DECISION SCIENCES INSTITUTE

Who Subscribe to Identity Theft Protection Service? An Exploration of Antecedent Factors

Yuan Li
University of Illinois at Springfield
Email: yli295@uis.edu

Jingguo Wang
University of Texas at Arlington
Email: jwang@uta.edu

H. Raghav Rao
University of Texas at San Antonio
Email: hr.rao@utsa.edu

ABSTRACT

This study explores the antecedent factors that influence a person's adoption of identity theft protection service. Fifteen factors related to the demographic, behavioral, and psychological attributes of individuals are examined. Data gathered from 600 consumers (including 145 subscribers and 455 non-subscribers) were tested using Logistic Regression, Decision Tree, Neural Network, and Support Vector Machines (SVM) methods. The accuracy and sensitivity measures suggest that SVM performs the best. Further analysis based on SVM suggests that past victimization related to personal information and credit cards, and also gender, are the three most important antecedents to subscription to identity theft protection service.

KEYWORDS: Identity theft, Identity theft protection service, Classification, Logistic Regression, Decision Tree, Neural Network, Support Vector Machines

INTRODUCTION

Identity theft is the misuse of another individual's personal information to commit fraud (Gonzales & Majoras, 2007, p.2). It has become a critical issue in today's networked environment due to the extensive and intensive use of electronic communications in one's daily work and life. Identity theft protection service, referring to the combination of identity monitoring, recovery, and insurance services, helps to address identity theft. Past research on identity theft has focused on technological and legislative solutions (Lai et al., 2012), but limited attention has been paid to identity theft protection service (Kim & Kim, 2016). This study addresses the issue by examining the demographic, behavioral, and psychological factors of individuals that may influence their adoption of identity theft protection service. The objective of the study is to develop classification models from these factors to predict (or prescribe) future subscription behavior.

Data from 600 U.S. consumers, including 145 subscribers to identity theft protection service and 455 non-subscribers, were collected via an online survey using the Qualtrics research panel. Demographic factors examined include a person's age, gender, education, and income; behavioral factors include a person's daily email loads, number of credit cards, online transaction experience, and past victimization related to the misuse of the person's credit cards,

online accounts, and personal information; psychological factors include a person's perceived vulnerability and severity of phishing attack and identity theft (since phishing is a common threat to identity theft), perceived phishing detection efficacy, anxiety about phishing attack, and dispositional optimism. These factors are examined via a number of classification models commonly used in classification tasks in recent literature (Delen, 2010; Thammasiri et al., 2014), including Logistic Regression, Neural Network, Decision Tree, and Support Vector Machines.

We used accuracy and sensitivity measures (Thammasiri et al., 2014) to compare the performance of the models. The results suggest that SVM exhibits the best predictive performance. Further analysis based on SVM shows that of the predictive factors, past victimization related to personal information and credit cards, and also gender, are the three most important antecedents. The study is reported as follows.

LITERATURE REVIEW

The proliferation of online transactions has led to a large number of incidents of identity theft that incurred expensive costs to consumers and e-commerce providers (Lai et al., 2012). To effectively deal with identity theft, various solutions have been proposed, such as task-focused coping and emotion-focused coping (Anandarajan et al., 2012), conventional coping and technological coping (Lai et al., 2012), and subscription to identity theft protection service (Kim & Kim, 2016). To date, most research has focused on technological, legislative, and even individual behavioral approaches to identity theft protection, but limited attention has been paid to identity theft protection service despite the existence of the service for years (Furnell, 2007).

A study on Korean consumers shows that perceived usefulness of identity theft protection service has a direct impact on the adoption of the service (Kim & Kim, 2016), while other psychological factors, such as perceived vulnerability, have indirect impact. However, further empirical research on other antecedent factors is lacking. We address this issue by exploring three broad categories of factors, including a person's demographic, behavioral, and psychological attributes. These factors from different perspectives influence a person's decision to adopt identity theft protection service. Specifically, demographic factors include a person's age, gender, education, and income, as these factors influence a person's risk-taking behavior in general (Chen et al., 2011). Behavioral factors include a person's daily email load, number of credit cards, intensity of online transactions, and past victimization related to misuse of one's credit cards, online accounts, and personal information. These factors were selected as they reflect one's online activities that may expose one to potential identity theft threats, and also actual victimization experience. Finally, psychological factors examine one's perceptions regarding identity theft and protection, including perceived vulnerability and severity of phishing attack and identity theft, perceived phishing detection efficacy, anxiety about phishing attack, and dispositional optimism. In the following sections, the characteristics of these factors and their impacts on the adoption of identity theft protection service are examined.

Demographic factors

Demographic factors constitute the foundation of many customer profiling and classification techniques. These factors have potential impacts on the risk-raking behavior of individuals (Chen et al., 2011), and are commonly examined as control variables in information security and identity theft protection literature (Chen & Zahedi, 2016). A literature reviewer on information privacy, for instance, suggests that women are more concerned about their information privacy than men, and that age in general has a positive impact on privacy concerns (Li, 2011). Similar

findings were reported in the identity theft protection literature. For example, Holt and Turner (2012) show that males differ from females in terms of the types of protective factors they use to remain resilient for identity theft, and that compared to females, males face a greater risk of fraud victimization. Similarly, Tsai et al (2016) find that females are in general more willing to take secure online behavior than males, although the effect does not reach significance.

The effect of age on victimization and protective behavior was also examined, but in most studies the effect remained insignificant (Holt & Turner, 2012; Lai et al., 2012; Tsai et al., 2016). Similarly, the effects of education and income on online safety behavior were also examined (Lai et al., 2012; Tsai et al., 2016), and the effects remained insignificant, too. A potential reason is that prior research on information security and identity theft seldom considered other aspects, such as the cost, of taking the behaviors, so that factors such as income played limited roles. When cost becomes a decision variable (Boss et al., 2015), as in identity theft protection service subscription, the person would compare the cost of identity theft protection service to the benefit, and such factors may play a role. We therefore include these factors in this exploratory study.

Behavioral factors

It is generally accepted that past behavior and experience have impact on future behavior, as they strengthen personal understanding of the consequences of the behavior such as victimization and identity theft. In this study, we examine behaviors and experiences that may expose a person to potential identity theft, and also past experiences with identity victimization.

For behaviors that may expose a person to potential identity theft, we study the person's daily email loads, number of credit cards, and intensity of online transactions. As mentioned above, extensive and intensive use of online communications such as emails elevates the person's vulnerability to identity theft, and one typical way of identity theft is through phishing email attack (Vishwanath et al., 2011). Despite technological solutions to filtering phishing emails, some phishing emails can still bypass the filters and land in one's email in-box, posing security threat. On the other hand, outgoing emails that contain sensitive information may be eavesdropped on, resulting in identity theft. Therefore, if one uses emails extensively and intensively, identity theft protection service may be a viable option to prevent potential identity theft.

Similarly, the extensive and intensive use of credits cards and online transactions may expose one to identity theft. Credit card fraud represents one of the most common incidents of identity theft, and is the main focus of identity theft protection service especially offered by financial firms. Intensity of online transactions also increases the chances that a person's identity is stolen, especially through the data breach at merchants' websites (Roberds & Schreft, 2009). All these suggest that the online transactional behaviors of individuals may make them more alert to identity theft and enhance their adoption of identity theft protection service.

In addition to the behaviors that may lead to identity theft, we also examine a person's actual victimization experience. We particularly focus on victimization experiences related to the misuse of one's credit cards, online accounts, and personal information, as these are typical incidents of identity theft. We expect that these experiences will facilitate the adoption of identity theft protection service.

Psychological factors

Psychological factors represent the underlying drivers of one's intention to use identity theft protection service. Early, Kim and Kim (2016) examined five psychological factors (or individual perceptions) that may influence the adoption of the service, and only perceived usefulness and social influence were significant. The antecedents were mostly developed from the Protection Motivation Theory (Rogers, 1975) perspective, although perceived response efficacy in the theory was operationalized as perceived usefulness. We adopted the original items in the context of phishing detection because, as mentioned above, phishing attack is a major vehicle for identity theft. Specifically, we examine perceived vulnerability and severity of phishing attack and identity theft, and also perceived phishing detection ability. We expect perceived vulnerability and severity to enhance the adoption of identity theft protection service, and perceived phishing detection ability to reduce the need for such service.

RESEARCH METHOD AND RESULTS

We conducted an exploratory study to examine the effects of the fifteen antecedents on identity theft protection service adoption. This study is part of a larger research project that addresses one's ability to deal with phishing attack and identity theft. Specifically, an online survey was conducted via the Qualtrics research suite, and 600 U.S. consumers participated in the study. The survey questionnaire asked the participants to assess items that measure each of the antecedent factors, and also report whether they subscribed to identity theft protection service. The measurement and data types of the variables are summarized in Table 1.

Table 1: Measurement and Data Types of the Variables

No.	Variables	Measurement	Data Type
1	Age	Single item	Numeric
2	Gender	Single item	Categorical
3	Education	Single item	Ordinal
4	Income	Single item	Ordinal
5	Daily email loads	Single item	Numeric
6	Number of credit cards	Single item	Numeric
7	Intensity of online transactions	Multiple items	Ordinal
8	Past victimization-credit cards	Single item	Categorical
9	Past victimization-online accounts	Single item	Categorical
10	Past victimization-personal information	Singe item	Categorical
11	Perceived vulnerability	Multiple items	Ordinal
12	Perceived severity	Multiple items	Ordinal
13	Perceived phishing detection ability	Multiple items	Ordinal
14	Perceived anxiety	Multiple items	Ordinal
15	Dispositional optimism	Multiple items	Ordinal
16	Subscription to identity theft protection service	Single item	Categorical

The whole sample consists of 145 subscribers and 455 non-subscribers. The number of subscribers is less than the number of non-subscriber, but it is consistent with prior findings (Kim & Kim, 2016). It also highlights the need to convert non-subscribers to subscriber to deal with identity theft.

Data analysis and results

To examine how the factors influence one's adoption of identity theft protection service, we employed four classification models that were commonly used in literature (Delen, 2010; Thammasiri et al., 2014), including Binary Logistic Regression (LR), CHAID-based Decision Tree (DT), Multilayer Perceptron Neural Network (NN), and Support Vector Machines (SVM). To compare the performance of the models, we adopted the accuracy and sensitivity measures, the formulas of which can be found in Thammasiri et al (2014, p. 327). Accuracy refers to the overall percentage of correct classifications, i.e., the number of correctly predicted subscribers and non-subscribers out of all the subjects (i.e., 600), and sensitivity, also called true positive rate, refers to the percentage of subscribers who are corrected predicted from all actual subscribers (i.e., 145). We focus on the sensitivity measure in order to find out which model is more capable of identifying subscribers (instead of non-subscribers). Other measures of classification performance (Thammasiri et al., 2014) are of no interest to this study and are ignored.

The analysis is done through SPSS v.24. The results based on the whole sample are reported in Table 2. Although the overall accuracy is high for each model, the sensitivity measure revealed the problem of imbalanced class distribution (Thammasiri et al., 2014). This is, because the majority (76%) of the subjects are non-subscribers, all four classification models had bias toward non-subscribers, leading to high accuracy but low sensitivity (i.e., the ability to identity subscribers).

Table 2. Classification Results based on the Whole Sample

	LR	DT	ANN	SVM
Accuracy	77.3%	80.7%	80.0%	78.8%
Sensitivity	16.6%	30.3%	27.7%	12.4%

To address the problem of imbalanced data, we followed the technique by Thammasiri et al (2014). Specifically, all the 145 subscribers were isolated from the sample, and then for the remaining non-subscribers, 145 of them were randomly selected and combined with the subscribers. This yielded a new balanced sample of 145 subscribers and 145 non-subscribers. To reduce potential bias in sampling, the procedure was repeated 10 times (Thammasiri et al., 2014) and the performances measures across the 10 estimates were aggregated (i.e., means). The results are show in Table 3. The table shows that SVM performs the best among the competing models, with a mean accuracy of 76.6% and mean sensitivity of 70.4%. The improved sensitivity (70.4%) suggests that the SVM model is capable of correctly identifying the subscribers, as compared to random classification (i.e., 50-50 or 50%) or based on sample distribution (i.e., 24%).

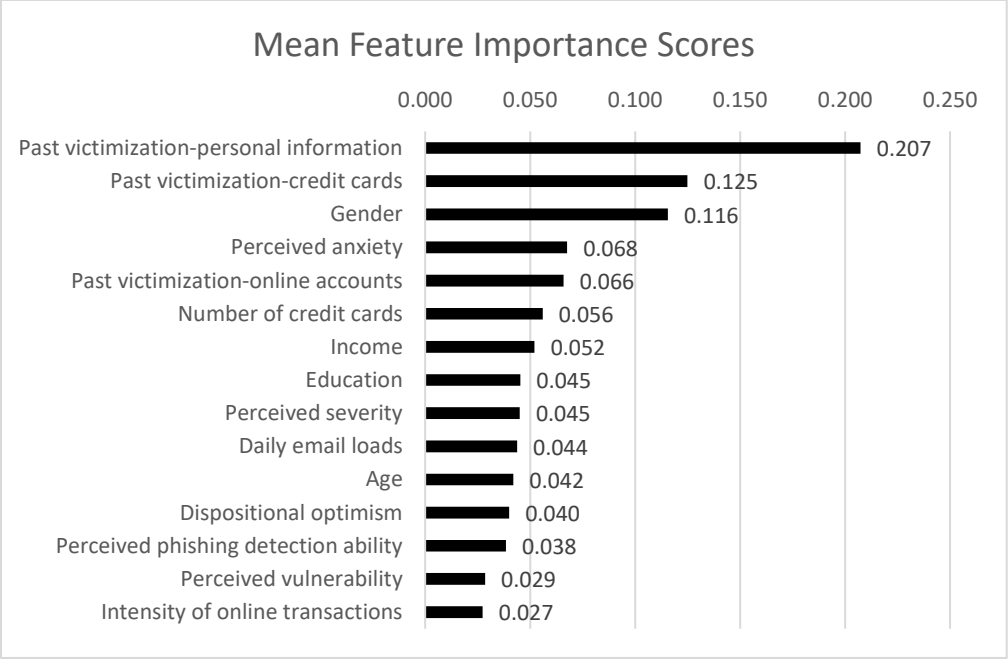
Based on SVM, we further examined the impact of each antecedent on subscription using their feature importance scores (Thammasiri et al., 2014). The scores reflect how important a factor is in the classification, and were derived from the SVM module in SPSS Modeler (v.18) since the SVM add-on in SPSS v.24 (based on R) does not generate such scores. The mean values of the feature importance scores of the antecedents based on the 10 samples are presented in Figure 1 in descending orders. The figure suggests that past victimization experience regarding one's personal information is the most important determinant of subscription to identity theft protection service, followed by past experience with credit card victimization. The third most

importance factor is gender, with women more likely to subscribe to the service than men. Compared to the above three factors, other factors have relatively less important impacts.

Table 3. Classification Results based on the Balanced Samples

Accuracy											
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	Mean
LR	63.8%	63.1%	67.9%	61.0%	65.9%	62.1%	64.8%	65.2%	66.6%	60.3%	64.1%
DT	64.8%	70.0%	66.2%	64.8%	64.5%	65.5%	66.2%	66.2%	63.4%	61.0%	65.3%
ANN	66.3%	61.5%	66.7%	67.5%	63.0%	65.3%	67.4%	60.7%	57.7%	61.4%	63.8%
SVM	76.2%	76.9%	76.9%	77.2%	76.9%	76.9%	76.9%	75.9%	77.6%	74.8%	**76.6%**
Sensitivity											
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	Mean
LR	59.3%	58.6%	64.1%	56.6%	62.1%	55.2%	59.3%	62.8%	61.4%	57.9%	59.7%
DT	65.5%	65.5%	51.0%	64.1%	60.7%	64.1%	66.9%	57.2%	57.9%	25.5%	57.8%
ANN	65.4%	66.7%	60.9%	67.4%	57.5%	66.7%	65.9%	55.3%	47.1%	44.7%	59.8%
SVM	70.3%	71.0%	64.1%	69.7%	74.5%	71.0%	72.4%	73.1%	70.3%	67.6%	**70.4%**

Figure 1. Feature importance scores of the antecedents



DISCUSSION AND CONCLUDING REMARKS

To the best of our knowledge, this study is the first to employ a multi-faceted data to examine a person's subscription to identity theft protection service. This addresses limitations in prior literature that examined psychological factors only (Kim & Kim, 2016). The application and comparison of multiple classification models also helps to avoid limitations of using one model and therefore enhances the acceptance of the results.

For an applied research, our study has implications for practice. First of all, our survey reveals that only 24% of the subjects subscribed to identity theft protection service, suggesting the need to further promote this service. Second, the recognition of important factors that include subscription behavior provides guidance for service providers (banks, credit card companies, and third-party providers, etc.) to target those potential customers.

In terms of research, our study highlights the importance of past experience on one's subscription to identity theft protection service. It suggests including these factors in future research on information security behavior. On the other hand, the lack of importance of the psychological factors suggests that these factors may deserve further investigation, as such factors have been typically examined for their impact on behavioral intentions (such as willingness to adopt the service) rather than actual behavior.

REFERENCES

Anandarajan, M., Paravastu, N., Arinze, B. & D'Ovdio, R. (2012). Online Identity Theft: A Longitudinal Study of Individual Threat-Response and Coping Behaviors, Journal of Information System Security 8(2), 43-69.

Boss, S.R., Galletta, D.F., Benjamin Lowry, P., Moody, G.D. & Polak, P. (2015). What Do Systems Users Have to Fear? Using Fear Appeals to Engender Threats and Fear That Motivate Protective Security Behaviors, MIS Quarterly 39(4), 837-864.

Chen, R., Wang, J., Herath, T. & Rao, H.R. (2011). An Investigation of Email Processing from a Risky Decision Making Perspective, Decision Support Systems 52(1), 73-81.

Chen, Y. & Zahedi, F.M. (2016). Individuals' Internet Security Perceptions and Behaviors: Polycontextual Contrasts between the United States and China, MIS Quarterly 40(1), 205-A212.

Delen, D. (2010). A Comparative Analysis of Machine Learning Techniques for Student Retention Management, Decision Support Systems 49(4), 498-506.

Furnell, S. (2007). Identity Impairment: The Problems Facing Victims of Identity Fraud, Computer Fraud & Security 2007(12), 6-11.

Gonzales, A.R. & Majoras, D.P. (2007). "Combating Identity Theft: A Strategic Plan." The President's Identity Theft Task Force.

Holt, T.J. & Turner, M.G. (2012). Examining Risks and Protective Factors of on-Line Identity Theft, Deviant Behavior 33(4), 308-323.

Kim, A.-Y. & Kim, T.-S. (2016). "Factors Influencing the Intention to Adopt Identity Theft Protection Services: Severity Vs Vulnerability," in: Proceedings of the 2016 Pacific-Asia Conference on Information Systems (PACIS).

Lai, F., Li, D. & Hsieh, C.-T. (2012). Fighting Identity Theft: The Coping Perspective, Decision Support Systems 52(2), 353-363.

Li, Y. (2011). Empirical Studies on Online Information Privacy Concerns: Literature Review and an Integrative Framework, Communications of the Association for Information Systems 28(28), 453-496.

Roberds, W. & Schreft, S.L. (2009). Data Breaches and Identity Theft, Journal of Monetary Economics 56(7), 918-929.

Rogers, R.W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change, Journal of Psychology 91(1), 9-114.

Thammasiri, D., Delen, D., Meesad, P. & Kasap, N. (2014). A Critical Assessment of Imbalanced Class Distribution Problem: The Case of Predicting Freshmen Student Attrition, Expert Systems with Applications 41(2), 321-330.

Tsai, H.-y.S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N.J. & Cotten, S.R. (2016). Understanding Online Safety Behaviors: A Protection Motivation Theory Perspective, Computers & Security 59(138-150.

Vishwanath, A., Herath, T., Chen, R., Wang, J. & Rao, H.R. (2011). Why Do People Get Phished? Testing Individual Differences in Phishing Vulnerability within an Integrated, Information Processing Model, Decision Support Systems 51(3), 576-586.