O&O Study: Data Protection on Used Storage Media

Data Data Everywhere 2005

Olaf Kehrer, O&O Software, Berlin – May 2005

In today's internet age, having protection while online is a necessity. More and more, well-intentioned users have to cope with the ever growing presence of internet thieves, questionable characters dead set on wreaking havoc on the net and scooping up any confidential bank or personal data to which they can gain access. In response, both the IT industry and a number of governmental institutions have attempted to makes lasting steps to battle this unsettling trend. And what happens to private or company data when a computer is no longer needed and to be retired? Are these data at risk of being reconstructed or exploited? The results of this study show exactly how data can be given away by unsuspecting users through the resale or disposal of a used hard disk.

As part of last year's study, we purchased 100 used hard disks by way of online eBay auctions. The goal was to see whether or not those hard disks purchased were securely cleared of any previous data. What we found was staggering [1]. 88% [1] of the functional hard disks were not correctly cleared of old data, leaving private and personal data open for restoration. In addition, patient information from an insurance company, internal documents from a pharmaceutical company, and multiple documents from a wide variety of business sources were also discovered. Gaining the attention of the media, press and television reports followed, educating a wider array of PC users than ever before. The message began to solidify among those who simply assumed that the formatting of a hard disk was the only sure way to securely delete one's data forever. Due to the dubious description of this function in Windows giving users a very false sense of security, such an assumption is understandable, regardless of how mistaken it is. (See Figure 1).

Attacks by viruses and other malicious software are not the only threats to data security on a PC. The disposal of a computer after it is no longer needed can also pose a serious threat as well and should be carefully planned and carried out correctly.

A year after the release of our initial study of the same title, we wanted to test whether the general awareness regarding this issue among private and business PC users had improved. As part of our evaluation, we once again purchased a series of hard disks and scanned them for deleted data using standard retail data recovery software.

2004 Results

While purchasing the hard disks for the 2004 study, we could never have imagined what sorts of data we would be finding. In fact, one of the first hard disks we received contained patient information from a health insurance company. How was this possible that a hard disk containing such data would be auctioned off to the public?

[1] In the study, 100 hard disks were analyzed, of which 15 were found to have physical hardware errors. On the remaining 85 hard disks, 75 produced restorable data. Therefore, data was restored on about 88% of the functioning hard disks or 75% of all tested hard disks [1]

But the surprises just kept on coming. Hard disks containing extremely personal and private data, including scanned charge cards, pin numbers, contracts, worker evaluations, and a court-ordered notice of termination, to name just a few. In addition, private family vacation photos and videos as well as pornography were also found. All in all, over a half a million files could be restored and accessed through this study.

It is important to once again note that all of this data restoration was accomplished with the help of standard data restoration software easily accessible to most Windows users without any prior skills or training in this field.

Reactions

After the study's release, we received a wide response in the media. Reports on television and in the press followed this study's lead and illustrated this problem even further through real-life applications. It was shown in these reports that most users are unaware of these issues.

This "deletion", considered by most users to be permanent, is merely the reformatting of data on the hard disk, a process that does not actually get rid of any data. If one were to compare a hard disk to a book, the process of formatting data on a hard disk would be analogous to tearing out the book's table of contents. Although the references to the data are removed, the actual corresponding data remain saved on disk. From the user's perspective, these data are no longer visible, nor accessible. This is, of course, also the case when files are deleted. Files are "moved" into the Recycle Bin, where they can be permanently deleted at a later time. The results, however, are the same as before. Should these files be permanently deleted by the user, only their entries on the disks table of contents are removed and not the data themselves.

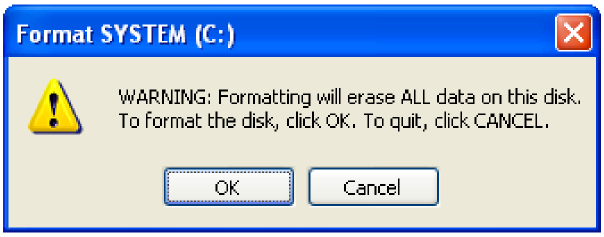
Using specialized data recovery software such as O&O DiskRecovery or O&O UnErase, data can easily be made visible and accessible again.

Such software can be extremely useful when files are mistakenly deleted. With only little disruption in productivity or game play, important files can be restored for use within seconds as if nothing had ever happened in the first place. On the other hand, this ease of restoration could also be misused by unauthorized individuals to access data on storage volumes disposed of, sold, or given away by their previous user.

The 2005 Study

Within the context of the issues described above, we asked ourselves whether or not the behavior of the PC user or salesperson concerning this data security issue has improved over the last year. Is more hard disk data securely deleted now than before or are private and corporate secrets still floating around on hard disks completely unbeknownst to their owners?

Figure 1: The warning message when formatting a partition in Windows is misleading.



Format SYSTEM (C:)

WARNING: Formatting will erase ALL data on this disk.
To format the disk, click OK. To quit, click CANCEL.

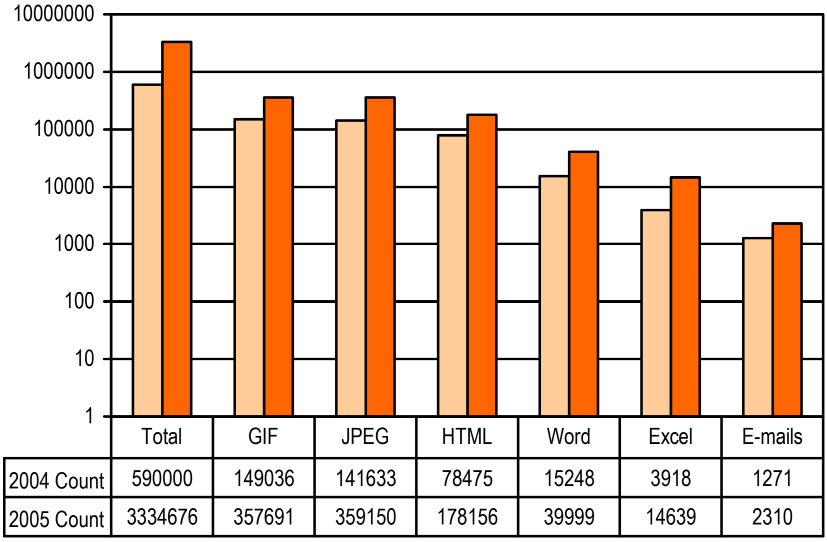
Figure 2: Types of Files Found

In total, over 3.3 million files could be restored (2004: 590,000). Once again, the majority of these files were graphic files or internet pages (GIF, JPEG, HTML).

About 40,000 Word and almost 15,000 Excel documents were readable (2004: around 15,000 and 4,000 respectively).

In addition, about 50 Outlook mailboxes (PST files) containing about 2,000 e-mails all together were found.

Note: The Y-axis of the graph is logarithmically distributed.



	Total	GIF	JPEG	HTML	Word	Excel	E-mails
2004 Count	590000	149036	141633	78475	15248	3918	1271
2005 Count	3334676	357691	359150	178156	39999	14639	2310

2005 Results

In order to tackle these questions, we analyzed 200 hard disks that were bought on eBay in the early months of 2005.

In total, 200 hard disks with a collective storage capacity of 3.18 Terabytes (3,255 GBytes) were purchased. The average volume of each of the hard disks was around 16.27 GBytes, more than triple that of last year's study, 5.26 GBytes.[2]

Of these 200 hard disks purchased, 42 were defective, around 21%. This shows a 6% increase from the results of the 2004 study. As was the case last year, defective hard disks were simply disqualified from this study due to the increased resources that would have been required to repair them. Because such resources would not be readily available to normal PC users, those defective test drives were not considered relevant to the study.

Of the 158 remaining hard disks, 45 were securely deleted, prohibiting any reconstruction of old data. This represents 28.5% of the functional drives tested this year and double the percentage from the 2004 study, 11.7%. Is this a sign that this problem is no longer a real issue among PC users? The answer: No, because the remaining 113 hard disks were either not erased at all or just simply formatted.

This means that 71.5% of the hard disks contained personal and company data that could be reconstructed - still a staggering trend.

In this study, over 3.3 million files were reconstructed, about 746 GBytes in total. Included in this restored data were 40,000 Word Documents and around 15,000 Excel Spreadsheets. In addition, around 50 complete e-mail mailboxes were found with their entire compliment of e-mail traffic from the previous users.

One of most dramatic findings of this year's study was a hard disk apparently from a federal governmental body in Germany. Along with reports and minutes of meetings, this disk

[2] In the 2004 study, 100 hard disks with a collective capacity of 526 GBytes were purchased and analyzed.

contained correspondences concerning details of a running external investigation. Also found was a hard disk once belonging to a prominent German bank that contained information regarding the credit ratings of other banks.

Getting to Know the Previous User

The results in 2004 were alarming. In 2005, the results show a small improvement. However, even when the percentage of recoverable data on the tested hard disks has fallen, the "quality" of recovered data has improved dramatically.

As more and more daily activities these days are completed with the aid of computers and the internet, recovering old data can give one an even clearer view into the daily lives of the previous user. From e-mail correspondences concerning a recent divorce, arrangements for a future rendezvous, or even "erotic" fantasies; from work or school evaluations to terminations of employment.

And that is just the beginning. With the age of digital photography now upon us, one can even find a photo of the unsuspecting previous user. From family photos from a previous year's holiday celebration or vacation photos to private bedroom images. And should that not be enough, there is always the chance of finding the previous user's complete collection of adult imagery.

In short, it is very easy to get to know the intimate details of a previous user by examining the data on their hard disk. Information that would take a private investigator days, perhaps weeks, to find out, we can purchase online for $40 and have it delivered to our doorstep. Shortly thereafter all intimate stories, anecdotes, e-mails, and photos can be reconstructed, easily and automatically.

Being a Government or Company Insider

It would be wrong to say that only the private users are careless with their data. The same trends occur among governmental institutions, non-governmental organizations, companies, and even banks.

Government Institutions

We were easily able to reconstruct data from a hard disk that, without a doubt, was once operated by a government institution. In addition to internal memos, data pertaining to specific persons were found. In particular, we were able to easily recognize legal correspondences between a government employee and a lawyer regarding a court procedure. Had those documents fallen into the wrong hands, the results might have been disastrous.

Travel Documents

Another hard disk came from a well-known German travel organization. Without much effort it was very easy to read up on the invoicing procedures and conditions specific practiced by travel agencies in their everyday operations.

In addition, it was possible to access documents concerning the termination of a specific client company's contract as well as the new contract with their replacement.

"Credit ratings, if you please?"

Another "Goldmine" came from the ever-sensitive world of finance. In the analysis of one of our test hard drives, our researchers stumbled across numerous internal documents from a major German bank. The information found included the credit ratings of other banks and this bank's potential for doing business with them. For competing banks, such data would be quite informative in developing new competitive strategies for success.

"The Onion Drive"

Especially interesting was a hard drive that had apparently been owned by three previous users before it came to us. All three users did not seem to be of any relation. This "onion" with all of its layers contained documents from a private detective who apparently had a series of correspondences with the authorities concerning his license. After "peeling" off more of its layers, documents were discovered containing information from a manufacturer of office machinery. Included in this information were customer information and billing calculations.

The most interesting documents of all, however, came from a former Managing Director of a middle-sized company, who meticulously recorded every aspect of his discord with the other managing directors. Starting with normal mail correspondences concerning blame all the way to accusations of fraud and embezzlement with specific shareholders, one could very easily trace the path of this company to probable bankruptcy. Other information included deferment agreements regarding payments and a list of potential providers of capital and their respective acquisition statuses. All in all, this was extremely explosive material that should never have made it out into the public domain.

How does this carelessness happen?

Lack of Knowledge

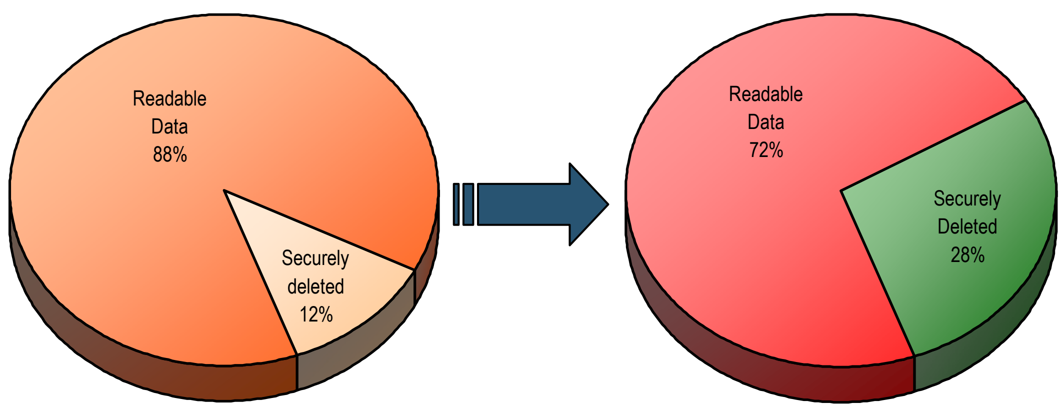
We assume that most users believe that their data are deleted when Windows confirms it to be. When formatting a partition in Windows, a warning message appears explaining that all data will be permanently deleted from the hard disk. This is misleading. Here it should be suggested that additional measures must be taken to ensure the data are securely destroyed.

It is also astounding to hear from Administrators working in companies who believe that, since all of their data are saved on a central server, no important data are stored locally. They believe, therefore, that special measures to securely delete data on client storage volumes are not necessary. Nothing could be more dangerous than such an assumption for a number of reasons.

For one, it is very difficult to confirm that no data are stored locally on client workstations. Such controlling efforts are resource intensive and, most of the time, never done completely, if at all.

Secondly, there is the temporary memory reserve known as Cache that is always created when surfing the internet. Even when data are

Figure 3: Percent Breakdown of Functional Test Hard Disks



In 2004, 88% of the functional test hard drives were readable (75% of all hard disks purchased). The percentage of defective hard disks in 2004 was 10% less than in 2005. In 2005, 72% of the hard disks were functional.

access on a company's intranet, a copy is made in the cache. Such data can be highly sensitive and must be dealt with securely.

Lastly, one must not overlook temporary files that are created when text, spreadsheets, or databases are edited on a client workstation. Through such files it is possible to reconstruct the original files at a later time.

Unwary Handling of Hard Disks

Another common problem is the lack of attentiveness on the part of the user concerning their hard disks. In today's age of service agreements and warrantees, most computers are returned to their corresponding retail outlets when an error occurs. When returning their merchandise, most users do not remove their hard disks.

Although it is usually not a conscious decision, such a step is placing private and sensitive data into unknown hands. What happens when an entire computer is defective? Perhaps "only" the hard disk has a problem and needs be replaced. What happens to all of the data stored on it? Sometimes these parts get "Recycled", as a report in the German TV magazine "EXTRA" on March 7, 2005 illustrated. In this report, an owner of a PC system was able to reconstruct old data on his allegedly "new" hard disk and contacted its previous owner using the data he found. It is, of course, possible that this is merely an exceptional case. Nevertheless, it should always be assumed that important data could fall into the wrong hands whenever a PC system is serviced.

Prevention

In the previous study we suggested different methods for avoiding the unauthorized reconstruction of deleted data [1]. From the encryption of the hard drive or the physical destruction of the hardware to the utilization of specific secure deletion software, there are many options from which to choose, all with their specific pros and cons.

Judging from the reader response to last year's study, however, we have come to the conclusion that the use of special software to securely delete sensitive data is by far the most common method in use. For this very reason, other methods will not be described in-depth in this paper. We invite interested readers desiring more information to take a look at the previous study containing more descriptions regarding the saving processes of hard disks.[1]

Formatting is not enough!

As was mentioned before, many PC users still believe that the formatting of a hard disk will mean the permanent deletion of its data. This is a false assumption because formatting simply removes a hard disk's table of contents, leaving all of the data intact.

In order to be on the safe side and ensure that old data are destroyed, the use of external software is required. This software should be able to delete data according to at least one of several internationally accepted deletion standards. Each of these standards corresponds to a specific method of overwriting data. Each method is executed in a differing order and repeated a number of times so as to make the restoration of data difficult or impossible. Using such software is generally very easy and efficient.

Standardized Processes for Secure Data Deletion

These processes include standards used by the US Department of Defense (DoD) and that of New Zealand-based expert, Peter Gutmann, who favors overwriting data up to 35 times to permanently annihilate them.[2][3]

The German Federal Office for Information Security (BSI) has also defined its own specific deletion standard [4]. Since not every software solution can use this standard, it is necessary to confirm this before making a purchase.[3]

Conclusions

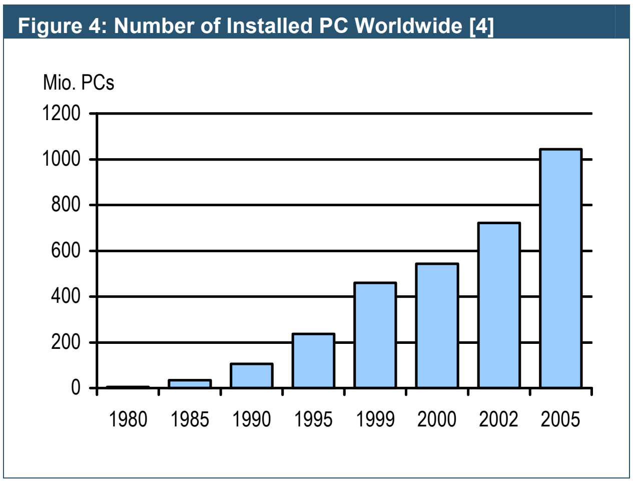
Humans strive to preserve their memories. Whenever possible, one tries to hold on to as many memories as possible, for as long as possible. Not so long ago, people would keep memories alive by simply keeping old letters, postcards, and photographs, or simply by continually retelling old stories to others. Today, maintaining memories is a very different matter. People write e-mails or text messages, create digital photos or videos, or simply write down their thoughts on their own internet-based "Blog" for the entire world to see.

These new technologies have brought with them a variety of new opportunities and many people today would be reluctant to do without such modern advances of information technology. This is, in itself, not at all the problem. One has to be, however, aware of the risks involved. Daily, stories appear in magazines, newspapers, and online describing recent virus attacks, Phising attacks[4], and Trojans, that attempt to gain access to user passwords. Software solutions do exist to combat such threats, although minimizing the risks of such attacks has a lot with user behavior as well. For example, users must be extremely attentive when using online banking access and providing personal data to an online source. In addition, the proper disposal of data storage devices at the end of their lifespan is just as important. This does not only includes hard disks, however, but diskettes, USB Sticks, memory cards, and all other data storage devices as well. As is the case with hard disk drives, their disposal should always include the secure deletion of all saved data they are carrying. What are unimportant data to one can be extremely useful to another; a situation with potentially disastrous consequences.

Summary and Comparison with 2004

According to Gartner, 189 million PCs were sold worldwide last year. This is an 11.8% increase from 2003 and 2005 promises to offer another increase from 2004.[6]

Figure 4: Number of Installed PC Worldwide [4]



The new PC systems bought today are most likely not the first for their users, but rather replacements for their current systems. Faster and more efficient storage technologies are constantly being developed and brought to the market, surpassing old storage standards by leaps and bounds.

Today, mobile telephones are able to save e-mails, contact information, photographs, and much more. Likewise, USB Sticks enable extraordinary storage capabilities with an extremely compact delivery system, easily able to store an entire encyclopedia on a keychain. It is also important to mention the omnipresent trend of MP3 Players

[3] O&O SafeErase V2.0 supports all aforementioned deletion methods as well as additional functions for securely deleting files, hard disks, and entire computers.

[4] "In computing, **phishing** is the act of attempting to fraudulently acquire sensitive information, such as passwords and credit card details, by masquerading as a trustworthy person with a real need for such information in a seemingly official electronic notification or message (most often an email, or an instant message)." [7]

that can store and play a lot more than just music these days. Everywhere one can find data storage devices. Sometimes these devices include personal data about the user; data sometimes unknown to user.

The steady fall in hardware costs in the IT branch is also enabling more people to purchase higher power systems to replace their existing PC systems. As often is the case, the data from the old system is quickly transferred from the old machine to the new so that the older machine can be sold or given to others.

Where the system actually "lands" is completely unknown. Another unknown is whether or not a user's old data can be misused by others. How do most people feel when strangers mull though their trash looking for letters, bank statements, bills, or alike?

The Dangers of Misuse

If someone gains and exploits the access data of another, entire identities of innocent individuals can be stolen. Such thieves can shop using other people's credit cards, auction items, or simply write e-mails posing as someone else. The consequences of such theft can be staggering, not the least of which being the time and resources it costs the victim to prove that wrongdoing did occur.

Companies and governmental agencies also need to improve their efforts to take control of this issue. As the study results show, sometimes sensitive data that should never get out, does and probably more often than one might believe. Although the secure deletion of data is an important element of some companies' data security policies, it appears that for many IT administrators this issue still remains either unknown or simply underestimated.

This is primarily a management issue because the failure to secure sensitive data means a failure to protect the liability of the customers and their stakeholders. The embarrassing publicity and damage lost data can deliver can quickly lead to a company's potential downfall.

While the misuse of private data can lead to user aggravation, the misuse or theft of company data can mean fatal damage to a business. The publishing of a company's internal data to the public can easily lead to civil or even criminal consequences. This year, we found data concerning specific individuals and specific clients, information that is certainly not public knowledge. Had this information made it to the public, it would have potentially meant irreparable damage to the company's reputation; a veritable field of opportunity to all sorts of people, not the least of which are this company's competitors.

Everyday another Secret

On eBay everyday, over 10,000 hard disks are auctioned off to the highest bidder. In examining the results of this study, it is safe to say that a large number of those hard disks contain information that does not belong in the hands of strangers.

Delete! Delete! Delete!

There is only one sure way to avoid the misuse of personal or private data from an old storage volume: Delete all data storage devices with specialized software before you let them leave you possession. Other than that, the only other option is to physically destroy the storage volume. Whatever option you choose, it is necessary to realize that the simple deletion in Windows is never enough! *[OK]*

Afterword

Acknowledgements

I would like to take this moment to thank my colleagues Frank Witter, André Weiß, Matthias Günther, and Fatihelyasin Erdas for their endless support in completing this study. Not only did they spend weeks purchasing hard disks online, but they were also invaluable in reconstructing the data found on the disks and recording the statistics for the study. In addition, I would like to thank my colleague Frank Alperstaedt for his constructive criticism while writing the first drafts of this study.

About the Author

Olaf Kehrer is Managing Director of the Berlin-based O&O Software GmbH, specializing in the secure deletion and recovery of data. His area of work includes the development of new technologies and products in the area of data security.

Products include O&O BlueCon, O&O DiskRecovery, O&O FormatRecovery, O&O UnErase, and O&O SafeErase. Aside from the deletion functions described in this study, these software solutions also enable Windows users to perform data restoration and repairs.

About O&O Software GmbH

O&O Software GmbH has been developing Tools for Windows since 1997. These products are now sold in more than 80 countries in multiple languages. Customers range from private users, small to medium-sized companies, and to public institutions and international corporations. The product portfolio includes applications for performance optimization, data restoration, and the secure deletion of data. O&O products are continuously tested and judged again and again as some of the most technologically advanced products on the market today.

Further information may be found on the internet or directly from us:

O&O Software GmbH
Am Borsigturm 48, 13507 Berlin, Germany

Internet: http://www.oo-software.com/
E-mail: info@oo-software.com

Tel: +49 (0)30 4303 43-00
Fax: +49 (0)30 4303 43-99

Bibliography

[1] OLAF KEHRER, O&O SOFTWARE GMBH, *"Data Data Everywhere",* April 2004; http://www.oo-software.com/en/study/

[2] DEPARTMENT OF DEFENSE, DEPARTMENT OF ENERGY, NUCLEAR REGULATORY COMMISSION, CENTRAL INTELLIGENCE AGENCY, *"National Industrial Security Program Operating Manual",* 1995, 1997, 2001; http://www.dss.mil/isec/nispom.htm

[3] PETER GUTMANN, *"Secure Deletion of Data from Magnetic and Solid-State Memory"*, Usenix Assoc., 1996; http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html

[4] GERMAN FEDERAL OFFICE FOR INFORMATION SECURITY, *"IT Baseline Protection Manual"*, BSI, 2003; http://www.bsi.bund.de/english/gshb/manual/index.htm

[5] EGIL JULIUSSEN, PH.D., "COMPUTERS-IN-USE FORECAST", eTForecasts, June 2000, http://www.etforecasts.com/products/ES_cinuse.htm

[6] NETZZEITUNG, *"PC-Absatz wuchs 2004 zweistellig"*, January 19, 2005; http://www.netzeitung.de/spezial/globalvillage/321693.html

[7] WIKIPEDIA, *"Phising"*; http://en.wikipedia.org/wiki/Phising