Computers in Human Behavior 66 (2017) 75–87

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Priming and warnings are not effective to prevent social engineering attacks

M. Junger [a, *], L. Montoya [b], F.-J. Overink [a, 1]

[a] University of Twente, Faculty of Behavioral, Management and Social Sciences (BMS), The Netherlands
[b] University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science, The Netherlands

ARTICLE INFO

Article history:
Received 6 June 2016
Received in revised form
26 August 2016
Accepted 7 September 2016

Keywords:
Priming
Warning
prevention
Social engineering
Phishing
Disclosure of personal information

ABSTRACT

Humans tend to trust each other and to easily disclose personal information. This makes them vulnerable to social engineering attacks. The present study investigated the effectiveness of two interventions that aim to protect users against social engineering attacks, namely priming through cues to raise awareness about the dangers of social engineering cyber-attacks and warnings against the disclosure of personal information. A sample of visitors of the shopping district of a medium-sized town in the Netherlands was studied. Disclosure was measured by asking subjects for their email address, 9 digits from their 18 digit bank account number, and for those who previously shopped online, what they had purchased and in which web shop. Relatively high disclosure rates were found: 79.1% of the subjects filled in their email address, and 43.5% provided bank account information. Among the online shoppers, 89.8% of the subjects filled in the type of product(s) they purchased and 91.4% filled in the name of the online shop where they did these purchases. Multivariate analysis showed that neither priming questions, nor a warning influenced the degree of disclosure. Indications of an adverse effect of the warning were found. The implications of these findings are discussed.

1. Introduction

The present study investigates whether users can be helped to protect their personal information against direct requests. It tests the effectiveness of two interventions that aim to protect users against social engineering attacks, namely priming through cues to raise awareness about the dangers of online activities and warnings against the disclose of personal information.

Social engineering has been defined as '*The science of using social interaction as a means to persuade an individual or an organization to comply with a specific request from an attacker where either the social interaction, the persuasion or the request involves a computer-related entity*' (Mouton, Leenen, Malan, & Venter, 2014). The success of a social engineering attack often depends on a target either being willing or tricked into disclosing personal information. Many cyber-attacks begin with users who unknowingly or mistakenly disclose personal information to attackers. For instance, attackers send users phishing emails containing a link to a webpage that requests the user to fill in personal information (Hong, 2012; Purkait, 2012). Today, it is estimated that there are thousands of unique phishing emails send to users on a daily basis. For instance, in March 2016 229,265 unique phishing e-mail reports (campaigns) were received by Anti-Phishing Working Group (APWG) from consumers and 123,555 unique phishing websites were detected (APWG., 2016). Most targeted industries were the retail industry (42.71% of all mails) and the financial industry (18.67% of all mails), meaning that attackers impersonated for instance a retail store or a bank.

Research has shown that the online and offline worlds are connected (Mesch, 2012). Offline trust, such as trust in social institutions and trust in individuals, is associated with trust online (Mesch, 2012). This is true for cybersecurity as well: from a security perspective physical and digital security are interconnected (Dimkov, 2012). Montoya, Junger, & Hartel, (2013) argued that crime can be situated on a continuum from only traditional or physical crime to completely digital crime. Many types of crime today have aspects of both. For instance, in a random sample of

* Corresponding author. University of Twente, Faculty of Behavioral, Management and Social sciences (BMS), Po Box 217, 7500 Ae Enschede, The Netherlands.
E-mail address: M.Junger@UTwente.nl (M. Junger).
[1] F.- J. Overink was a student of the University of Twente at the time that this study.

http://dx.doi.org/10.1016/j.chb.2016.09.012
0747-5632/

crimes reported to the police, 41% of all frauds and 16% of the threats have in part a digital modus operandi (MO). To commit burglaries, offenders hardly ever use ICT. But in 2.9% of the residential burglaries, however, bank cards were stolen which were later used to steal money from a bank account (Montoya et al., 2013). It is therefore important to understand how users react offline to understand the online threats. This is also illustrated with evidence on identity theft. Information necessary to execute a crime such as phishing or other forms of identity theft often come from the victims themselves. For instance, in a random survey of the Australian population on identity theft, victims reported that the information originated from email (18.3%), from information placed on a website, such as an online shopping website (15.7%), from information placed on social media (e.g. Facebook, Linked-in) (6.9%) and/or from text messages (SMS) (6.4%) (R. G. Smith & Hutchings, 2014). Information also originated from direct contact with the victim, namely from a face-to-face meeting (e.g. a job interview or a door knock appeal) (7.5%) or a telephone conversation (10.5) (R. G. Smith & Hutchings, 2014; pp., table 18). Although the percentages cannot be summed up (because attackers may have used several methods) these figures show that in about half of the incidents of identity theft, information used was provided voluntarily by victims, and in a sizable proportion of these cases, information was provided in a direct, not online contact. Consequently, user's perceptions about what constitutes sensitive personal identifiable information (PII), and their reactions to requests for sensitive PII matters for computer security. A better understanding and quantification of privacy and security perceptions is needed (Cranor, 2016).

Phishing attacks become increasingly sophisticated. For instance, spear-phishing mails or 'targeted attacks', are an increasingly popular (Hong, 2012; Wueest, 2014). In targeted phishing attacks, attackers attempt to better mimic genuine emails by using personal details from customers. Genuine emails from online shops usually mention the name of the customer and what was purchased. They often refer to the bank account number, and mention only the last three digits of the account (in the Netherlands), for safety reasons. The more an attacker knows from his potential victim, the better he can mimic genuine emails. Examples of phishing emails can be found on http://www.social-engineer.org/wiki/archives/Phishing/Phishing-eBay.html.
Although getting this additional information takes time, an advantage of targeted attacks is that they are relatively successful (Jagatic, Johnson, Jakobsson, & Menczer, 2007; Rocha Flores, Holm, Nohlberg, & Ekstedt, 2015). An experiment using social network information, showed that people are 4.5 times more likely to fall for a phishing message sent from an existing contact compared to standard phishing attacks. Out of 512 students at the Corps of Cadets at West Point receiving a spear phishing email mentioning a problem with their Grade Report, 80% clicked on the link in the email (Ferguson, 2005).

Many security-related organizations have lamented on the vulnerability of users to social engineering attacks and their tendency to disclose information (Adams & Sasse, 1999; Kirlappos & Sasse, 2012). To overcome this situation, customers are informed about the occurrence of cybercrime and phishing emails and they receive tips and instructions on how to protect themselves against disclosing PII. For example, most banks and online shops have webpages devoted to security. Often, they use leaflets and warnings to inform users or customers about what they should and should not do to protect themselves against possible attacks (see for instance: http://pages.ebay.com/help/account/recognizing-spoof.html).
However, it is not well known how effective these warnings messages are. The present study tests warning and priming for cybercrime to study their effectiveness in a sample of shoppers selected in a shopping area.

1.1. Disclosure of personal information

Trust can be defined as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or the behavior of another" (Rousseau, Sitkin, Burt, & Camerer, 1998). Trust determines the way in which an individual approaches other people (Glaeser et al., 2000; Fehr et al., 2003; Kosfeld et al., 2005).

Humans tend to conform and are relatively trustworthy by nature: trust has evolutionary survival value e.g. children need to trust others in order to be able to learn (Dawkins, 1993; Morgan & Laland, 2012; Ostrom, 1998; Penner, Dovidio, Piliavin, & Schroeder, 2005). (Fetchenhauer & Dunning, 2009; Glanville & Paxton, 2007; P. L.; Harris et al., 2012; P. L.; Harris & Corriveau, 2011). Most researchers believe that adults start with the presumption of truth (Burgoon & Levine, 2010; Mills, 2013). In general, having trust in others has positive outcomes for individuals (Dohmen, Falk, Huffman, & Sunde, 2012; Fetchenhauer & Dunning, 2009; Frattaroli, 2006; Glanville & Paxton, 2007; Ostrom, 1998).

Disclosure or self-disclosure can be defined as the process of communicating information about oneself verbally to another person (Cozby, 1973). Besides having relatively high trust, humans seem to have low thresholds for disclosing personal information and do it relatively often. Most of the studies on disclosure have been done in the field of psychology and mental health. These studies showed that self-disclosure has positive outcomes (Dindia, Allen, Preiss, Gayle, & Burrell, 2002; Cozby, 1973; Omarzu, 2000; Sprecher, Treger, & Wondra, 2013; Worthy, Gary, & Kahn, 1969).

Although in general trusting others has positive outcomes (Dindia et al., 2002), personal information can be abused relatively easily (Acquisti, Brandimarte, & Loewenstein, 2015; Gross & Acquisti, 2005; Hann, Hui, Lee, & Png, 2002). Research has been done to investigate the degree to which users are prepared to disclose personal information online and the situations in which disclosure increases or decreases.

John, Acquisti, and Loewenstein (2011) presented four experiments in which disclosure was measured by investigating whether subjects answered questions on deviant behavior such as 'Having sex with the current husband, wife, or partner of a friend' or 'Making a false insurance claim'. Three experiments showed that users disclosed more personal information on unprofessional looking websites, which are arguably more likely to misuse it than on professional looking websites which were less likely to misuse it (John et al., 2011, p. 868). In other words, 'individuals are prone to disclose in contexts that downplay privacy concerns—ironically, even when such contexts are likely higher in both objective and perceived disclosure danger' (John et al., 2011, p. 868). Interestingly, in an experiment in which users were cued to think about privacy, these contextual differences i.e. type of website-disappeared and all users had similar rates of disclosure. John et al. (2011, p. 868) concluded that their results 'stand in contrast to the considerable body of privacy research that is premised on the assumption of rational choice', which states that people make trade-offs between privacy and other concerns, implying that disclosure is the result of this rational weighing of costs and benefits, in which objective costs – such as an unprofessional looking website - should prevent or at least decrease disclosure.

Joinson, Reips, Buchanan, and Schofield (2010) combined a survey with an experiment. Their findings differ from John et al. (2011), and are more in line with rationality. They report that self-disclosure was reduced when the context involved a weak privacy policy and low trust. In all other combinations of trust (high versus low) and privacy (high versus low), self-disclose was higher.