

September 2001

Warenkorbanalyse im Online-Handel

Frank Bensberg

Westfälische Wilhelms-Universität Münster, awfrbe@wi.uni-muenster.de

Follow this and additional works at: <http://aisel.aisnet.org/wi2001>

Recommended Citation

Bensberg, Frank, "Warenkorbanalyse im Online-Handel" (2001). *Wirtschaftsinformatik Proceedings 2001*. 10.
<http://aisel.aisnet.org/wi2001/10>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISEL). It has been accepted for inclusion in Wirtschaftsinformatik Proceedings 2001 by an authorized administrator of AIS Electronic Library (AISEL). For more information, please contact elibrary@aisnet.org.

In: Buhl, Hans Ulrich, u.a. (Hg.) 2001. *Information Age Economy*; 5. Internationale Tagung
Wirtschaftsinformatik 2001. Heidelberg: Physica-Verlag

ISBN: 3-7908-1427-X

© Physica-Verlag Heidelberg 2001

Warenkorbanalyse im Online-Handel

Frank Bensberg

Westfälische Wilhelms-Universität Münster

Zusammenfassung: Dieser Beitrag thematisiert den Einsatz des Data Mining-Konzepts zur Warenkorbanalyse in Online-Shopsystemen. Aufbauend auf den konzeptionellen und methodischen Grundlagen wird eine Anwendungsstudie aus dem Elektronik-Versandhandel beschrieben. Auf Basis der empirischen Ergebnisse werden Adaptionspotenziale für den Online-Handel identifiziert und Implementierungsaspekte herausgestellt.

Schlüsselworte: Warenkorbanalyse, Online-Handel, Data Mining

1 Problemstellung

Eine Vielzahl von Unternehmen sucht den Einstieg in die New Economy durch die Realisierung von Online-Shopsystemen, die den marktlichen Transaktionsprozess von der Informationsphase über die Vereinbarungs- und Abwicklungsphase bis hin zur After Sales-Phase unterstützen. Zwar kann auf diese Weise das Sortiment ad hoc im Internet angeboten werden, doch verfügen Shopsysteme per se nicht über Informationssysteme, die durch Analyse der Warenkörbe eine adäquate Informationsversorgung des Handelsmanagementprozesses sicherstellen. Dabei eignen sich gerade Online-Bestelldaten aufgrund ihrer hohen Aktualität und elektronischen Verfügbarkeit zur Fundierung handelsbetrieblicher Entscheidungstatbestände. Zu diesem Zweck können Kennzahlen wie z. B. der durchschnittliche Umsatz, Preis und Postenzahl ermittelt werden [Fisc93, S. 40f.]. Des Weiteren stellen Warenkorbdaten eine empirische Basis zur Analyse von Kaufverbundeffekten dar. Diese Verbundphänomene stellen eine zentrale Informationsbasis zur Ausgestaltung handelsbetrieblicher Aktionsparameter dar [Mich95, S. 37f.]:

- Verbundbeziehungen sind zunächst für sortimentspolitische Entscheidungstatbestände relevant. Herrscht ein starker Nachfrageverbund zwischen zwei bestimmten Artikeln, so kann sich die Auslistung eines Artikels über diese Verbundbeziehung negativ auswirken.
- Im Rahmen des Store Layouts können Artikel, die gemeinsam gekauft werden, in Gruppen als Themen- oder Sonderplatzierung angeordnet werden. Diese Anordnung kann auch bei der Gestaltung kommunikationspolitischer Maßnahmen berücksichtigt werden.

- Nachfrageketten liefern eine Entscheidungsgrundlage zur Optimierung der logistischen Prozesse. Durch die Berücksichtigung von Verbundbeziehungen bei der Nachdisposition von Primärartikeln können Präsenzlücken von abhängigen Sekundärartikeln vermieden werden.

Diese Zusammenhänge werden in Abbildung 1 zusammenfassend dargestellt.

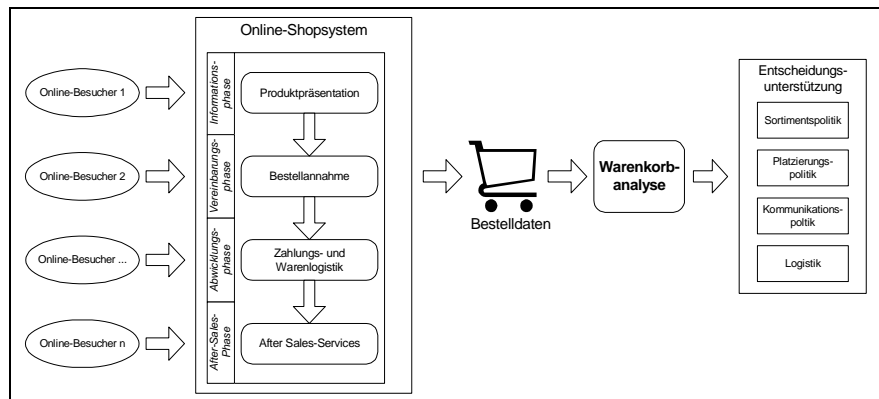


Abbildung 1: Online-Shopsysteme als Datenlieferanten für die Warenkorbanalyse

Zur Analyse von Online-Bestelldaten eignen sich datengetriebene Ansätze, die dem dynamischen Charakter und dem großen Volumen der Datenbasis gerecht werden. Ein geeigneter Ansatz stellt daher das Data Mining-Konzept dar, das aufgrund seines umfangreichen Methodenarsenals und der Eignung für großvolumige Datenbestände zunehmend Bedeutung gewinnt.

2 Das Data Mining-Konzept

Data Mining wird als automatischer oder semiautomatischer Prozess verstanden, der aus einer gegebenen Datenmenge implizit vorhandene, aber bisher unentdeckte, nützliche Informationen extrahiert [Mert94, S. 740] (Abbildung 2).

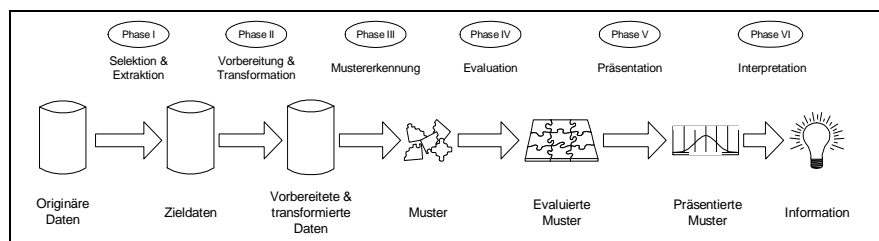


Abbildung 2: Data Mining-Prozess

In der ersten Phase des Prozesses findet die *Selektion* und *Extraktion* der analyse-relevanten Daten statt. Im Zuge der *Selektion* wird eine Menge von Datensätzen (vertikale Selektion) und Attributen (horizontale Selektion) ausgewählt. In der Phase der *Vorbereitung* und *Transformation (Preprocessing)* wird zunächst die Datenqualität des selektierten Datenpools untersucht. Damit die Daten von dem in der nächsten Phase anzuwendenden Mustererkennungswerkzeug verarbeitet werden können, ist außerdem die *Transformation* in ein geeignetes Datenformat notwendig. In der Phase der *Mustererkennung* erfolgt die Methodenauswahl und -parametrisierung sowie deren Anwendung auf der Basis des vorbereiteten Datenbestands. Die Methodenauswahl wird von der Art der gesuchten Zusammenhänge bestimmt und weist daher eine starke situative Problemorientierung auf. Dabei stehen zahlreiche Methoden zur Verfügung, die der Klassifikation, Segmentierung, Prognose, Abhängigkeits- und Abweichungsanalyse dienen [AINi00, S. 9]. Da im Zuge der Methodenanwendung eine große Anzahl von Mustern generiert werden kann, sind diese hinsichtlich ihrer Interessanztheit für den Anwender zu bewerten (Phase IV). Die Muster werden anschließend visualisiert (Phase V) und stehen dem Anwender zur Interpretation zur Verfügung (Phase VI).

Der Data Mining-Prozess stellt einen zeitlich-logischen Bezugsrahmen für die weitgehend automatische Wissensgewinnung dar, der im Wesentlichen durch die Methodenauswahl für den betriebswirtschaftlichen Anwendungskontext ausgestaltet wird. Zur Warenkorbanalyse von Kassendaten im Einzelhandel wurde die Assoziationsanalyse entwickelt [Agra93, S. 207].

3 Assoziationsanalyse

Zielsetzung der Assoziationsanalyse ist die Ermittlung von Abhängigkeiten zwischen Teilmengen eines Datenbestands. Die Darstellung dieser Abhängigkeiten erfolgt in Form von Assoziationsregeln, die die Struktur und die Signifikanz der Abhängigkeit beschreiben. Die strukturellen Komponenten einer Assoziationsregel sind die *Prämisse* und die *Konsequenz*. Diese beiden Komponenten, die die Musterstruktur beschreiben, können beliebig viele Elemente umfassen, die durch logische Operatoren verknüpft werden. Die Güte von Assoziationsregeln wird durch den Support- und den Konfidenzfaktor erfasst. Der *Supportfaktor* trifft Aussagen über die Häufigkeit der festgestellten Regel. Der *Konfidenzfaktor* ist ein Maß für die Zuverlässigkeit bzw. die Strenge der Regel. Im Folgenden wird ein formales Modell zur Ableitung von Assoziationsregeln für den Anwendungsfall der Warenkorbanalyse dargestellt [Agra96, S. 307f.].

Das Sortiment einer Einkaufsstätte sei durch die Menge der Produkte $I = \{i_1, i_2, \dots, i_m\}$ gegeben. Die historischen Kaufhandlungen werden in einer Transaktionsdatenbank $L = \{l_1, l_2, \dots, l_n\}$ fortgeschrieben. Für eine Transaktion $l_j \in L$ gilt somit $l_j \subseteq I$. Die Menge $A \subset I$ bildet die Prämisse und die Menge $B \subset I$ die Konsequenz

der Assoziation. Damit keine tautologischen Assoziationsregeln erfasst werden, dürfen die Prämisse und die Konsequenz der Assoziation keine Überschneidungen aufweisen, d. h. es handelt sich um disjunkte Mengen ($A \cap B = \emptyset$). Eine Transaktion l_j erfüllt die Prämisse A , wenn $A \subseteq l_j$. Analog besitzt eine Transaktion l_j die Konsequenz B , wenn $B \subseteq l_j$. Eine Assoziationsregel wird mit den Sprachelementen der Aussagenlogik als Implikation der Form $A \Rightarrow B$ erfasst. Assoziationsregeln werden daher auch als Wenn-dann-Regeln (if-then-rules) im Sinne von "Wenn A , dann B " bezeichnet.

Um die relative Häufigkeit einer Assoziationsregel zu erfassen, wird der Supportfaktor (support) als normiertes Maß über dem Intervall $[0;1]$ gebildet:

$$\text{support}(A) := \frac{|\{l \in L \mid A \subseteq l\}|}{|L|} \quad (1)$$

$$\text{support}(A \Rightarrow B) := \text{support}(A \cup B)$$

Der Konfidenzfaktor (confidence) gibt an, mit welcher Wahrscheinlichkeit die Konsequenz einer Assoziationsregel eintritt, wenn die Prämisse erfüllt ist. Es handelt sich um die bedingte Wahrscheinlichkeit, die als normiertes Maß über dem Intervall $[0;1]$ wie folgt gebildet wird:

$$\text{confidence}(A \Rightarrow B) := \frac{|\{l \in L \mid (A \cup B) \subseteq l\}|}{|\{l \in L \mid A \subseteq l\}|} \quad (2)$$

Um eine geschlossene Beschreibung von Assoziationsregeln zu ermöglichen, ist die Darstellung der strukturellen Elemente (A , B) und der Gütekriterien (confidence, support) notwendig:

$$A \Rightarrow B^{\text{confidence;support}} \quad (3)$$

Auf objektsprachlicher Ebene werden im Kontext der Warenkorbanalyse Aussagen über Artikel oder Warengruppen gebildet. Eine Assoziationsregel kann z. B. folgendermaßen dargestellt werden:

$$\text{Brot} \rightarrow \text{Butter}^{c=0,8;s=0,2} \quad (4)$$

Diese Assoziationsregel liefert die Aussage, dass ein Kunde, der Brot kauft, mit einer bedingten Wahrscheinlichkeit von 80% auch Butter erwirbt. Diese Assoziationsregel lässt sich in 20% der Transaktionen nachweisen. Bei der Interpretation ist allerdings zu beachten, dass aufgrund der beiden Maßgrößen Support und Konfidenz nicht festgestellt werden, ob die Prämisse für das Auftreten der Konsequenz förderlich ist [Boll96, S. 260]. Existiert beispielsweise die Assoziationsregel $\text{Brot} \rightarrow \text{Butter}^{c=0,8;s=0,2}$, so kann anhand der Gütekriterien nicht bestimmt werden, ob es sich um einen positiven oder negativen Zusammenhang handelt. Kaufen z. B. 90% sämtlicher Kunden Butter, so handelt es sich um einen negativen Zusam-

menhang, d. h. die Käufer von Brot erwerben Butter seltener als der durchschnittliche Käufer. Zur Vermeidung von Fehlinterpretationen ist daher die bedingte Wahrscheinlichkeit der Regel (Konfidenz) in Bezug zur unbedingten Wahrscheinlichkeit der Konsequenz zu setzen [Brin97, S. 260; Boll96, S. 260].

Aufgrund des kombinatorischen Potenzials von Transaktionsdatenbanken können i. d. R. sehr viele Regeln generiert werden. Bei einer gegebenen Transaktionsdatenbank mit einer Anzahl von $|I|$ verschiedenen Produkten können $2^{|I|}$ Teilmengen erzeugt werden. Bei einem Sortimentsumfang von 1.000 Artikeln sind somit $2^{1000} \approx 10^{301}$ Kombinationen zu überprüfen. Infolgedessen ist der Suchraum durch die Definition von Constraints adäquat einzuschränken. Durch Anwendung *semantischer Constraints* kann die Menge der Regeln inhaltlich beschränkt werden. Interessieren den Anwender z. B. nur Regeln, die bestimmte Artikel betreffen, so werden nur diejenigen Transaktionen berücksichtigt, die diese Artikel aufweisen. Semantische Constraints können die Regelmenge zwar inhaltlich einschränken, doch werden hierbei auch Regeln mit geringer Signifikanz generiert. Lässt sich eine Regel nur in wenigen Transaktionen nachweisen, so stellt sie für den Anwender meist kein interessantes Wissen dar. Um diesem Problem gerecht zu werden, ist ein *quantitativer Constraint* in Form eines *kritischen Supportfaktors* minsupport notwendig, so dass nur Regeln generiert werden, die diesen kritischen Wert aufweisen bzw. überschreiten. Der Supportfaktor gibt zwar einen Hinweis darauf, wie „wichtig“ Assoziationsregeln in Bezug auf den Datenbestand sind, doch wird sich ein Anwender nur für Regeln interessieren, die auch einen bestimmten *kritischen Konfidenzfaktor* minconfidence aufweisen. Infolgedessen werden im Rahmen der Assoziationsanalyse nur Regeln generiert, die einen bestimmten Konfidenz-Constraint erfüllen.

Aus ökonomischer Perspektive stellt die Assoziationsanalyse der handelsbetrieblichen Marketingforschung ein leistungsfähiges Instrument zur Identifikation von Verbundphänomenen zur Verfügung. Um die Potenziale dieses Instrumentariums im Rahmen des Online-Handels zu identifizieren, werden zunächst empirische Anwendungsergebnisse erörtert.

4 Warenkorbanalyse im Elektronik-Versandhandel

4.1 Zielsetzung

Im Rahmen eines Pilotprojekts wurden die Bestelldaten eines Online-Shopsystems aus dem Elektronik-Versandhandel ausgewertet. Aus methodischer Perspektive standen dabei folgende Fragestellungen im Erkenntnismittelpunkt:

- Welche Verbundbeziehungen existieren zwischen den Warengruppen?
- Welche Verbundbeziehungen existieren zwischen einzelnen Artikeln?

Während die erste Fragestellung zur Informationsversorgung der strategischen Sortimentspolitik dient, bilden elementare Verbundbeziehungen auf Artikelebene eine Grundlage für operative Sortimentsentscheidungen, wie z. B. Bundling. Da keine Hypothesen über diese beiden Teilaspekte des Online-Käuferverhaltens verfügbar waren, wurde ein datengeleiteter Forschungsansatz verfolgt, dessen Aktivitäten nach dem Prozessmodell des Data Mining strukturiert wurden. Als Data Mining-Werkzeug wurde Intelligent Miner for Data 6.1 (IBM) eingesetzt.

4.2 Selektion und Extraktion

Als empirische Basis für die erforderliche Warenkorbanalyse standen insgesamt 3245 Online-Bestellungen zur Verfügung, die im Rahmen der Selektion und Extraktion vollständig in einen Zieldatenbestand überführt wurden. Diese enthielten folgende Merkmale:

- Kundenbezogene Merkmale (Kundennummer, Name, Vorname, Anrede, etc.)
- Auftragsbezogene Merkmale (Artikelnummer, Anzahl, Bestellkennzeichen)
- Abwicklungsbezogene Merkmale (Lieferbedingung, Zahlungsbedingung)
- Kaufzeitpunktbezogene Merkmale (Datum, Uhrzeit)

Im Zuge der Selektion und Extraktion wurden diese Dateien in ein relationales Datenbanksystem überführt und mit Artikelstammdaten angereichert. Die Anreicherung mit den Artikelstammdaten war aus zwei Gründen notwendig:

- In den Bestelldaten wurde die Warengruppe der bestellten Produkte nicht erfasst, so dass eine Analyse der Verbundbeziehungen auf Warengruppenebene nicht hätte erfolgen können.
- Die Bestelldaten enthielten keine Artikelbezeichnung, so dass zur Interpretation von Verbundbeziehungen lediglich die Artikelnummern zur Verfügung gestanden hätten. Durch Integration des Artikelstamms konnten die Artikelkurzbezeichnungen integriert werden.

Die resultierende Datengrundlage wurde in Form eines Stern-Schemas abgelegt.

4.3 Vorbereitung und Transformation

Im Zuge der Vorbereitung und Transformation wurden die Warenkörbe für die Assoziationsanalyse vorbereitet. Als relevante Attribute der Bestelldaten wurden die *Artikelnummer* und die *Anzahl* der bestellten Produkte erfasst. Zur Steigerung der Datenqualität wurde zunächst eine Datenbereinigung (data scrubbing) durch-

geführt. Zu diesem Zweck wurden Wertebereichskontrollen der relevanten Attribute durchgeführt. Durch Abgleich mit dem Artikelstamm konnten Bestellungen mit ungültigen Artikelnummern identifiziert werden ($n=47$). Diese waren darauf zurückzuführen, dass der Online-Käufer einen relativ hohen Freiheitsgrad bei der Eingabe seiner Bestelldaten hat. So können im Online-Bestellformular z. B. Artikelnummern eingegeben werden, für die kein entsprechender Artikel existiert. Auf diese Weise konnten Unzulänglichkeiten des Online-Shopsystems identifiziert werden, deren technische Korrektur nachhaltig zur Steigerung der Bestelldatenqualität beitragen kann. Anschließend wurden die Bestelldaten in ein Datenbankschema überführt, das den Anforderungskriterien des eingesetzten Data Mining-Werkzeugs Intelligent Miner for Data entspricht. Dieses Datenbankschema strukturiert die Kaufakte anhand einer Transaktionsnummer (TID) und umfasst die Attribute *Artikelbezeichnung* und *Warengruppe* aus dem Artikelstamm (Tabelle 1).

TID	Artikelnummer	Artikelbezeichnung	Warengruppe
1	976830	EINBAUWINKEL 2,5"	11
1	345286	HP TINTE 51649A	24
2	486746	MEMORY CARD	16
2	757536	NETZTEIL-GEHAEUSE	02
3	424098	ANTENNE	15
...

Tabelle 1: Vorbereitete und transformierte Daten

4.4 Mustererkennung und Visualisierung

Auf der vorbereiteten Datengrundlage erfolgte die Anwendung der Assoziationsanalyse. Zu diesem Zweck wurde der Apriori-Algorithmus eingesetzt, der eine weitgehend hypothesenfreie Analyse von Warenkorbdaten gestattet. In einem ersten Schritt erfolgte die numerische Belegung des kritischen Konfidenz- und Supportfaktors. Aufgrund der relativ geringen durchschnittlichen Warenkorbgröße (4,16 Artikel/Warenkorb) und einem Sortimentsumfang von mehr als 50.000 Artikeln musste der kritische Supportfaktor sukzessive reduziert werden, damit Regeln generiert werden konnten. Um Fehlinterpretationen zu vermeiden, wurde die Konfidenz der Regeln vor dem Hintergrund der unbedingten Wahrscheinlichkeit der Konsequenz interpretiert. Dabei konnten ausschließlich positive Zusammenhänge identifiziert werden.

Auf Artekebene konnten bei einem kritischen Supportfaktor von 0,1% und einer kritischen Konfidenz von 50% insgesamt 58 Regeln identifiziert werden. Zur Regelpräsentation wurde eine Visualisierung in Form gerichteter Grafen vorgenommen. In Abbildung 3 wird ein Set von 15 Regeln präsentiert, das Zusammenhänge

zwischen unterschiedlichen Werkzeugartikeln erfasst. Dabei geben Kantenstärke und -farbe Aufschluss über die Güte des ermittelten Zusammenhangs. Je stärker die Kante, desto höher die Konfidenz und somit die Zuverlässigkeit der Assoziationsregel. Der Supportfaktor der Assoziationsregel wird durch die Kantenfarbe abgebildet. Beispielsweise führt der Erwerb einer *Beisszange* immer auch zum Kauf der Artikel *Spitzzange gerade*, *Spitzzange gebogen* und *Seitenschneider* (Konfidenz=100%). Diese Regeln bieten einen Anhaltspunkt für sortimentspolitische Maßnahmen. So können die Verbundbeziehungen z. B. zur Gestaltung eines Bundles genutzt werden, das sämtliche stark assoziierten Produkte umfasst.

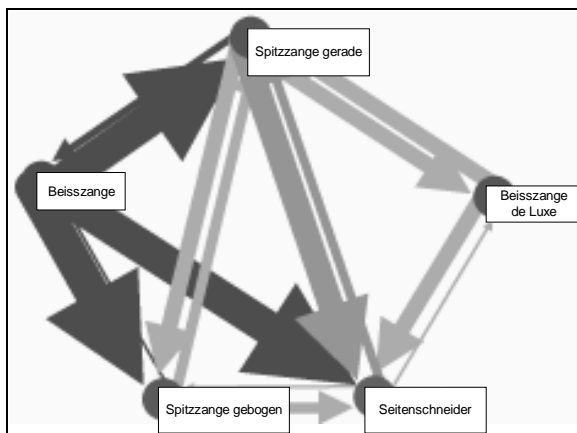


Abbildung 3: Verbundbeziehungen zwischen Artikeln

Neben diesen Zusammenhängen aus der Warengruppe *Werkzeug* konnten weitere Verbundeffekte abgeleitet werden. Ein starker Zusammenhang herrschte z. B. zwischen Software-Sonderangeboten aus der Warengruppe *Schnäppchen*. So führt der Kauf eines Softwareprodukts mit einer Konfidenz von 75% zum Kauf eines anderen Software-Sonderangebots.

Aus diesen Ergebnissen wird deutlich, dass auf Articlebene keine warengruppenübergreifenden Verbundphänomene identifiziert werden konnten. Um sortimentspolitische Entscheidungen jedoch auch auf strategischer Ebene zu unterstützen, wurden in einem anschließenden Schritt die Verbundeffekte zwischen den Warengruppen untersucht (Abbildung 4).

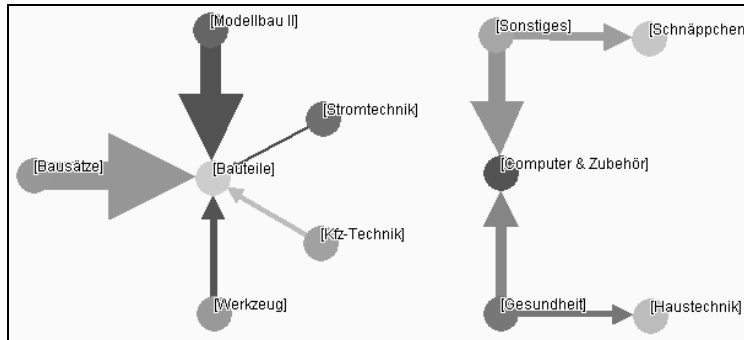


Abbildung 4: Verbundbeziehungen zwischen Warengruppen

Warengruppenübergreifende Verbundphänomene treten fokussiert bei den beiden zentralen Warengruppen *Bauteile* und *Computer & Zubehör* auf:

- Die Käufer von Artikeln aus den Warengruppen Modellbau II und Bausätze erwerben mit einer relativ hohen bedingten Wahrscheinlichkeit (43% bzw. 49%) auch Bauteile. Diese Zusammenhänge treten in den Bestelldaten mit einem Support von 0,7% bzw. 2,1% auf. Werkzeugkäufer erwerben Bauteile nur mit einer bedingten Wahrscheinlichkeit von 34%, jedoch weist diese Regel den stärksten Support auf (4%).
- Die Käufer der Warengruppen Sonstiges und Gesundheit erwerben mit einer Konfidenz von 40% bzw. 37% auch Artikel der Warengruppe Computer & Zubehör. Der Support beträgt dabei 2% bzw. 0,9%.

Im Anschluss an die Mustererkennung und Visualisierung wurden die identifizierten Verbundphänomene der Informationsverwendung zugeführt.

4.5 Informationsverwendung

Insgesamt konnte mit den generierten Hypothesen ein Beitrag zur informatischen Versorgung der Sortimentspolitik geliefert werden. Während die artikelbezogenen Verbundbeziehungen als Informationsgrundlage zur Konstruktion von Bundles eingesetzt werden können, zeigt die warengruppenbezogene Analyse deutliche Hinweise auf warengruppenübergreifende Käuferpräferenzen, die durch weiterführende Segmentierungsstudien zu präzisieren sind. Dabei ist allerdings die hybride Struktur des betrachteten Handelsunternehmens zu berücksichtigen. So verfolgt das Unternehmen eine Multi-Kanal-Strategie, die den Online-Handel mit traditionellen Handelsstrukturen (z. B. Katalog-Versandhandel) kombiniert. Zur Koordination dieser Kanäle ist im Rahmen weiterführender Warenkorbanalysen ein methodisch gestützter Vergleich des Online-Kaufverhaltens und des Kaufverhaltens über klassische Kanäle zu leisten. Signifikante Unterschiede zwischen die-

sen beiden Kaufakttypen können dazu führen, dass sich die Gültigkeit der identifizierten Zusammenhänge auf den Online-Handel beschränkt. Eine ad hoc-Informationsverwendung zur Ausgestaltung der traditionellen Kanäle (z. B. zur Gestaltung von Print-Katalogen) ist daher kritisch zu bewerten. Neben der informatorischen Fundierung handelsbetrieblicher Entscheidungen eröffnet die Analyse von Verbundbeziehungen im Online-Handel die informationstechnologische Möglichkeit, spezifische Aktionsparameter kundenorientiert auszugestalten.

5 Adaptionspotenziale im Online-Handel

Die Verbundbeziehungen, die mit Hilfe der Assoziationsanalyse auf der Grundlage von Online-Bestelldaten ermittelt werden können, beziehen sich auf das realisierte Käuferverhalten und reflektieren somit Präferenzstrukturen, die zur Steuerung von Shopsystemen eingesetzt werden können. Infolgedessen eröffnet diese Informationsbasis die Möglichkeit, *adaptive Shopsysteme* zu gestalten, deren Mensch-Maschine-Schnittstellen sich an die beobachteten Käuferpräferenzen anpassen [Brus96]. In Bezug auf den marktlichen Transaktionsprozess ist dabei vor allem eine Adaption in der Phase der Informations- bzw. Produktsuche sinnvoll. Können dem Online-Besucher gezielt Produkte präsentiert werden, die den subjektiven Präferenzstrukturen entsprechen, gehen hiermit potenziell absatzfördernde Effekte einher. Ansatzpunkte für eine adaptive Gestaltung von Shopsystemen sind daher diejenigen Schnittstellen, über die die Suchaktivitäten der Online-Besucher erfolgen.

Um eine effiziente Suche in komplexen Sortimenten zu gestatten, unterstützen Online-Shops meist zwei unterschiedliche Suchmechanismen (shopping metaphors, [Lee00, S. 22]). Im Rahmen des *Browsing* navigiert der Online-Besucher in der Produkthierarchie, die von dem elektronischen Produktkatalog über geeignete Schnittstellenelemente visualisiert wird. Durch entsprechende Drill Down-Funktionen gelangt der potenzielle Käufer über die gewünschte Warengruppe zum gesuchten Produkt (z. B. über die Warengruppe *Werkzeuge* zum Produkt *Beisszange*). Da das manuelle Navigieren in komplexen Informations- bzw. Produkthierarchien zeitintensiv ist, bieten Shopsysteme die Möglichkeit, mittels *Textsuche* gezielt nach bestimmten Begriffen (z. B. *Spitzzange*) in der Produktdatenbank zu suchen. Die Visualisierung passender Produkteinträge erfolgt dabei über Ergebnislisten, die Verweise auf weiterführende Produktinformationen enthalten.

Im Rahmen dieser Suchmechanismen können Verbundbeziehungen eingesetzt werden, um eine bedürfnisorientierte Informations- und Produktpräsentation zu erzielen. Dabei können sowohl warengruppenbezogene als auch artikelspezifische Verbundbeziehungen zum Einsatz kommen (Abbildung 5).

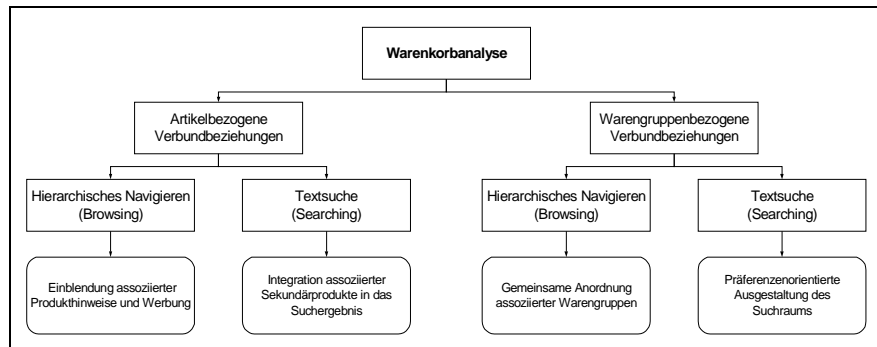


Abbildung 5: Adaptionspotenziale von Verbundbeziehungen

Im Rahmen der hierarchischen Navigation können *artikelbezogene Verbundbeziehungen* genutzt werden, um Produkt- oder Werbeinformationen über assoziierte Artikel einzublenden. So kann beispielsweise bei Auswahl des Produkts *Beisszange* ein Verweis auf das Produkt *Seitenschneider* eingeblendet werden. Diese Adaptionstechnik wird beispielsweise im Internet-Buchhandel eingesetzt, um den Online-Besucher bei der Präsentation detaillierter Produktinformationen auf stark assoziierte Produkte hinzuweisen. Bei einer textuellen Suche können assoziierte Produkte mit in das Suchergebnis aufgenommen werden. Allerdings ist dabei zu beachten, dass diese lediglich nachrangig angezeigt werden dürfen, damit keine Produkte verdrängt werden, die dem Primärinteresse des Käufers entsprechen. *Warengruppenbezogene Verbundbeziehungen* können dagegen eingesetzt werden, um die Präsentation komplexer Warengruppenhierarchien kundenorientiert auszugestalten. So sorgt eine gemeinsame Anordnung assoziierter Warengruppen (z. B. *Sonstiges* und *Computer & Zubehör*) für eine Sortimentswahrnehmung, die den Präferenzen des Online-Besuchers entspricht. Bei textuellen Suchprozessen bieten warengruppenbezogene Verbundbeziehungen die Möglichkeit, den Suchraum an die Präferenzstrukturen anzupassen. Wird beispielsweise in der Warengruppe *Bausätze* nach einem bestimmten Begriff gesucht, so kann eine automatische Ausdehnung der Suche auf die assoziierte Warengruppe *Bauteile* sinnvoll sein.

Die Anwendung adaptiver Techniken auf der Grundlage von Verbundbeziehungen kann auch in der After Sales-Phase erfolgen. Beim Abruf von produktspezifischen Serviceinformationen (z. B. Wartungs- und Reparaturinformationen) können Verbundbeziehungen eingesetzt werden, um aktiv auf assoziierte Produkte hinzuweisen. Auf diese Weise eröffnet sich die Möglichkeit, zusätzliche Cross Selling-Potenziale auszuschöpfen.

6 Implementierungsaspekte

Als konzeptionelle Grundlage für die Realisierung adaptiver Shopsysteme ist der Data Mining-Prozess zugrunde zu legen. Durch informationstechnologische Ausgestaltung der einzelnen Phasen ist es möglich, Hypothesen über das Konsumentenverhalten zu generieren. Bei Anwendung der Assoziationsanalyse kann diese *Hypothesengenerierung* weitgehend automatisiert erfolgen. Dabei ist allerdings zu beachten, dass die Datengrundlage dynamischen Charakter besitzt. So können Veränderungen im Online-Käuferverhalten dazu führen, dass historische Verbundbeziehungen keine Gültigkeit mehr besitzen. Da eine Adaption auf Basis ungültiger Zusammenhänge nicht den Konsumentenpräferenzen entspricht, sind die Adaptionmechanismen so zu gestalten, dass eine kontinuierliche *Hypothesenprüfung* erfolgt. Damit wird sichergestellt, dass nur bewährte Hypothesen über das Konsumentenverhalten als Informationsgrundlage für die Adaption genutzt werden und somit eine evolutorische Anpassung der Online-Einkaufsstätte an die Käuferpräferenzen erfolgt. Zur Automatisierung der Hypothesengenerierung und –prüfung ist eine Datenbank notwendig, die die entdeckten Verbundbeziehungen zeitbezogen (z. B. täglich) speichert. Auf Grundlage dieser *Hypothesenbank* ist die Identifikation von Zusammenhängen möglich, die mit einer hinreichenden Regelmäßigkeit und Güte auftreten. Diese bewährten Verbundbeziehungen können schließlich als Informationsbasis zur Anpassung des Shopsystems an die Käuferpräferenzen genutzt werden.

Zur Integration von Informations- und Aktionsseite sind Mechanismen erforderlich, die eine automatische Adaption des Shopsystems sicherstellen. Ein derartiger Mechanismus kann z. B. in Form eines *Konfigurators* realisiert werden, der auf der Grundlage bewährter Verbundbeziehungen und exogen vorgegebener *Marketingregeln* die Anpassung spezifischer Elemente des Shopsystems vornimmt. Im Rahmen von Marketingregeln hat der Entscheidungsträger zunächst zu spezifizieren, welche Bedingungen erfüllt sein müssen, damit eine Adaption des Shopsystems erfolgt. Handelt es sich bei der Verbundbeziehung *Beisszange* → *Seitenschneider* beispielsweise um eine bewährte Hypothese, so kann als Bedingung der Abruf von Informationen über das Produkt *Beisszange* aus dem Produktkatalog formuliert werden. In Abhängigkeit von dieser Bedingung ist die entsprechende Aktion zur präferenzorientierten Ausgestaltung des Shopsystems zu definieren. Entsprechende Aktionen können beispielsweise das Einblenden von Werbebotschaften oder das Anzeigen eines Verweises zum Produkt *Seitenschneider* sein, die die kommunikative Präsenz des assoziierten Produkts im Wahrnehmungsfeld des Käufers in unterschiedlichem Ausmaß steigern. Bei der Auswahl der Aktionen ist die Konfidenz der Verbundbeziehung zu berücksichtigen. So kann bei Assoziationsregeln mit hoher Konfidenz die Anzeige eines Verweises zur Erzielung einer kaufentscheidenden Wirkung hinreichend sein. Demgegenüber können bei schwach assoziierten Produkten Werbebotschaften eingeblendet werden, die eine stärkere Informationswirkung besitzen.

Die technologischen Makrostrukturen einer adaptiven Marktapplikation werden in Abbildung 6 zusammenfassend dargestellt.

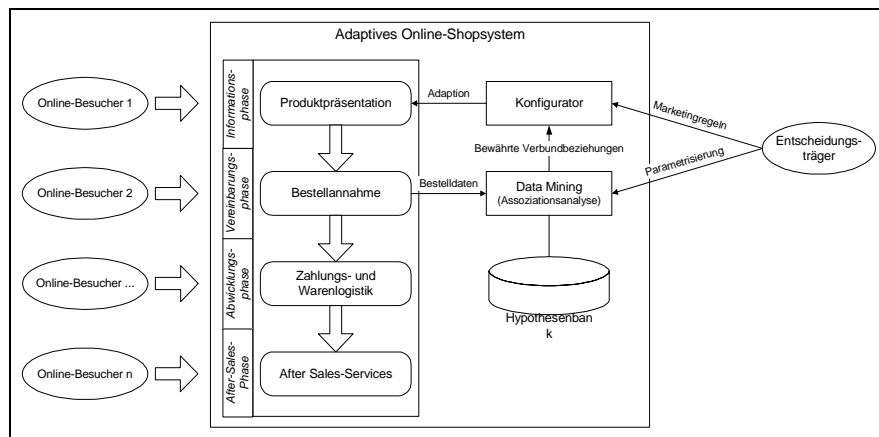


Abbildung 6: Technologische Makrostrukturen eines adaptiven Shopsystems

Durch Anwendung von Adaptionmaßnahmen können zunächst die Gütekriterien der zugrunde gelegten Verbundbeziehungen positiv beeinflusst werden. So geht mit der Einblendung von Werbung bei schwach assoziierten Produkten eine Steigerung der Kaufneigung einher, die bei der Analyse der Kaufakte schließlich in einer höheren Konfidenz zum Ausdruck kommt. Zum Zweck der Entscheidungsunterstützung sind die erzielten Adaptionseffekte systematisch zu erfassen. In einem ersten Schritt ist die Absatzwirkung der Adaption zu bewerten. Zu diesem Zweck sind auf der Grundlage der historischen Bestelldaten zunächst Mengeneffekte zu ermitteln, die durch Anwendung artikelbezogener Adaptionstechniken induziert wurden. Zur Entscheidungsunterstützung sind diese mengenmäßig ausgewiesenen Adaptionswirkungen monetär zu bewerten. Zu diesem Zweck kann beispielsweise der zusätzliche Rohgewinn ermittelt werden, der durch Adaption induziert wurde. Diese Informationsgrundlage erlaubt schließlich eine zielsetzungsgerechte Planung und Kontrolle von Adaptionmaßnahmen.

7 Ausblick

In diesem Beitrag wurde die Anwendung der Assoziationsanalyse im Online-Handel thematisiert. Dabei konnte gezeigt werden, dass sich dieses Data Mining-Verfahren zur Entscheidungsunterstützung im Rahmen des handelsbetrieblichen Managementprozesses eignet. Des Weiteren liefert dieses Verfahren eine informatorische Grundlage zur Gestaltung adaptiver Shopsysteme, die eine evolutorische Anpassung der Online-Einkaufsstätte an die Käuferpräferenzen leisten. Dies

erschließt dem Online-Handel den kanalspezifischen Aktionsparameter *Adaptionspolitik*, von dem absatzfördernde Effekte ausgehen. Darüber hinaus liefert die Analyse von Online-Kaufakten einen Informationsbeitrag zur Theoriebildung im Objektbereich des Online-Konsumentenverhaltens. So verfügt das Internet über Protokollmechanismen, die eine detaillierte Erfassung des konsumentenbezogenen Informationsverhaltens leisten [BeWe99, S. 426ff.]. Folglich kann eine empirische Basis geschaffen werden, die eine integrierte Analyse von Informations- und Kaufverhalten gestattet. Durch die Verfügbarkeit von Data Mining-Techniken zur Hypothesengenerierung und -prüfung eröffnen sich somit auch der Grundlagenforschung neue Möglichkeiten.

Literatur

- [Agra93] Agrawal, R. et al.: Mining Association Rules between Sets of Items in Large Databases. In: ACM SIGMOD Record 22 (1993) 2, S. 207–216.
- [Agra96] Agrawal, R. et al.: Fast Discovery of Association Rules. In: Advances in Knowledge Discovery and Data Mining, Hrsg.: U. M. Fayyad et al., Menlo Park, Cambridge, London 1996, S. 307–328.
- [AlNi00] Alpar, P.; Niedereichholz, J.: Einführung zu Data Mining. In: Data Mining im praktischen Einsatz. Hrsg.: P. Alpar; J. Niedereichholz. Braunschweig, Wiesbaden 2000, S. 1-27.
- [BeWe99] Bensberg, F.; Weiß, T.: Web Log Mining als Marktforschungsinstrument für das WWW. In: Wirtschaftsinformatik 41 (1999) 5, S. 426–432.
- [Boll96] Bollinger, T.: Assoziationsregeln – Analyse eines Data Mining Verfahrens. In: Informatik-Spektrum 19 (1996) 5, S. 257–261.
- [Brin97] Brin, S. et al.: Dynamic Itemset Counting and Implication Rules for Market Basket Data. In: ACM SIGMOD Record 26 (1997) 2, S. 255-264.
- [Brus96] Brusilovsky, P.: Methods and Techniques of Adaptive Hypermedia. In: User Modeling and User Adapted Interaction 6. (1996) 2–3, S. 87–129.
- [Fisc93] Fischer, T.: Computergestützte Warenkorbanalyse. Reihe "Schriften zu Distribution und Handel", Band 11, Frankfurt a. M. et al. 1993.
- [Lee00] Lee, J. et al.: Understanding Merchandising Effectiveness of Online Stores. In: Electronic Markets 10 (2000) 1, S. 20-28.
- [Mert94] Mertens, P. et al.: Datenmustererkennung in der Ergebnisrechnung mit Hilfe der Clusteranalyse. In: Die Betriebswirtschaft 54 (1994) 6, S. 739–753.
- [Mich95] Michels, E.: Datenanalyse mit Data Mining: Kassenbons – die analysierbaren Stimmzettel der Konsumenten. In: Dynamik im Handel (1995) 11, S. 37-43.