Abteilung für Entwicklungspsychologie und Pädagogische Psychologie des Instituts für Psychologie der Uni Wien

Brigitte Wagner:

Online Research in der Psychologischen Forschung

Teil 5: Nonreaktive Erhebungsmethoden

Skriptum zur Lehrveranstaltung
Online Research in Theorie und Praxis

Leitung:
Ass.Prof. Dr. Marco Jirasko
(Überarbeitung der vorliegenden Textbasis)

Studienjahr 2001/02

5. Nonreaktive Untersuchungsverfahren

Wie bereits darauf hingewiesen wurde, beinhalten nonreaktive wie auch reaktive Forschungsmethoden sowohl qualitative als auch quantitative Ansätze. Während bei den reaktiven Verfahren der Schwerpunkt auf quantitative Verfahren gelegt wurde, gehen hier Kapitel 5.1 und 5.2 verstärkt auf qualitative Ansätze ein. Weiters werden die Möglichkeiten und Einsatzbereiche von Beobachtungsstudien dargestellt, wobei darauf hinzuweisen ist, dass Beobachtungen selbstverständlich auch reaktiv erfolgen können, was jedoch meist nicht im Sinne der Untersuchung wäre.

5.1 Definition nichtreaktiver Untersuchungsverfahren

Bereits 1975 widmeten sich Webb et al. in ihrer Arbeit den nichtreaktiven Meßverfahren bzw. verhaltensorientierten Untersuchungsmethoden (unobtrusive measures). Sie wollten bereits damals Alternativen zu den reaktiven Methoden aufzeigen und deren Forschungsartifakten entgehen (Fehler durch die Respondenten, z. B. Antworthaltungen; Fehler durch Forscher, z. B. Interviewereffekte). Das Bekanntwerden dieser Art von Forschung ist jedoch sicherlich auf Milgram zurückzuführen, dessen Lost-Letter-Technik im Laufe dieses Kapitels erörtert werden wird.

Bortz und Döring (1995, S. 300) definieren nichtreaktive Verfahren als Datenerhebungsmethoden, *„die im Zuge ihrer Durchführung keinerlei Einfluß auf die untersuchten Personen, Ereignisse oder Prozesse ausüben“*.

Weiters treffen sie über die Interaktion zwischen Forscher und Untersuchungsobjekt folgende Aussage, aus der bereits die Vorteile dieser Erhebungsmethode hervorgehen: „*Bei nichtreaktiven Verfahren treten der Beobachter und die untersuchten Personen nicht in Kontakt miteinander, so dass keine störenden Reaktionen, wie Interviewer- oder Versuchsleitereffekte, bewußte Testverfälschung oder andere Antwortverzerrungen auftreten können*“ (Bortz & Döring, 1995, S. 300).

Darüber hinaus stellt diese Unwissenheit der Untersuchungsobjekte darüber, Teil einer Studie zu sein, ein wesentliches Charakteristikum nichtreaktiver Erhebungsmethoden dar. Die Versuchspersonen wissen folglich nicht, dass sie bzw. ihr Verhalten exploriert wird. Wie groß das Spektrum dieser Forschungsmethode ist bzw. welche Grundlagen für nichtreaktive Forschung herangezogen werden können, führt Friedrichs (1990) wie folgt auf:

- *Physische Spuren*

  So können beispielsweise abgetretene Teppichspuren in Museen Aufschluß über häufig gewählte Besucherwege geben, eingestellte Sender in Autoradios können registriert werden und als Indikator für die Popularität einzelner Radiostationen dienen etc.

- *Schilder, Wegweise etc.*

  z. B. können etwa „Spielen verboten“-Schilder als Indiz für Kinderfeindlichkeit interpretiert werden.

- *Lost-Letter-Technik*

  im Kapitel 5.2.1 „Anwendungen ethnographischer Forschung im Internet“ wird ausführlich über diese, durch Milgram bekannt gewordene, Technik berichtet.

- *Archive und Verzeichnisse*

  z. B. können Adoptionsstatistiken Hinweise auf schichtspezifische Bevorzugungen von Jungen oder Mädchen geben; Unfall-, Krankheits-, oder Fehlzeitenstatistiken können zur Ermittlung von Arbeitsplatzzufriedenheit oder Betriebsklima herangezogen werden u. v. m.

- *Verkaufsstatistiken*

- *Einzeldokumente*

  z. B. Leserbriefe, Tagebücher, Sitzungsprotokolle, Gebrauchsanweisungen usw. können als Unterlagen für Inhaltsanalysen herangezogen werden.

Weiters können folgende Materialien zur Analyse herangezogen werden:

- *Bücher, Zeitschriften, Filme und andere Massenmedien*

  z. B. Anzahl der Nennungen weiblicher Personen in Geschichts- oder Schulbüchern als Indikator für eine männerorientierte Sichtweise.

- *Symbole*

  z. B. Autoaufkleber, Abzeichen, Buttons und Anstecker als Hinweise auf Gruppenzugehörigkeit bzw. soziale Identität des Trägers.

5.2 Freie Beobachtung

Bortz und Döring (1995, S. 248) definieren freie Beobachtung folgendermaßen:

„Bei einer freien (offenen, unstandardisierten, qualitativen) Beobachtung verzichtet man in der Regel auf die Vorgabe von Beobachtungsrichtlinien. Sie kommt vor allem für Untersuchungen in Betracht, mit denen ein bislang weitgehend unerforschtes Gebiet erkundet werden soll“.

Sie sprechen weiters davon, dass ausgearbeitete Beobachtungspläne hier nicht sinnvoll wären, da diese die Aufmerksamkeit des Untersuchungsobjektes auf Details lenken könnte und somit nur die Offenheit reduzieren würde, was nicht im Sinne dieser Forschungsmethode wäre.

Somit handelt es sich hierbei eher um subjektive Erfahrungsberichte als um systematische, wissenschaftliche Forschung, was Döring (1999, S. 173) wie folgt beschreibt: „*Bei der freien Beobachtung ist die Aufmerksamkeit nicht von vornherein auf bestimmte Merkmale oder Ereignisse fokussiert, sondern richtet sich je nach momentaner Befindlichkeit, individuellen Interessen und Vorannahmen auf jeweils unterschiedliche Dimensionen des aktuellen sozialen Geschehens. Prozesse der Selektion, Dokumentation und Interpretation der beobachteten Netzkommunikation verlaufen unkontrolliert, die Erlebnisse werden unter Rückgriff auf Alltagstheorien gedeutet*“.

Oftmals verlangt eine (psychologische) Fragestellung jedoch nach anderen methodischen Kriterien, als die Methode der freien Beobachtung je im Stande wäre zu erbringen. Dennoch hat diese Methode ihre Berechtigung, nicht nur innerhalb eines Methodenmixes als vielmehr als heuristische Funktion der Generierung von Forschungshypothesen aber auch zur Illustration theoretischer Modelle.

5.3 Feldbeobachtung (ethnographische Beobachtung)

Im Gegensatz zur eben beschriebenen freien Beobachtung, erfolgt die Feldbeobachtung systematisch, basiert bereits auf methodischen Regeln und kann somit als kontrollierte Datenerhebungsmethode gewertet werden, die sich wie folgt charakterisieren läßt:

„*Sie findet in der Regel teilnehmend und verdeckt statt. Die Forschenden agieren demnach beispielsweise in Mailinglisten, Newsgroups, Chat-Channels oder MUDs, wobei sie versuchen, sich dem Feld möglichst unauffällig anzupassen und die sozialen Prozesse durch ihre Forschungstätigkeit nicht zu beeinträchtigen*“ (Döring, 1999, S. 174).

Oftmals greift man in diesem Kontext zusätzlich auf Inhaltsanalysen zurück um die Resultate der Feldbeobachtung zu komplettieren. Gegenstand dieser Analyse können beispielsweise FAQs sein, als auch Geschehnisse in MUDs oder Chat Rooms.

Die Bedeutung dieser Variante des Methodenspektrums liegt nach Döring (1999) in der Möglichkeit „*größere soziale Gebilde wie etwa einzelne Mailinglisten, Newsgroups, Chat-Channels und MUDs ganzheitlich in ihren Strukturen und Prozessen aus der Perspektive der Beteiligten zu beschreiben“.*

5.3.1 Exkurs: Klassisches Lost-Letter-Experiment

Bevor auf die internetspezifischen Anwendungsmöglichkeiten eingegangen wird, soll kurz das wohl bekannteste Experiment dargestellt werden, das 1965 erstmals durchgeführte Lost-Letter Experiment von Milgram: Dabei wurden in einem Stadtgebiet bereits adressierte und frankierte Briefe ausgelegt, wobei die jeweiligen Finder den Eindruck gewinnen sollten, diese wären verloren gegangen. Adressaten dieser Briefe waren entweder fiktive Organisationen oder Institutionen (z. B. Parteien, Tierschutzvereine etc.). Werden diese Brie-

fe von Passanten schließlich in einen Postkasten geworfen, erreichen diese natürlich nicht die angegebene Adresse sondern die Versuchsleiter. Milgram wollte auf diesem Weg anhand der jeweiligen Rücklaufquote das Image der genannten Institutionen eruieren (Bortz & Döring, 1995). Was jedoch zu überlegen bleibt, ist das Faktum, dass hier Verhalten mit Einstellung konfundiert ist.

5.3.2 Lost-Letter-Technique im Internet

Basierend auf Milgrams Idee, Einstellungen zu beobachten anstatt diese zu messen, nutzten 1996 Stern und Faber (1997) das Medium e-Mail und versuchten anhand dessen, die Methode des Lost-Letters online zu replizieren. Stern und Faber führten im Rahmen dieser Untersuchung zwei Experimente durch. In beiden experimentellen Designs werden die studentischen Adressaten in dem Glauben gelassen, sie würden die e-Mails zweier fremder Studenten zu lesen bekommen, hätten folglich fälschlich zugestellte e-Mails empfangen. Der Inhalt dieser elektronischen Mitteilung variiert hinsichtlich der Wertschätzung bezüglich des entsprechenden Colleges respektive der Dringlichkeit einer Antwort. Registriert wurde schließlich die Anzahl der an die vermeintlich intendierte e-Mail Adresse weiter gesandten Nachrichten bzw. ob und welche Kommentare beigefügt worden sind.

Untersuchungsteilnehmer waren 79 Studenten eines kleines Colleges in Philadelphia; folgende Hypothesen wurden überprüft:

- *Enthält das vermeintlich falsch zugestellte e-Mail sozial angemessene Aussagen über das College wird dieses e-Mail eher dem „wahren Empfänger“ zugestellt.*

  Sozial angemessene Aussagen wurden operationalisiert als Grad der anerkennenden Äußerungen über das College:

  (1) „I’m really glad that I ended up staying in school. I think I would have made a big mistake by dropping out“ vs. **(2)** „Things haven’t gotten much better. I’m just not sure that school is for me“.

- *Sind im e-Mail Angaben über Dringlichkeit von Antworten gegeben, werden diese eher weiter gesandt als andere:*

  (1) „We are going to the hospital next week“. vs. **(2)** „We are going to the hospital this afternoon“.

Der Absender des e-Mails war mittels des verwendet Student-Accounts als Student des entsprechenden Colleges erkennbar.

Folgende Resultate wurden beobachtet: Zwischen den verschiedenen Untersuchungsbedingungen konnten keine signifikanten Unterschiede gefunden werden. Insgesamt retournierten nur 19 % der Versuchspersonen (n=15) das empfangene e-Mail an den Absender, keine einzige Nachricht wurde dem angegebenen Empfänger weitergeleitet.

Die beiden oben beschriebenen Hypothesen konnten also nicht bestätigt werden. In allen vier Versuchsbedingungen wurden den retournierten e-Mails jedoch eigene Kommentare zugefügt.

Auch im zweiten experimentellen Design wurden die Adressaten im Glauben gelassen, fälschlicherweise die Konversation zweier Fremder zugestellt zu bekommen. Diesmal wurde versucht die Einstellung zum damaligen (sehr unpopulären) amerikanischen Präsidentschaftskandidaten Ross Perot zu erheben. Die e-Mails mit entsprechendem positiven vs. negativen Aussagen zu Perot wurden an insgesamt 200 Versuchspersonen gesandt, welche aus den sogenannten „white papers“ (entspricht einer großen Sammlung von e-Mail Adressen im WWW) rekrutiert wurden. Das experimentelle Design stellte sich wie folgt dar: Eine Versuchsgruppe erhielt ein e-Mail mit einem Spendenaufruf für den Wahlkampf von Perot, die Kontrollgruppe erhielt die entsprechenden e-Mails ohne diese Erwähnung. Anders als im ersten Experiment wurde hier die e-Mail Adresse des vermeintlichen Adressaten nicht genannt, gemessen wurde lediglich die Anzahl der an den Absender retournierten e-Mails.

Diesmal wurden insgesamt 29 % (n=58) Nachrichten zurückgesandt. Aber auch in diesem Experiment zeigten sich keine signifikanten Unterschiede zwischen den beiden Versuchsbedingungen (Anzahl der retournierten e-Mails: 31 % (Kontrollgruppe) vs. 27 % (Versuchsgruppe).

In weiterer Hinsicht wurden die Kommentare in den retournierten e-Mails klassifiziert und gerated (in eher positive, neutrale und eher negative Aussagen über Perot, die Politik im allgemeinen oder das e-Mail im speziellen). Es überrascht jedoch kaum, dass negative Kommentare über den Präsidentschaftskandidaten überwogen haben.

Wie auch schon die Autoren des genannten Experimentes schlußfolgerten, waren die polarisierenden Inhalte der e-Mails zu schwach um entsprechende Effekte erzielen zu können. Stern und Faber (1997) begnügten sich folglich mit der Aussage, dass diese Methode zur Exploration von Verhalten und Erwartungen sich wesentlich von der ursprünglichen Methode der Lost-Letter Technik unterscheidet und erklären sich die Ergebnisse unter anderem damit, dass das Weiterleiten von e-Mails mehr Aufwand bedeutet als das simple Betätigen des „Reply“- Buttons des e-Mail Programms.

Dieses Experiment stellt jedoch einen kreativen Ansatz dar, alte Untersuchungstechniken in Verbindung mit neuen Erhebungsmedien zu bringen.

5.4 Automatisierte Beobachtung (geplante Beobachtung)

In Bezug auf die Anwendung dieser Methode im Internet stellt Döring (1999) folgende Definition auf: „*Bei der automatischen Beobachtung von computervermittelter Kommunikation macht man es sich zunutze, dass im Netz sämtliche zwischenmenschlichen Kommunikati-*

on- und Interaktionsakte digital erfolgen und somit automatisch und vollständig registrierbar sind".

Folglich bedient sich die automatische Beobachtung jener digitalen Spuren, die jeder Internet Nutzer im WWW hinterläßt.

Warum die nichtreaktive Erhebung gerade im Internet einen besonderen Stellenwert genießt ist wohl darauf zurückzuführen, dass sich dieses Medium besonders für diese Forschungsvarianten anbietet. Einerseits fallen automatisch Datenmengen an (z. B. Logfiles), andererseits ist das Erheben von technischen Umgebungsvariablen (wie Browser-Typ, Betriebssystem, Bildschirmauflösung etc.) bzw. das Aufzeichnen von personenbezogenen Datenmengen (z. B. via Cookies) nicht allzu aufwendig. Hinsichtlich psychologisch relevanter Fragestellungen gestaltet sich die Aufbereitung der Daten jedoch schwieriger als die anfängliche Euphorie über den doch sehr leichten Datenzugang vermuten läßt. Schließlich handelt es sich speziell bei Logfiles um eine große Datenmenge die zumeist nicht auf individueller Ebene auszuwerten ist, wie das in der psychologischen Forschung häufig der Fall ist. Vielmehr hat man es mit aggregiertem Datenmaterial zu tun. Anders ist es jedoch beim Einsatz von Cookies, auf die in Kapitel 5.4.3 näher eingegangen wird.

5.4.1 Logfile Analyse

Im folgenden werden nichtreaktive Erhebungsverfahren beschrieben, die sich der automatisch anfallenden Daten im Internet bedienen, den *Logfiles*. Diese Dateien enthalten, wie Janetzko (1999, S. 169) beschreibt, Kennwerte „*mit denen Zugriffe auf Web-Sites beschrieben werden. Dazu zählen IP-Adresse des aufrufenden Rechners, die Uhrzeit des Abrufs oder bei der Anfrage verwendete Methode. Logdaten lassen sich auf der Seite des Servers oder auf der Seite des Clients auf nicht-reaktive Weise erheben*".

Betreiber eines Servers verfügen jederzeit über diese anfallenden Daten und nutzen diese zumeist dazu, Aufschluß über die Funktionalität ihres Systems zu gewinnen bzw. können mittels entsprechender Software (sog. *Logfile Analyzers*) Häufigkeitsstatistiken samt der ungefähren Verweildauer der besuchten Seiten erstellt werden. Warum dies nur annäherungsweise möglich ist wird später erläutert werden. Wie ein solcher Eintrag eines Logfiles aussieht und was seine einzelnen Komponenten aussagen, wird im folgenden näher erläutert. Neben der Differenzierung zwischen *serverseitigem* und *clientseitigem* Logfile werden verschiedene Formate unterschieden (Janetzko, 1999):

- Common-Log-Format (CLF)
- Combined-Log-Format
- Agent-Log-Format
- Refferer-Log-Format
- Microsoft-IIS-Log-Format

Die angegebenen Formate unterscheiden sich in den von ihnen registrierten Daten. Im weiteren wird nur auf das CLF näher eingegangen, da dieses das gebräuchlichste darstellt. Für weitere Informationen zu den übrigen genannten Formaten siehe Janetzko (1999, S. 173 ff.).

Common-Logfiles registrieren also:

- eine allfällige **Authentifizierung** (z. B. bei paßwortgeschützten Web-Sites)
- die **Anzahl der vom Server an den Client übertragenen Bytes**
- das **Datum mit/ohne Uhrzeit der Anfrage**
- den **anfragenden Host**, sprich den Domain-Namen des Clients bzw. seine **IP-Adresse**
- die **Identifizierung des Users** auf Seiten des Clients
- die **Anfrage des Clients** in Anführungszeichen
- den **angefragten Dienst**

Die Einträge eines Common-Log-Files (CLF) können wie folgt aussehen:

124.86.56.25 - - [26/NOV/2000:20:56:08 +0200] ‘‘GET /START.HTML HTTP/1.0‘‘ 200 2048

Bedeutung dieser Angaben:

Der Rechner (Client) mit der IP Adresse 124.86.56.25 hat am 26. November 2000 um 20:56 Uhr und 8 Sekunden die Datei start.html aufgerufen. Der Aufruf dieser Seite erforderte keine Authentifizierung, deshalb findet sich an der dafür vorgesehenen Stelle jenes Zeichen (-), das für fehlende Angaben steht. Der zweite diesbezügliche Eintrag weißt auf die nicht vorliegende Client-Identifizierung hin, wobei unter dem Begriff nicht die Identifizierung einer Person oder ähnliches gemeint ist, sondern lediglich der frei gewählte Name eines Rechners (Clients) verstanden wird. Schließlich wurde bei diesem Seitenaufruf eine Datenmenge von 2048 Bytes übertragen wobei anhand des Eintrages 200 erkenntlich ist, dass dieser Transfer erfolgreich war.

Weitere diesbezügliche Status-Codes wären beispielsweise (Janetzko, 1999):

- 304 - Dokument wurde seit dem letzten Aufruf nicht modifiziert, Abruf vom Browser-Cache.
- 3xx bezeichnet Fehler die darauf hinweisen, dass weitere Aktionen zur Bearbeitung der Anfrage erforderlich sind.

- 401 - Unauthorisierter Zugriff (ID und Paßwort erforderlich).
- 404 - Dokument wurde nicht gefunden.
- 4xx steht generell für einen clientseitigen Error (z. B. Syntaxfehler in der Anfrage).

- 500 - Interner Server Error
- 5xx weißt auf einen serverseitigen Error hin (z. B. Overload des Servers).

Ist man jedoch nicht in der Lage um beispielsweise auf die Server Logfiles zurückgreifen zu können, kann man entsprechende Tools anwenden, welche ebenfalls in der Lage sind, oben genannten Kenngrößen zu erheben. Die derart gewonnenen Einträge sehen in der Regel etwas anders aus als im oben beschriebenen CLF.

Grundsätzlich muß man sich bewußt sein, dass beim Aufruf einer WWW-Seite die auch Grafiken (gif, jpg etc.) beinhaltet, diese jeweils einen gesonderten Eintrag im Logfile hinterlassen. Erst nach der Bereinigung des Rohfiles von eben dieser Datenmenge, kann eine Analyse sinnvoll durchgeführt werden.

Eine solche Analyse von Logfiles wird aber nicht nur von Systemadministratoren oder auch Forschern genutzt, sondern stellt vor allem auch im kommerziellen Bereich (und hier besonders in der Werbung oder auch im Marketing) einen wesentlichen Bestandteil der Nutzung des Internets dar. So unterscheidet man beispielsweise im Rahmen der Kontaktmessung, welche ihrerseits Relevanz bezüglich Werbeeinnahmen durch Banner etc. besitzt, zwischen folgenden Begriffen (Janetzko, 1999 & Werner, 1999):

- *PageImpressions (früher: PageViews)*

  Diese zählen zu den am leichtesten meßbaren Größen. Sie beschreiben die Anzahl der Sichtkontakte beliebiger Nutzer mit einer potentiellen werbeführenden HTML Seite. Sie liefert ein Maß für die Nutzung eines Angebotes. Welche Probleme bei der Messung dieser Größe anfallen (hinsichtlich des Einsatzes von Frames, Proxy-Server etc.) ist für interessierte Leser bei Werner (1999) nachzulesen.

- *Visits*

  Ein Visit bezeichnet einen zusammenhängenden Nutzungsvorgang (Besuch) eines WWW-Angebotes. Er definiert beispielsweise den Werbeträgerkontakt. Als Nutzungsvorgang zählt ein technisch erfolgreicher Seitenzugriff eines Internet-Browsers auf das aktuelle Angebot, wenn er von außen erfolgt.

  Oder weniger werbetechnisch formuliert: Während es sich bei PageImpressions bzw. PageViews um den Aufruf und das Betrachten von nur einer einzelnen WWW-Seite handelt, bezeichnet der Visit den Durchgang mehrerer WWW-Seiten einer WWW-Site und ist somit beispielsweise in Hinblick auf das Navigationsverhalten aussagekräftiger.

  Weiters beschreibt der Begriff des Visits keine „Besucher“ sondern lediglich „Besuche“, was Werner (1999) wie folgt problematisiert: Gemessen wird ein Übertragungszeitpunkt jedoch kein Zeitpunkt für das Verlassen einer Seite. Internet-Adressen werden mehrfach vergeben (dynamische Vergabe), somit können hinter einer IP-Adresse mehrere verschiedene Nutzer stehen.

  Somit hat auch diese eher aus der Technik heraus entstandene Definition ihre Probleme.

- *ViewTime*

  Zeitliche Dauer, in welcher sich ein Nutzer mit einer Page oder einer Website beschäftigt. Was hier gemessen wird entspricht eher einer „Betrachtungszeit“ als einer „Beschäftigungszeit“ denn man kann aufgrund einer erhobenen Zeitspanne kaum Aussagen darüber treffen, ob sich ein Besucher tatsächlich während dieser Zeit mit dem Inhalt der Seite beschäftigt hat oder einfach anderen Tätigkeiten nachging.

  Vorschläge die ViewTime zu messen stammen ebenfalls, wie die meisten dieser Begriffe und Methoden aus der Werbeforschung und sind ausführlich auf den Seiten der „Informationsgesellschaft zur Feststellung der Verbreitung von Werbeträgern e.V“. (unter http://www.ivw.de) wie auch bei Werner (1999) beschrieben.

- *AdClickRate*

  Beschreibt die Anzahl der Clicks auf einen Werbebanner im Verhältnis zu den Sichtkontakten einer Page.

- *Clickstream*

  Dieser Begriff steht für die Seitenaufrufe während eines Visits. Besonders innerhalb der Werbung wird viel investiert in die Analyse dieser Clickstreams, zumal sie Aufschluß geben sollen über die Nutzungsweisen von Besuchern respektive Kunden. Im Zusammenhang mit den Begriffen „Data-Mining“[1] bzw. „Profiling“[2] gibt es unzählige diesbezügliche Ansätze den Besuchern einer WWW-Site möglichst maßgeschneiderte Inhalte liefern zu können und somit das eigene Internet Angebot möglichst attraktiv gestalten zu können.

Durch obige Differenzierung der gebräuchlichen Begriffe ist vielleicht die Komplexität und auch Problematik der solcherart gewonnenen Daten ersichtlich geworden. Darüber hinaus, stellen sich neben diesen Definitionsfragen noch wesentlich grundlegendere Probleme die es zu beachten gibt bevor überhaupt erst an eine sinnvolle Aufbereitung des Datenmaterials zu denken ist.

Probleme bei der Interpretation von Logdateien

Wie bereits weiter oben erwähnt wurde, liefern diese automatisch anfallenden Daten vorrangig Systemadministratoren Hinweise über den Besucherverkehr auf ihren Seiten: Angaben über fehlgeschlagene Seitenaufrufe, Verweildauer, Ladezeiten etc. Um jedoch an den wirklichen inhaltlichen Gehalt dieser Datenmenge heranzukommen, der für eine wissenschaftliche Interpretation unerläßlich ist, muß man sich über die möglichen Fehlerquellen im Klaren sein. Janetzko (1999) listet diesbezüglich folgende Punkte auf:

[1] Bezeichnet das "Schürfen" nach bzw. Sammeln von relevanten Daten in Datenbanken.

[2] Dient beispielsweise der Zielgruppen-Segmentierung.

- Ein Benutzer kann *unzählig viele Hits* erzeugen, indem er mehrfach zu einer Seite zurückkehrt.
- Das Problem des *Cachings*:

  Darunter ist das (lokale) Zwischenspeichern von WWW–Dokumenten zu verstehen. Einerseits können häufig frequentierte Internetseiten aus Gründen der Datenreduktion auf Proxy-Servern abgelegt werden, andererseits können entsprechende Files auch lokal beim Client verbleiben und erzeugen dadurch keine wiederholten Einträge in ein Logfile.

  Es sei jedoch an dieser Stelle gleich angemerkt, dass dieses Problem technisch lösbar ist, wenn man sich grundsätzlich dessen bewußt ist. Nur kurz sei an hier ein möglicher Lösungsvorschlag umrissen: Mittels einem geeignetem HTTP header z. B. „Cache-control: proxy-revalidate“ oder „Expires: <Nennung eines bereits vergangenen Datums>. Mittels solcher Anweisungen wird der Client gezwungen, das File vom Server neu zu laden. Diese Technik ist unter dem Begriff „cache busting“ bekannt (Jackl, 1999).
- *Abschätzung der Zugriffshäufigkeiten* verschiedener Internet Domains.
- *Einsatz von „Robotern“ (Suchmaschinen etc.)*

  Diese Programme durchforsten das Internet nach Informationen und hinterlassen daher auch ihre Spuren in den Logfiles und verzerren somit das Datenmaterial.
- *Zeitverzögerte Statistiken*

  Wesentlich zu wissen ist, dass üblicherweise Logfiles nicht in Echtzeit aktualisiert werden.
- *Zeitmessungen*

  Was im Rahmen von Logfiles automatisch festgehalten wird, ist der Zeitpunkt der Übertragung eines aufgerufenen Internet Dokumentes. Wann jedoch die entsprechende Seite verlassen wurde, darauf gibt dieses File keinen Aufschluß.

Janetzko (1999) zieht in seiner Arbeit folgendes Fazit bezüglich der Brauchbarkeit von Logfiles: „*Da bereits die grundlegenden Daten in einer schwer abzuschätzenden Weise verzerrt sind, können auch komplexere statistische Analysen (z. B. regressionsanalytische Vorhersagen von Click-Streams) kaum erfolgreich durchgeführt werden*“.

Dennoch gibt es zahlreiche Untersuchungen, die sich auf Logfile-Analysen stützen, einige davon sollen auch im weiteren dargestellt werden (vgl. Kapitel 5.4.2), um die beschriebenen Probleme und Möglichkeiten anschaulich demonstrieren zu können. Bevor jedoch auf diese Arbeiten eingegangen wird, sollen die herkömmlichen Auswertungstools von Logfiles vorgestellt werden:

Logfile Analyzer

Wie dargestellt liefern Logfiles eine Unmenge an aggregiertem Datenmaterial, das jedoch nicht auf Personenebene ausgewertet werden kann. Begnügt man sich jedoch mit deskriptiven Statistiken (z. B. Häufigkeit der Zugriffe pro Wochentag bzw. Uhrzeit etc.) kann man auf eine Fülle von Auswertungssoftware zurückgreifen. Diese Tools findet man in unterschiedlichsten Preisklassen. Höherpreisige Anwendungen funktionieren in Zusammenhang mit Datenbankanbindungen, kostenlose bzw. preisgünstige Programme liefern dementsprechend weniger aussagekräftige Analysen. Zumeist beschränken sich diese Tools auf die Aufgliederung folgender Kennwerte (Janetzko, 1999):

- Identifikation der Benutzer, Länder etc. mit den jeweils häufigsten Zugriffen
- Angabe der Anzahl der Hits
- Angabe der übertragenen Bytes
- Wahlweise monatliche, wöchentliche, tägliche bzw. stündliche Zusammenfassungen der Übertragungsstatistik
- Gesamtzahl der Übertragungen ausgewählter Clientdomänen (z. B. at, de oder com)
- Refferer-Informationen
- Statistik über den vom Besucher verwendeten Browser
- Betrachtungsdauer pro Seite
- Falsche URLs, die zu Fehlern geführt haben
- Tägliches, wöchentliches Neuanlegen von Logdateien und Erstellung eines Log-Reports (log file rotation)
- Ergebnis-Export im HTML – Format
- Zustellen von Log-Statistiken per Mail an definierte Benutzer

Zumeist bieten bereits kostenlose bzw. günstige Logfile Analyzer ansprechende graphische Aufbereitungen (einfache Häufigkeitstabellen etc.) der oben angeführten Befunde an.

Wie sich jedoch bereits aus den vorangegangenen Ausführungen herausgestellt haben sollte, ist eine wissenschaftliche Verwertung dieses Datenmaterials nicht so einfach möglich, wie die anfängliche Euphorie über das sehr einfach zu erhaltende Datenmaterial einen vermuten lassen möchte.

Exemplarisch einige dieser Anwenderprogramme:

Hitlist Pro (www.marketwave.com)
Netstats (www.netstats2000.com)
Websuccess (www.websuccess.de)
Webtrends (www.webtrends.com)
Webalizer (http://www.mrunix.net/webalizer/)
Analog (http://www.download.com)

5.4.2 Beispiele für Untersuchungen

Zu den oft zitierten Klassikern zählen die Arbeiten von Catledge und Pitkow (1995) bzw. Tauscher und Greenberg (1997). Für den deutschsprachigen Raum sind Wandke und Hurtienne (1999) und Berker (1999) zu nennen. Die beiden erstgenannten Studien wurden von Forschern mit technischem Hintergrund durchgeführt (Relevanz der Ergebnisse für Design und Usability Kriterien für WWW-Seiten) während die Beiträge von Wandke und Hurtienne bzw. Berker auf psychologischem bzw. soziologischem Hintergrund basieren.

Die Logfile-Analyse von Catledge und Pitkow basiert auf clientseitigen Datenfiles, die nicht in einem WWW-Experiment erhoben wurden. Vielmehr fällt diese Forschungsarbeit in die Kategorie der Online-Laborforschung. Dennoch sollen hier kurz ihre Ergebnisse bzw. ihre Vorgehensweise beschrieben werden. Die von ihnen 1994 durchgeführte, insgesamt drei Wochen dauernde Untersuchung ging der Frage nach Hypertext-Navigation nach bzw. welche Konsequenzen das für die Gestaltung von Internetseiten nach sich zieht. Aufgezeichnet wurden die User Aktionen mittels des Programms Xmosaic (clientseitig). Die Analyse der so gewonnenen Daten erfolgte mittels einem „Pattern Detection Module (PDM) algorithm“ aufgrund dessen die Häufigkeiten von wiederholten Seitenzugriffen bzw. Zugriffen auf Dokumente eruierbar wurden, speziell jedoch unterteilte dieses Programm die Anzahl der Besuchssequenzen in Pfade.

Auch Tauscher und Greenberg bedienten sich des Programms XMosaic um die gewünschten Daten festhalten zu können. Da jedoch auch diese Untersuchung bereits Jahre zurückliegt und auch wenig Bezug zur psychologischen Forschung aufweist, wird im folgenden gleich auf die Untersuchung von Wandke und Hurtienne eingegangen.

Vorausgeschickt werden muß nur noch folgendes: Für eine Vielzahl von heute durchgeführten Experimente lieferten Chen und Rada (1996) wesentliche Grundlagen. In ihrer Metastudie über das Navigieren in Hypertexten führen sie die Unterscheidung von „browsing“ und „searching“ ein, auf die sich heute fast alle diesbezüglichen Arbeiten beziehen.

So eben auch Wandke und Hurtienne, die in ihrer Untersuchung einer Stichprobe von 25 Psychologiestudenten (ohne Interneterfahrung) jeweils 15 verschiedene Suchaufgaben vorlegten und deren Navigationsverhalten festhielten. Sie vermuteten, dass die Seiten-Navigation von Erstbenutzern durch Vorherrschen der Bottom-Up Strategie charakterisiert werden würde und prüften diese Annahme anhand folgender Kriterien: Abweichungen vom optimalen Weg, subjektive Sackgassen, Rückkehrstrategien, Aufsuchen bekannter Zustände, unmittelbare Hin- und Zurückbewegungen, Beschränkung der visuellen Suche auf Links, Wahl von irrelevanten Links bei Sichtbarkeit relevanter und Wahl sichtbarer Links bzw. Verzicht auf Rollen („scrollen“) von Fensterinhalten. Für jede der gestellten 15 Aufgaben (gezielte Informationsrecherche) wurde der kürzeste Weg ermittelt. Wie sich zeigte, wechseln sich bei Internet Anfängern (sogenannten „newbies“) Bottom-Up und Top-Down Prozesse ab.

Wandke und Hurtienne beschreiben dies so: *„Aus der Analyse: Wo bin ich?, Wie bin ich hierher gekommen?, Wohin kann ich gehen? werden Handlungsschritte abgeleitet. Top-Down Prozesse unterstützen vermutlich die Bildung von Ziel- und Zwischenzuständen, z. B. wenn vorhandenes Vorwissen aktiviert wird, um Entscheidungen über den nächsten Schritt zu treffen“.* (1999, S. 29).

Leider geben die genannten Autoren in ihrer Arbeit keinen Hinweis darauf, nach welcher Methode sie die Logfiles analysiert haben.

Berker (1999) analysierte serverseitig anfallende Logfiles. Auch in dieser Arbeit zeigte, dass sich eine Auswertung dieses Datenmaterials (Logfiles) aufwendig gestaltet und man mitunter dazu gezwungen ist, auf Auszählen per Hand zurückzugreifen, aber auch auf Plausibilitätsprüfungen angewiesen ist.

Aufgrund oben genannter Aspekte scheint schließlich klar, dass eine sinnvolle, psychologische Analyse von Logfiles nur unter enormen Aufwand zu bewerkstelligen ist bzw. sieht man sich derzeit noch gezwungen auf wenig komfortable Auswertungsmethoden zurückzugreifen. Bleibt abzuwarten ob Pfadanalysen, Baumdiagramme oder ähnliche statistische Verfahren hier Abhilfe versprechen.

5.4.3 Cookies

Cookies sind eine clientseitige Zusammenstellung von Daten und dienen eigentlich dazu, benutzerbezogene Informationen (z. B. bezüglich Seitenaufrufe) erhalten zu können. Diese Datenpakete stellen demnach Informationen dar, die zwischen Anwendungen ausgetauscht werden (Janetzko, 1999). Gesetzt werden Cookies vom Browser, wobei man diesen Vorgang durch Einstellungen im entsprechenden Browsermenü auch unterbinden kann. Hat man jedoch die Annahme von Cookies im Browsermenü aktiviert, so werden diese kleinen Datenpakete im Arbeitsspeicher des Besuchers (Client) abgelegt. Erst nach dem Beenden des Browsers erfolgt die Speicherung auf der Festplatte. Bei einem nächsten Besuch der entsprechenden Web-Site, welche das entsprechende Cookie gesetzt hat, schickt der Client-Browser dem aufgerufenen Server diese Informationen. Somit ist ein Datentransfer möglich, als ob eine ständige Verbindung zwischen Server und Client bestehen würde. Aufgrund der Tatsache, dass auf diesem Wege nichtreaktiv einfache technische Daten des Benutzers erhoben werden können, stellt der Einsatz von Cookies einen wesentlichen Bestandteil beispielsweise der Online-Marktforschung bzw. Kundensegmentierung dar. Im Rahmen der Online-Forschung kommen Cookies vor allem dann zum Einsatz, wenn es beispielsweise um die Ermittlung von Mehrfachteilnahmen geht. Abschließend sei angemerkt, dass Cookies auf der Festplatze des Benutzers weder eigenständige lesen, noch schreiben odere andere Funktionen ausüben können. Darüber hinaus sind sie sowohl in ihrem Datenvolumen begrenzt (4 KB) als auch in ihrer Anzahl pro Domain (maximal 20) (vgl. Janetzko, 1999).