

The Economy of Phishing: A Survey of the Operations of the Phishing Market

By Christopher Abad, Research Scientist

Table of Contents

Abstract	2
Introduction	3
The phishing landscape	3
The Marketplace.....	9
Conclusion	13
Notes.....	14
References.....	14

Abstract

Phishing has been defined as the fraudulent acquisition of personal information by tricking an individual into believing the attacker is a trustworthy entity¹. Phishing attacks are becoming more sophisticated and are on the rise. In order to develop effective strategies and solutions to combat the phishing problem, one needs to understand the infrastructure in which phishing economies thrive.

We have conducted extensive research to uncover phishing networks. The result is detailed analysis from 3,900,000 phishing e-mails, 220,000 messages collected from 13 key phishing-related chat rooms, 13,000 chat rooms and 48,000 users, which were spidered across six chat networks and 4,400 compromised hosts used in botnets.

This paper presents the findings from this research as well as an analysis of the phishing infrastructure.

Introduction

"I have ways of making money that you know nothing of." — John D. Rockefeller

Phishing has been defined as the fraudulent acquisition of personal information by tricking an individual into believing the attacker is a trustworthy entity¹. However, this definition is over-simplified and the fact is that phishing has become a thriving economic infrastructure.

Online phishing can be traced as far back as 1996² and has escalated to today's levels of an estimate of more than 250,000 phishing attempts in a single day against a given financial institution and its customers.

Phishing does not occur in isolation but rather operates within a complex network. In fact, individuals involved in phishing do not typically understand how to orchestrate an entire phishing attack.

Phishing economies are self-organized merchants and consumers governed only by the laws of supply and demand. These laws govern all aspects from the global economics to the launch and spread of attacks. These systems are formally known as scale-free networks and researchers have only just begun to understand how they operate.

The major effects of phishing are identity theft and monetary loss.

Identity theft is a major fear for most consumers as a result of the awareness created by the media through second and third hand accounts on its effects.

Monetary loss impacts both consumers and the corporate brands. For consumers, phishing usually translates to direct monetary loss via theft. For the label, it leads to rising costs of prevention and remediation, and soft monetary loss as a result of brand erosion and undermined consumer trust.

Cloudmark's SpamNET blocks close to 500 million e-mail threats per day. Less than 1/10th of a percent of that number is estimated as phishing. However, the impact of phishing on consumers as well as on organizations is exponentially more severe than other e-mail threats, such as spam.

The phishing landscape

Phishing networks

In an environment of overloaded information people tend to congregate in locations that are easy to find³. Phishing networks present information efficiently. Channels often have naming conventions that connect them clearly to the phishing underground. Newcomers can easily become acquainted with the basics by joining chat rooms or forums with simple names such as "credit cards."

Chat room content

On one particular chat network node-breaking occurred constantly because of per-server chat room and namespace restrictions. Node breaking is a process that causes channel interconnectivity to be severed or significantly weakened. This results in disconnected islands of chatrooms; it becomes difficult for users to efficiently navigate through the community and engage in phishing transactions. Another chat network had network-wide channel restrictions which created more problematic namespace restrictions, making it difficult to navigate the potential phishing community. The most viable public network had an official channel services infrastructure to support enforcement and authority in registered chat rooms.

However, unregistered chat rooms are unregulated. As long as participants in the phishing economy maintain authority over unregistered chat rooms and block attempts at node-breaking from unsanctioned attacks, they remain viable hubs in the phishing community. Smaller peripheral and completely unregulated networks are used and endorsed by loosely affiliated fraud groups. Newcomers eventually discover these networks through wide area communications channels.

Personal interconnectivity

We have made an attempt to analyze individual interconnectivity through consistent monitoring of user idle times and by using a modified version of the tool, PieSpy⁴. The main premise of the tool is that individual interconnectivity can be inferred by analyzing the frequency of communications between individual users and by then comparing the time gaps in communication to other individuals in a monitored group of users.

Botnets

Although a peripheral aspect of phishing, bots play a major role in the Internet chat community. A bot is an automated chat program that can perform certain tasks autonomously and can be remotely controlled by chat participants. A botnet is essentially a cluster of bots running on many computers, which can be centrally controlled to carry out any task it has been programmed for, whether *en masse* or individually. Botnet tasks vary, but a few common uses are synchronizing distributed denial of service attacks (DDoS), virus-like self-propagation through exploiting vulnerable computer systems, as well as the more benign task of chatroom management.

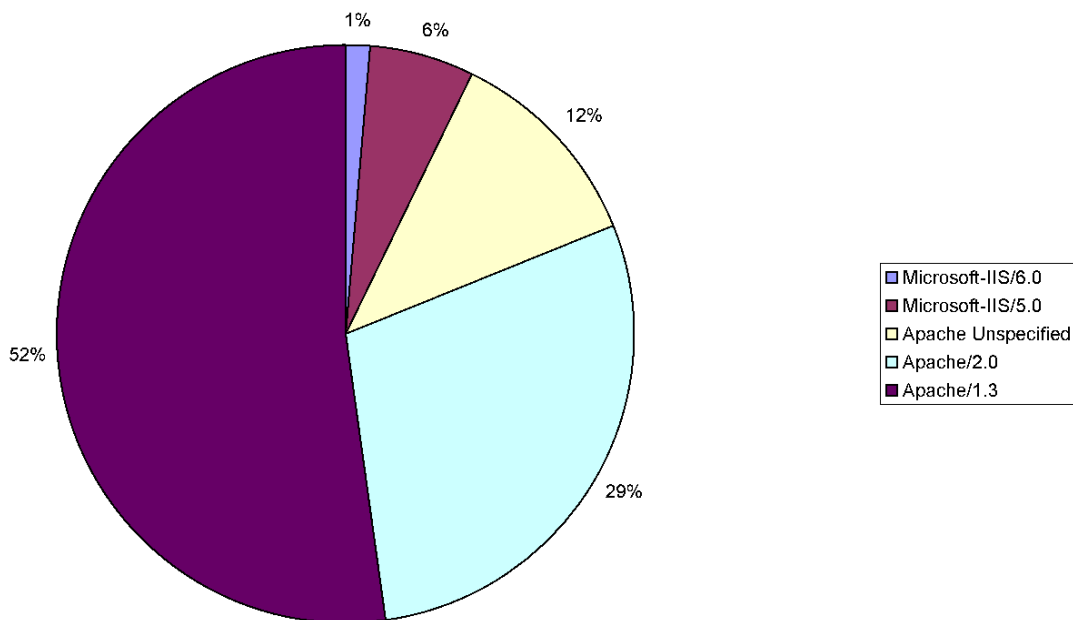
The illicit nature of phishing chat rooms creates the need for alternative means to maintain authority and control on public servers. This is achieved through networks of bots. Their main role in Internet chat is to form a mercenary authority infrastructure in otherwise unregulated chat rooms. Although botnets need not be composed of compromised computers, the most powerful ones typically are.

Botnets are also used for organizing massive attacks, such as flooding DDoS. These botnets usually consist of up to 10,000 compromised zombie computers. During this study, one of our own research bots was attacked on two separate occasions by another botnet comprised of over 4,000 computers and with close ties to phishing chat rooms.

Phishing scam pages and system analysis

Cloudmark conducted an analysis of 143 machines used to launch phishing scam pages. We observed that the ratio of Apache versus IIS based scam pages is 133:10 and Netcraft's Web server survey's findings of Apache's market share ratio to IIS of 33:10⁵. Our research confirms that Apache servers running Linux is the preferred platform for launching phishing scams. One possible reason for this is that it is easier to remotely manage and stage a phishing scam on a Linux-based computer than on a Windows-based computer. The relative difficulty in actually compromising these two platforms is questionable and difficult to compare; however, the high rate at which Windows computers are compromised is attested to by the prevalence of Windows-based worms.

Webservers Used in Phishing Attacks



Infrastructure robustness

Scale-free networks have proved to be highly resistant to random failure⁶. Eighty percent of participants can be removed in a social network before the core infrastructure completely collapses⁷. However, scale-free networks are vulnerable to attacks on hub nodes. Hub nodes are essentially well-known participants and key forums within the community, which fuel the constant influx of newcomers and connect peripheral systems together. Successful attacks on hubs sometimes trigger what are known as "Node-breaking Avalanches." When a key hub is successfully attacked, it can cause a chain reactions that successively causes channel connectivity to be severed a few degrees of separation away from the key hub

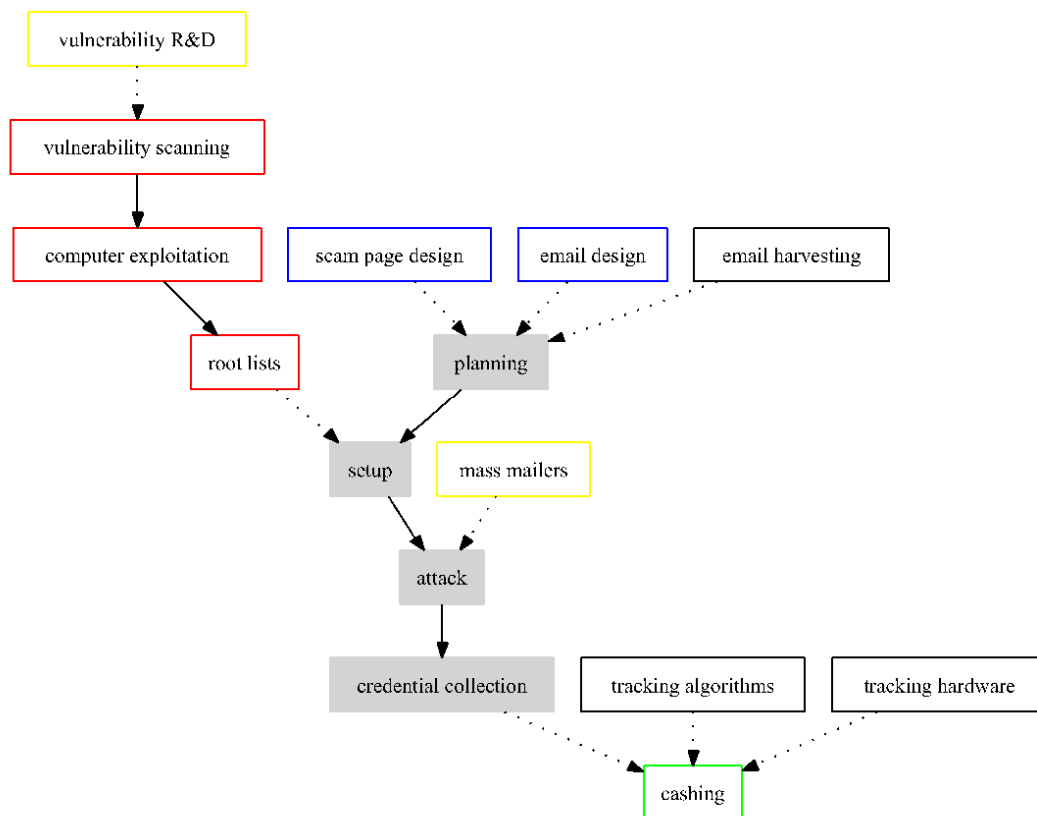
node. Within certain chat communities, measures have been put into place by chat network administrators to successfully block the creation of viable phishing hubs.

The strategy of eliminating hubs is unlikely to be effective because the environment allows for them to be easily rebuilt. Smaller disconnected chat room islands simply reconnect when they can reassert their authority.

Although it did not impact the phishing economy as a whole, an example where the phishing infrastructure was successfully dismantled on a single network occurred on a cluster of public chat servers collectively known as EFNET. Server administrators simply elected to prohibit users with fraud-related names from joining chat rooms.

Process flow of a phishing attack

The phishing process is easily represented in a flowchart with input required from different players at various stages. Each step involves specialized skills from other members of the Internet community.



Planning

The planning step requires information to be collected, such as target e-mail lists, *scam page* templates, as well as demanding knowledge from consumers of phishing credentials.

Detailed information such as target e-mail lists and scam page templates needs to be collected. The phisher need not be adept at Web design, but instead can obtain an already used scam page. Scam pages and e-mail templates are widely available within the community. If more advanced templates techniques are preferred, skilled Web designers who advertise on known fraud-related forums can be hired.

The trade of compromised computers, also known as *Roots*, is a thriving economy. Computers are easily compromised through various public exploits and Trojans with security holes in network software. The security community readily provides proof-of-concept exploits which can be used to gain access to vulnerable computers so there is a constant supply of compromised hosts. Phishers do not require technical knowledge of how to compromise hosts and can purchase access to compromised hosts from hackers.

Setup

The next step involves ensuring the proper scam page infrastructure on the compromised hosts used in the phishing attack.

Mapping a process to send back credentials to an anonymous e-mail address or a chat room.

This step requires minimal technical expertise and consists of little more than uploading Web site content, and setting up what is known as an eggdrop or a simple mail script. An eggdrop bot is a scriptable automated program that connects to a chat room, allowing the user to remotely control it and issue commands. It is also the basic unit of a botnet (a centrally controlled cluster of these chat room bots). An eggdrop bot used for the purpose of harvesting credentials from a scam page machine could be programmed to relay the collected information back to a chat room or sent to a specific user on demand.

Attack

Numerous programs have been written to handle mass mailings, and there are commercial appliances which also generate mass mails. As with the preceding steps the phisher does not need specialized knowledge to send out e-mails *en masse* and only needs to acquire the right tool.

The primary protocol responsible for sending e-mail, SMTP, inherently lacks a mandatory authentication facility as with most protocols designed during the same era. Many tutorials, including the HACK FAQ, provide simple explanations on how to send forged e-mail via a Telnet program. Although some authentication schemes now exist to counter this problem, sender authentication is not universally deployed. The Sender Policy Framework (SPF) attempts to address one aspect of sender verification issues.

Basic knowledge of HTML enables the phisher to copy the formatting and style of valid e-mails from the banking institution.

Collection

The collection phase of phished credentials is often performed anonymously; for example, a process on the scam-page hosting machine periodically sends back phished credentials to anonymous Web-based e-mail accounts. These accounts are then accessed via a proxy server or sent to a chat room by an eggdrop chat bot. It is also possible to place the credentials into a readable directory on the Web server and download new credentials directly from a browser pointed to the proper directory where the credentials are stored.

Cashing

This is often the end of the line for the phisher. At this point phishers are now providers of credential goods with a limited supply of customers. Consumers of financial institution credentials are known as Cashers. The Cashers' main role is to take the phished credentials and obtain currency directly from the accounts attached to the credentials. Phishing and Cashing are distinct and often separate roles.

The Marketplace

The phishing marketplace

The phishing marketplace is a loosely tied group of forums where participants can trade goods, services, and money. The key goods are credentials. Credentials are then valued according to the level of detail. It is currently difficult for cashers to extract monetary value from credentials based on different measures each financial institution has in place.

Advertisements used in the phishing economy were collected and we observed a wide range of credential valuations, noting the relative demand for a number of banking institutions. We gained some insights into the characteristics that govern credential valuation. Some examples of preventative measures in place by banking institutions, namely security advancements in ATM card encoding processes, have resulted in greatly diminished demand for those institutions.

Cashers and the credential trade

Cashers often play no role in obtaining the credentials from customers but instead purchase credentials in bulk through bartering various goods and services, or taking commission on extracted funds.

Purchase rates on a per account basis range from US\$0.50 for supplier-selected credit card samples and up to US\$100 for high balance verified fulls, which consists of full cardholder information, banking and routing numbers, credit card numbers, expiration dates, the cvv2 code (the three-digit number on the back of a credit card directly following the credit card

number), current account balance, and the ATM pin. Commission rates for cashers are as high as 70 percent of cash outs. Commission-based cash outs are often wired over Western Union to the credential supplier because of its international presence and relative anonymity for the pick-up party.

Following are some typical casher advertisements:

```
<SuperCash | #cctradez> I can cashout Washington Mutual ,  
Key bank , Money access , HouselHold, CitizensBank,  
Mellon Bank, Sky Bank, BankNorth , Zip Network ,  
Commerce Bank, PNC bank (443071) , Regions Bank ,  
Banknorth , Bank One (478200), Capital One (517805 ,  
529149, 493422, 412174) PEOPLE'S BANK , Bank Of America  
(440893 , 549105, 550535), The Hungtington National Bank  
, JAPAN , MBNA (549035, 426429, 549198), Republic Bank
```

```
<ebayinfos | #WAMU> I can cashout all wamu bins, PNC,  
TCF, Citibank South Dakota, MBNA, Summit Visa Sweden  
Banknorth, Canadian Imperial Bank VSB International B.V,  
M& I BANK OF SOUTHERN WISCONSIN, and more bins.....  
your share is 60%.
```

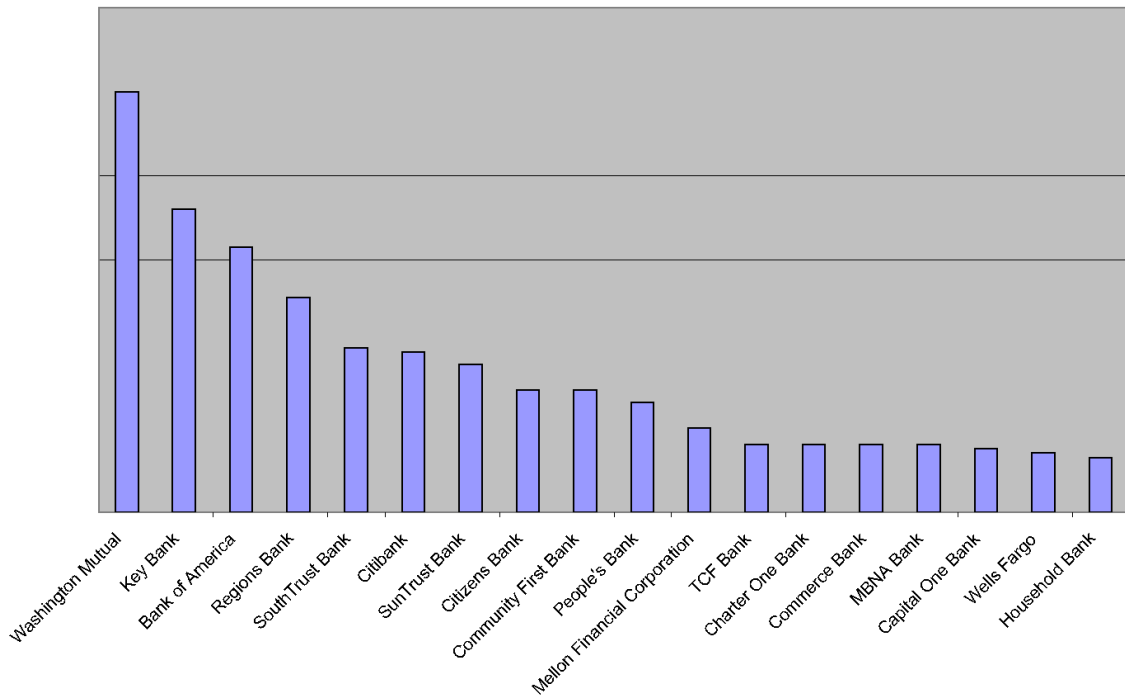
What is seen here, is that the user, *SuperCash*, is a member of the chatroom, *#cctradez*. They are advertising that they have the ability to cash a number of banks. Additionally, some banks have a list of BINs in parentheses to the right of the bank name. These BINs are listed because only a subset of the BINs associated with those banks can be cashed. One reason why only certain BINs for a bank can be cashed is that banks sometimes have BINs associated with management regions where the bank accounts were opened and each region uses a different tracking method. Another reason is that while an encryption-based tracking method is used, a different encryption key is associated with each BIN and certain keys have been successfully cracked.

Credential supply and demand

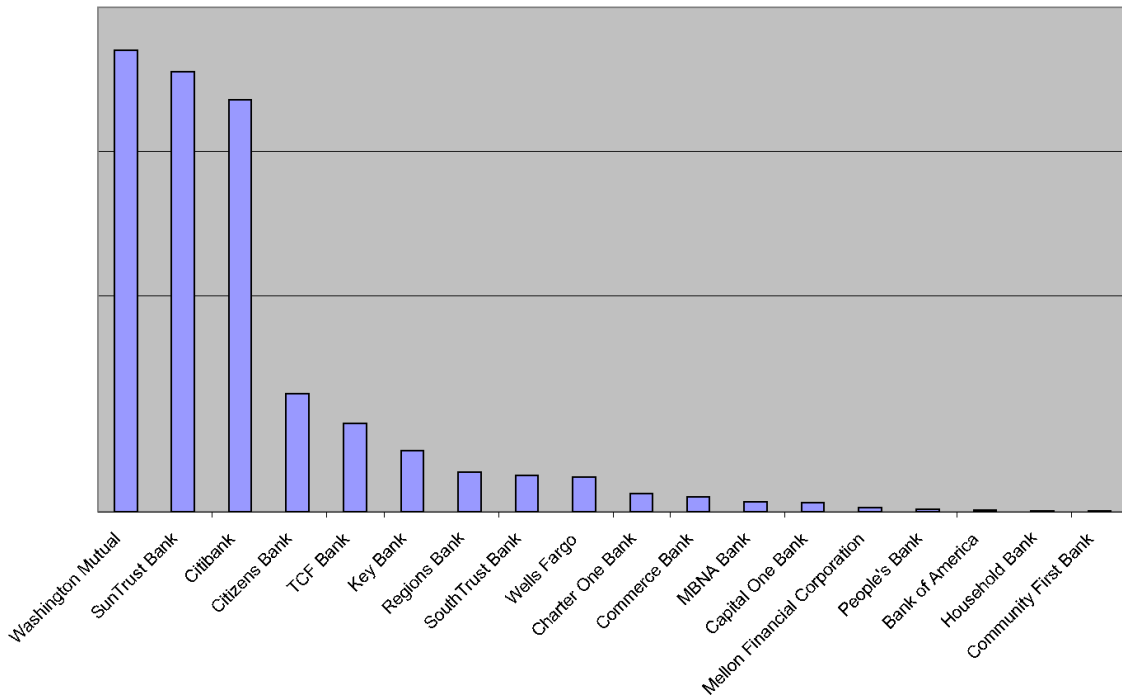
As expected, phishing reports and casher advertisements for specific financial institutions follow power-law distribution, but the largest institutions are not necessarily the most targeted or reported. The report rate data was collected from Fraud Watch International.

The reasons for targeting a specific brand may partly be due to demand for a given brand's customer information. Information was collected about the consumers of phished credentials from various chat rooms and online forums. A graph was developed to show the relative demand of given financial institution credentials based on advertisements by the consumers of phished credentials. There are some similarities in ranking between credential demand and phishing reports.

Institutions by Advertisement Rate



Institutions by Report Rate



Tracking and credential demand

There are varying degrees of difficulty to cashing out certain credentials. For banking credentials, the preferred though more difficult method is ATM fraud, where the casher actually encodes the banking information (tracking) onto an ATM card and withdraws the maximum daily funds from that account. The popularity of tracking has grown because it has become increasingly difficult to ship purchased goods to countries where credit card fraud is a major problem.

The main difficulty with tracking is the encoding of bank data to the ATM card. The preferred hardware used to encode information onto magnetic stripe cards is the MSR-206. Although the MSR-206 hardware most preferred by cashers can be easily obtained, each bank uses a specific encoding algorithm to translate the credentials into the encoded data written to an ATM card. The tracking algorithm may be as simple as appending the expiration date and cvv2 code along with a fixed numeric value to the end of a check card number, or as complex as encrypting the information with a secret key and then encoding the encrypted block to the card.

It is no surprise that Washington Mutual, Key Bank, and various other institutions are at the top of phishers' lists. The tracking algorithms for these financial institutions are easily obtained from within the phishing economy, while Bank of America, a huge financial institution, is nearly off phishers' radar because their encoding algorithm is very hard to obtain or crack. According to statements by phishers, it may be based on Triple-DES, a strong encryption algorithm.

Only a subset of accounts from certain banks can be cashed out as the above advertisement shows. The names of the banks in the advertisement are followed by a number called a BIN. A BIN is a bank identification number which associates an account number with a banking institution. Some large institutions have many BINs due the large number of banking customers they have. Casher advertisement of BINs lets potential credential sellers know that only accounts associated with these BINs can be cashed out.

The rate at which a financial institution is targeted for phishing scams is affected more by the pressure cashers put on the financial institutions and customer credentials, and less by the phisher's ability to supply. Additionally, demand for credentials is created by the cashers' ability to cash out on a given financial institution. This is dependent on the level of difficulty used in the credential tracking.

Conclusion

The findings of this study illuminate both the activities and social environment of the phishing economy and some successful methods in undermining phishing-related activities.

Chat network administrators can help thwart the phishing community from establishing a presence on a chat network simple by restricting chatroom names.

The phishing economy is comprised of many participants that play specialized roles which overlap with other online communities. Phishing participants are not an isolated set of individuals.

Keeping computer systems up-to-date with the latest security fixes will lessen the chance that they are compromised and possibly used in botnets and fraudulent activities.

Although all banks are at risk of having their users' private information phished, the impact of phishing can be reduced by using encrypted magstripe tracking methods which will thwart ATM fraud.

Phishing e-mails are only a small aspect of the overall phishing economy and until now, the only aspect seen by the most people. The phishing economy is a decentralized and self-organized social network of merchants and consumers governed by laws of supply and demand. This clearer picture of the landscape, the players, and insight into phishing operations will hopefully assist in the fight against online fraud.

About the author

Christopher Abad is a research scientist with Cloudmark (<http://www.cloudmark.com/>), a San Francisco-based spam filtering service provider.

Notes

1. Financial Services Technology Consortium, "Understanding and Countering the Phishing Threat," at <http://fstc.org/projects/counter-phishing-phase-1/>, last accessed 16 August 2005.
2. Next Generation Security Software Ltd., "The Phishing Guide: Understanding and Preventing Phishing Attacks," at <http://www.ngssoftware.com/papers/NISR-WP-Phishing.pdf>, last accessed 16 August 2005.
3. John Robb, "Scale-Free Networks," at http://globalguerrillas.typepad.com/globalguerrillas/2004/05/scalefree_terro.html, last accessed 16 August 2005.
4. PieSpy Social Network Bot, "Inferring and Visualizing Social Networks on IRC," at <http://www.jibble.org/piespy/>, last accessed 16 August 2005.
5. Netcraft, "Web Server Survey Archives," at http://news.netcraft.com/archives/web_server_survey.html, last accessed 16 August 2005.
6. Liang Zhao Kwangho Park and Ying-Cheng Lai, 2004. "Attack vulnerability of scale-free networks due to cascading breakdown," *Physical Review*, E 70, 035101(R), and at http://chaos1.la.asu.edu/~yclai/papers/PRE_04_ZPL.pdf, last accessed 16 August 2005.
7. John Robb, "Cascading System Failure," at http://globalguerrillas.typepad.com/globalguerrillas/2004/05/cascading_syste.html, last accessed 16 August 2005.

References

Reka Albert and Albert-László Barabási, 2002. "Statistical mechanics of complex networks," *Reviews of Modern Physics*, volume 74, number 1 (January), pp. 47–97, ; see also <http://arxiv.org/abs/cond-mat/0106096>, accessed 16 August 2005.

Albert-László Barabási, Réka Albert, and Hawoong Jeong, 2000. "Scale-free characteristics of random networks: The topology of the World Wide Web," *Physica A: Statistical Mechanics and its Applications*, volume 281, issues 1–4 (15 June), pp. 69–77.

Cloudmark SpamNet, at <http://www.cloudmark.com/>, accessed 16 August 2005.

Financial Services Technology Consortium, "Understanding and Countering the Phishing Threat," at <http://fstc.org/projects/counter-phishing-phase-1/>, last accessed 16 August 2005.

Fraud Watch International, "Phishing Scams: Understanding the Latest Trends," at <http://www.fraudwatchinternational.com/internetfraud/phishing/report.pdf>, last accessed 16 August 2005.

Graphviz, "Graph Visualization Software," <http://www.graphviz.org/>, last accessed 16 August 2005.

IRC Spider, at <http://www.the-mathclub.net/site/code/ircspider.tar.gz>, last accessed 16 August 2005.

Adilson E. Motter and Ying-Cheng Lai, 2002. "Cascade-based attacks on complex networks," *Physical Review*, E 66, 065102(R), and at http://chaos1.la.asu.edu/~yclai/papers/PRE_02_ML_3.pdf, last accessed 16 August 2005.

Netcraft, "Web Server Survey Archives," at http://news.netcraft.com/archives/web_server_survey.html, last accessed 16 August 2005.

Next Generation Security Software Ltd., "The Phishing Guide: Understanding and Preventing Phishing Attacks," at <http://www.ngssoftware.com/papers/NISR-WP-Phishing.pdf>, last accessed 16 August 2005.

Liang Zhao Kwangho Park and Ying-Cheng Lai, 2004. "Attack vulnerability of scale-free networks due to cascading breakdown," *Physical Review*, E 70, 035101(R), and at http://chaos1.la.asu.edu/~yclai/papers/PRE_04_ZPL.pdf, last accessed 16 August 2005.

PieSpy Social Network Bot, "Inferring and Visualizing Social Networks on IRC," at <http://www.jibble.org/piespy/>, last accessed 16 August 2005.

John Robb, "Cascading System Failure," at http://globalguerrillas.typepad.com/globalguerrillas/2004/05/cascading_syste.html, last accessed 16 August 2005.

John Robb, "Scale-Free Networks," at http://globalguerrillas.typepad.com/globalguerrillas/2004/05/scalefree_terro.html, last accessed 16 August 2005.