, keine Lautsprecher). Professionelle Inhalte werden aus diesem Grund auf Workstations erstellt. Besonders interessant ist es allerdings wie „normale“ Mobilfunknutzer Inhalte auf dem mobilen Endgerät erstellen, und wie die dazu nötigen Programme konzipiert sein müssen, um dem Nutzer einen größtmöglichen Nutzen zu bieten [2].

Soziale Aspekte

Ein weiteres interessantes Gebiet des „Mobile Multimedia Messaging“ sind die sozialen Aspekte im Zusammenhang mit mobilen Endgeräten [3,4, 21].

Im folgenden Kapitel unter Punkt 3.1 werden einige davon aufgezeigt. Ein Unterpunkt beschäftigt sich mit dem Verlust von Emotionen bei der Kommunikation über Nachrichten. Die Veränderungen der Interaktion zwischen Personen, die über Messaging kommunizieren, durch die Möglichkeit multimediale Daten einzubinden, wird ebenfalls erörtert. Der letzte Aspekt der Betrachtung findet ist die Privatsphäre im Zusammenhang mit Nachrichten auf mobilen Endgeräten [3,4, 21].

Anpassung von Inhalten (Adaption)

Die Anpassung der abgerufenen Inhalte an das gerade genutzte Endgerät und seine Situation spielen eine immer größere Rolle bei multimediale Inhalten. Unter diesen Bereich der Adaption fallen sowohl „Location-Based-Services“ (standortbezogene Dienste, z.B. Ortung) , als auch die „Context-Awareness“ (Situationsbewußtsein, z.B. Mobiltelefon kennt eigene Lage im Raum) [5, 19].

Punkt 0 wird zunächst darauf eingehen, wieso Adaption einen wichtigern Forschungsbereich darstellt, anschließend wird das Thema tiefergehend behandelt.

Beispiele aus dem Bereich des „Mobile Multimedia Messaging“ (SMS, MMS)

Dem Gebiet des „Mobile Multimedia Messaging“ lassen sich zwei bekannte Mobilfunkdienste zuordnen, die SMS und ihr Nachfolger die MMS. Beide Nachrichtensysteme stellen dem Mobilfunknutzer die Möglichkeit zur Verfügung mobil, per Funk, multimediale Daten in Form von Nachrichten zu übertragen und zu empfangen. Auch wenn sich bereits die SMS, welche nur Text enthalten kann, diesem Forschungsbereich zuordnen lässt, so passt doch die MMS noch besser in dieses Schema, da mit ihr wirkliche multimediale Nachrichten (z.B. Bilder) übertragen werden können.

MMS steht für „Multimedia Messaging Service“ und stellt eine Lösung zum Versenden und Empfangen von multimedialen Daten auf mobilen Endgeräten (Handy, PDA, Smartphone) dar [1]. Die MMS ist die logische Weiterentwicklung der SMS („Short Message System“) und stellt erweiterte Funktionalität zur Verfügung. Dazu gehört, dass neue Arten von Inhalten (z.B. Klingeltöne) übertragen werden können [2]. Außerdem wurde die Erstellung von Inhalten, ebenso wie deren Anzeigemöglichkeiten verbessert. Im Folgenden werde ich noch etwas genauer auf die MMS und ihre Eigenschaften eingehen, da sie das Paradebeispiel für „Mobile Multimedia Messaging“ ist [1, 20].

Die MMS ist ein trägerunabhängiger Service, das bedeutet die MMS kann die moderne Infrastruktur der dritten Generation (3G = 3rd Generation) aber auch die der zweiten Generation (2G) zur Übertragung nutzen. Ebenso verhält es sich bei der SMS. MMS und SMS arbeiten beide offline, d.h. es wird keine Verbindung zwischen dem versendenden und dem empfangenden Gerät aufgebaut. Die Übertragung gelangt über mehrer Zwischenstationen (Hops = Sprünge) zum Empfänger.

Den „store-and-forward“-Ansatz haben sowohl die SMS als auch die MMS mit der Email gemein. Dabei wird die Nachricht an den eigenen Dienstanbieter (Service Provider) verschickt, dieser speichert und leitet diese dann zum Empfänger weiter (Email: Pull, SMS & MMS: Push). Um die Kommunikation per MMS zu ermöglichen, muss der Service Provider des Mobilfunknetzes diesen Service zur Verfügung stellen. Je nachdem ist es auch möglich, per MMS mit anderen Nachrichtensystemen, wie Email, zu kommunizieren und umgekehrt [1].

Da die MMS als globaler Service gesehen wird, ist sie sowohl durch 3GPP (3rd Generation Partnership Project), als auch durch das WAP (Wireless Application Protocol) Forum standardisiert worden. Obwohl sich die Implementierung der beiden unterscheidet sind sie dennoch kompatibel und komplementär zueinander [1].

Obwohl die MMS als Vorzeigeobjekt des „Mobile Multimedia Messaging“ viele positive neue Eigenschaften gegenüber der SMS besitzt, ist dies kein Garant für den

Erfolg am Markt (wie die SMS hat). Die mit dem Verschicken von MMS verbundenen Kosten stellen hierbei einen wichtigen Punkt dar. Außerdem spielen Faktoren wie die Usability oder die Verfügbarkeit eine wichtige Rolle [1,20].

Nachdem ich nun die wichtigsten Bereiche des „Mobile Multimedia Messaging“ angeschnitten habe und die MMS als Beispiel vorgestellt habe, werden im Folgenden Kapitel besonders interessante Bereiche genauer beleuchtet. Abschließend sei noch darauf hingewiesen, dass die in Abbildung 1 dargestellten Bereiche sich nicht immer klar trennen lassen. Es gibt eine Vielzahl von Berührungspunkten zwischen den verschiedenen Gebieten, die sich auch oftmals überschneiden. Ein gelungenes Zusammenspiel all dieser verschiedenen Aspekte stellt die größte Herausforderung in diesem Forschungsbereich dar.

3 Aktuelle Forschungsbereiche

3.1 Soziale Aspekte

Der Einfluss von „Mobile Multimedia Messaging“ auf das soziale Verhalten soll Thema dieses Kapitels sein. Da diese Thematik jedoch sehr umfangreich ist beschränkt sich diese Arbeit im Folgenden auf drei Aspekte [3,4, 21].

Der erste davon ist, wie sich ein Mobiltelefon und insbesondere die Nachrichten darauf in die Privatsphäre einer Person einordnen lassen, und welche Regeln und Normen hierbei gelten [3].

Das Fehlen von Emotionen bei Nachrichten ist der zweite Aspekt, der im Folgenden betrachtet wird. Dies ist besonders interessant da Emotionen beim direkten Gespräch zwischen Personen einen wesentlichen Anteil dazu beitragen, dass Gesprochene richtig zu interpretieren [4].

Die Auswirkungen auf die Interaktion über Nachrichten durch die Möglichkeit multimediale Daten einzubinden wird dann das Kapitel schließen. Es wird betrachtet, wie Bilder die Kommunikation bereichern und diese auch stark beeinflussen können [21].

Privatsphäre

Durch die immer stärkere Nutzung von Messaging auf mobilen Endgeräten hat sich dessen Nutzung in die Gesellschaft integriert und automatisch haben sich verschiedene Regeln und Normen in diesem Bereich gebildet. Das Messaging ist Teil einer Art eigenen, mobilen Kultur (Sprache: „SMS Slang“). Dies ist insbesondere bei Teenagern, die mobile Endgeräte stark nutzen, zu erkennen. Mobiles Messaging ist hier ein wichtiger Teil des sozialen Netzes geworden und hilft dieses zu erhalten. Nachrichten werden gegenüber der Telefonie aufgrund folgender Punkte bevorzugt: Billiger, schneller und einfacher [3].

Ziel der weiteren Ausführung ist es aufzuzeigen, wie Normen und Regeln, welche die Privatsphäre von Nachrichten auf mobilen Endgeräten betreffen, aussehen und wie verbreitet diese Auffassungen sind [3].

Um das Verhalten der Nutzer in diesem Bereich besser verstehen zu können, ist es zunächst interessant wofür Personen Mobiltelefone (beispielhaft für verschiedene mobile Endgeräte) die mobiles Messaging unterstützen, benutzen. Eine Studie, siehe [3], kam zu dem in Abbildung 2 dargestellten Ergebnis. Es kann deutlich abgelesen werden, dass soziale Motive der Hauptgrund für den Besitz von Mobiltelefonen sind [3].

Reasons for having a mobile phone (N=119)	Primary reason ($\Sigma > 100\%$)	Secondary reasons ($\Sigma > 100\%$)
Social reasons	71.4 %	26.1 %
Security/Emergency	16.8 %	48.7 %
Work	8.4 %	23.5 %
Fun	2.5 %	19.3 %
Image	1.7 %	10.1 %
Information Access	0.8 %	5.0 %
Something else	3.4 %	9.2 %

Abbildung 2. Gründe für den Besitz eines Mobiltelefons [3].

Zusätzlich wurde jedem Teilnehmer der Studie noch die Frage gestellt, wie privat das eigene Telefon eingestuft wird. 82,3% hielten es für ein privates Gerät [3].

Die nahe liegende Vermutung, dass ungeschriebene Regeln und Normen in dieser mobilen Kultur existieren, wurde in der selben Studie untersucht. Das Ergebnis zeigte einige Regeln auf, die nahezu jede Testperson angab. Dazu zählen:

- Nutzung von allgemein bekannte Abkürzungen („SMS Slang“).
- Einsatz von so genannten Emoticons, z.B. :-).
- Nachrichten, wie SMS und MMS, sind absolut privat.

Insbesondere der letzte Punkt war für alle Teilnehmer der Studie wichtig und Aussagen, wie dass das Lesen einer Nachricht vergleichbar ist mit dem Öffnen eines fremden Briefs, macht die Sensibilität in diesem Punkt deutlich.

Obwohl es verschiedene Möglichkeiten gibt Nachrichten auf dem eigenen Mobiltelefon vor fremdem Zugriff zu schützen, wie bspw. Verschlüsselungen, sehen nahezu alle Testpersonen so etwas als unnötig an. Dies wird dadurch begründet, dass

für jedermann klar ist, dass es sich hier um ein privates Gerät handelt und damit weitere Maßnahmen unnötig sind.

Es kann also festgestellt werden, dass Regeln, die als SMS-Etikette bezeichnet werden können, existieren und das Mobiltelefon als privat, vergleichbar mit einem Brief, einstufen und somit schützen. Diese Tatsache ist insbesondere interessant, da die Nachrichtenübermittlung technisch unsicher ist.

Übertragung von Emotionen

Trotz verschiedener Vorteile des Messaging (günstig, schnell, vertraulich, einfach) gegenüber der Telefonie, gibt es doch verschiedene Schwachstellen. Der größte ist hier das völlige Fehlen von nicht-verbaler Kommunikation, welche im direkten Gespräch einen wichtigen Teil darstellt. Dazu zählen Dinge, wie der Gesichtsausdruck oder Gesten, welche den Gesprächspartnern helfen die andere Person besser zu verstehen. Durch das Fehlen dieser Eindrücke beim Messaging ist es für den Empfänger oft schwer die Nachricht richtig zu interpretieren [4].

Emoticons (= Icons die Emotionen ausdrücken) werden schon jetzt genutzt, um beim Messaging dem Empfänger die eigene Gefühlslage und die Intention dieser näher zu bringen. Diese Art von Kontextinformation stellt schon heute einen wichtigen Aspekt des Nachrichtenverkehrs dar [4].

Ein interessanter Ansatz wird in [4] verfolgt. Dieser beschäftigt sich mit der Erstellung eines Konzepts für die Erweiterung der SMS. Vier Informationen die bei einem Gespräch von Person zu Person jedem zur Verfügung stehen, sollen in das Messaging integriert werden, dabei ist es das Ziel Missverständnisse zu verhindern und das Messaging mehr an die sprachliche Kommunikation anzunähern [4].

- Die Identität des Gesprächspartners ist bei einem direkten Gespräch eindeutig zu erkennen. Um dies auch beim Messaging zu gewährleisten, sollen Avatare (Bilder, wie Zeichentricksfiguren oder Karikaturen) genutzt werden um den Gegenüber zu repräsentieren und damit eine eindeutige schnelle Identifizierung zu ermöglichen.
- Die Personalisierung soll eine größere Gestaltungsfreiheit bei Nachrichten ermöglichen. Dadurch kann sich der Nachrichtenersteller besser selbst verwirklichen. In Fall der SenseMS wird es durch verschiedene Formatierungen des Texts realisiert.
- Der Gesichtsausdruck verrät im Gespräch die Intention der Person und stellt einen entscheidenden Teil der nicht-verbale Kommunikation dar. Um diesen Part zu transportieren soll die momentane Stimmung des Versendenden den Gesichtsausdruck des Avatars entsprechend beeinflussen.
- Die Situation in welcher Kommunikation stattfindet ist ein weitere Teil der normalen Kommunikation. Ein Bild zusammen mit der Nachricht zu übertragen soll es dem Empfänger ermöglichen sich besser in die Lage des Versendenden zu versetzen.

Abbildung 3 zeigt die Funktionen die eine SenseMS gegenüber einer normalen SMS voraus hat [4].

Das linke Bild zeigt eine Briefmarke, welche für den Absender gebündelt verschiedene Informationen über den Empfänger darstellt. Beim Öffnen einer SenseMS wird dieses Bild der eigentlichen Nachricht vorgelagert. Neben Informationen, die auch bei einer normalen SMS angezeigt werden, wie der Name des Absenders und Empfangszeit und –datum, kommen noch weitere hinzu. Der Hintergrund der Briefmarke wird durch ein Bild der momentanen Umgebung des Senders geschmückt. Der Avatar rechts unten auf der Briefmarke repräsentiert den Absender und dessen momentane Gefühlslage durch seinen Gesichtsausdruck [4].

Auf der rechten Seite sieht man eine Autovervollständigingsfunktion. Zusätzlich zu den bereits existierenden werden dem Nutzer hier, abhängig von der Eingabe, dazu passende Emoticons angeboten. Entsprechende der damit verbundenen Intention kann außerdem zwischen verschiedenen Formatierungen gewählt werden, welche die Bedeutung des Wortes unterstreichen sollen.

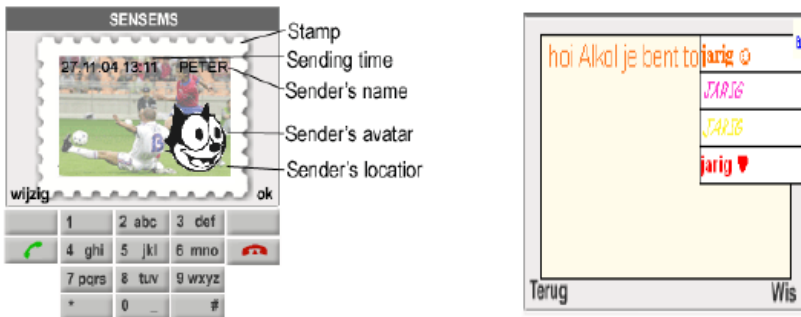


Abbildung 3. SenseMS: Erweiterte Funktionalität gegenüber der SMS [4].

Eine abschließende Evaluierung des SenseMS-Prototypen ergab, dass es damit dem Empfänger leichter möglich ist die Intention des Senders richtig zu interpretieren. Vor allem negative Emotionen waren leichter zu erkennen. Die Testpersonen gaben an, dass das Verschicken von SenseMS wegen der Möglichkeit Kontext und Emotionen zu übertragen Spaß und Sinnmacht. Allerdings wurde festgestellt, dass SMS und SenesMS für verschiedene Zwecke eingesetzt werden. Eine SMS wird bevorzugt, wenn keine weiteren Informationen als der Text nötig sind [4].

Dieser Ansatz, fehlende Teile der Kommunikation beim Messaging in die Nachricht einzubauen, bereichert das „Gespräch“ per SMS. Die SenseMS stellt also eine Schritt hin zu einer besseren Kommunikation dar [4].

Interaktion

Wenn neue Technologien Einzug in den Alltag erhalten, dann entstehen neue Verhaltensmuster, wie Personen diese bedienen. Zwei Aspekte haben hierauf Einfluss. Zum Einen wird Verhalten von ähnlichen, vertrauten Anwendungen teilweise übertragen und zum Anderen entstehen neue Verhaltensweisen, welche die

Neuerung hervorruft. In [21] wird beispielhaft untersucht, welche Auswirkung die Einführung der MMS auf die Interaktion per Messaging hat. Der entscheidende Unterschied zur SMS und damit auch zu deren Kommunikation ist, dass auch multimediale Daten übertragen werden können. Am Beispiel von Bildern, die per MMS versendet wurden, wird dargestellt, dass diese „Neuerung“ (im Vergleich zur SMS) durchaus starken Einfluss auf den Verlauf des „Gesprächs“ haben kann [21].

Auch wenn sich die Ausführung in [21] lediglich auf die MMS und den Einfluss von Bildern konzentriert kann man doch erkennen wie neue Möglichkeiten (durch neue Technologien) das Verhalten der Personen, bei der Interaktion untereinander, beeinflussen.

Den Einfluss von Bildern auf den Kommunikationsverlauf werde ich nun an einem Beispiel deutlich machen. Abbildung 4 zeigt die Nachricht welche die Kommunikation startet. Betrachtet man den Text für sich alleine, kann ist die Interpretation der Nachricht nahezu unmöglich. Das Bild alleine hingegen gibt schon mehr Informationen preis als der Text, welcher eventuell gar nicht benötigt wird um die Intention zu verstehen. Die Interpretation der komplette Nachricht hingegen fällt leicht, der Sender hat geheiratet [21].

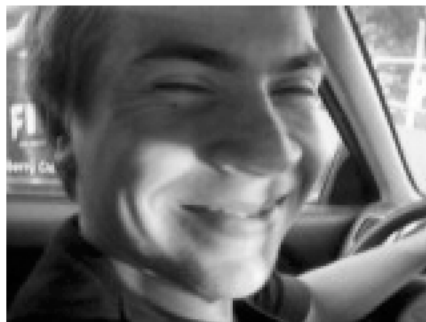
From Thomas to all: Tuesday 16:42 (+ 0:00 h)



It took 15 years ! But good things come to those who wait.

Abbildung 5. Erste Nachricht der Kommunikation [21].

From Jarkko to Thomas: Tuesday 21:27 (+ 4:45 h)



Congratulations! Jarkko, Leo and Topi

Abbildung 4. Antwort auf die Nachricht aus **Abbildung 5** [21].

Die in Abbildung 4 zu sehende Antwort enthält eine Information doppelt. Sowohl das Bild, als auch der Text „Congratulations!“, zeigen die Freude des Senders über die Heirat. Der Text fügt jedoch noch weitere Personen der Kommunikation hinzu, die durch das Bild nicht einbezogen worden sind [21].

Nachdem nun klar sein sollte, dass die Bilder einen Einfluss auf die Kommunikation haben, sollen im Folgenden die wichtigsten Ergebnisse der Analyse aus [21] vorgestellt werden:

- Bilder sind interaktiv: Die Bedeutung eines Bildes ist nicht eindeutig und kann so für verschiedene Intentionen genutzt werden. Außerdem ermöglichen Bilder an sich schon einen gewissen Grad an Kommunikation.
- Zu Zeigen bedeutet zu berichten und zu erzählen: Mit Bildern können Dinge ausgedrückt werden, die mit Text nicht darstellbar sind.
- Bilder sind co-gegenwärtige Objekte: Bilder können müssen aber nicht den Fokus der Nachricht enthalten. Es kann durchaus sein, dass der Text alleine die volle Information enthält.

Nachdem nun darauf eingegangen wurde, wie Bilder die Kommunikation verändern können, soll nun die Verbindung zum „Mobile Multimedia Messaging“ hergestellt und Ergebnisse darauf übertragen werden.

Ebenso wie Bilder können andere multimediale Daten die Kommunikation bereichern und Dinge darstellen, die mit reinem Text nicht möglich wären. Die Möglichkeit multimediale Inhalte zu übertragen hat erst dadurch Einfluss auf die Interaktion, weil es eine weitere Annäherung des Messaging an ein „normales“ Gespräch von Gesicht zu Gesicht ermöglicht. Ein Bild per Nachricht wäre beispielsweise damit vergleichbar jemandem ein Foto zu zeigen. Bei beiden Varianten kann das Bild/Foto das Gespräch (die Kommunikation) Einfluss auf die weitere Kommunikation nehmen [21].

Abschließend lässt sich feststellen, die Möglichkeit multimediale Daten in Nachrichten mit zu übertragen, hilft die Kommunikation zu verbessern, indem neue Elemente diese bereichern.

3.2 Adaption

Die Adaption oder Anpassung von Inhalten an das momentane Endgerät spielt erst im Zusammenhang mit der MMS eine wirkliche Rolle. Da die SMS nur Text enthalten kann, hier auch bei der Interoperation bspw. zwischen Mobiltelefon und Email keine größeren Probleme auf. Ganz anders verhält es sich im Bezug auf die MMS, da diese zusätzlich zu Text noch eine Vielzahl andere Datentypen enthalten kann. Die MMS ist die Basis für verschiedene Typen von Applikationen, dazu gehören [20]:

- „Mobile to Mobile“-Anwendungen (Empfangen und Versenden von Photos)
- „Web applications to mobile“ (Nachrichten, Push-Dienst)
- Programme für den Internetzugriff auf und von mobilen Geräten (Emailempfang und – versand)

Alle diese Anwendungen werfen bei dem nötigen Zusammenspiel verschiedener Geräte Probleme auf und dies wird durch die große Divergenz und den schnellen Fortschritt in diesem Bereich noch verschärft [20].

Gerade die Interoperabilität spielt jedoch die entscheidende Rolle bei der Nutzung und hat somit einen entscheidenden Einfluss auf den Erfolg dieser neuen Dienste. Serverseitige Adaption von multimedialen Inhalten (MMA = Multimedia Messaging Adoption) ist eine Technologie, die zum Ziel hat Probleme bei der Interoperabilität zu beheben. Die folgenden Ausführung beschränken sich auf eine High-Level-Ansicht eines solchen Systems [20].

Die Anpassung von Inhalt kann über zwei Wege realisiert werden. Dies ist zum Einen das Transcoding (= Codeübersetzung) und zum Anderen die „Content Selection“ (= Auswahl von Inhalt) [20].

Betrachten wir zunächst das Transcoding. Die Idee dahinter ist einfach die, den angeforderten Content für das empfangende Gerät entsprechend anzupassen und zwar indem man den Inhalt umwandelt. Als Beispiel kann hier die Veränderung der Auflösung und die anschließende Konvertierung in ein unterstütztes Format genannt werden. Diese Lösung des Adaption-Problems birgt jedoch einige Schwächen. Die Codeumwandlung ist sehr rechenintensiv und belastet daher die CPU. Dadurch sinkt dann auch deutlich die Anzahl der Nachrichten die in einer gewissen Zeit abgearbeitet werden können. Durch die automatische Umwandlung der Inhalte kann es außerdem passieren, dass der Inhalt danach nicht mehr brauchbar ist. Weiterhin können Nebeneffekte der Transformation die Qualität mindern [20].

Der zweite Lösungsansatz, die „Content Selection“, versucht das Problem anders anzugehen. Die Idee ist, dass ein und der selbe Inhalt in verschiedenen Versionen zur Verfügung gestellt werden und dann der passende zur Wiedergabe ausgewählt wird. Ein Bild könnte beispielsweise hochauflösend und farbig oder klein und schwarz-weiß vorhanden sein. Je nach Gerät wird dann der besser passende Inhalt ausgewählt und zur Verfügung gestellt. Gegenüber dem „Transcoding“ besitzt die „Content Selection“ einige Vorteile. Bessere Kontrolle über die zur Verfügung stehenden Inhalte und sehr viel geringerer Rechenaufwand seinen in diesem Zusammenhang genannt. Es gibt jedoch eine gravierende Schwäche und zwar den großen Aufwand der im Vorfeld betrieben werden muss um verschieden Versionen eines Inhalts zu erstellen. Momentan bieten nur webbasierte Applikationen die Eigenschaft an ihren Inhalt der „Umgebung“ (Browserabhängig, ohne/ mit Javascript) entsprechend anzupassen [20]. (Die SMIL „Content Control“-Funktion ist ein Beispiel für eine Technologie, welche dieses Konzept unterstützt.)

Heutzutage wird trotz der offensichtlichen Vorteile des „Content Selection“-Ansatzes zur Adaption von multimedialen Inhalten fast ausschließlich das „Transcoding“ in den MMSC's der Service Provider eingesetzt. MMSC steht für „multimedia message service center“ und bezeichnet den Server, der vom Nutzer über das kabellose Netz angesprochen wird und welcher die Adaption der Inhalte realisiert [20].

Das Design eines Systems, welches „Transcoding“ oder „Content Selection“ unterstützt ist komplex. Insbesondere muss nicht nur die Entwicklung der Adaptionskomponente beachtet werden, sondern auch die Nachrichtenlogik. Regeln,

die den Prozess am Server steuern müssen alle denkbaren Fälle mit dem bestmöglichen Ergebnis zusammen bringen [20].

Zusammenfassend ist also festzustellen, dass es bereits Systeme und Lösungen für die Adaption gibt, diese aber noch nicht ausgereift sind. Meist wird für die Interoperabilität auf das „Transcoding“ zurückgegriffen welches es ermöglicht Inhalte wiederzuverwerten. Abschließend sei noch darauf hingewiesen, dass die OMA (Open Mobile Alliance) die Arbeit an einem „Transcoding“-Interface ins Leben gerufen hat, um eine einheitliche Schnittstelle zwischen Server und Client zu schaffen [20].

Ein interessanter Lösungsansatz um multimediale Inhalte adaptiv anzupassen ist die M³-Box (Multi Media Message Box) aus [5]. Dieses Forschungsprojekt hat zum Ziel Content an einer zentralen Stelle, in der M³-Box, abzulegen. Verschiedene mobile und standortgebundene Geräte können dann auf die dort hinterlegte Information zugreifen. Abbildung 6 zeigt das grundsätzliche Vorgehen einer solchen Lösung [20].

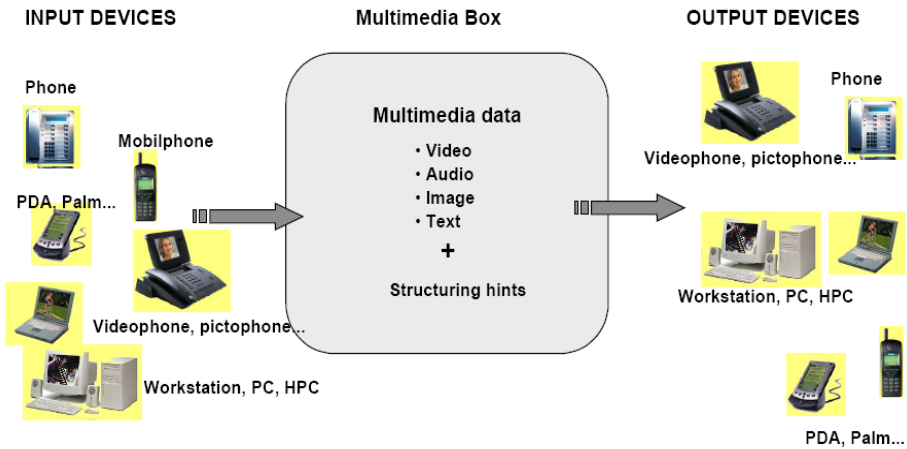


Abbildung 6. Anwendungsszenario für die M³-Box [20].

Die M³-Box bedient sich der MPEG-7 Technology. Dabei spielen vor allem die MPEG-7 Descriptions (=Beschreibungen) eine wichtige Rolle. Sie sind die Grundlage um die Daten systematisch in der M³-Box abzulegen, und um auf diese später wieder effektiv zugreifen zu können. Die Strukturinformation kann von den verschiedenen Endgeräten anhand des MPEG-7 Standards erstellt werden. Sie können Daten, wie das Gerät von welchem dieser Inhalt stammt, enthalten. In der Mitte von Abbildung 6 sieht man, dass sowohl der Inhalt an sich, als auch die wichtigen zusätzlichen Informationen (Descriptions) in der M³-Box abgelegt werden. Wird nun eine Anfrage für einen bestimmten Content an die Box gestellt, wird bei der Initialisierung der Verbindung (Handshake) von dem anfragenden Endgerät Information über die darstellbaren Daten übertragen. Diese werden dann in der M³-Box verarbeitet und ermöglichen es an das Endgerät angepassten Content auszugeben. Hierbei wird der Ansatz des „Transcoding“ ebenso wie der der „Content Selection“ verwendet. Zunächst wird nach einer geeigneten Repräsentation des Inhalts gesucht,

ist keine vorhanden, wird er Content entsprechen umgewandelt. Es bietet sich hierbei beispielsweise an, exotische Formate gleich offline in Verbreitete umzuwandeln, um die aufwändige Prozedur nicht erst bei der Anfrage machen zu müssen. Die Funktion der M³-Box wird durch vier Hauptkomponenten realisiert, siehe dazu Abbildung 7 [20].

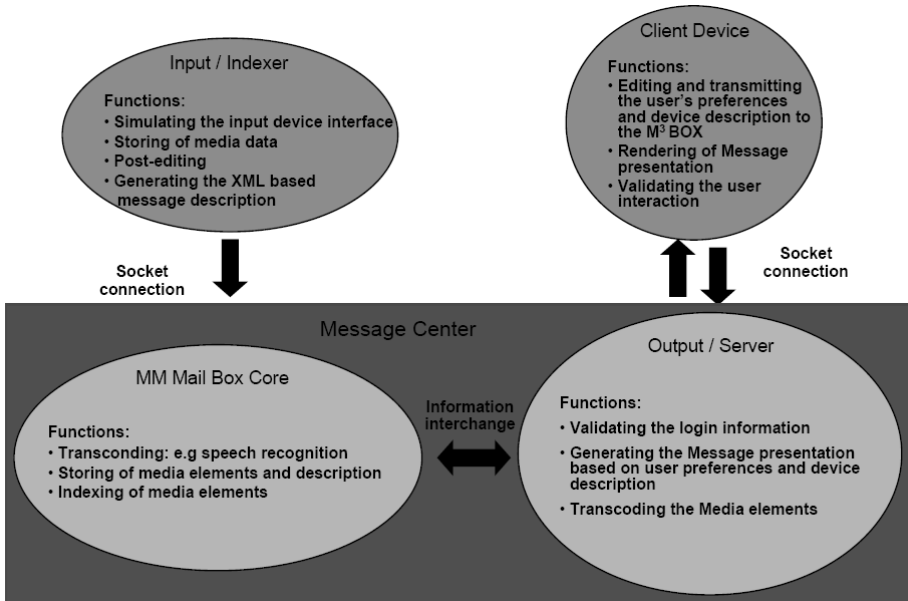


Abbildung 7. Funktionale Elemente der M³-Box [20].

Abschließend kann also festgestellt werden, dass es zwei Ansätze zur Adaption, den der „Content Selection“ und den der „Transcodierung“, gibt. Systeme zur Anpassung verwenden hauptsächlich Zweiteren. Die M³-Box stellt ein gutes Beispiel für solch eine Lösung dar.

4 Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde der Bereich des „Mobile Multimedia Messaging“ untersucht. Nachdem eine Einteilung des Forschungsgebietes in verschiedene thematische Bereiche stattgefunden hat, wurden im zweiten Teil der Ausarbeitung dann zwei besonders interessante Aspekte ausführlich dargestellt.

Bei der Identifizierung der verschiedenen Teilgebiete des „Mobile Multimedia Messaging“ haben sich viele interessante Themen finden lassen. Vor allem die Push-

Dienste wären ein weiterer interessanter Bereich, konnten aber leider im Rahmen dieser Arbeit nicht wie die zwei Ausgewählten näher betrachtet werden. Das Paper [6] bietet hierzu weitergehende Informationen.

Der Hauptteil beschäftigte sich dann mit zwei Themengebieten: Soziale Aspekte und Anpassung von Inhalt (Adaption).

Zunächst wurde der Einfluss von mobilem Messaging auf soziales Verhalten beleuchtet. Dabei wurde zunächst auf Aspekte der Privatsphäre eingegangen. Probleme und Lösungen für das Verlorengelassen von Emotionen bei der Kommunikation per SMS bzw. MMS wurden dann im zweiten Part erörtert. Geschlossen wurde dieses Kapitel dann durch einen Blick auf die Auswirkungen die multimedialen Inhalte auf die Interaktion über Nachrichten haben [3,4].

Der letzte Teil der Arbeit beschäftigt sich dann damit, wie Inhalte den genutzten Endgeräten entsprechend angepasst werden können (Adaption von Inhalten) [5, 20].

Es ist davon auszugehen, dass alle in dieser Arbeit genauer betrachteten Bereiche Einfluss auf zukünftige Messaging-Lösungen für mobile Endgeräte haben werden. Als besonders sinnvoll erachte ich die in [4] vorgestellten Funktionen der SenseMS, um Emotionen in die Kommunikation einzubinden. Kritisch würde ich hingegen Lösungen wie die M³-Box aus [5] einstufen. Die Idee dahinter, Inhalte zentral abzulegen, um von verschiedenen Geräten darauf zugreifen zu können, ist zwar meiner Meinung nach grundsätzlich sinnvoll, aber momentan lassen sich die damit verbundenen Aufwendungen (Kosten, Einrichtung, Wartung) sicher nicht rechtfertigen. Außerdem stellt sich die Frage, für wie viele Personen diese Entwicklung überhaupt interessant ist, insbesondere die normalen Mobiltelefonnutzer werden kaum Verwendung dafür haben. Denkbar sind hier jedoch Szenarien, in denen bspw. verschiedene Mitarbeiter einer Firma Daten einspeichern und auslesen. Ein Beispiel wäre, dass ein Versicherungsvertreter den entstandenen Schaden mit dem Handy als Bild aufnimmt, zentral ablegt, und der Kollege im Büro diese Bilder dann gleich einem Sachverständigen zuführt.

Man kann zwar nicht vorhersagen in welche Richtung sich das „Mobile Multimedia Messaging“ entwickelt, aber es wird sich so oder so in der Zukunft verändern und erweiterte Funktionalität bieten.

Literatur

1. Mostafa, M.-E.. MMS - the modern wireless solution for multimedia messaging. In: Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on, Volume: 5, 15-18 Sept. 2002 Page(s): 2466 -2472 vol.5.
2. Jokela, T.. Authoring tools for mobile multimedia content. In: Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on, Volume 2, 6-9 July 2003 Page(s): II - 637-40 vol.2.
3. Häkkinlä, J., and Chatfield, C. (2005). 'It's Like If You Opened Someone Else's Letter' -- User Perceived Privacy and Social Practices with SMS Communication. In: Proc. of MobileHCI 2005.
4. Amin, A. K., Kersten, B. T. A., Kulyk, O.A., Pelgrim, E., Wang, J., Markopoulos, P. The SenseMS: Enriching the SMS experience for Teens by Non-verbal Means. In: Proceedings of Tenth IFIP TC13 International Conference on Human-Computer Interaction (INTERACT), Lecture Notes in Computer Science, Volume 3585, 2005, pp. 962-965.
5. Heuer, J., Casas, J.L., Kaup, A. Adaptive Multimedia Messaging Based on MPEG-7 - The M3-Box. In: Proc. Second International Symposium on Mobile Multimedia Systems & Applications, pp. 6-13, Delft, Nov. 2000.
6. Aalto, L., Göthlin, N., Korhonen, J. & Ojala, T. Bluetooth and WAP Push based location-aware mobile advertising system. In: Proc. Second International Conference on Mobile Systems, Applications and Services, Boston, MA, 49 – 58, 2004.
7. Aust, J. Mobile Interaktion mit der realen Umwelt. In: Mobile Interaction and Mobile Media. München, Germany, Juli 2005.
8. Wikipedia. Mobile Endgeräte. 16.06.2006.
http://de.wikipedia.org/wiki/Mobile_Endger%C3%A4te
9. Wikipedia. Mobilfunknetz. 16.06.2006.
<http://de.wikipedia.org/wiki/Mobilfunknetz>
10. Message-oriented vs. Stream-oriented Communication. 16.06.2006.
http://www.der1.at/fileadmin/documents/teaching_archive/acsp0405/02_Witting_Ausarbeitung.pdf
11. Wikipedia. Mobility Management. 16.06.2006.
http://en.wikipedia.org/wiki/Mobility_management
12. Yuguang, F. Movement-based mobility management and trade off analysis for wireless mobile networks. In: Computers, IEEE Transactions on Volume 52, Issue 6, June 2003 Page(s): 791 – 803.
13. Wikipedia. Quality of Service. 16.06.2006.
<http://de.wikipedia.org/wiki/QoS>
14. Quality of Service. 16.06.2006
<http://www.voip-information.de/voip-qualitaet.html>
15. Ling, R., Julsrud, T. The development of grounded genres in multimedia messaging systems (MMS) among mobile professionals. In: The Global and the Local in Mobile Communication - Places, Images, People, Connections. Conference of Hungarian Academy of Sciences. Budapest, Hungary, June 2004.
16. Podnar, I., Hauswirth, M., Jazayeri, M. Mobile Push: Delivering Content to Mobile Users. In: Proceedings of the International Workshop on Distributed Event-Based Systems in conjunction with the 22nd International Conference on Distributed Computing Systems, 2002.
17. Pan, P., Kastner, C., Crow, D., Davemport, G. M-Studio: An Authoring Application for Context-Aware Multimedia. In: Proceedings of the 2002 ACM Workshops on Multimedia, pp. 351-354, 2002.
18. Java Specification Request 135. Mobile Media API. 16.06.2006.
<http://www.jcp.org/en/jsr/detail?id=135>

19. Priyantha, N., Miu, A., Balakrishnan, H., Teller, S. "The Cricket Compass for Context-Aware Mobile Applications". In: ACM SIGMOBILE (2001): "Mobile Computing and Networking", 7th Annual International Conference on Mobile Computing and Networking, Rome 2001, New York: Association for Computing Machinery, S. 1-14.
20. Coulombe, S., Grassel, G. Multimedia Adaptation for the Multimedia Messaging Service. IEEE Communications Magazine, vol. 42, no. 7, pp. 120-126, 2004.
21. Kurvinen E. Only when miss universe snatches me: teasing in MMS messaging, Proceedings of the 2003 international conference on Designing pleasurable products and interfaces, June 23-26, Pittsburgh, PA, USA, 2003.

Autorenverzeichnis

Attenberger, Andreas	54
Drasch, Lucie	153
Emrich, Tobias.....	35
Fritsche, Torsten	95
Hilkert, Daniel	303
Jüngst, Johannes	77
Kauntz, Gert.....	249
Lederer, Thomas	229
Lindner, Maximilian.....	9
Pahre, Christoph	285
Rabinowitsch, benjamin	267
Richter, Hendrik	115
Ruch, Claudia	191
Seitz, Stefan	135
Ticaloiu, Ana Monika.....	175
Zumbruch, Ina.....	211

