Privacy Risk of Personal Information Exposure

N. Yahya[1], N. Mansor[2,3], N. Z. Nizam[2,3] and R. Ruslan[1]

[1]*Fakulti Pengurusan Teknologi dan Perniagaan, Universiti Tun Hussein Onn.*
[2]*Fakulti Pengurusan Teknologi dan Teknousahawanan, Universiti Teknikal Malaysia Melaka.*
[3]*Centre for Technopreneurship Development (C-TeD), Universiti Teknikal Malaysia Melaka*

nusaibah@utem.edu.my

Abstract—**Customers individual private information are collected to gain valuable information. Through business analytics, companies gain competitive advantage in making decisions. The fact that some irresponsible companies sell customer's data to third party for illegal activities. This caused distrust and lack of confidence among consumers in revealing their private information. The purpose of this study is to investigate consumer sensitivity towards data privacy. This study adopted a research framework based on 'Concern for Information Privacy' (CFIP) dimensions constructed into five independent variables that is extracted from questionnaires distributed to 384 respondents, while the regression approach was utilized to analyse the relationship between among identified variables. Result found that collection, unauthorized secondary internal use, unauthorized secondary external use, improper access and errors have a positive influence on consumer sensitivity towards data privacy. Furthermore, the effects of unauthorized secondary internal use on consumer sensitivity towards data privacy is found to be significantly greater.**

Keywords—data privacy; consumer sensitivity; risk

I. INTRODUCTION

CONSUMER today beginning to concern about their information privacy which had been utilized by retailers, manufacturers, marketers, site and web. They concern about how these parties observe their action by exploitation their information privacy. Due to technology advancement and the convenience of the web today, causes the problems and cases of privacy breech arising that is that they're going to misuse the data. However, all this crime and problems are going to be straightforward to try to through the web.

Marketers currently will track what's precisely a consumer purchase, wherever the buyer buys it, once the buyer buys it, what quantity consumer buys it and the way typically the buyer buys it. Using grocery discount cards and store credit cards allow the marketers to use, record and even can sell information relating to consumer's card-paying patterns. But, how do we measure consumer sensitivity towards marketers who use data gathered from like discount and loyalty purchases. Sometimes, consumer got to provide the customer's basic information data to end the trade-off or group action. Somehow consumer is requested by the marketers to stock up a lot of data regarding their information personal data like hobbies and private interest so as to receive a lot of regarding personalized services.

With technology advancement, the internet could be a potential tool in promoting products and services in marketplace. The facts, various of existing internet promoting activities will however have negative effects on consumer's attitudes towards information privacy. Some consumers are ignorant on the usage of junk e-mails and internet cookies in mining their personal data. Their illiteracy on internet usage caused their personal information at risk to be exposed to internet predators. These differences impact several aspects of selling, as well as the perception of consumer privacy and therefore the use of consumer data.

Article history: Manuscript received 31 August 2019; received in revised form 4 October 2019; Accepted 7 October 2019.

II. LITERATURE REVIEW

The digital technology advancement and e-commerce has raised client privacy concern all around the world [1]. Consumer who are concern regarding their information info is unwilling to disclose personal information on websites and unwilling to form any transaction as a result of they need sturdy influence on privacy concern [2]. A found that ladies are more concern over general privacy issues and identity data speech act than men [3]. Finally, information privacy is a vital factor that concern to the knowledge privacy as a result of the extent of risk that will be raising up.

Collection is the amounts of identifiable information to be collected and stored in database [4]. An individual may become resentful if their data are collected in great quantity based on personal preferences and background or in other words, personal information. Collection also is defining as the degree to which an individual is concerned about the amount of individual specific data possessed by other relative to the value of benefits received [5]. A study by Phelps et al. found that majority of the respondents wanted to limit the collection of personal information by marketers in terms of direct marketing [6].

A. Unauthorized Secondary Internal Use

Unauthorized secondary internal use is when an organization collected information from individuals for one single use but used it for secondary purposes without authorization from the individuals [7]. Fortes and Rita, also define similarly unauthorized secondary internal use which refers to the concern that the data collected for a specific purpose is used by the corporations for other purposes without the consent of the individual [8]. Hence, an organizational use personal information for undisclosed purposes for the internal unauthorized secondary.

B. Unauthorized Secondary External Use

According to Smith et al., they have found that the concern about the secondary use is arising when personal information is disclosed to another organizational which is an external party [7]. Consumer that concern about the information is collected for one purpose but is used for another, secondary purpose after disclosure to an external party which is not the collecting organization. This is supported by Smith et al., which refers unauthorized secondary external user as the concern that the information collected for a specific purpose is used for another purpose after its disclosure to an external organization.

C. Improper Access

Anyone who is working at the organization is not allowed to access personal information of others. This is supported by Fortes and Rita, which is improper access is referred to the concern that personal data be made available to the person who is not properly authorized to do so [8]. An individual should have the need to know before access to personal information that is granted. According to Tang and Lin, improper access is defined as when individual's personal data is without appropriate protection and some 'illegal parties or users' may gain access your data personal [4]. Consumers may decide to avoid purchase online as they may face with the possibility of improper access and their personal information being sold to others for wrongful reason. The risk of their personal information being improper access and being sold to others for unlawful activities might cause the possibilities of financial losses.

D. Errors

Errors are referring to the improper protection and violations of an individual's personal information which is data privacy [4]. Most individuals know that an organization is not taking enough step to lower the problem or errors in data personal information. Although some errors might be intended such as dissatisfied employee maliciously falsifying data. Errors also refer to which pointing to concern that the protection against accidental and deliberate errors in personal data would not be appropriate [8]. Meanwhile, according to Smith et al., they mention that errors are the concern that protections against deliberate and accidental errors in personal data are inadequate [7].

The following Fig. I shows the adoption of research framework to guide the research process. This study has come out with the research framework as show at Fig. I. The independent variable is the dimension of Concern for Information Privacy (CFIP) and the sub variables is collection, unauthorized secondary internal use, unauthorized secondary external use, improper access and errors. The dependent variable is the consumer sensitivity towards data privacy. The foundation of this research framework is developed based on a few researchers such as Smith et al. which original work proposed the CFIP dimensions, Fortes and Rita and Tang and Lin [4], [7], [8].

III. METHODOLOGY

To achieve the aim of this study, quantitative method is chosen. Quantitative research is to measure numerically and analyze using statistical technique which is to test the relationship between variables [9]. Quantitative research needs to be well-structured the qualitative design. In general, quantitative research is related with survey research strategy and conducted through questionnaires. The requirement of quantitative questionnaires was public social respond and could provide enough data for research to make a generalization regarding the finding [10].

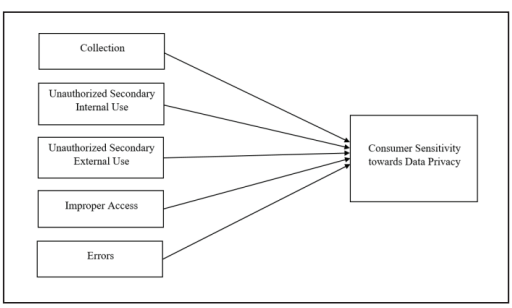

Fig. 1 Research Framework

In this research, respondents are targeted those who live in Melaka area, central state of Malaysia. According to Department of Statistics Malaysia, the population of Melaka estimated about 931,210 in 2016 with an average annual population growth of 2%. There were 400 respondents are selected randomly as the source of investigation to respond to the questionnaires [11], however 384 respondents send the feedback towards the survey conducted. A simple random sampling method was used in conducting the study. The type of questionnaires was close ended questions and was distributed to targeted sample.

Reliability test was conducted in this research by using Cronbach's alpha coefficient. The test was opted as it is one of the commonly used to test inter-item consistency reliability. A total of six constructs undergo the test and resulted an overall Cronbach's Alpha value of 0.94 which is more than acceptable.

IV. RESULT

A. Demographic Analysis

Table I showed the detail about respondent's demographic background related to gender, race, age, education level, employment status and income.

Based on 384 respondents, approximately more than 45% of the respondents are male, meanwhile about 52% of the respondents are female. The survey indicated that Malay encompasses 63%. While 25% (97 respondents) of the respondents are Chinese and 12% (44 respondents) of the respondents are Indian. Whiles based on the age of respondents, the age group of 18-25 years old represented about 46% (176 respondents) from the survey 30% (114 respondents) of them under age group pf 26-33 years old, 34-42 years old and 42-50 years old are 14% (56 respondents) and 7% (27 respondents respectively). The rest 3% (11 respondents) of the respondents are above 50 years old. The highest rate of respondents was age between 18-25 years old. This is because the teenager is easier to expose with the data privacy compare to others age group.

TABLE I. DEMOGRAPHIC ANALYSIS

Factors	Item	Number of Respondents	Percentage
Gender	Male	186	48.4
	Female	198	51.6
Race	Malay	243	63.3
	Chinese	97	25.3
	Indian	44	11.5
Age	≤ 25 years old	176	45.8
	26-35 years old	114	29.7
	36-45 years old	56	14.6
	≥ 46 years old	38	9.9
Education level	SPM and below	82	21.4
	STPM/Matriculation /Diploma	144	37.5
	Degree	128	33.3
	Master and above	30	7.8
Income	≤ RM1,000	123	32.0
	RM1,001 – RM3,000	191	49.7
	RM3,001 – RM5,000	54	14.1
	≥ RM5,001	16	4.2

From the Table I the range of income shows clearly of all the respondents income categorized into four categories. The highest came from those with range of income from RM1001-RM3000, which contributed to 50% of respond (191 out of 384 respondents). While the rest are between RM3001 to RM5000, above RM5001, and below RM1001. With 14% (54 respondents), 4% (16 respondents) and 32% (123 respondents) respectively.

B. Inferential Analysis

Regression analysis is used to predict the value of independent variables for given the value of one or multiple dependent variables by using a regression equation (Saunders et al, 2012). The multiple regression analysis was carried out between the identified dependent variables of Collection, Unauthorized secondary internal use (B), Unauthorized secondary external use (C), Improper Access (D), and Errors (E) and dependent variables (Consumer sensitivity towards data privacy). Table III showed the result of multiple regression analysis of measuring the effect of independent variables toward dependent variables.

TABLE II. DESCRIPTIVE ANALYSIS

Variables	Mean	Std Dev.
Collection	4.0703	0.4381
Unauthorized Secondary Internal Use	4.1380	0.3445
Unauthorized Secondary External Use	4.2604	0.4610
Improper Access	4.3047	0.5359
Errors	4.3255	0.3307

TABLE III. MODEL SUMMARY

Model	R	R Square	Adjusted R square	Std. Error of the Estimate
	0.760[a]	0.577	0.571	0.66492

The value of R square indicated that 0.577 that consists 57.7% of variance affected towards consumer sensitivity towards data privacy can be determined by the variable of Collection (A), Unauthorized secondary internal use (B), Unauthorized secondary external use (C), Improper Access (D), and Errors (E) (Table III). The remaining 42.3% was explained by other factors which were not taken into this research.

Furthermore, the ANOVA test is used to find out the overall probability of relationship between dependent variables and all independent variables. Based on the ANOVA test, the result F-statistic was 103.105 with the significant 0.000. By using p-value of 0.05 to the significant relationship, the null hypothesis is rejected due to values is less than 0.05 between consumer sensitivity towards Collection, Unauthorized secondary internal use, Unauthorized secondary external use, Improper Access, and Errors.

TABLE IV. MODEL SUMMARY

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	-0.041	.196		-.207	.836
A	.292	.057	.270	5.098	.000
B	.431	.052	.401	8.269	.000
C	.186	.068	.157	2.724	.007
D	.054	.070	.044	.771	.441
E	-.013	.068	-.011	-.193	.847

Three variables identified are Collection, Unauthorized secondary internal use and Unauthorized secondary external use significantly contribute to consumer sensitivity towards data privacy. Except for improper access and errors, both are not significant when considered together in predicting the consumer sensitivity.

V. CONCLUSION

In conclusion, this study provides some insight view on how consumer responded on their concern when private information is at risk of being exposed. Of all five determinant factors, three factors have indicated strong relationship with consumer sensitivity towards data privacy such as collection, Unauthorized secondary internal use and Unauthorized secondary external use. Collection refers to the concern that large amounts of personal data are collected and stored because nowadays people are always related to big data and internet of things in fourth industrial revolution.

As the consumer, they don't want their data privacy such as personal information is collected by any other third parties or company to use their personal information in wrongful ways. When personal information is collected by more sophisticated machines, the availability of data is used for lawful activities but might be abused as well. Sometimes, company may request for additional information that might not seem necessary. This surely elevate user's concern on information are being overly collected. Therefore, to reduce their concern companies should proactively inform consumer that their data will not be collected in anyways. Furthermore, consumer data is now being protected under Personal Data Protection Act 2010 whereby individuals or organizations dealing with processing personal data must comply with the rule.

Even though consumers are aware of unintentional unauthorized secondary use of personal information and feel unsafe, they tend to give it away when the benefits outweigh the perceived risk. In order to reduce the concern about data privacy, consumer must ensure data security practices of the third-party organization are adequate. Consumer has to educate themselves as well of not revealing sensitive information to companies. Perhaps more should be done on penalizing companies which put their customer's information at risk.

ACKNOWLEDGMENT

Authors wish to acknowledge Universiti Teknikal Malaysia Melaka and SuITE, Center of Tecnopreneurship Development (C-TeD) for the support.

REFERENCES

[1] Z., Gao and S., O'Sullivan-Gavin, "The development of consumer privacy protection policy in China: a historical review," *Journal of Historical Research in Marketing, 7*(2), pp.232-255. 2015

[2] K.W. Wu, S.Y. Huang, D.C. Yen and I. Popova, "The effect of online privacy policy on consumer privacy concern and trust. *Computers in human behavior*, 28(3), pp.889-897. 2012

[3] I. Arpaci, K. Kilicer, and S. Bardakci, "Effects of security and privacy concerns on educational use of cloud services". *Computers in Human Behavior, 45*, pp.93-98. 2015

[4] J.H. Tang and Y.J. Lin, "Websites, data types and information privacy concerns: A contingency model". *Telematics and Informatics, 34*(7), pp.1274-1284. 2017

[5] N.K. Malhotra, S.S. Kim and J. Agarwal, "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model". *Information systems research, 15*(4), pp.336-355. 2004

[6] Phelps, Joseph, Glen Nowak, and Elizabeth Ferrell. "Privacy concerns and consumer willingness to provide personal information." *Journal of Public Policy & Marketing 19*(1): 27-41.2000

[7] H. J. Smith, S.J. Milberg and S.J. Burke, "Information privacy: measuring individuals' concerns about organizational practices". *MIS quarterly*, pp.167-196.1996

[8] N. Fortes and P. Rita, "Privacy concerns and online purchasing behaviour: Towards an integrated model". *European Research on Management and Business Economics, 22*(3), pp.167-176. 2016

[9] N.K Saunders, Mark and Philip Lewis. "Doing research in business & management: An essential guide to planning your project". Pearson, 2012.

[10] N.K. Malhotra and David Birks. "Marketing Research: an applied approach: 3rd European Edition". Pearson education, 2007.

[11] D.W. Morgan and R. V. Krejcie. "Determining sample size for research activities." *Educational and psychological measurement* 30.3: 607-10. 1970