Title	Social Networking and Personal Data Security: A Study of Attitudes and Public Awareness in Ireland
Author(s)	Lang, Michael; Devitt, Jonathan; Kelly, Seán; Kinneen, Andrew; O'Malley, John; Prunty, Darren
Publication Date	2009
Publication Information	Lang, M., Devitt, J., Kelly, S., Kinneen, A., O Malley, J. & Prunty, D. (2009) Social Networking and Personal Data Security: A Study of Attitudes and Public Awareness in Ireland. In Wan, C. et al. (eds) Proceedings of International Conference on Management of e-Commerce and e-Government (ICMeCG), Nanchang, China, September 16-19. IEEE Computer Society, pp. 486-489.
Publisher	IEEE
Item record	http://hdl.handle.net/10379/403

Social Networking and Personal Data Security: A Study of Attitudes and Public Awareness in Ireland

Michael Lang, Jonathan Devitt, Seán Kelly, Andrew Kinneen, John O'Malley, Darren Prunty
J.E. Cairnes School of Business & Economics
NUI Galway, Ireland
Michael.Lang@nuigalway.ie

***Abstract*— This paper reports the findings of a study of attitudes towards data security issues and awareness of the potential risks of social networking sites, most notably the possibility of identity fraud. The population for this study is people in the 18-24 age group living in Ireland, for which the students of one of the country's main universities were used as a proxy sample. The main findings is that many people have a very casual attitude towards data backup and password protection, a considerable number are not adequately aware of the threats posed by viruses transmitted via portable devices and other such means, and a lot of young people are openly publishing personal information of a sensitive nature that could potentially be maliciously exploited.**

Keywords-component; social networking; data security; public awareness

I. INTRODUCTION

Social networking is a new form of interacting on-line where participants in a virtual network can share information and communicate with one another [1]. However, due to the phenomenal growth in popularity of sites such as Bebo and Facebook, potential security issues have arisen because personal information which could possibly be misused by computer criminals for such purposes as identity theft or impersonation has suddenly become much more readily available. An added possibility is that personal information displayed on publicly-available Web pages might be used by savvy password crackers because many people use weak passwords based on easy-to-remember personal facts, such as one's date of birth, which are relatively easy to uncover.

A recent study of computer crime in Ireland revealed that, of 42 organisations surveyed, 90% experienced losses due to viruses and malicious software, 88% suffered the misuse of systems, while 63% and 56% had been subjected to asset theft and phishing respectively [2]. Passive attitudes towards the pervasiveness of information technology into almost every aspect of peoples' daily lives in Western society has led the UK Information Commissioner to comment on the dangers of sleepwalking into a surveillance society [3]. As an example of how ignorance of the perils of careless disclosure of personal information can result in financial loss, the recent case of an opinionated British journalist is notable; so convinced was he that a public scandal, which involved the loss by the government of CDs containing personal information about millions of citizens, was a "storm in a keycup" that he printed his own bank account and electoral register details in a national newspaper, asserting he had "nothing to fear". The outcome was that an unidentifiable person promptly managed to set up an untraceable direct debit from his bank account [4]. The possibility of identity fraud using information recovered from social networking sites such as Facebook has become a particularly serious problem for financial institutions [5,6], some of whom have seen their names emblazon the headlines for recent data protection breaches e.g. [7].

Against this recent background of well publicized incidents, we determined to conduct some preliminary research into the current state of public awareness of data security and social networking risks amongst the young population (18-24 years) of Ireland. The objectives of our study were two-fold:

- Firstly, we set out to explore the level of general knowledge about data security issues such as passwords, data backups, viruses and phishing. A lack of awareness of risks and threats means that a person is more likely to become the unwitting victim of a security breach.
- Secondly, with particular regard to the security of personal data published on social networking Web sites, we aimed to investigate the vulnerability of typical Bebo and Facebook personal profiles. Our focus was on identifying information which could potentially be misused by somebody acting with malicious intent (e.g. date of birth, address, email, sensitive photographs).

II. RESEARCH METHOD

We chose to use a dual research approach, collecting data concurrently using two different strategies: primary data was elicited using a Web-based survey, and secondary data was collated from an analysis of personal Web pages on social networking sites (Facebook and Bebo).

A. Primary data: Web-based survey

The Web-based questionnaire was implemented using SurveyMonkey, an on-line survey hosting service. The population comprised students and staff of NUI Galway, the largest higher education institution in the west of Ireland. We acknowledge that this sample is not representative of the general public at large, but it is likely to be a reasonable reflection of attitudes and awareness of data security issues amongst the youth of Ireland, most of whom are computer

literate, as evidenced by national census data pertaining to domestic PC usage and Internet penetration [8]. We also recognize that the younger generation (18-24 years) have been found in previous studies of data protection public awareness in Ireland to be slightly less concerned about personal data than people in the 25-49 years age group [9].

The questionnaire was designed in accordance with the established principles of survey research [10,11]. Prior to distribution, it was initially pilot tested with a small number of selected users, after which some minor revisions were implemented. Email invitations were then sent to students at NUI Galway, but because of an institutional policy of restricting broadcast mailings, the entire student body could not be reached. Instead, we mainly concentrated on large undergraduate groups which we could access via class distribution lists that were made available by individual lecturers. To further expand the research sample, we also used the "snowballing" technique as well as some "convenience sampling" of personal contacts. After a number of follow-up rounds, a total of 351 responses was received. Of these, the majority (54%) came from first year undergraduates, but responses were received from students across all four years of the undergraduate cycle, as well as from postgraduates, visiting students, and staff. The respondents were also distributed in terms of their disciplinary affiliation, with Business/Law (61%), Science (25%), Arts/Humanities (4%), Medicine/Health Sciences (5%) and Engineering/IT (4%) being represented.

B. *Secondary data: meta-analysis of social networking Web sites*

To gather secondary data from personal profiles on social networking Web sites, we used a free Web-based email account to invent a fictitious character who passed himself off as a disk jockey (DJ). We then created a group on Bebo and Facebook called "Irish Nightlife", managed by this fictitious DJ, on the supposition that a group of this name was likely to be attractive and interesting to a substantial cohort of the general public, especially the younger generation. A personal profile with minimal information was created on Facebook and Bebo for this DJ. We then randomly selected 120 subjects (60 each from Facebook and Bebo), ensuring to get an equal split of males/females and public/private profiles. These subjects were sent invitations by the fictitious DJ to join his "friend list", the objective being to ascertain the extent to which social network users are willing to permit total strangers to see their personal data. For those users who accepted this unsolicited invitation from an on-line stranger, we analyzed the contents of their personal profile pages, categorizing various data items as "sensitive", "personal", etc. This experiment was run over the course of two months, after which time the fictitious accounts were closed down. Although this research deliberately employed a misleading practice, we did so purely for the purposes of simulation, and at all times we were conscious of our legal and ethical responsibilities neither to lure anybody into unwillingly disclosing personal information nor to compromise the integrity of personal data.

We classified the type of profiles as "public" (i.e. open to everyone to see) or "private" (i.e. visible only to admitted profiles). We expected public profiles would be more willing to accept the invitation from the fictitious DJ, as well as having more sensitive data on their profile, because we presumed that persons with a public profile would be much less conscious or wary of the threats posed by fraudsters.

III. Discussion of Findings

A. *Analysis of Web survey responses*

1) *Profile of respondents*

About half of respondents said that they use the Internet for 1-2 hours each day, with a further third indicating a daily usage level of 3-4 hours. As regards the nature of Internet use, 87% use email regularly, 72% regularly browse the Web, and 66% regularly use social networking sites (e.g. Bebo, Facebook). On the other hand, much fewer respondents indicated that they are regular users of on-line banking (21%), on-line shopping (10%) or Skype/VoIP (6%), with 50%, 34% and 68% respectively responding that they *never* use these services. These patterns therefore present an indicative profile of the respondent sample. The relatively low usage of on-line banking and on-line shopping is significant, because presumably persons who do not have an on-line bank account or a credit card are less fearful of suffering financial loss through on-line fraud.

2) *Password security*

Over 83% of respondents indicated that they use the same password for multiple accounts. The results also revealed that 74% of people never change their passwords. Astonishingly, 23 of the 40 respondents who indicated that they consider it "very likely" that their password could be stolen also indicated that they use the same password for multiple accounts, and that they never change this password. This is potentially a serious security risk because in the case of a user with numerous on-line accounts which are all protected by the same password (which may be a weak one), a breach of any of those accounts could place an entire network at risk. For example, if an employee registered to gain access to some "free" e-book or on-line forum, and in so doing used his normal email address as a username and his normal email password as a password, it would be very easy for an intruder to hack into that email account and gain access to sensitive or confidential messages.

3) *Attitudes towards the risk of data loss*

Another interesting result is that half of the respondents *never* backup their data files, which is all the more alarming given that 85% think it is likely or very likely they will lose files, the same number expect that they may lose an external memory device (e.g. a USB stick), and 68% are of the opinion that their computer is likely to crash. A possible explanation for this surprising finding is that students may expect that the university's campus network is regularly backed up, thereby placing the onus of responsibility on computer services to recover data in the event of a loss. What this tells us is that it is important for all organizations to have explicit policies and procedures on data backup, as most

people seem to have a relaxed attitude in this regard and "expect the system to look after it", which in fact it may not.

4) Awareness of viruses and similar threats

Another finding that stood out is that half of respondents did not know that Bluetooth devices, CD's and DVD's can carry viruses, while over a third were unaware that USB flash drives can also be carriers. As these are very common ways of sharing data, it is important that all organizations should have appropriate mechanisms in place to safeguard against data loss or virus propagation caused by rogue devices. 39%, 44% and 56% of respondents indicated that that they were completely unaware of "Trojans", "worms", and "malware" respectively, while in response to another question that asked if "you have ever experienced a virus, worm or other form of intruder on your computer?", 22% responded that they were unsure. What these findings reveal is that, even amongst a reasonably computer-literate population, there is quite a high level of non-awareness of common data security threats.

5) Security of personal data on social networking sites

With regard to social networking, we asked a few questions to investigate if people use such sites (e.g. Facebook, Bebo), and if so to see if they use the privacy settings, accept invitations from people they do not know, or post sensitive information about themselves. The results revealed that 28% of people don't use the privacy settings provided, and 10% are unsure, so we presume they don't. 234 respondents (87%) indicated that they don't post sensitive information about themselves, although of course the issue of what a person may consider "sensitive" is subjective depending on one's perspective. This high percentage should therefore be interpreted as the proportion of people who *believe* that the information displayed on their personal profiles is not sensitive.

A similar number of respondents (88%) said that they would not accept an invitation from a person they do not know. Adding a person to your "friends list" in a social networking site means that they can see your profile and personal information, so the finding that people are generally cautious in expanding their on-line networks suggests a growing public awareness of the potential dangers and nuisances attached to these new communication technologies.

Interestingly, 35 (12%) of respondents who said that they regarded the chances of identify theft occurring as "likely" or "very likely" also indicated that they do not use the privacy settings in their social networking profiles. This may suggest a general lack of awareness of how identity fraud is perpetrated. Of course, so numerous are the users on social networking sites like Bebo or Facebook that the probability of any given user being the victim of *attempted* identity fraud, let alone *actual* identity fraud (which would mean circumventing various authentication checks), is rather low. Nevertheless, would-be con-men are constantly "phishing" for vulnerable individuals who innocently divulge personal data, being unaware of how such data may be used to commit unlawful acts.

B. Analysis of social network experimental data

120 personal social networking profiles were chosen at random, 60 from Bebo and 60 from Facebook, with an equal division of males/females and public/private profiles (see Table I). Interestingly, for both males and females, 40% of those who had private profiles accepted the request from a stranger to join his network of friends, which meant that whatever information was previously private (and therefore invisible to the general public) could now be seen by the enigmatic impostor that we artificially created for the purposes of this experiment.

TABLE I. OUTLINE OF SOCIAL NETWORKING SAMPLE

Profile Type	**Table Column Head**		
	Male	***Female***	***Total***
No. of private profiles on Bebo	10	20	30
No. of private profiles on Facebook	15	15	30
No. of public profiles on Bebo	20	10	30
No. of public profiles on Facebook	15	15	30
Totals	60	60	120

Contrary to our initial expectations that we would receive more acceptances from the public profiles than from the private profiles, the proportions of those with public profiles who accepted our invitation were considerably lower (males 29%, females 12%, overall 22%). This may be because public profiles have enough friends already (the median value for the number of friends in our Bebo public profiles sample was 335), or perhaps it is because public profiles are targeted with more random friend requests and are therefore more likely to disregard them.

Overall, the percentage of males that accepted the stranger's invitation (33%) is slightly higher than the percentage of females (28%), but caution must be exercised not to read too much into these differences because they are based on small sample sub-sets (N=60) and therefore are not statistically significant. However, what is of significance is that the number of Facebook public profiles that accepted the invitation request (10%) was substantially less than the number of Bebo public profiles who did so (33%). This may possibly be explained by the fact that Bebo and Facebook have different orientations and different user bases, the former generally being recognized as the social networking site of choice for teenagers, with the latter typically having a slightly older demographic. Hence, the novelty factor of "interesting" strangers may be greater on Bebo.

What is also interesting is the disparity between our survey findings, where just 9% of respondents (N=296) said they would accept an invitation from a stranger, and our experiment findings, where 31% of the participants (N=120) actually did accept an invitation from a stranger.

For all the profiles to which we could gain access, we analyzed the content of the users' Web pages. Many of the pages contained photographs which could possibly, depending on who was looking at them, be regarded as "sensitive"; for example, students may think little of sharing photographs of their exploits at a party but if such photographs were to fall into the wrong hands, they could

lead to a person's reputation being damaged or, in extreme cases, blackmail. To assure rigor and reliability in our analysis, two members of the research team separately coded the photographs and resolved any inter-rater inconsistencies regarding the interpretation of "sensitive" photographs (for the most part, they related to consumption of intoxicating substances). Out of a total of 58 profiles for which we could access photographs, exactly half contained images that we regarded as sensitive. No differences were found between males and females here, in so far as they are both equally likely to have such images on their personal profiles. However, amongst our Bebo sample, we found that males were less cautious about publishing photographs in the publicly accessible space, whereas females were more reluctant to do so, preferring to keep their sensitive photographs within the private zone.

Similarly, we also looked for "sensitive" information, which we defined as data that might be of some potential use to a fraudster (based on our general knowledge of common Internet scams). This included not just potentially compromising accounts of one's life (as documented in on-line diaries), but also personal details such as date of birth, educational history, description of physical attributes and various other data that could, for example, be used by password crackers (e.g. to try to guess the answers to "secret questions"), social engineers (e.g. who might seek to infiltrate a trusted circle of friends for the purposes of personal gain), or impersonators.

Although a lot of sensitive data was found in peoples' blogs, it was the Bebo applications and comment boxes that held the most sensitive information. For example, we found a birthday reminder application and a horoscope application that would reveal a person's date of birth. In order to use this application, a Bebo user must permit it to access their personal information. Much of what we classified as sensitive information was found in one particular application called the "Best Profile Survey" which invited users to complete a set of questions about hobbies, hair colour, date and place of birth, type of vehicle driven, etc. Of 57 profiles for which we could access information, about a quarter of them (15) were found to contain sensitive content; of these, they were almost equally divided between male and female users.

We noticed that Facebook users are more likely to display sensitive information on their profile, but this may simply be because the Facebook sign-up process prompts the user to enter such information. For example, of the 30 public Facebook profiles that we visited, 26 of them displayed the user's date of birth, 25 showed educational details, 21 disclosed a home address, and 8 included an email address. Even the users with private profiles openly revealed personal information: of the 11 private Facebook accounts which accepted our invitation, 10 displayed a date of birth, 6 provided details of educational history, and 4 featured the user's home address.

IV. Conclusions, Limitations, and further Study

This study had a number of limitations which we would seek to redress if replicating it again in the future. Firstly, we would enlarge our social networking sample because 120 profiles is not a sufficient number to be able to conduct meaningful statistical correlations and tests for differences. Most of the participants in our questionnaire were students from NUI Galway, and most of the social networking profiles analyzed were also those of Irish students. In future, we would like to distribute the questionnaire to a wider international population, and also to expand the reach of the social networking experiment.

The implications of this study are alarming for potential employers. Many of the students who responded to the questionnaire would have, as part of their curriculum, covered the basics of personal data security. As such, they ought to be aware of potential risks and preventative measures. Nevertheless, there appears to be a culture of "it will never happen to us", even though in fact a substantial number of the respondents had experienced the adverse impacts of a computer virus infection or system crash. As evidenced by the study mentioned in the introduction of this paper [2], many Irish organizations have suffered serious losses of productivity because of data security breaches. In recent years, there has been considerable talk of the potential benefits of using social networking and other Web 2.0 collaborative technologies in business. However, there has been relatively little focus on the potential security risks of using these technologies. One of the key lessons of this paper is that greater security awareness and training is necessary, and that organizations must have enforceable data protection regulations in place, always bearing in mind that human folly is the weakest link in the security chain.

References

[1] I. O'Murchu, J. G. Breslin and Stefan Decker, Online Social and Business Networking Communities, Working Paper, Digital Enterprise Research Institute, NUI Galway, Ireland, 2004.

[2] ISSA/UCD, Irish Cybercrime Survey 2006: The Impact of Cybercrime on Irish Organisations. School of Computer Science and Informatics, University College Dublin, Ireland, 2006.

[3] Information Commissioner (UK), A Report on the Surveillance Society, September 2006.

[4] N. Paris, "Jeremy Clarkson eats his words over ID theft", Daily Telegraph, January 8, 2008. http://www.telegraph.co.uk [downloaded November 18, 2008].

[5] H. Wallop, "Fears over Facebook identity fraud", The Telegraph, July 4, 2007. http://www.telegraph.co.uk [downloaded January 28, 2008].

[6] Anonymous, "Fear and loathing online: How the nightmare of identity theft is coming true", Irish Independent, June 2008. http:// http://www.independent.ie/business/technology/fear-and-loathing-online-1422286.html [downloaded June 28, 2008].

[7] G. Brennan and G. Cunningham, "Bank alert as details of 10,000 files stolen", Irish Independent, April 28, 2008. http://www.unison.ie [downloaded April 22, 2008].

[8] CSO, Report on Information Society and Telecommunications. Cork, Ireland: Central Statistics Office, 2006.

[9] Data Protecion Commissioner (Ireland), Report on Public Awareness of Data Protection and Privacy Issues, January 2006.

[10] D. A. Dillman, Mail and Internet Surveys: The Tailored Design Method, 2nd ed. New York: Wiley, 2000.

[11] A. Oppenheim, Questionnaire Design, Interviewing and Attitude Measurement, 2nd ed. London: Pinter, 1992.