Analysis of new threats to online banking authentication schemes

Oscar Delgado¹, A. Fúster-Sabater¹, and J.M. Sierra²

Abstract—In this work it is shown how current online banking and payment methods are highly insecure. All the countermeasures currently used by financial institutions are vulnerable to some kind of attack, mainly to the recent and powerful last generation trojans, which are able to defeat mechanisms as strong as SSL sessions or two-factor authentication schemes. We point out the problems of these models and we propose a mobile phone OTP scheme with a new method for the challenge generation.

Index Terms—Online banking, authentication, trojan horses.

I. INTRODUCTION

WITH the arrival of traditional mafias to Internet, fraud has been professionalized and the old reconnaissance desire has become into the possibility of making much money. For example, in Spain, according to the periodic polls about Internet use, the concern to fraud is usually the first reason of distrust about e-commerce [8]. In fact, this is not surprising, because according to these reports, 72% of the analyzed computers was infected by some type of malware, half of it classified as “dangerous”, either for the computer integrity or for the privacy of its data. Other interesting conclusion, that will be analyzed in detail in section III-B, makes reference to the inability of antivirus software, due to different reasons, for reliable detecting and neutralizing these kind of threat.

Clearly, the reason of this growth is the online fraud evolution. For example, phishing is one of the most sophisticated, organized and innovative technological crime waves ever faced by online business. The term was coined in 1996 by crackers attempting to steal AOL account credentials. Since the beginning of 2004, phishing has remained the most widespread online fraud technique, but in the last years this has changed. Fraudsters have new tools at their disposal and are able to adapt more rapidly than ever. These new tools are mostly malware in general and trojan horses in particular.

The rest of the paper is organized as follows. Most recent and important threats to the e-commerce development, including phishing, pharming and last generation trojans, are analyzed in sections 2 and 3. Section 4 analyzes the relation between malware and cryptography, and some misconceptions about it. Section 5 presents and analyzes the current online-banking authentication schemes and their related problems. A new method for generating mobile phone OTP tokens and other solutions are introduced in section 6. Finally, conclusions are presented in section 7.

II. RECENT THREATS TO ONLINE BANKING

Some of the most recent threats to the development of e-commerce include phishing and pharming attacks and, overall, last generation trojan horses. These are supplied with new and powerful capabilities, which make them able to even defeat the most sophisticated authentication protocols such as two-factor schemes.

A. Phishing

Phishing has probably been the most important and technically complex wave of organized crime that e-commerce fraud has faced in recent years. Although the first cases of phishing were reported in 1996 against AOL user passwords, the first financial institution, eGold, was not attacked till 2001.

Since then, the number of reported cases has grown, with annual increases of nearly 300%, till reach, for example, the 1,184 reported cases in Spain in 2006. Roughly speaking, the 85% of them was directed against financial institutions, although a small tendency to drop can be appreciated due mainly to the improvement in the management attacks and fraudulent sites shutdowns. The new objectives include, among others, ISP’s, NGO’s and even government institutions, such as the Spanish National Statistics Institute.

B. Pharming

Given the limited success of these efforts, fraudsters began to use other methods, such as pharming. Pharming is an attack to the name resolution system of a operating system, in such a way that this provides a false IP address for a specific domain. There is even a more sophisticated and dangerous attack, whose objectives are name servers instead of the operating system. In this case, the attack is often called DNS poisoning and it affects to all customers who use the attacked name server.

This attack becomes particularly dangerous if it is performed by, for example, a trojan. In this case, the website of any banking institution can be redirected to a fraudulent domain, where the fraudster previously placed a copy of the original site.

In some curious cases, the phishers went as far as to even buy a digital certificate to a recognized certification authority
in order to be able to establish copies of some bank web pages. In this way, the page could be served even under HTTPS, without the browser warning message appearing to the user. On the other hand, this message is almost unintelligible for most users, so that, in practice, it is not very useful.

III. LAST GENERATION TROJANS

Traditional methods used by phishers, as massive spam, similar domains\(^1\) and other somewhat crude tricks, have been gradually losing effectiveness, as the problem has reached the media and is already known by the general public. As a result, criminals have been forced to seek new ways to carry out their frauds.

In this sense, and due to their powerful features, trojan horses are the perfect tool. Although they never were the most popular code among malware creators, with the recent fraud professionalization and the resulting change in the objectives pursued, their use has greatly increased: since 2000 the amount of malware created every year has grown 25,000%. On the other hand, the type of malware most frequently used by cyber-crooks has also changed over the years. Whereas in 2000 viruses accounted for 81% of all new malware detected, in 2007 this figure barely reached 5%. In the same period, trojans have increased from 14 to more than 90%, reaching the amazing volume of 15,292 new malicious programs each month. Table I summarizes these data.

This change is clearly driven by the new objectives of malware creators. Viruses were closely related to massive epidemics and the quest for notoriety, while trojans are the weapon of choice for those looking for profit. This is clearly indicated by the fact that almost 70% of the new written trojans are specifically designed for capturing online banking credentials.

Figure 2 shows detailed data year-by-year and confirms this tendency. Data fit surprisingly, even suspiciously, well to an exponential behavior. Despite the fact that there are clear suspicions that security vendors over-report the volume of incidents and malware [2], these figures can be useful for getting a general idea and confirm, in any case, the undeniable growth in the malware creation.

A. Infection vectors

Other important issue with malware is the so called infection vector. The infection vector is just the way these programs infect the host computer. The great majority of this kind of malware does not have autoreplication capability, like worms, so they must use other ways to propagate.

\(^1\)For example, registering a domain like www.foolbank.com instead of www.foolbank.com, trying to cheat the user.
sites, if is possible with a great number of visits. In order to do so, they usually use automatic tools, that are able to collect in a short period of time hundreds of infected machines. Other surprising and worrying method, recently discovered, uses the so called affiliate marketing programs, which pay web site operators to install malware on its visitor’s computers using exploits. In return, the webmaster receives a commission ranging from $0.08 to $0.50 per infection [2], [5].

B. (In)efficiency of antivirus software

Other reason for the spectacular growth of malware is the inability of the antivirus software to successfully detect it. It is true that this is not an easy task, because of trojans are increasingly complex programs, that include many auto-protection and stealth techniques [4].

But the main difficulty for detecting this new kind of malware is its very high mutation rate, up to six in the same day. In order to do so, attackers use automatic tools, that allow to easily and quickly generate these mutations. The changes are minimal, often without changes in the behavior of the trojan, but differently enough to avoid the antivirus detection.

All these reasons mean that antivirus software, even of high quality and often updated, is no longer a guarantee of security against this new kind of threat.

C. Malware economics

A long time ago, creation of malware stopped being an exclusive activity for teenagers freaks. According to the antivirus companies, 75% of malware detected in the last quarterly of the last year were designed to obtain money fraudulently. Nowadays this is a very lucrative business, which moves many million dollars around the world, and that has gone progressively being occupied by the traditional organized crime, which can easily use its established networks for money laundering.

Specialization has reached online fraud too, in such a way that the chain of the fraud process has fragmented in parts clearly differentiated. The first link of the chain is composed by the coders of malware: people with technical knowledge, who sell their creations using all type of marketing strategies, e.g. in specialized forums, where this type of malware is sold, it is easy to find special offers, discounts by volume or warranties of that the trojans are not detected by any antivirus software.

Once the buyer has chosen the product, he/she contacts the seller, often using the ICQ protocol, and a final price is negotiated. Eventually the transaction is made, usually through WebMoney, the new preferred system of electronic payment for the fraudsters. In fact, often the prices are directly announced in WMZ, the official currency of WebMoney, equivalent to 1 dollar USA.

D. Last Trojan Techniques: HTML Injection

This is one of the state-of-the-art techniques used by the last generation of malware, specifically by an increasingly number of banking trojans. It consists in modifying the HTML code of the attacked web page (typically, the login page of an online bank), before this is presented to user. Figure 4 shows a real attack to a well-known American bank, and how the page’s original look and HTML code are modified to trick the user.

 Till the present, this technique has been almost exclusively implemented by using the Browser Helper Object (BHO) interface, defined by Microsoft for Internet Explorer in 1997. Therefore, the first cases for other browsers, as Firefox, have been recently discovered. The original goal of the interface is, of course, legitimate and a large number of known applications, like the PDF reader plug-in or the Google toolbar, make use of it. On the other hand, it has a very powerful side effect: it provides access to the traffic in the clear of the browser, before and after it is encrypted by, for example, an SSL session.

As it can be easily seen, if the purpose of the program that uses the interface is fraudulent, the consequences are potentially very dangerous. The main reason is that this attack is almost undetectable for the majority of non-expert users because it does not break the usual protection measures proposed by the financial institutions:

- The ‘lock’ in the browser window is closed, that is, the encryption has not been defeated (but the attacker has access to all communications in the clear).
The source URL is correct. The user has not been redirected to the source forged. The origin of the data has not been manipulated, just its presentation to the user. Despite the fact that this is a vulnerability that mainly affects Internet Explorer, its scope is broader. In fact, it has nothing to do with the old and useless controversy about the intrinsic security of each browser. It seems logic that if Internet Explorer is the most attacked browser is due to the fact that it is the most widely used and, therefore, the most attractive to attackers.

The core of the matter is that, taking into account that this attack manipulates date in the application layer, a SSL session, that operates protecting only lower levels, is not a security warranty against this kind of threats. Therefore, it can be said that a browser is not a secure platform to manage sensitive information anymore.

As a result of this fact, most of the current banking trojans are written using this interface, because they are able to easily bypass the current countermeasures taken by online banking, as virtual keyboards or, even, the encryption of the communication. As we will see in next section, this malware can even defeat more sophisticated authentication schemes, like two-factor, by performing a real-time man-in-the-middle attack.

IV. Malware and Cryptography

As it is known, the security of a system can be seen as a chain of protection methods, and the global level is just as strong as its weakest link. From this point of view, cryptography is no more than another component.

 Fortunately, most of the times it is the strongest link, so cryptography is usually considered like the philosopher’s stone, that could provide the definitive protection against all the threats to electronic commerce.

In theory, this could be probably so, but cryptography now on the market does not provide the level of security it advertises. Most systems are not designed and implemented in concert with cryptographers, but by engineers who think of cryptography just as another component.

A. What cryptography can and cannot do

In any case, practical cryptography is rarely broken through the mathematics; other parts of systems are much easier to break. The best protocol ever invented can fall to an easy attack if no one pays attention to the more complex and subtle implementation issues. Netscape’s bug in the random-number generator is an example [7]. Flaws can be anywhere: the threat model, the system design, the software or hardware implementation, and the system management. On the other hand, fatal bugs may be far from the security portion of the software too; a design decision that has nothing to do with security can nonetheless create a security flaw.

In addition to implementation issues, there are also consequences when cryptographic protocols are not designed by experienced cryptologists. The well-known flaw in the WEP protocol could be, maybe, the best example [6]. Any cryptographer knows the birthday paradox and, thus, that a 24-bit initialization vector (IV) would be completely insufficient, because, if randomly chosen, it would have 50% chances of collisions after less of 5000 encryptions.

In any case, there are only 16 million IV values and, in heavy-loaded networks, these could run out in a few minutes. When no implementation or design error has been committed, and the cryptographic protocols have been carefully planned, you still have to face malware and the type of trojans we have analyzed in the previous sections.

We think this new class of malware, with so powerful characteristics, is not only a perfect tool for fraudsters, but also probably the most serious threat to the development of the electronic commerce. If these trojans are able, as we are seeing, to easily bypass every cryptographic protection, then these lose all their effectiveness. It is important to make notice that these problems had nothing to do with mathematics or design flaws, but to implementation and, overall, the environment where these protocols are carried out, as we will analyze in the next section.

For example, a SSL channel is no longer a protection against this last-generation trojans. This, of course, does not mean that SSL protocol has been broken or it is not necessary anymore. Due to the place where the fraudsters are hooked the trojans, they have access to all the traffic interchanged between the browser and the web server in the clear. So if you are infected, all the classical protection measures do not offer protection any longer.

V. Current Online-Banking Authentication Schemes

Maybe the most representative example of the relation between malware and cryptography are the authentication mechanisms. In addition to the fact that a PC is an inherently insecure environment, the situation gets worse for online banking. The reason is that schemes of authentication provided by the banking organizations are not balanced, in the sense that the authentication process is performed exclusively at one end of the architecture: the user’s computer.

Furthermore, this process is still carried out in many cases using weak schemes, like the traditional user/password. Even with two-factor schemes all the process takes place, exclusively, in the user side of the equation. Due to this, if this authentication process is hacked, by using a trojan or a simple credentials theft, the bank has no way to distinguish between legitimate and illegitimate clients.

A. Virtual keyboards

The first cases of this type of trojans used old viral techniques, as capturing keyboard keystrokes, looking for words such as “password”. To fight them, banks began using virtual keyboards, so common today, although they are completely useless against some types of trojans. Figure 5 shows a real example from a Spanish bank.

In any case, just as expected the next movement of criminals came promptly. They provided trojans with the ability to make screenshots and save small video clips. After that,
next generation trojans, which are situated inside the browser itself, completely defeated these countermeasures and, at the moment, they should be considered completely useless.

**B. Two-factor authentication**

The most prevalent approach to reducing the impact of data compromise is known as two-factor authentication. This refers to requiring proof of two out of the known three criteria (something you are, you have or you know) to permit a transaction to occur.

Second-factor authentication has been implemented in several ways in online banking, ranging from matrix cards or hardware tokens, to the use of mobile phones. Table II summarizes currently used two-factor authentication schemes, discussing advantages and disadvantages of each approach.

It is important to remark that almost all of these schemes can be easily defeated by last generation trojans. Till the present, just one variation on the mobile phone scheme could be considered secure against this threat. But, unfortunately, even this proposal has some drawbacks too, which will be analyzed in detail in section V-D.

**C. Attacking current two-factor schemes**

In this section it is shown how a last generation trojan can defeat a two-factor authentication scheme. The attack could be carried out as follows:

1) The user has been infected by a banker trojan through one of the ways shown in section III-A. Victim is user of a online bank that uses some two-factor authentication scheme, for example, a matrix card. Strong authentication process can be performed at login time or when a transfer is requested. The attack works in any case.

2) Trojan waits the user initiates a transfer process. At that moment, the user will have to provide the bank with the next valid one-time password from his matrix card. The OTP is then captured, with other sensitive data, as user’s name, login password or account number.

3) Trojan silently changes the destination number account to other of its election and allows the transfer process to continue. The amount to transfer will obviously be changed too. At the same time, the trojan shows the user the correct account number. If this attack is well implemented, it is almost undetectable.

This attack, with slightly variations for adapting it to every environment, could be performed against any two-factor authentication carried out in the user’s computer exclusively.

The cause of the problem is not in the authentication scheme, which is perfectly secure in theory, but in the fact that it is performed in a inherently insecure environment: a personal computer with a general purpose operating system.

It is often forgotten that cryptographic protocols cannot be isolated of their environment. In the real world, operating systems have not been created with security as a crucial design criteria and, therefore, have critical vulnerabilities.

As a new example of this problem we can cite the new developed techniques for recovering private keys from a process memory [10]. Again, the problem does not stem in the public key paradigm, but in storing, even temporarily, a private key in memory without other protection.

1) A real example: Silentbanker: The scale and sophistication of this recently discovered banking trojan is worrying. It is able, for example, to targeting over 400 banks and having the ability to circumvent two-factor authentication [11].

The ability of this trojan to perform man-in-the-middle attacks on valid transactions is what is most worrying. The trojan can intercept transactions that require two-factor authentication and it can then silently change the user-entered destination bank account details to the attacker’s account details instead.

Since the user does not notice anything wrong with the transaction, they will enter the second authentication password, in effect handing over its money to the attackers. The trojan intercepts, as we have been in previous section, all of this traffic before it is encrypted, so even if the transaction takes place over SSL the attack is still valid.

**D. Mobile phones**

Recently, some banks have started to use mobile phones as authentication tokens. The goal is to defeat the classic man-in-the-middle attack carried out by a last generation trojan, since authentication token travels along an independent path – an SMS mobile phone network.

Since the communication now uses two channels, this method is also known as a two-channel authentication or out-of-band mechanism. Although majority of banks have implemented this solution by using mobile phones, it can be used with any telephone, not just mobile devices.

On the other hand, the banks apply this mechanism in different stages of the online banking process. Some of them, use the mechanism at log-in, and others just when a money transfer is requested by the user.

Roughly speaking, the scheme is simple and it can be described as follows:

- The user requests log-in or a money transfer.
- The banks sends a SMS to the user’s mobile phone with a one time password, usually in the range of 4 to 8 digits.
- The customer must then manually copy the code from their phone to confirm the transaction.

The advantages of this scheme are clear and it is supposed effectively a more secure solution. It is attractive for both the bank and the customer too, given that there is no additional hardware token to issue or carry. On the other hand, it has some drawbacks:

- Mobile phone cannot have coverage or run out of battery. It can be forgotten or lost by the user. In some situations, the user cannot have mobile phone at all.
<table>
<thead>
<tr>
<th>Scheme</th>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Something you are | Keystroke dynamics | • No extra hardware needed  
• High user’s accept-ability | • Not reliable.  
• High enrollment time |
| Something you know | Transfer password | Simple to use | Can be easily captured as login credentials. |
| Something you have | Matrix card | Cheap and relatively simple to use | • Can be lost or forgotten.  
• Difficult to share |
| Hardware tokens | Difficult to be tampered | | • Expensive  
• Can be lost or forgotten  
• Difficult to share |
| Mobile phone | No extra hardware needed. | | • No coverage or run out of battery  
• Cost for bank |

TABLE II
TWO-FACTOR AUTHENTICATION SCHEMES USED IN ONLINE BANKING

- The SMS delivery has a considerable cost for the bank, which increases with the number of users and performed transactions. In turn, this cost is usually charged to the user.
- Some accounts, as a corporate account, need to be managed by several persons, probably since different physical locations. In this case, this scheme is extremely difficult to be used.

Other important issue, which is necessary to have into account, is that information confirmation must include details of the transaction itself. Otherwise an attacker could perform a session hijacking attack and alter a transaction that the user would confirm. Real trojans which already exploit this fact were shown in section V-C1. In next section, a proposal to securely include this information is presented and analyzed.

On the other hand, even if details are included in the SMS, some authors have pointed out other interesting issues that can still be applied. In [1] authors performed a study that finds that one out of five online transactions were vulnerable to attack despite the adoption of SMS passwords (more due to human error than to a technical security problem). To simulate a trojan MITM attack, the bank account number in the SMS message was not the same as the intended account number and some customers failed to notice this.

The research simulated an online bank and it was used by more than 90 participants to undertake more than 700 financial transactions using an SMS authorization code. Two types of attacks were simulated - an obvious attack that involved altering five or more digits in the account number, and a stealth attack altering only one digit. The study showed that 21 per cent of obvious attacks were successful, and 61 per cent of stealth attacks succeeded.

It is clear that security includes usability, not just the technical aspects of security and that banks must improve it in their services. Therefore, a user interface should be used in such a way that the user actively selects a transaction to confirm. In any case, transaction analysis and confirmation, when well implemented, can be effective and possibly the best current mitigation countermeasure available.

VI. FUTURE SOLUTIONS

In this section some solutions to these problems are outlined and a new method for generating OTP tokens using mobile phones is presented and analyzed. Sections VI-B and VI-C analyze solutions already found in the literature, that we believe to be a effective way of dealing with the problems of online banking authentication.

A. Phone OTP Token

As shown in previous section, it is important that information confirmation sent to the user includes details of the transaction itself, to avoid a real-time MITM attack. Together with this information, an OTP challenge is sent in the SMS message. This challenge has to be correctly repeated, trough the Web channel, to successfully confirm the operation.

This challenge can be generated in different ways. Many banks use just a random number, but this approach has several disadvantages. The main is that it is necessary to store information (the issued challenge, the account number, the amount to transfer, etc...) for each requested operation and every user. This scheme can be also prone to denial-of-service attacks.

In this situation, it would be more adequate a state-less protocol. Therefore, our goal is twofold:
- To obtain a state-less protocol, which makes unnecessary to store information for each user and provides independently verifiable challenges.
- To obtain a time-varying scheme, i.e., it should provide different OTP values for the same user at different moments.

According to our proposal, the complete SMS message would be composed of:
- Critical information to be protected, as account number, amount to transfer and time of transaction. Optionally,
other information could be included, as the name of the destination bank.

- OTP token, generated by performing a HMAC over the previous critical information with a server secret key, $K$. For convenience reasons, the full output (e.g. of 160 bits) can be reduced to only 24 bits, 6 hexadecimal characters, which are user-friendly and large enough to defeat brute force attacks. The process is shown in figure 6.

In such a way, the OTP value is cryptographically binded to the financial information. In addition, this scheme allows the server not to store per-user information. This reduces the server load and avoids programming errors that can lead to the compromise of information of other users [12].

B. Logical Countermeasures

We claim than any method that carries out the authentication process entirely at the client side is not completely safe. At this point, it seems clear that the only actors who can really mitigate the problem until marginal levels are the own banking organizations. This should be achieved through detection approaches that, in any case, the banks have been using for years for other types of frauds, such as the credit card one.

One approach to reducing the risk is to concentrate on online transactions that may be fraudulent. This is analogous to the risk management measures that banks take in the physical world: every credit card transaction is evaluated, and suspicious transactions are checked with the customer.

An analysis of online transactions may be performed by using a variety of metrics, such as the user’s IP address, the presence of authentication information such as a cookie on the user’s machine, the amount of a transaction, the destination bank account, characteristics of the destination bank account, and cross-account analysis of transactional patterns.

Other non-technical methods may include the capacity to recover the transactions in a fast and effective way, although have spent several days, or even months, since they were carried out. In this sense, there are positive examples, like PayPal, which considers the payments made by individuals provisional for 180 days [3].

C. Dealing with Financial Benefit

Financial institutions have stated delays in certain types of money transfers to allow the detection of accounts used in a fraudulent way. If a fraudulent receiving account is identified during the holding period, the transfer can be voided and the economic damage can be avoided. Of course, this is useful to law enforcement too. Often, attackers can be caught by tracing the flow of money from the use of stolen credentials.

Real world examples include Barclays Bank, HBOS, NatWest, and Royal Bank of Scotland, which have instituted delays of between several hours and one day on online transfers between two accounts within the same bank. The delays, which apply the first time a transfer is attempted, are intended to give the banks time to detect suspicious activity, such as a large number of transfers from multiple accounts into a single account.

The money-transfer delay was adopted in response to a wave of phishing incidents in which thieves transferred funds from victims’ bank accounts into accounts owned by “mules”—people duped through e-mail solicitations into opening accounts, usually under the guise of a business proposal. From the mule accounts, the thieves withdraw cash, open credit cards, or empty the account.

VII. Conclusions

In this work, we prove that current online banking is highly insecure. All the currently used countermeasures implemented by banks are vulnerable to some kind of attack, mainly to the recent and powerful last generation trojans. A mobile phone OTP scheme with a new proposal for the challenge generation is presented.

In any case, the security of a system can be seen as a chain of protection methods, and the global level is just as strong as its weakest link. From this point of view, cryptography is no more than other link in the chain. Most of the times, it is probably the strongest link, but there are many others weak links which decrease the security level until unacceptable level for carrying out electronic commerce. In last term, exclusively technical protection methods seem not to be able to completely solve the situation. The focus should be changed towards trying to increase the resistance of the system, instead of its invulnerability that, in any case, cannot be completely achieved in general purpose operating systems.

In the financial world there exist tools, like user behavior profiling, which are starting to be considered, and that in fact, have been used by the banks for many years in the “real world” in form of back-end controls. From the dot com boom, banks have tried to push the entire business logic, authentication included, to the front end (the user’s computer), as it is easier for them. Possibility of ubiquitous transactions and distinction between ‘safe’ and ‘dangerous’ ones vanished too. All of these get worse the cases of user’s identity or credential theft, which need more advanced measures and, curiously, a return of traditional back-end controls should be expected.

References

[10] Tobias Klein. All your private keys are belong to us, 2006.