Lastdrager *Crime Science* 2014, **3**:9
http://www.crimesciencejournal.com/content/3/1/9

Crime Science
a SpringerOpen Journal

SYSTEMATIC REVIEW Open Access

Achieving a consensual definition of phishing based on a systematic review of the literature

Elmer EH Lastdrager

Abstract

Background: Phishing is a widely known phenomenon, but currently lacks a commonly accepted definition. As a result, many studies about phishing use their own definition. The lack of a common definition prevents knowledge accumulation and makes analysing studies or aggregating data about phishing a difficult task.

Method: To develop a definition, we used existing definitions as input and combined them using crime science theories as the theoretical framework. A systematic review of the literature up to August 2013 was conducted, resulting in 2458 publications mentioning the word phishing. All journal articles, together with both highly cited and recent conference papers were selected, giving a total of 536 peer-reviewed publications (22%) to be manually reviewed. This resulted in 113 distinct definitions to be analysed.

Results: An analysis identified key concepts that were found in most definitions and formed the building blocks for a consensual definition. We propose a new definition that is based upon current ones, which defines phishing in a comprehensive way and - in our opinion - addresses all important elements of phishing: *'phishing is a scalable act of deception whereby impersonation is used to obtain information from a target'*.

Conclusions: A consensual definition allows future research to be aligned and it facilitates the interpretation and comparison of existing research. The findings suggest that the routine activity approach can be applied to the digital world. Finally, the 'scalability' concept of our definition provides a new theoretical notion to digital crime that is independent of the employed channel.

Keywords: Phishing; Definition; Cybercrime; Crime; Internet; Digital crime

Background

The term *phishing* is currently widely used with thousands of mentions in the scientific literature, lots of media coverage and widespread attention from organisations such as banks and law enforcement agencies. However, this prompts a question: what exactly is phishing? In some publications, the phenomenon of phishing is explicitly defined; in some, it is described by means of an example, while others assume that the reader already knows what phishing is. Many authors propose their own definition of phishing, leading to a large number of different definitions in the scientific literature.

With no scientific consensus, other sources could provide a standard definition. The first point of reference for finding the definition of a word would be a dictionary. Four definitions from prominent English dictionaries are shown in Table 1. Additionally, it lists the definition of the Anti-Phishing Working Group (APWG), a non-profit foundation that keeps track of phishing. The APWG definition is rather lengthy compared to the dictionary definitions. The five definitions vary in the level of detail and the scope of the phenomenon. For example, whereas the American Heritage definition includes phone calls, the others do not. In addition, the goal of phishing differs in the definitions, ranging from financial account details (Collins, APWG) to the more general personal information (Oxford, Merriam-Webster, American Heritage). There is greater consensus about the origin of the term phishing; it was first used around 1995–1996 (James 2005; Khonji et al. 2013; Press 2013; Purkait 2012) and is a variation on the word 'fishing', something hackers commonly did (James 2005; McFedries 2006; Press 2013; Purkai 2012).

Correspondence: e.e.h.lastdrager@utwente.nl
Services, Cybersecurity and Safety Group, University of Twente, Drienerlolaan 5, Enschede, OV, Netherlands

Table 1 Definitions of phishing from four dictionaries and the APWG

Source	Definition
Oxford University Press (2014), UK	The fraudulent practice of sending emails purporting to be from reputable companies in order to induce individuals to reveal personal information, such as passwords and credit card numbers, online.
Collins English Dictionary (2013), UK	The practice of using fraudulent e-mails and copies of legitimate websites to extract financial data from computer users for purposes of identity theft.
Merriam-Webster (2013), USA	A scam by which an e-mail user is duped into revealing personal or confidential information which the scammer can use illicitly.
American Heritage Dictionary (2013), USA	To request confidential information over the Internet or by telephone under false pretenses in order to fraudulently obtain credit card numbers, passwords, or other personal data.
Anti-Phishing Working Group (2013)	Phishing is a criminal mechanism employing both social engineering and technical subterfuge to steal consumers' personal identity data and financial account credentials. Social engineering schemes use spoofed e-mails purporting to be from legitimate businesses and agencies, designed to lead consumers to counterfeit websites that trick recipients into divulging financial data such as usernames and passwords. Technical subterfuge schemes plant crimeware onto PCs to steal credentials directly, often using systems to intercept consumers online account user names and passwords – and to corrupt local navigational infrastructures to misdirect consumers to counterfeit websites (or authentic websites through phisher-controlled proxies used to monitor and intercept consumers' keystrokes).

In common with fishing, phishing is about setting out 'hooks', hoping to get a 'bite'.

The lack of a standard definition of phishing has been observed previously (Abu-Nimeh et al. 2007; Al-Hamar et al. 2011; Khonji et al. 2013). This causes several problems for scientists, practitioners and consumers. For scientists, it is difficult to compare research on phishing in a meaningful way. Aggregating research consists of classification (in which attacks are considered phishing), and identification (measuring how often it occurs). Furthermore, countermeasures against phishing cannot be effectively evaluated without knowing the extent of the phenomenon. Additionally, having no standard definition is an indication of the immaturity of the field with researchers refining their own definitions over the years (e.g., (Kumaraguru et al. 2010, 2007), (Moore 2007; Moran and Moore 2010) and (Hong 2012; Xiang and Hong 2009; Xiang et al. 2011)). Institutions, such as banks or governments, face problems understanding one another if their definitions of phishing are different. For example, one bank may consider a fraudulent phone call to be phishing, whereas another bank will not, making a comparison of victimisation or countermeasures difficult. Consumers may also experience the downside of a lack of a standard definition. Persons who are less computer literate, for example, may become confused when several awareness campaigns describe phishing differently.

We aim to clarify the definition of the phishing phenomenon by analysing existing definitions, in contrast to most standard definitions, which are developed using expert panels. The resulting definition is based on consensus drawn from literature, and is sufficiently abstract to support future developments. To the best of our knowledge, no previous attempt has been made to synthesise a definition of phishing.

In order to interpret existing definitions of phishing in the right context, one needs a theoretical framework. An initial exploration revealed that phishing contains elements from criminal activities. Crime science theories are used for crime in the physical world, which raises the question of their applicability in the digital world. Previous research supports the idea of applying crime science theories to digital crime (Pratt et al. 2010; Reyns et al. 2011; Yar 2005) and there is limited evidence of its applicability to phishing (Hutchings and Hayes 2009). Therefore, crime science theories are used to achieve a better understanding of phishing and to provide us with concepts to analyse it. The focus of crime science is on the opportunity for a crime, rather than on the characteristics of the criminal. Three theories on crime opportunity form the foundation of crime science (Clarke 2009; Felson and Clarke 1998): the Rational Choice Perspective; Crime Pattern Theory; and the Routine Activity Approach. Each of these theories takes a distinctly different approach to crime (Clarke 2009). The rational choice perspective offers a view on offender's decision-making, assuming bounded rationality (Cornish and Clarke 2008). An offender is assumed to make a rational decision and commit a crime if the perceived benefit outweighs the perceived cost. Crime pattern theory (Brantingham and Brantingham 1993, 2008) focuses on the relation between crime and the physical environment, in particular the crime opportunities that emerge in the daily lives of the offender. According to crime pattern theory, crime is not randomly distributed in time and space. For example, a potential offender may come across opportunities for crime during his regular daily commute. Finally, the routine activity approach (Cohen and Felson 1979) states that a crime occurs when a likely offender and a suitable target converge in the absence of a capable guardian. Routine activity theory can be interpreted broadly (Pratt et al. 2010; Reyns et al. 2011) to include crime without direct contact. For example, in the case of cyber bullying an

online chat room can be the location where an offender and victim "meet". The focus on offender decision making within the rational choice perspective makes this theory less suited for reasoning about phishing, since the offender is mostly unknown. Similarly, applying crime pattern theory is difficult for phishing, since it often occurs on the Internet. The routine activity approach however, is applicable to phishing (Hutchings and Hayes 2009) with concepts such as offender and target, especially useful.

To elaborate upon the routine activity approach, crime scripts (Cornish 1994; Schank and Abelson 1975) can be used. Crime scripts describe the sequential steps that lead to an offence, much like a film script. Using crime scripts allows for interpretation of definitions of phishing in such a way that the act of phishing is decomposed into several steps. An example of such a step is "Victim receives an email". To fully understand each definition, we decompose each step into several key concepts. To structure the identification and classification of these concepts, we use the 3A model El (Helou et al. 2010). The 3A model is an activity-centric framework that provides three categories: Actors, Assets and Activities. In the context of phishing, actors are humans (e.g., the offenders) who conduct activities (e.g., send a message) to achieve their goal. The goal itself could be to obtain an asset (e.g., credentials). The routine activity approach together with the tools of crime scripts and the 3A model, are used to identify relevant concepts within each definition.

The goal of the literature search is to find scientific definitions of phishing. We formulated the following research question: *How is phishing defined in the research community?* Three steps are taken to generate a definition. Firstly, relevant literature is selected and definitions of phishing are extracted. Secondly, the concepts of phishing are extracted and scored according to their occurrence. Finally, concepts that are found in most definitions are selected and a standard consensual definition is developed from these concepts.

Method

Selection of literature

To obtain data on the existing definitions of phishing, a systematic study of the peer-reviewed scientific literature was performed, following the guidelines of Kitchenham (Kitchenham and Charters 2007). Three digital libraries were selected for the search: ACM digital library, IEEEXplore and Scopus. The fields relevant to phishing, such as computer science and various social sciences (i.e., psychology or criminology), are covered by these three databases. The literature search (see Figure 1) resulted in 2458 publications up to August 2013 that used the word 'phishing' in the title, abstract or keywords. We filtered the publications based on our exclusion criteria: studies had to be written in English to be included in our selection, so that we could run a syntactical analysis on them, and had to be peer-reviewed.

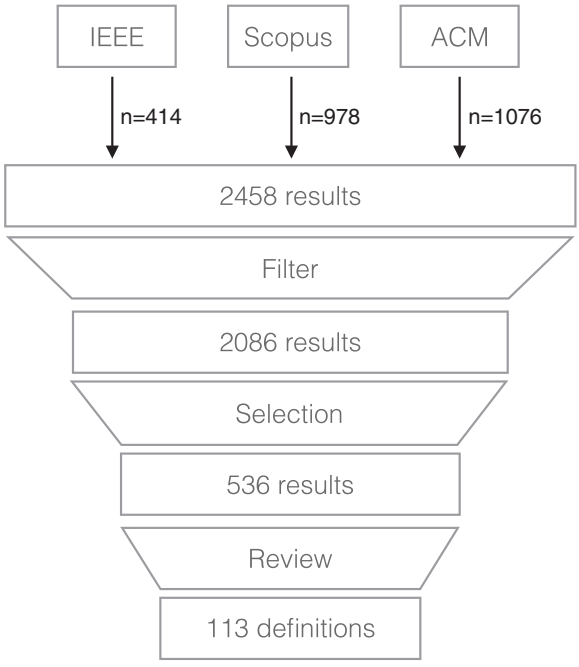

Figure 1 Search, selection and review of the results.

After filtering, the literature set was narrowed down to 312 journal articles and 1774 conference papers. Since it was not feasible to read all publications, we created a subset of the literature to be reviewed manually. Journals generally have less strict review deadlines than conferences, resulting in longer reviews and possibly higher quality. In addition, generally journals have higher limits on the number of pages, resulting in more in-depth articles. Therefore, we included all 312 journal articles in the review. Turning to the 1774 conference papers, we note that in the field of computer science, publishing in conference proceedings is generally favoured (Freyne et al. 2010), whereas journals are preferred in other fields. For the conference papers, we used the number of citations as an indication of quality and based our selection on this criterion. This resulted in the inclusion of 135 conference papers with more than 10 citations each. However, the selection based on citation count may exclude high quality conference publications that have recently been published and thereby have not yet received many citations. Therefore, we included all 69 recent conference papers from 2013 (from January to August) and the 20 newest from 2012.

All 536 eligible publications were manually searched for definitions of phishing by performing a case-insensitive search for the word 'phish', so that abbreviations within the paper would also be covered. If a definition was

present, it was extracted for further analysis. Studies were excluded if they: (1) did not include a definition, or at least a clear and concise description, of the word phishing; or (2) merely cited a definition of others. If an included paper cited the definition from another peer-reviewed publication (7 occurrences), the cited publication was included in our dataset. The approach involved considering not only explicit definitions but also descriptions of phishing in terms of concepts. Definitions had to be one or two sentences in length, but longer definitions were included if they were clear and to the point. However, publications giving only a specific example, such as an anecdote, were not included.

Since the search was performed by a single researcher, the extraction of definitions was re-evaluated by a second researcher by randomly selecting 100 publications from the dataset. The second researcher then manually reviewed each publication to identify a definition. The two sets of results were compared and the inter-rater reliability (Cohen's Kappa) was found to be $K = 0.70$ ($p < 0.001$) with a 95% confidence interval of (0.561, 0.839), indicating substantial agreement and supporting the feasibility of the method.

Careful analysis of the 118 extracted definitions resulted in the exclusion of five of them as non-cited duplicates. Among the duplicate definitions, we selected the definition that had been published the earliest and excluded the others. This reduced our dataset to 113 unique definitions, all of which can be found in the Additional file 1.

Identification of common words

We initially analysed the definitions in a purely syntactical way (i.e., without context) to obtain an overview of the most commonly used words. The analysis consisted of a simple frequency count of all words to establish which ones occur most often. Although a frequency count removes all contextual information from the individual words, it does give an indication of the relative importance of each word compared to all the others. In addition, words that appear throughout all definitions are probably important to phishing. All definitions were first processed by removing all punctuation, putting all words in singular form and merging different spellings. For example, 'creditcard' became 'creditcard', 'ID theft' became 'identity theft', and 'web page' became 'webpage'. Multiple occurrences of a single word were counted only once per definition to avoid biasing the frequency count. All adverbs were removed, since they give no additional information in a frequency count. Finally, the word *phishing* itself was removed from all definitions, as counting its occurrences would not give any insights. The resulting list of definitions contains normalised words (i.e., singular form, one spelling, no punctuation), which was analysed to get some basic understanding of the concept of phishing. The result of the frequency count was plotted in a 'word cloud' (McNaught and Lam 2010), where the font size of the words represents the number of occurrences relative to other words, i.e., the word that is mentioned the most, is set in the largest font.

Identification of concepts

In order to make sense of the set of definitions, we need to identify concepts by combining words with common meaning. This is required since the results of the frequency count are insufficient for words that refer to the same concept. For example, an attacker, criminal, crook, conman and variations thereof are all types of offender. In a simple frequency count, such as a word cloud, these individual words would occur in low frequencies, but the overall concept (offender) would occur less frequently.

Firstly, we drew a random sample ($n = 20$) from the set of definitions. By analysing this sample and highlighting words, we established which of them were relevant in each definition. We used the theoretical framework (crime science, crime scripts, 3A-model) to determine whether a word is relevant to phishing. The routine activity approach states that phishing requires a motivated offender, a suitable target and the absence of a capable guardian. In the context of phishing, the motivated offender initiates the phishing attack, the suitable target is the intended target, and no capable guardian (such as a phishing filter) is present (Hutchings and Hayes 2009). For each definition, we tried to identify these actors. Then, we identified the phases of phishing that each definition assumes. Hong (2012) identifies three phases: (1) potential victim receives a message; (2) the victim takes the suggested action; (3) offender monetises the information. Others identified phases of phishing from the viewpoint of the offender (Bose and Leung 2008), or with more detail about the methods (Forte 2009). Essentially, these phases are all high-level crime scripts. Using the phases of phishing as a framework, we identified in what way the definitions structure a phishing attack. In each definition, we highlight the words that could relate to a particular phase of phishing, even when the authors do not identify the phases explicitly. For example, Herzberg (2009) defines phishing as *'Password theft via fake websites'*, whereas Amin et al. (2012) state that phishing is *'email soliciting personal information'*. Herzberg (2009) focuses on the way passwords are stolen, not on how potential targets are drawn to the websites. Amin et al. (2012), on the other hand, identify the method of attracting potential targets, but do not explicitly state to whom the personal information is sent, or how this is done. Furthermore, after having highlighted words from the theoretical framework and words relating to the phases of phishing, any remaining words (i.e., nouns, verbs or adjectives) used to define the process of phishing are highlighted as well.

The result of the identification of important words in the sample of 20 definitions is a list of nouns, verbs and adjectives. In several iterations, synonyms and words referring to the same concept are merged. For example, the words 'creditcard numbers', 'credentials' and 'sensitive data' refer to the concept 'information'. In each iteration, we tried to find which words were related in an attempt to merge them into one concept. This resulted in 18 concepts, categorised as 3 actors, 1 asset and 14 activities (Table 2). All 93 remaining definitions were analysed using these 18 concepts to see whether they can be described as a subset of them. A second rater re-evaluated the extraction of concepts. Since the data are based on the output of the raters, Kappa is not the correct statistic to calculate the level of agreement (Feinstein and Cicchetti 1990). In this case, the proportion of agreements (agreements divided by non-agreements) was used, which was 0.78. This substantial agreement supports the applicability of the method and indicates the clarity of the theoretical framework for the raters.

The results of the frequency count, as shown in the word cloud, together with the theoretical framework, were used to label the concepts with the most commonly used terminology.

Analysis of concepts

All definitions were scored on the 18 identified concepts that were extracted. Together with the meta-data for each definition (i.e., year of publication, field and country of affiliation of first author), the results were entered into a data file. Frequency analysis was used to determine which concepts were the most important. This frequency analysis consists of establishing whether there is consensus within the set of definitions on whether to include or exclude a concept. For each concept, we determined whether the definitions agree on either inclusion or exclusion by calculating whether the number of definitions that use the concept differs significantly ($p < 0.05$) from 50% by using Pearson's chi-square test, the results of which can be found in Table 2. This results in three categories: (1) concepts that are used in significantly fewer than 50% of the definitions; (2) concepts where there is no clear consensus; (3) concepts that are mentioned in significantly more than 50% of the definitions. Concepts where there is consensus are either included (category 1) or excluded (category 3). The remaining concepts from category 2, where there is no consensus, are considered in the discussion section.

Finally, we calculate the Pearson's correlation between the year of publication and each concept, to identify evolution of the definitions with respect to the emerging concepts.

Validity

One of the threats to the validity of our study is that the review was conducted by a single researcher. However,

Table 2 Concepts used in the phishing definitions: χ^2-tests are used to determine whether the frequency of use of a concept is significantly more or less than 50% of all definitions

Type	Extracted concept	Occurrence (N)	χ^2	p	
Asset	**Mentioning information***	105	83.27	.00	Consensus
Actor	**Mentions a target***	87	44.61	.00	
Activity	Phishing is digital*	87	32.93	.00	
Activity	Phishing is Internet-based*	84	26.77	.00	
Activity	**Using deception***	79	17.92	.00	
Activity	Communication from target to offender	64	1.99	.16	No consensus
Activity	Communication from offender to target	62	1.07	.30	
Activity	Phishing is a criminal activity	61	0.72	.40	
Activity	**Using impersonation**	60	0.43	.51	
Activity	Phishing uses websites	56	0.01	.93	
Activity	Phishing uses messages	51	1.07	.30	
Actor	Mentions a trusted third party	50	1.50	.22	
Activity	Phishing is fraud*	43	6.45	.01	Consensus
Actor	Mentions an offender*	40	9.64	.00	
Activity	Using persuasion*	30	24.86	.00	
Activity	Mentions the later abuse of information*	22	42.13	.00	
Activity	Related to identity theft*	20	47.16	.00	
Activity	Related to social engineering*	19	49.78	.00	

χ^2-test with df = 1. N = 113. Boldfaced concepts are included in standard. *$p < 0.05$.

subjective decisions are mitigated by following a systematic protocol and discussing this, and the results of the exercise, with senior researchers. Additionally, a second researcher replicated the method. Cases where the second rater disagreed with the initial rater were discussed, which led to the inclusion of six definitions that had previously not been included. For the extraction of concepts, differences were discussed, leading to no changes in the 18 included concepts.

By including peer-reviewed scientific literature only, we were able to search systematically for all publications on phishing in three digital libraries. Due to the goal of this research, i.e., finding out how phishing is defined in the research community, only scientific research was included. Our design suffers from a publication bias, since all included definitions are peer-reviewed. There may be very comprehensive definitions beyond the scientific domain. If this were to be the case, we assume that a large number of research papers would reference this definition.

Although our approach of selecting publications covers a large set of the available literature, there is the possibility of not including a relevant publication. However, we minimise this potential bias by selecting based on citation count (i.e., 10 or more), source (i.e., all journals) and including recent conference papers (i.e., from 2013 and the latest 20 from 2012). If a definition of high importance to the field has been established, it is likely to have been cited by many. In addition, if an included paper cites a definition from another publication, the cited publication is included in our dataset, thereby further decreasing the potential of missing of a key definition. Finally, due to the large number of definitions, it is unlikely that the results would have been different by including a small number of additional definitions.

The extraction of concepts was based on a sample of the definitions, which could result in certain concepts not being included. We mitigated this by comparing all definitions against the identified concepts, to find out whether any definition had a different concept. Additionally, as mentioned before, another researcher reviewed a random sample of the publications. A consequence of a consensual definition is that is it based on concepts that are used in the majority of the source definitions. We did not conduct any quality assessment of the publications. The quality control was implicitly performed by including all journal articles and highly cited conference papers.

Results

The total sample of selected publications consisted of roughly 22% (n = 536) of the available peer-reviewed literature. This subset of the literature covers highly cited publications, journal articles and recent publications. The selection covers, in our opinion, most of the important literature on phishing. After review, 113 distinct definitions were extracted from the peer-reviewed literature. The definitions were analysed at the level of words and concepts.

The word cloud (Figure 2) shows the results of the frequency analysis that was used to analyse the words. The five most-used words are *information*, *website*, *user*, *personal* and *email*. From the figure, we can identify the actors, assets and activities. Actors are *user*, *victim*, *attacker*, *bank* and *business*. The assets that were found are *information, website, email, password, creditcard, username* and *account*. Finally, activities such as an *attack*, *social engineering, identity theft* or *spoofing* are most often used.

Eighteen concepts were extracted from the definitions (Table 2). Two of these concepts are common to the routine activity approach: an offender and a target. There is a weak relationship between usage of the concept social engineering in the definition and the year of publication ($r(105) = .23$, $p = .015$). This indicates that recent publications refer to social engineering more often than older publications. The presence of other concepts and the year of publication were not related, giving no evidence of evolution of the definitions with regard to other concepts.

The concepts that are used most frequently in the definitions lead to the following phishing crime script. First, the offender sends a communication to the target, which 62 of the definitions state. Typically, the offender sends the target an email (n = 30) or sends a message using a method that is not specified (n = 22), occasionally

Figure 2 A word cloud of the phishing definitions. A graphical representation of the most-used words in the definitions called a 'word cloud'. The font size represents the number of occurrences; words in a larger font are used more than words in a smaller font.

using other methods such as websites (Hodgson 2005; Levy 2004; Olurin et al. 2012), social spaces (Piper 2007), instant messages (Ali and Rajamani 2012; Verma et al. 2012), text messaging (Hinson 2010) or even letters (Workman 2008). Then, the target may reply by sending information to the offender, which is mentioned in 64 of the definitions, mostly through the use of a website (n = 40). The information that is transmitted, according to 113 definitions, can be categorised as: (1) authentication credentials (n = 13); (2) identity information (n = 5); (3) sensitive information (n = 23); or (4) personal information (n = 24). Variations or combinations account for the remaining types of information.

The results of the analysis of concepts are shown in Table 2. In the literature, there is a consensus that the concepts of deception (n = 79), a target (n = 87), information (n = 105), being digital (n = 87) and Internet-based (n = 84) should be mentioned in a definition. Furthermore, the concepts of fraud (n = 43), an offender (n = 40), persuasion (n = 30), the abuse of information (n = 22), identity theft (n = 20) and social engineering (n = 19) should not be included according to a significant majority of the definitions. There is no consensus for the remaining concepts.

Figure 3 shows the number of publications per year that define phishing, indicating several peaks in the number of definitions within particular years. Partly, this is due to the criteria used in the literature selection. For example, the peak in 2013 is due to the inclusion of all recent conference papers. However, that does not explain the decrease of definitions in 2008, and the increase thereafter. Such changes could indicate emerging consensus about the definition, so that authors start citing earlier definitions they consider useful, or, where there is a rise in the number of definitions, a change in the phenomenon might be developing, requiring redefinition.

The research field and affiliation of the first author show that mostly researchers located in the USA (n = 53) or in the field of Computer Science (n = 88) define phishing. Other countries in which the first author is located include the UK (n = 9), China (n = 8), India (n = 7), Canada (n = 7) and Australia (n = 6). There is a significant correlation between the year of publication and the first author being affiliated within the USA ($r(105) = -.46$, $p < 0.001$)), indicating that recent definitions originate more often from countries other than the USA. Almost no definitions originate from research fields other than Computer Science, with Psychology (n = 4) or Law (n = 3) as largest contributors. For 14 authors, it was not possible to establish the research field (for example, when the first author is a journalist). A possible reason for the large number of computer scientists who produce their own definition of phishing, is that they feel more inclined or capable to define phishing, whereas researchers from other fields would rather use another author's definition, or none at all.

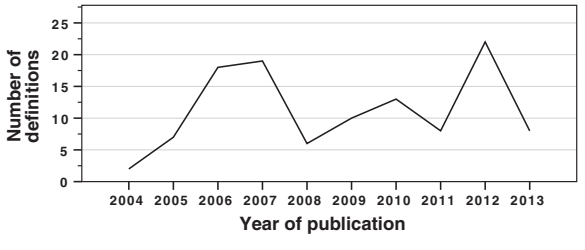

Figure 3 Number of publications with a definition of phishing, till August 2013.

Discussion

The present study identified concepts of phishing according to the peer-reviewed literature. There is a consensus on most concepts, with seven concepts present in approximately half of the definitions. We discuss each of these concepts and consider whether they should be included in the definition. However, we first observe that the concept 'Internet-based' is a subset of the concept 'digital' and therefore, one is redundant. As Internet-based is the most precise concept, arguably it should be included in the definition. This, however, leads to the discrepancy that instant messaging through an Internet-based application on a phone can be phishing, whereas a regular text message on a phone cannot (not Internet-based), even though both methods are essentially the same. In our view, phishing was made possible due to the ability to mass-distribute messages. Whereas the Internet has served as a catalyst, in facilitating communication cost efficiently, it is not the only way to do so. We propose to replace the concepts of Internet-based and digital with *scalability*. Being scalable refers to the ease of scaling from a single occurrence to hundreds, thousands or millions. Whereas digital specifies the encoding used for the channel (in bits, '0' or '1') and Internet-based is a specific channel, scalability only requires the channel to support mass-distribution.

We decided to exclude the concept of 'mentioning a trusted third party' (included in 50 of the definitions) in favour of impersonation (n = 60), since deception through impersonation by abusing the target's trust implies the existence of a trusted third party. The communication between a target and an offender is mentioned in slightly over half of the definitions (n = 62 and n = 64). However, we decided to exclude the explicit mentioning of communication, as this follows from the exchange of information from a target to an offender. Using websites (n = 56) or messages (n = 51) as specific channels for phishing were not included since these are absent from a significant

majority of the definitions. Phishing as a criminal activity is not included in the list of essential concepts, even though 61 of the definitions mention this, as it is included in deception and furthermore depends on legislation in a particular jurisdiction.

Consequently, the concepts of deception, impersonation, target, information and scalability are the most important aspects of a phishing definition. Therefore, we propose a definition of phishing that comes out of the synthesis of literature and includes all the important concepts that existing definitions have in common:

> *Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target.*

A first observation is that our definition provides a high level of abstraction, compared to most alternatives. This derives from the method used. The consequence of this is that there are no details about specific methods (such as email or websites) required to perform a phishing attack. By comparing our definition to those in Table 1, it can be seen that our definition is sufficiently abstract to be compatible with the dictionary and APWG definitions. The Oxford, Collins and Merriam-Webster definitions can be mapped entirely onto our definition, as they are more specific. For example, our definition does not include the offender's misuse of the obtained information, such as identity theft. The APWG definition is compatible as well, although it is much more specific to what is considered phishing. For example, the APWG definition specifically mentions 'technical subterfuge' schemes that tamper with a target's PC, such as installing a virus, whereas our definition – being broader – states that deception and impersonation are used. Whether or not this is followed by, or consists of, technical subterfuge, is not mentioned. Therefore, we consider the APWG definition to be compatible to ours. Finally, the American Heritage definition is the only one that is not completely compatible, since it mentions the use of a telephone, which does not scale well.

The methods employed in phishing could be used long before the Internet became popular. However, the term phishing only arose around 1995–1996 (James 2005; Khonji et al. 2013; Press 2013; Purkait 2012), indicating that mass-communication is one of the foundations of phishing. Another factor contributing to the success of phishing on the Internet is that it is cost-effective for mass-communication (i.e., spreading millions of messages). Although both are potential forms of mass-communication, letters and telegraph messages are more costly to employ on a large scale, whereas sending emails over the Internet is cheaper. This contributed to the success of the Internet as a channel for phishing. Other channels, such as telegraph messages or text messages, can be scalable, apart from the potentially high costs of sending millions of messages.

Only one indication of the evolution of phishing definitions was found: the tendency to refer to Social Engineering in papers that are more recent. However, there could still have been evolution within the literature on the act of phishing. For example, authors may have identified specific methods of phishing throughout the years, which in our analysis were mapped onto the same concept. Additionally, recent publications that define phishing more often have a first author with an affiliation not in the USA, whereas early definitions originate mainly from the USA. This could indicate that authors from outside the USA feel the need to redefine phishing because of local differences, or indicate more international interest in phishing. However, this could also be a result of the inclusion criteria (i.e., publication in English), or more interest or funding in the United States for phishing research.

Conclusions

The goal of this research was to identify a consensual definition of phishing from the literature. In the literature search, 113 different definitions were found, indicating that many researchers have thought about a definition of phishing. We identified the core concepts which the research community agrees are part of phishing, resulting in a consensual definition: *'Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target'.*

The principles of phishing were used by offenders long before the advent of the computer and the Internet. Before computers became a consumer product, these principles were considered a type of fraud. Digitalisation and mass-communication through networks provide new channels to exploit the same human vulnerabilities on a larger scale. The Internet opened many opportunities for new types of fraudulent behaviour, such as phishing. Phishing on particular channels is sometimes named differently, such as SMSishing (channel is SMS). We consider these types of phishing if they fit the consensual definition that we developed.

The implications for other definitions are mainly caused by the concepts scalable, deception and impersonation. Phishing must use deception by impersonation in order to be called phishing. When no impersonation is used, for example just asking for information, the act cannot be called phishing. Furthermore, it should be easy to scale, implying that one-to-one communication, such as a phone call, is not phishing. Spear phishing, which is phishing with a single target, is possible, as long as the employed method supports scalability.

The main theoretical contribution of this study is threefold. Firstly, we validated the findings of Hutchings and

Hayes (2009), Reyns et al. (2011) and Pratt et al. (2010) that the routine activity approach, developed for explaining crime in the physical world, can be applied to the digital world. Within the context of phishing, routine activities include, for example, giving one's email address away, time spent on the Internet, time spent on email. Such routine activities could lead to more opportunity for victimisation. Additionally, we suggest the notion of crime facilitation to be relevant to cybercrime, and specifically phishing. People can deliberately, negligently or unconsciously facilitate their own victimisation by placing themselves at special risk Sparks (1982). The second theoretical contribution of this research is the development of a consensual definition of phishing. Yar (2012) states that networked communications act as a force-multiplier and that the impact is further increased by a space-time compression, whereby actions can occur almost instantly in different locations. Therefore, he argues that new theoretical notions are required for theorising about cybercrime. We believe these notions are manifested in the concept 'scalability' of the consensual definition and therefore constitute the third theoretical contribution.

This research adds a consensual definition of phishing to the body of existing definitions so that others can be weighed against the concepts with consensus within the research community. Research can be aligned by using a common definition, thereby avoiding misinterpretations. Researchers who define phishing differently can relate their definition to the consensual one, thus positioning better which actions they consider phishing. Furthermore, meta-studies on phishing are better facilitated with our definition. Institutions, such as the police or banks, benefit from a consensual definition as well. Collaboration and data sharing between different organisations is easier if both have a common vocabulary. Organisations labelling phishing incidents according to a consensual definition will find it easier to compare the effectiveness of countermeasures.

Future research could focus on translating and interpreting the consensual definition into other languages. The consensual definition can be related to the definitions that practitioners use, thereby extending this study into the non-scientific domain. Furthermore, a discussion in the research community should establish more clarity on the concepts where there is no consensus at this moment. We believe that the lessons learned in crime science and the theories and tools that crime scientists developed, should be applied to phishing. In particular, we suggest studying the notion of crime facilitation in cybercrime, in addition to crime opportunity. Ultimately, a collaboration of crime science and computer science could help in reducing phishing victimisation and avoid reinventing the wheel.

Additional file

Additional file 1: List of all definitions. This file includes a table with all included definitions.

Competing interests
The author declares he has no competing interests.

Acknowledgements
The author would like to thank Lorena Montoya, Eleftheria Makri and John Horn for their feedback on earlier versions of this manuscript. Furthermore, he would like to thank the reviewers and editorial staff for their insightful comments and suggestions. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318003 (TREsPASS). This publication reflects only the author's views and the Union is not liable for any use that may be made of the information contained herein.

Received: 21 August 2013 Accepted: 4 June 2014

References

Abu-Nimeh, S, Nappa, D, Wang, X, Nair, S (2007). *A comparison of machine learning techniques for phishing detection* Vol. 269, (pp. 60–69). New York, NY, USA: ACM.

Al-Hamar, M, Dawson, R, Al-Hamar, J (2011). The need for education on phishing: a survey comparison of the uk and qatar. *Campus-Wide Information Systems*, *28*(5), 308–319.

Ali, M, & Rajamani, L (2012). Deceptive phishing detection system: from audio and text messages in instant messengers using data mining approach. In *Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME)* (pp. 458–465). Salem, Tamilnadu: IEEE.

American Heritage Dictionary (2013). Phishing, 5th edn.: Houghton Mifflin Harcourt Publishing Company. http://www.ahdictionary.com/word/search.html?q=phish. Accessed 20 December 2013.

Amin, R, Ryan, J, van Dorp, J (2012). Detecting targeted malicious email. *IEEE Security Privacy*, *10*(3), 64–71.

Anti-Phishing Working Group (2013). Phishing Activity Trends Report, 2nd Quarter 2013. http://docs.apwg.org/reports/apwg_trends_report_q2_2013.pdf. Accessed 23 December 2013.

Bose, I, & Leung, A (2008). Assessing anti-phishing preparedness: a study of online banks in Hong Kong. *Decision Support Systems*, *45*(4), 897–912.

Brantingham, P, & Brantingham, P (1993). Environment, routine and situation: toward a pattern theory of crime. In RV Clarke & M Felson (Eds.), *Routine Activity and Rational Choice: Advances in Criminological Theory,* volume 5 (pp. 259–294): Piscataway: Transaction Press.

Brantingham, P, & Brantingham, P (2008). Environmental criminology and crime analysis. In R Wortley & L Mazerolle (Eds.) Devon: Willan Publishing.

Clarke, RV (2009). Situational crime prevention: theoretical background and current practice. In Krohn M D, Lizotte A J, Hall G P (Eds.), *Handbook on Crime and Deviance,* Handbooks of Sociology and Social Research chapter 14 (pp. 259–276). New York: Springer New York.

Cohen, LE, & Felson, M (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, *44*(4), 588–608.

Collins English Dictionary (2013). Phishing. http://www.collinsdictionary.com/dictionary/english/phishing. Accessed 20 December 2013.

Cornish, D (1994). The procedural analysis of offending and its relevance for situational prevention. *Crime Prevention Studies*, *3*, 151–196.

Cornish, D, & Clarke, RV (2008). Environmental criminology and crime analysis. In R Wortley & L Mazerolle (Eds.) Devon: Willan Publishing.

El Helou, S, Li, N, Gillet, D (2010). The 3a interaction model: towards bridging the gap between formal and informal learning. In *Proceedings of the Third International Conference on Advances in Computer-Human Interactions (ACHI)* (pp. 179–184). Saint Maarten: IEEE.

Feinstein, AR, & Cicchetti, DV (1990). High agreement but low kappa: I. The problems of two paradoxes. In *Journal of Clinical Epidemiology*, 43 (pp. 543–549).

Felson, M, & Clarke, R (1998). Opportunity makes the thief: practical theory for crime prevention. *Home Office*, *98*, 1–36.

Forte, D (2009). Anatomy of a phishing attack: a high-level overview. *Network Security*, *2009*(4), 17–19.
Freyne, J, Coyle, L, Smyth, B, Cunningham, P (2010). Relative status of journal and conference publications in computer science. *Communications of the ACM*, *53*(11), 124.
Herzberg, A (2009). Why johnny can't surf (safely)? Attacks and defenses for web users. *Computers and Security*, *28*(1–2), 63–71.
Hinson, G (2010). There must be thirty ways to steal your id. *EDPACS*, *41*(5), 1–15.
Hodgson, P (2005). The threat to identity from new and unknown malware. *BT Technology Journal*, *23*(4), 107–112.
Hong, J (2012). The state of phishing attacks. *Communications of the ACM*, *55*(1), 74–81.
Hutchings, A, & Hayes, H (2009). Routine activity theory and who gets caught in the 'Net'? *Current Issues in Criminal Justice*, *20*(3), 433–451.
James, L (2005). *Phishing exposed*. Rockland: Syngress Publishing Inc.
Khonji, M, Iraqi, Y, Jones, A (2013). Phishing detection: a literature survey. *IEEE Communications Surveys & Tutorials*, *15*(4), 2091–2121.
Kitchenham, B, & Charters, S (2007). *Guidelines for performing systematic literature reviews in software engineering*: Technical Report EBSE-2007-01, Software Engineering Group, Keele University.
Kumaraguru, P, Rhee, Y, Acquisti, A, Cranor, L, Hong, J, Nunge, E (2007). Protecting people from phishing: the design and evaluation of an embedded training email system. In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 905–914). New York, NY, USA: ACM.
Kumaraguru, P, Sheng, S, Acquisti, A, Cranor, L, Hong, J (2010). Teaching johnny not to fall for phish. *ACM Transactions on Internet Technology*, *10*(2), 1–31.
Levy, E (2004). Interface illusions. *IEEE Security and Privacy*, *2*(6), 66–69.
McFedries, P (2006). Technically speaking: gone phishin'. *43*, *4*, 80.
McNaught, C, & Lam, P (2010). Using wordle as a supplementary research tool. *The Qualitative Report*, *15*(3), 630–643.
Merriam-Webster (2013). Phishing. http://www.merriam-webster.com/dictionary/phishing. Accessed 20 December 2013.
Moore, T (2007). Phishing and the economics of e-crime. *Infosecurity*, *4*(6), 34–37.
Moran, T, & Moore, T (2010). The phish-market protocol: secure sharing between competitors. *IEEE Security and Privacy*, *8*(4), 40–45.
Olurin, M, Adams, C, Logrippo, L (2012). Platform for privacy preferences (p3p): current status and future directions. In *Tenth Annual International Conference on Privacy, Security and Trust* (pp. 217–220). Paris: IEEE.
Oxford University Press (2014). Phishing. http://www.oxforddictionaries.com/definition/english/phishing. Accessed 18 July 2014.
Piper, P (2007). A newer, more profitable aquaculture. *Searcher: Magazine for Database Professionals*, *15*(9), 40–47.
Pratt, TC, Holtfreter, K, Reisig, MD (2010). Routine online activity and internet fraud targeting: extending the generality of routine activity theory. *Journal of Research in Crime and Delinquency*, *47*(3), 267–296.
Oxford University Press (2013). OED Online. Accessed 20 December 2013.
Purkait, S (2012). Phishing counter measures and their effectiveness—literature review. *Information Management & Computer Security*, *20*(5), 382–420.
Reyns, BW, Henson, B, Fisher, BS (2011). Being pursued online: applying cyberlifestyle-routine activities theory to cyberstalking victimization. *Criminal Justice and Behavior*, *38*(11), 1149–1169.
Schank, R, & Abelson, R (1975). Scripts, plans, and knowledge. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence* (pp. 151–157). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
Sparks, RF (1982). *Research on victims of crime: accomplishments, issues and new directions*. Rockville: National Institute of Mental Health: Crime and delinquency issues.
Verma, R, Shashidhar, N, Hossain, N (2012). Two-pronged phish snagging. In *Proceedings of the 7th International Conference on Availability, Reliability and Security (ARES)* (pp. 174–179).
Workman, M (2008). Wisecrackers: a theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology*, *59*(4), 662–674.
Xiang, G, & Hong, J (2009). A hybrid phish detection approach by identity discovery and keywords retrieval. In *Proceedings of the 18th International World Wide Web Conference (WWW)* (pp. 571–580). New York, NY, USA: ACM.
Xiang, G, Hong, J, Rose, C, Cranor, L (2011). Cantina+: a feature-rich machine learning framework for detecting phishing web sites. *ACM Transactions on Information and System Security*, *14*(2), 1–28.
Yar, M (2005). The novelty of 'cybercrime': an assessment in light of routine activity theory. *European Journal of Criminology*, *2*(4), 407–427.
Yar, M (2012). Sociological and criminological theories in the information era. In W Stol & R Leukfeldt (Eds.), *Cyber-Safety: An Introduction*. Utrecht: Eleven International Publishing.

doi:10.1186/s40163-014-0009-y
Cite this article as: Lastdrager: **Achieving a consensual definition of phishing based on a systematic review of the literature.** *Crime Science* 2014 **3**:9.