

Association for Information Systems AIS Electronic Library (AISeL)

AMCIS 2007 Proceedings

Americas Conference on Information Systems
(AMCIS)

December 2007

Knowledge Discovery on Consumer Trust in B2C Electronic Commerce

Dan Kim

University of Houston Clear Lake

Follow this and additional works at: <http://aisel.aisnet.org/amcis2007>

Recommended Citation

Kim, Dan, "Knowledge Discovery on Consumer Trust in B2C Electronic Commerce" (2007). *AMCIS 2007 Proceedings*. 101.
<http://aisel.aisnet.org/amcis2007/101>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2007 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

KNOWLEDGE DISCOVERY ON CONSUMER TRUST IN B2C ELECTRONIC COMMERCE

Dan J. Kim

Computer Information Systems
University of Houston Clear Lake
kimdan@uhcl.edu

Abstract

The purpose of this research is to discover knowledge (association rules) in consumer behavioral data with regard to their trusting intension. The main research question is: "what are there consumer perception profiles that tend to be associated with a consumer's perception of trust?" To answer the question, real-world data obtained from a customer trust survey have been collected and analyzed using a data mining association rule discovery algorithm (APRIORI). From a managerial point of view, the analyzed results will provide some insights into specific target groups and their relevant drivers for trustworthiness. For example, one of the association rules is, customers who have high perception of trustworthiness and high perception of convenience make a transaction with 95% confidence. This shows the importance of trustworthiness and convenience for completion of transactions.

Key words: knowledge discovery, consumer trust, Business-to-Consumer e-commerce, apriori algorithm, association rules, data mining

Introduction

Due to the nature of Internet transactions (i.e., blind transactions, borderless transactions, 24 hour transactions, and prior transactions), the issue of trust may be even more important in electronic transactions than it is in traditional off-line transactions. Since trust is important in exchange relations (Mayer et al. 1995), it has been identified as a key component of marketing and e-commerce literature (Ba et al. 1999; Beatty et al. 1996; Chang et al. 2005; Cheung et al. 2006; Czepiel 1990; Hoffman et al. 1999; Jarvenpaa et al. 2000; Kim et al. 2006; Kim et al. 2005; Lim et al. 2006; Morgan et al. 1994; Noteberg et al. 1999; Pavlou et al. 2004; Reichheld 1994). It has been found that the higher the levels of a consumer's trust, the higher the levels of consumer's commitment (e.g. purchase). Trust is a prerequisite for behavioral commitment (Morgan et al. 1994). Berry (1995) describes trust as the single most powerful relationship marketing tool. Grabosky (2001) supports the idea that the key to success in online business is the establishment of trusted processes. This fact mandates that online sellers engender an environment in which a prospective consumer can be relaxed and confident about any prospective transactions.

In order to create this trusted environment, it is necessary to understand the association relationships among factors that affect a consumer's trust and behavioral commitment (completion of purchase). Thus, it is important to identify the association

relationships on trust among factors related to a consumer's purchasing intentions. While a lot of research on applications of knowledge discovery have been reported -- including discovering affinities for market basket analysis and cross-marketing, catalog design, loss-leader analysis, store layout and customer segmentation based on buying patterns; and association rule mining in health insurance and in predicting telecommunications order failure and medical test results (Ali et al. 1997; Srikant et al. 1997) -- there is no study on trust from the data mining perspective. Thus, it is meaningful to discover knowledge on trust using data mining techniques. In this paper, we are especially interested in finding the association rules on consumers' trust and their behavior.

Theory Background

The Theory of Reasoned Action (TRA) (Fishbein et al. 1975) attempts to explain how a person's beliefs are translated into intentions and how intentions affect actual behavior. TRA is based on the assumption that human beings make rational decisions based on the information available to them. The theory hypothesizes that a person's behavioral intention to perform (or not to perform) a behavior is the immediate determinant of that person's actual behavior. The behavioral intention is the function of both personal and social influence. The personal influence is reflected in attitude toward the behavior. According to the theory, the most important determinant of a person's behavior is behavioral intent. The individual's intention to perform a behavior is a combination of two factors: the attitude toward performing the behavior and the subjective norm. The individual's attitude toward the behavior includes: behavioral belief, evaluations of behavioral outcome, subjective norm, normative beliefs, and the motivation to comply. Subjective norm refers to "the person's perception of social pressure put on him to perform or not perform the behavior in question" (Ajzen et al. 1980). The TRA is used to provide a sound theoretical framework for the study of causal relationships between attitudes and behaviors (Madden et al. 1992).

Research Question and Purpose

In practice, marketers are often interested in a subset of association rules to understand their customers better and to predict the future value of customers based on their demographic characteristics, life-styles, attitudes, behavior intentions, and previous behaviors. Sometimes, marketers may want only rules that contain a specific item or rules that hold a particular consequent or antecedent. For example, they may want any rules that describe either i) a consumer's high perception of trustworthiness, and ii) a consumer's high willingness purchase intention as a consequent (result) variable.

This study is an attempt to discover knowledge (association rules) from consumer survey data using a data mining technique. The main research question is: "are there any consumer characteristics or perception profiles that tend to be more associated with a consumer's perception of trust?" To answer the question, real-world survey data regarding consumer trust has been collected. Based on the collected survey data and background theory, three features (consumer characteristics, attitudes, and behavioral intentions and behavior) of consumers' profiles-related Internet purchasing behavior are focused in this study. They are summarized in Table 1.

Table 1: Consumer Characteristics, Attitudes and Behavioral Intention

Consumer Characteristics	Consumer Attitudes	Consumer's Behavioral Intention & Action (Behavior)
Age	Perceived trustworthiness	Willingness to exchange ¹
Gender	Privacy concern	Completion of transaction (behavior)
Computer expertise	Security concern	
Internet expertise	Perceived benefit	
Household income		
Amount spending		

¹ The word 'exchange' is used in the broader sense of the term, including purchase, buying, transactions, trade and information transfer.

For mining association rules on trust, a data mining association-rule discovery algorithm (Apriori) with item constraints is used to identify associations among characteristics of customers and their intention profiles. Association rule mining is a powerful method which aims at finding interesting association relationships or correlation relationships in a given data set. It explains what attributes or what items tend to appear together. In information-related marketing, association rules can help retailers do selective marketing and plan shelf space.

A typical example of association rule mining is market basket analysis. This process analyzes customer buying habits by finding associations between the different items that customers place in their shopping baskets. For example, the information below can be represented by the association rule R1.

"In at least 10% of all consumer transactions, a consumer buys milk and bread together, but whenever she/he buys milk, she or he also buys bread with an 80% chance at the same trip to the supermarket"	R1: buy (milk) \Rightarrow buy (bread)
--	--

This association rule states that if we pick a customer at random and know that he bought certain products (milk), we can be confident by a percentage (80%) that he also bought certain other products (bread). Since several association rules can be generated from large transaction databases, some weak and non-significant rules have to be filtered out. To eliminate spurious association rules, two measures have been used, called minimum support and confidence. The minimum support indicates the frequency of a pattern, i.e. how often items occur together. A minimum support is necessary if an association is going to be of some business value. The minimum confidence denotes the strength of an association, i.e. how much a specific item is dependent on another. In the example above, 80% is called the confidence (recall) of the rule, and 10% is the support (precision) of the rule. Detailed definitions of support and confidence follow.

Support (Precision): Given the association rule $X_1, \dots, X_n \rightarrow Y$, the support is the percentage of records for which X_1, \dots, X_n and Y both hold. Support is the statistical significance of the association rule. **Confidence (Recall):** Given the association rule $X_1, \dots, X_n \rightarrow Y$, the confidence is the percentage of records for which Y holds, within the group of records for which X_1, \dots, X_n hold. Confidence is the degree of correlation in the dataset between X and Y , and a measure of the rule's strength.

The problem of discovering association rules from the data has received considerable research attention in the data mining area, and several fast algorithms for mining association rules have been developed (Srikant et al. 1997). The main problem of association rule mining is that there are many possible rules coming from different aspects. For example, for the product range of a supermarket, which may consist of several thousand different products, there are billions of possible association rules. It is obvious that such a vast number of rules cannot be processed by inspecting each one in turn. Therefore efficient algorithms are needed that restrict the search space and check only a subset of all rules without missing important rules. One such algorithm is the Apriori, which was introduced by Agrawal, Imielinski, and Swami (Agrawal et al. 1993).

To identify the associational rules of attributes of consumer-related fields, the Apriori algorithm is used. The Apriori algorithm is one of the efficient algorithms that restricts the search space and checks a subset of all rules without missing important rules. Since the focus of this study is not on performance optimization, the Apriori algorithm is chosen since it is a well established, commonly used, and well-studied algorithm (Agrawal et al. 1996a; Agrawal et al. 1996b; Agrawal et al. 1994). One more important reason to choose the Apriori algorithm is accessibility of the source program. The Apriori program used in this study was developed by Christian Borgelt at University of Magdeburg in Germany and is freely available on the Internet under the terms of the GNU Lesser (Library) General Public License (<http://fuzzy.cs.uni-magdeburg.de/~borgelt/software.html>).

Mining Data Set and Constructs

The dataset used for this study was collected from a group of students enrolled in two public universities in the northeastern United States. To extend the dataset, another survey was conducted from another group of students enrolled in two universities in Korea. The number of samples used for this study is total 664 (including 468 from America and 196 from Korea).

As a consumer's belief, *perceived trustworthiness (TRUST)* regarding an online transaction is a construct that the seller entity i.e. a firm or Website (e.g. buy.com) fulfills its obligations as understood by the consumer. At the time of a transaction, online sellers collect the names, e-mail addresses, phone numbers, and home addresses of buyers, and often pass on the information to telemarketers. *Privacy Concern (P_Concern)* refers to a consumer's perception that the Internet vendor will not protect consumers' personal information which is collected during electronic transactions from unauthorized use or the disclosure of confidential information. *Security Concern (S_Concern)* refers to a consumer's perception that the Internet

vendor will not fulfill security requirements, such as authentication, integrity, encryption, and non-repudiation. *Perceived Benefit (Benefit)* refers to a consumer's belief about the extent to which he or she will become better off from the online transaction with a certain Website. *Willingness to Exchange (WE)* refers to the degree to which a consumer intends to exchange from a certain Website. The Theory of Reasoned action (TRA) presumes that volitional behavior is determined by intentions to act. For example, Ajzen and Fishbein (1980) pointed out that behavior intention (willingness to exchange or purchase) is to be a predictor of actual behavior (completion of purchase). *Completion of Transaction (CT)* is a dichotomous trusting behavior variable (purchasing or not purchasing) in this study.

The constructs of the study were measured by at least three observable indicators based on the recommendation by Anderson and Gerbing (1984) and Bentler and Chou (1987). All observable indicators for each construct were developed by a panel of experts as a result of a literature review on the topics. Table 4 shows the measurement items for constructs. The indicators, except completion of transaction, were written in the form of statements or questions. Most of the scales used a 7-point scale Likert rating system with end points such as strongly disagree/strongly agree, extremely unlikely/extremely likely, and not at all confident/completely confident.

Constructs Measurement

Most of the items were adapted from previous research and modified to fit the context of this research. Table 2 shows the descriptive statistics, the Cronbach reliability coefficients, and the literature source of constructors. The reliability coefficients of all variables are higher than the minimum cutoff score of 0.65 (Lee et al. 1999).

Table 2: Descriptive Statistics and Reliability Coefficients for Constructs

Constructors	Mean	S.D.	Reliability (alpha)	Scales adapt from
Perceived Trustworthiness (Trust)	5.33	1.01	(.74)	(Gefen 2000; Jarvenpaa et al. 2000; Portz 2000)
Privacy Concern (P_Concern)	3.82	1.51	(.89)	(Chen 2000)
Security Concern (S_Concern)	2.81	1.12	(.86)	(Gefen 2000; Swaminathan et al. 1999)
Perceived Benefit (Benefit)	5.42	1.21	(.85)	(Davis 1989; Moore et al. 1991; Swaminathan et al. 1999)
Willingness to Exchange (WE)	5.03	1.26	(.79)	(Gefen 2000; Jarvenpaa et al. 2000)

Note: N= 664

To examine convergent validity, an exploratory factor analysis of pooled constructs was conducted. Table 3 shows the results of factor analysis to measure the construct validity of the five factors. The items for each construct loaded into only one factor with eigenvalues greater than 1.0, which is an indication of convergent validity. The total cumulative percentage of variance explained by the five factors is 74.6%.

Table 3: Rotated Component Matrix (Factor Loadings)

Constructors	Items	Component				
		1	2	3	4	5
Perceived Trustworthiness (Trust)	PT1	-.064	.214	.344	.260	.715
	PT2	-.045	.218	.285	.225	.777
	PT6	-.178	.093	-.044	.107	.742
Privacy Concern (P_Concern)	PP2	.877	-.010	-.065	-.138	-.085
	PP3	.871	-.007	-.061	-.097	-.103
	PP4	.808	-.060	-.062	-.185	-.074
	PP5	.812	-.035	-.059	-.161	-.057
Security Concern (S_Concern)	PS4	-.189	.093	.142	.842	.191
	PS5	-.232	.185	.190	.794	.216
	PS8	-.244	.203	.237	.754	.162
Perceived Benefit (Benefit)	PB1	-.050	.812	.191	.165	.151
	PB2	-.050	.861	.152	.052	.215

	PB3	.005	.827	.198	.195	.081
Willingness to Exchange (WE)	WT1	-.030	.187	.786	.121	.131
	WT2	-.088	.162	.792	.131	.228
	WT3	-.114	.181	.783	.231	.020
Eigenvalue		3.052	2.369	2.287	2.280	1.948
Percent of Variance		19.078	14.809	14.292	14.253	12.175
Cumulative (%) Explained Variance		19.078	33.887	48.179	62.432	74.608

Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization.

Coding Scheme for Data Mining

Apriori is one of the most popular algorithms for mining frequent item sets for categorical association rules. To identify simple and powerful Boolean association rules (e.g. high and low trustworthiness), Likert 7 scale data need to be converted into Boolean variables. Using the mean values of each construct, the data were recorded as high and low cases. For instance, the mean value of the perceived trustworthiness is 5.33. Trust(L) and Trust(H) are recorded for the cases when the mean value of perceived trustworthiness is less than 5.33 or greater than 5.33, respectively. The ambiguous cases which have mean value were eliminated, since they can be interpreted as both high and low cases. Table 4 summarized the data coding scheme.

Table 4: Data Coding Scheme

Variable	Recoded Variable	Variable	Recoded Variable
Age	AGE(L), if Age < mean AGE(H), if age > mean	Perceived trustworthiness	Trust(L), if perceived trustworthiness < mean Trust(H), if perceived trustworthiness > mean
Gender	Male Female	Privacy concern	P_Concern(L), if privacy concern < mean P_Concern(H), if privacy concern > mean
Household Income	Income(L), if household income < 60,000 Income(H), if household income > 60,001	Security concern	S_Concern(L), if security concern < mean S_Concern(H), if security concern > mean
Money spent on this purchase	MSpend(L), if money spent < \$50 MSpend(H), if money spent > \$51	Perceived benefit	Benefit(L), if perceived benefit < mean Benefit(H), if perceived benefit > mean
Expertise on computer	CExp(L), if expertise on computer < mean CExp(H), if expertise on computer > mean	Willingness to exchange	WE(L), if willingness to exchange < mean WE(H), if willingness to exchange > mean
Expertise on the Internet	IExp(L), if expertise on the Internet < mean IExp(H) if expertise on the Internet > mean	Completion of purchase	Purchase Not_purchase

The Result and Analysis

Given the recorded dataset of the survey, the problem is to find all association rules that satisfy specified minimum support and minimum confidence with certain latent construct constraints. The Apriori program tries to generate all satisfied rules for the study with 10% of minimal support and 80% of minimal confidence.

The program begins with a minimum support of 100% of the data items and decreases this in steps of 5% until there are rules that satisfy the required minimum confidence, or until the support has reached a lower bound of 10%, whichever occurs first.

The Apriori program can also find association hyperedges (Han et al. 1998). Hyperedges are the set of items that are strongly predictive with respect to each other.

Using the Apriori algorithm, all association rules were selected with 10% minimal support and 60% of minimal confidence. For example, an association hyperedges, CExp(H) IExp(H) Trust(H) (23.4%, 88.0%), can be interpreted to mean that a consumer having high computer expertise and high Internet expertise has a high perception of trustworthiness, as well as 23.4% of support and 88% of confidence. From this result we know that there are strong predictive relationships among high computer skill, high Internet skill, and high perception of trustworthiness.

One of the main problems with the association rule induction is that there are so many possible rules. Of course, markets do not want just association rules. They want good rules, rules that are expressive, reliable, and applicable. However, good rules (rules that are often true) are not always interesting rules (rules that reveal something about the interdependence of the items). For example, it is easy to find out from a medical database that the rule “female \Rightarrow pregnant” true with a confidence of 100%. Even if it is a perfect rule, this is not an interesting rule. Another example of a not-applicable rule is the rule R2: purchase WE(H) \Rightarrow Trust(H). It is a common rule that the pre-condition of high willingness to exchange with high trustworthiness is associated with the consequent, purchase decision. But the revised case like R2 does not make sense in terms of a sequential behavior pattern. Therefore, even though the rule provides the association relationships among purchase, WE (H), and Trust (H), we can not consider it as an applicable rule. Thus, based on the purpose or motivation of the marketer, every rule should be filtered into interesting and applicable rules.

Some applicable rules related to the consequents, Trust(H), Trust(L), WE(H), WE(L), and Not_purchase for Internet retailers are summarized in Table 5.

Table 5: Association rules

Antecedent (if)	Consequents (then)	Minimal (Support ,Confidence)
Association Rules for Trust(H) as Consequent		
P_Concern(L) S_Concern(L) Benefit(H)	Trust(H)	(23.2%, 88.0%)
P_Concern(L) IExp(H) S_Concern(L)	Trust(H)	(20.6%, 85.4%)
AGE(L) WE(H)	Trust(H)	(21.5%, 85.0%)
Male AGE(H) P_Concern(H) S_Concern(H)	Trust(L)	(11.0%, 80.4%)
P_Concern(H) S_Concern(H) MSpend(H)	Trust(L)	(12.7%, 74.6%)
AGE(H) P_Concern(L) S_Concern(H) Benefit(L)	Trust(L)	(10.8%, 74.0%)
Income(L) S_Concern(H) MSpend(H)	Trust(L)	(11.4%, 73.6%)
AGE(H) IExp(L) S_Concern(H) Benefit(L)	Trust(L)	(11.2%, 73.1%)
P_Concern(H) IExp(L) S_Concern(H) Benefit(L)	Trust(L)	(14.0%, 72.3%)
P_Concern(H) S_Concern(H) Benefit(L)	Trust(L)	(20.4%, 71.6%)
Male S_Concern(H) Benefit(L)	Trust(L)	(16.3%, 71.1%)
Association Rules for WE(H) as Consequent		
P_Concern(L) S_Concern(L) Benefit(H) Trust(H)	WE(H)	(20.4%, 81.1%)
Trust(H) Benefit(H) S_Concern(L)	WE(H)	(23.2%, 77.1%)
Trust(H) S_Concern(L) P_Concern(L)	WE(H)	(21.1%, 74.2%)
S_Concern(L) IExp(H)	WE(H)	(22.2%, 68.2%)
Trust(H) P_Concern(L)	WE(H)	(24.5%, 67.9%)
Benefit(H) IExp(H)	WE(H)	(20.4%, 66.9%)
S_Concern(L)	WE(H)	(34.6%, 66.8%)
MSpend(L) S_Concern(L)	WE(H)	(20.6%, 66.2%)
Trust(H) IExp(H)	WE(H)	(22.4%, 65.0%)
Benefit(H)	WE(H)	(33.8%, 64.1%)
Trust(H)	WE(H)	(37.2%, 62.0%)
Trust(H) MSpend(L)	WE(H)	(22.4%, 61.2%)
IExp(H) CExp(H)	WE(H)	(20.9%, 60.2%)

Trust(L) Benefit(L) S_Concern(H)	WE(L)	(20.9%, 88.7%)
Trust(L) S_Concern(H) CExp(L)	WE(L)	(20.4%, 86.3%)
Benefit(L) S_Concern(H) IExp(L)	WE(L)	(20.2%, 86.2%)
Benefit(L) S_Concern(H) P_Concern(H)	WE(L)	(20.4%, 85.3%)
Benefit(L) S_Concern(H) CExp(L)	WE(L)	(21.9%, 85.3%)
Trust(L) IExp(L) CExp(L)	WE(L)	(21.5%, 82.0%)
Trust(L) S_Concern(H) P_Concern(H)	WE(L)	(20.2%, 80.9%)
Association Rules for Purchase and Not_purchase as Consequents		
WE(H) P_Concern(L) S_Concern(L) Male Trust(H)	Purchase	(10.3%, 81.3%)
MSpend(H) WE(H) Trust(H)	Purchase	(14.8%, 81.2%)
MSpend(H) WE(H) S_Concern(L) Trust(H)	Purchase	(12.3%, 80.7%)
WE(H) P_Concern(L) Male Trust(H)	Purchase	(11.8%, 80.0%)
WE(H) S_Concern(L) Benefit(H) Male Trust(H)	Purchase	(11.8%, 80.0%)
IExp(L) P_Concern(H) WE(L) CExp(L)	Not_purchase	(11.2%, 69.2%)
P_Concern(H) WE(L) MSpend(L) CExp(L)	Not_purchase	(10.3%, 68.8%)
Trust(L) S_Concern(H) Male Income(L)	Not_purchase	(11.0%, 66.7%)
IExp(L) P_Concern(H) WE(L)	Not_purchase	(12.0%, 66.1%)
P_Concern(H) WE(L) MSpend(L)	Not_purchase	(13.3%, 66.1%)
CExp(H) IExp(H) WE(L)	Not_purchase	(13.8%, 65.6%)

Based on the output of the Apriori program, several good rules, i.e. interesting and applicable rules, can be identified. For example, R3: "P_Concern(L) S_Concern(L) Benefit(H) \Rightarrow Trust(H) (23.2%, 88.0%)" infers that a customer who perceived a low privacy concern, low security concern, and had a high perception of benefit had a high perception of trustworthiness with 23.2% of support and 88.0% of confidence. The support of an association rule is the percentage of those transactions in the set of all transactions under consideration which contain the item set. In other words, 23.2% of all transactions contain P_Concern(L), S_Concern(L), Benefit(H), and Trust(H). The confidence of a rule is intuitively the number of cases in which the rule is correct relative to the number of cases in which it is applicable. For instance, if a customer has low privacy concern, low security concern, and high benefit, then the rule is applicable and it says that he or she can be expected to have a high perception of trustworthiness. If she or he does not have low privacy, security concern or high benefit, then neither does she or he have high trustworthiness.

Probably, Internet retailers are more interested in the antecedents of a low level of trust. Since trust is a prerequisite for behavioral commitment, increasing consumers' trust will induce consumers' high level of purchase intention. The rule R4: Male AGE(H) P_Concern(H) S_Concern(H) \Rightarrow Trust (L) (11.0%, 80.4%), may be interpreted as follows. A consumer's low level of trustworthiness (Trust(L)) is associated with high privacy concern (P_Concern(L)), and high security concern (S_Concern(H)).

Another interesting finding from the results is the effect of antecedents on consequen ces. Suppose the minimal confidence and the minimal support are 80% and 10% respectively. From the rule, R4: Trust(H) \Rightarrow WE(H) (37.2%, 62.0%), for example, we know that there is an associational relationship between high trustworthiness and high willingness to exchange. But the rule itself is not strong enough since the confidence (62%) is less than the minimal confidence (80%). When we look at other rules (R5, R6, and R7) that include more antecedents (security concern, privacy concern, and perceived benefit) for the consequence (willingness to exchange), the confidence increases gradually and, finally, R7 crosses over the minimal support (80%).

R5: Trust(H) S_Concern(L) P_Concern(L) \Rightarrow WE(H) (21.1%, 74.2%)

R6: Trust(H) Benefit(H) S_Concern(L) \Rightarrow WE(H) (23.2%, 77.1%)

R7: P_Concern(L) S_Concern(L) Benefit(H) Trust(H) \Rightarrow WE(H) (20.4%, 81.1%)

From the finding above, we can interpret that high trustworthiness itself is not the only antecedent for consumers' high willingness to exchange. With other factors such as high perception of benefit, low privacy concern, and low security concern of a transaction with Internet retail, a consumer has a high degree of willingness to exchange.

Discussion and Conclusion

Using survey data items, this study has tackled the problem of identifying factors related to consumers' trust. We deal with qualitative data and expressions of opinions, i.e. survey items, rather than with transactional data. Most data mining studies (Adriaans et al. 1996) utilized transactional data based on actual behavior. However, according to the Theory of Reasoned Action (Fishbein et al. 1975), actual behavior follows from behavioral intention, which is captured in this study by the consumers' perceptions of the survey. TRA provides a framework to study attitudes toward behaviors. This study suggests that an analysis of perceived behavioral intention can be valuable in the context of trust management.

This study has both theoretical and practical implications. The findings of this study extend our knowledge of the association relationships among factors that affect a consumer's trust and behavioral commitment. The association rules selected by the Apriori algorithm highlight several trust-related antecedents that affect a consumer's purchase intention and finally influence the successful completion of an Internet transaction. Consumers' privacy and security concern are strongly associated with consumers' trust and purchase intention. This result provides evidence to support the research hypothesis: consumers' privacy concern (security concern) negatively affects the perceived trustworthiness. Thus, the result empirically suggests that Internet retailers should make efforts to better incorporate trust-building mechanisms by focusing on the impact of consumers' privacy and security concerns with online purchases. Another implication from a marketing perspective is that customers who have a high degree of trust might have a high probability of becoming loyal customers in the near future (Chow et al. 1997). We may infer association rules from a marketing perspective. For example, customers who have a low perception of trust may be good indicators of how the element of trust should be managed in order to assure a greater level of trust in future customers (Marcella 1999).

There are several limitations of the study. One of the limitations of the study is the relatively small amount of data. Even though there is no problem to use consumer's purchase intention data, the number of samples used for the study is relatively small for data mining techniques. Another limitation of the study is the self-reporting bias of the respondents. Since this study deals with latent constructs (e.g. trust), there is a potential possibility regarding consumers' intentionally incorrect responses to the survey items. In order to apply the Apriori algorithm, a simple Boolean association rule mining technique, the data was recorded in only high and low cases. Therefore, loss of information due to the recording process is one more limitation of the study.

As for antecedents of trust and willingness to exchange, only latent constructs (i.e. perceived trustworthiness, privacy concern, security concern, perceived benefit, and willingness to exchange) were considered in this study. I do not deny the importance of other factors as antecedents such as familiarity with a website, perception of system reliability, information quality, presence of a third party seal, and so on. Therefore, investigation of association rules, including other antecedents, is appropriate for future study.

References

- Adriaans, P., and Zantinge, D. *Data Mining* Addison-Wesley, 1996.
- Agrawal, R., Imielinski, T., and Swami, A. "Mining Association Rules Between Sets of Items in Large Databases," the ACM SIGMOD Conference on Management of Data, 1993, pp. 207-216.
- Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., and Verkamo, A. "Fast Discovery of Association Rules," in: *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, S. Amith, P. and R. Uthurusamy (eds.), MIT Press, Cambridge, 1996a, pp. 307-328.
- Agrawal, R., and Shafer, J. "Parallel Mining of Association Rules," *IEEE Transactions on Knowledge and Data Engineering*, (8:6) 1996b.
- Agrawal, R., and Srikant, R. "Fast Algorithms for Mining Association Rules," Very Large Data Bases Conference, 1994, pp. 487-499.
- Ajzen, I., and Fishbein, M. *Understanding Attitude and Predicting Social Behavior* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1980.
- Ali, K., Manganaris, S., and Srikant, R. "Partial Classification using Association Rules," the 3rd International Conference on Knowledge Discovery in Databases and Data Mining, 1997.
- Anderson, J.C., and Gerbing, D.W. "The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness-of-fit Indices for Maximum Likelihood Confirmatory Factor Analysis," *Psychometrika* (49) 1984, pp 155-173.
- Ba, S., Whinston, A.B., and Zhang, H. "Building Trust in the Electronic Market through an economic incentive mechanism," the 1999 International Conference on Information Systems, Charlotte, NC, 1999, pp. 208-213.
- Beatty, S.E., Mayer, M., Coleman, J.E., Reynolds, K.E., and Lee, J. "Customer-Sales Associate Retail Relationships," *Journal of Retailing* (72:3) 1996, pp 223-247.

- Bentler, P.M., and Chou, C.P. "Practical issues in Structural Modeling," *Sociological Methods & Research* (16) 1987, pp 78-117.
- Berry, L.L. "Relationship Marketing of Services-Growing Interest, Emerging Perspectives," *Journal of Academy of Marketing Science* (23:4) 1995, pp 236-245.
- Chang, M.K., Cheung, W., and Lai, V.S. "Literature Derived Reference Models for the Adoption of Online Shopping," *Information and Management* (42) 2005, pp 543-559.
- Chen, L.-d. "Consumer Acceptance of Virtual Stores: A Theoretical Model and Critical Success Factors for Virtual Stores," University of Memphis, 2000.
- Cheung, C.M.K., and Lee, M.K.O. "Understanding Consumer Trust in Internet Shopping: A Multidisciplinary Approach," *Journal of the American Society for Information Science and Technology* (57:4) 2006, pp 479-492.
- Chow, S., and Holden, R. "Toward an Understanding of Loyalty: the Moderating Role of Trust," *Journal of Managerial Issues* (15:3) 1997, pp 275-298.
- Czepiel, J.A. "Service Encounters and Service Relationships: Implications for Research," *Journal of Business Research* (20:1) 1990, pp 13-21.
- Davis, F.D. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3) 1989, pp 319-340.
- Fishbein, M., and Ajzen, I. *Belief, Attitude, Intention, and Behavior: an Introduction to Theory and Research* Addison-Wesley, 1975.
- Gefen, D. "E-commerce: the role of familiarity and trust," *The International Journal of management Science* (28) 2000, pp 725-737.
- Grabosky, P. "The Nature of Trust Online," in: *IT News*, 2001.
- Han, E.-H., Karypis, G., Kumar, V., and Mobasher, B. "Hypergraph Based Clustering in High-Dimensional Data Sets: A Summary of Results," *Data Engineering Bulletin* (21:1) 1998, pp 15-22.
- Hoffman, D.L., Novak, T.P., and Peralta, M. "Building Consumer Trust Online," *Association for Computing Machinery. Communications of the ACM* (42:4) 1999, pp 80-85.
- Jarvenpaa, S.L., Tractinsky, N., and Vitale, M. "Consumer Trust in an Internet Store," *Information Technology and Management* (1:1-2) 2000, pp 45-71.
- Kim, D., and Benbasat, I. "The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation," *Information Systems Research* (17:3) 2006, pp 286-300.
- Kim, D.J., Song, Y.I., Braynov, S.B., and Rao, H.R. "A Multi-dimensional Trust Formation Model in B-to-C E-Commerce: A conceptual Framework and Content Analyses of Academia/Practitioner Perspective," *Decision Support Systems* (40:2) 2005, pp 143-165.
- Lee, J.-N., and Kim, Y.-G. "Effect of Partnership Quality on IS Outsourcing: Conceptual Framework and Empirical Validation," *Journal of Management Information Systems* (15:4) 1999, pp 29-61.
- Lim, K.H., Sia, C.L., Lee, M.K.O., and Benbasat, I. "Do I Trust You Online, and If So, Will I Buy? An Empirical Study of Two Trust-Building Strategies," *Journal of Management Information Systems* (23:2) 2006, p 233.
- Madden, T.J., Ellen, P.S., and Ajzen, I. "A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action," *Personality and Social Psychology Bulletin* (18) 1992, pp 3-9.
- Marcella, A.J. *Establishing Trust in Virtual Markets* The Institute of Internal Auditors, Altamonte Springs, Florida, 1999.
- Mayer, R.C., Davis, J.H., and Schoorman, F.D. "An Integrative Model of Organizational Trust," *Academy of Management Review* (20:3) 1995, pp 709-734.
- Moore, G.C., and Benbasat, I. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information Systems Research* (2:3) 1991, pp 192-222.
- Morgan, R.M., and Hunt, S.D. "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing* (58) 1994, pp 20-38.
- Noteberg, A., Christaanse, E., and Wallage, P. "The Role Of Trust And Assurance Services In Electronic Channels: An Exploratory Study," the Twentieth International Conference on Information Systems, Omnipress, Charlotte, North Carolina, 1999.
- Pavlou, P.A., and Gefen, D. "Building Effective Online Marketplaces with Institution-Based Trust," *Information Systems Research* (15:1) 2004, pp 37-59.
- Portz, K.S. "The Effect of WebTrust on the Perceived Trustworthiness of a Web Site and the Utilization of Electronic Commerce," University of Nebraska, Lincoln, Doctoral Dissertation, 2000.
- Reichheld, F.F. "Loyalty and the Renaissance of Marketing," *Marketing Management* (2:4) 1994, pp 10-21.
- Srikant, R., Vu, Q., and Agrawal, R. "Mining Association Rules with Item Constraints," 3rd International Conference Knowledge Discovery and Data Mining, AAAI Press, 1997, pp. 67-73.
- Swaminathan, V., Lepkowska-White, E., and Rao, B.P. "Browsers or Buyers in Cyberspace? An Investigation of Factors Influencing Electronic Exchange," *Journal of Computer Mediated Communications* (5:2) 1999.