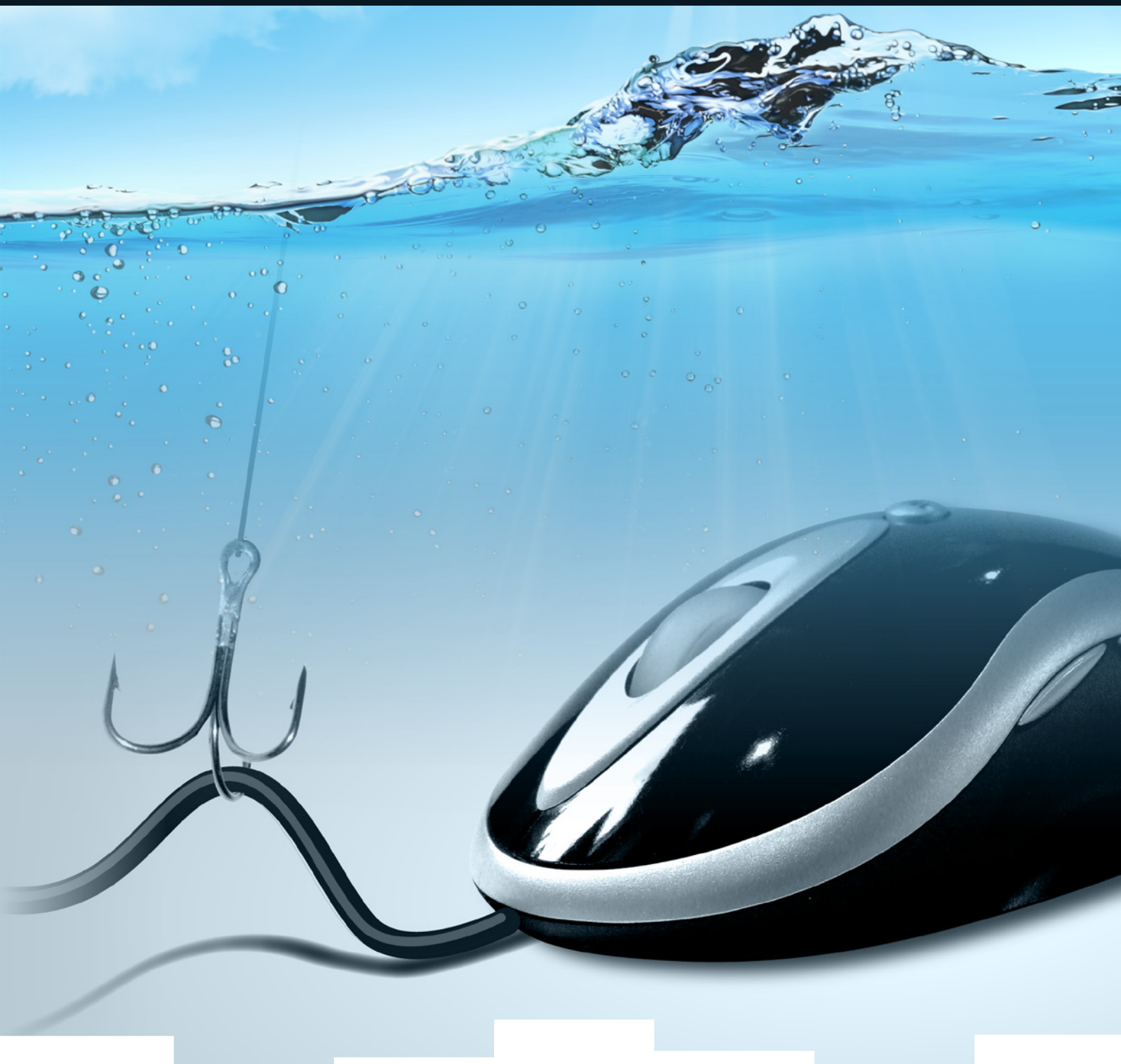
technicolor



THE SECURITY NEWSLETTER #19

SUMMER 2011

THE SECURITY NEWSLETTER #19

In this Issue

Published Quarterly By
Technicolor Security & Content Protection Laboratories
Part of Office of the CTO

Technical Editor: Eric Diehl

Editors: Sharon Ayalde

Contributors: Michael Arnold
Patrice Auffret
Peter Baum
Marc Joye
Mohamed Karroumi
Michel Morvan

EVP: Gary Donnan

Subscribe to the newsletter:
security.newsletter(at)technicolor.com

Report vulnerability:
security(at)technicolor.com

EDITORIAL

Since last issue, what a dreadful quarter! I do not remember a quarter with so many publicized attacks. This time, analyzing the associated reactions was interesting. Obviously, the attack that got the most airplay was the hack of Sony's game network. For several weeks, it affected gamers who could not access their accounts. Interestingly, when reading the forums, the gamers were more worried to learn when they will get again access to the network rather than which personal data may have been stolen. As the media coverage was huge (despite a lack of technical information), we will not address this attack, but we will rather focus on another less covered exploit: the leak of Sony's PS3 private root key used to sign firmware.

High profile hackers were extremely active and successful. Hackers stole sensitive data inside RSA servers. Although not confirmed by RSA, the security community anticipated that this stolen data may reduce the security of RSA's SecureID system. A few weeks later, the Lockheed Martin company announced that crackers used this information in a very sophisticated attack to try to penetrate its network. RSA announced that it will replace the tokens for business owners who use them to protect sensitive data.

This quarter, a new pirate "team" appeared: LulzSec. It penetrated many governmental sites (CIA, Senate), game sites, (Sony, Eve Online...) and banks, and published a large database of passwords they collected. They have a public presence. Their twitter account (@LulzSec) has more than 280,000 followers! LulzSec's objectives seemed to come back towards the old objectives of reputation. After fifty days of intense activity, LulzSec announced the end of its journey.

Apple also had an interesting quarter. First, we discovered that iOS™ traced and logged all your locations for about a one year period. Apple quickly issued a patch removing this indiscreet feature. The first serious malware, MacDefender, infected Mac OS X. A beta version of iOS5 was jail-broken in less than one day... Interestingly, with iTunes Match, Apple proposed a promising answer to audio piracy. For an annual subscription of $24.99, after an initial scanning of your hard drive, iCloud™ gives you legal access to every audio track present on your hard drive for eight devices, regardless of if it was legitimately acquired or not. This feature seems to be a first instantiation of the concept of digital locker.

E. DIEHL
Technical Editor

THE SECURITY
NEWSLETTER
#19

BE OUR GUEST

Helena Handschuh

Helena, you are working at Intrinsic-ID, a spin-off of Philips in the field of security. How did you get interested in this field?

I had discovered security as a grad-student in a thrilling course on cryptography and security. This led me to pursue a PhD while at Gemplus at that time. There, I was lucky enough to work with an extraordinary team of extremely competent people, which was very exciting and motivating. I never left the field since then, working on many challenging topics, with excellent teams, for different companies.

Can you tell us more about your current position?

Two years ago, I joined Intrinsic-ID, a company commercializing security products and new concepts around physically unclonable functions or PUFs. These are functions and features found in objects that uniquely identify them; the existence of those functions allows to protect these objects against counterfeiting and cloning.

Can you give some examples of PUFs?

The very early examples of PUFs were based on optics. Later were introduced so-called silicon PUFs. The exact speed a specific ring oscillator is oscillating at on a given device is unique. Measuring speed variations thus allows distinguishing among devices. Furthermore, there is no way to manufacture a device that will behave in a predictable way. Around 2007, new types of PUFs were invented at Intrinsic-ID including the ones we are focusing on in our products: volatile memory-based PUFs. SRAMs for instance offer a very stable source for producing unique identifiers.

What are your current interests with PUFs?

I have several topics of interest. One of them is to find ways to use PUFs in products on an industrial scale. Even though this is still challenging for most silicon PUFs (a.k.a. delay PUFs), for SRAM-based PUFs, the picture is completely different. As they rely only on standard design components, such PUFs are easily integrated into commercial devices. The aim now is to make them ever more efficient and cost-effective for mass-scale industrial applications. Besides that, we are also involved in European project UNIQUE. For this project, we are building a test-chip containing several existing but also newer types of PUFs (SRAM of course, flip-flops, latches, delay PUFs, ...) to study their adequacy for different use-cases.

Can you give an illustration?

Obviously, PUFs can be used to detect counterfeiting and cloning, but there is more. For example, at Intrinsic-ID, we propose to use PUFs to bind cryptographic keys to a device for digital asset protection. Actually, unique per-device keys are reconstructed by querying the unique start-up values of SRAM PUFs at power on. Keys are thus never stored on chip, but always available.

Thank you!

H. HANDSCHUH (CTO, Intrinsic ID)
Interview by M. JOYE

THE SECURITY
NEWSLETTER
#19

THE NEWS

Android Malware: Suspect Your Applications

The number of free Android applications exceeds the number of free iPhone ones. The total number of Android applications will probably get close to Apple score by the end of this year. This success comes with a drawback: the amount of malware contained in Android applications has drastically increased

Malware is mostly contained in malicious applications, but may also appear in custom Chinese firmware (`jSMSHider`), or even hidden within in-app advertisements (`GGTracker`). For instance, "Droid Dream Light" malware affected up to 26 apps targeting users' fields of interest: "Sexy Legs", "HOT Girls 4", "Beauty Breasts" but also "System Monitor", "Quick Uninstaller" or "Super Photo Enhance".

Malicious apps contain code that can steal the phone's unique ID, model number, bookmarks, and other private information but also install additional applications, email phishing scams, accept Trojans sent via SMS, or hijack Wi-Fi connections. They mostly exploit weaknesses in the implementation or in access control as well as naïve acceptance of service access by users. Antivirus or antimalware applications might help to detect some of them, but we recommend you to become suspicious when loading a new application on an Android smartphone.

June's malwares like `jSMSHider` , `GGTracker`, Plankton, or `YZHCSMS`, show that attacks are becoming more and more intrusive. `YZHCSMS`, spam users' phones with premium-rate text messages every 50 minutes without notifying users except on their bill...

M. MORVAN

TMG Affair: or Why You Should Have a Good Security Assessment Policy

The French company TMG (Trident Media Guard) recently suffered [1] from a breach leading to confidential data leak. TMG is the private company mandated by HADOPI to discover pirate users sharing copyrighted content. HADOPI is the legal entity which implements the French graduated response. TMG quickly announced [2] that the breach was only on a test server with fake data.

We will go in more technical details about what happened.

A simple directory listing issue has been found by Cult of the Dead HADOPI team and reported in hackerish style on the full-disclosure mailing list [3]. A directory listing is possible when you browse a Web site directory whose server has been badly configured. In that case, it shows every existing file.

By deeply analyzing the content of the vulnerable directory, the hacker team discovered much interesting data like a hard-coded password in a script, a list of content identified by their torrent hash, and last but not least: some infringer's IP addresses.

Be it a test server or not, the vulnerable Web site would have been easily detected by any quick security assessment. If this is truly a test server, the question is why is it accessible directly from the Internet? And if this test server is scanning valid torrent trackers on the Internet, it is not really a test server anymore.

Following this problem, HADOPI suspended [4] its connection (the link used to send infringing IP addresses to HADOPI) with TMG and asked for a security assessment of their processes and infrastructure. Some people could say it is never too late, but acting proactively in such a mission would have been a more responsible behavior.

P. AUFFRET

Cracking audio Captcha

Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHAs) are used at many Internet sites to prevent malicious scripts to send or post spam. A CAPTCHA is a test which should be easy to solve for a human, but difficult or impossible to solve for a computer.

The most popular CAPTCHAs involve deciphering distorted text on a noisy background. To allow visually impaired persons to participate to Internet sites, audio CAPTCHAs are used where either spelled characters and digits or continuous speech have to be recognized. Researchers claim to have broken, with the help of audio processing and machine learning techniques; the first type of these audio CAPTCHAs with a correct detection of for example 50% for Microsoft's system [5].

1 bluetouff, "Le honeypot de TMG," relets.info, May 13, 2011, http://reflets.info/le-honeypot-de-tmg/.
2 Jacques Franc de Ferrière, "Hadopi : TMG admet « une fuite de donnnées sur un serveur test »," ITespresso.fr, May 16, 2011, http://www.itespresso.fr/hadopi-tmg-admet-une-fuite-de-donnees-sur-un-serveur-test-42818.html.
3 "Full Disclosure: Too Many Gremlins for Trident MediaGuard (HADOPI)," Seclists.org, May 20, 2011, http://seclists.org/fulldisclosure/2011/May/434.
4 Eric Walter, "URGENT : l'Hadopi suspend ses liens avec TMG," Numerama, May 16, 2011, http://www.numerama.com/magazine/18803-urgent-l-hadopi-suspend-ses-liens-avec-tmg.html.
5 E. Bursztein et al., "The Failure of Noise-Based Non-Continuous Audio Captchas" (presented at the 32nd IEEE Symposium on Security & Privacy, Oakland, USA, 2011), ftp://ftp.computer.org/press/outgoing/proceedings/SP%202011/Papers/4402a019.pdf.

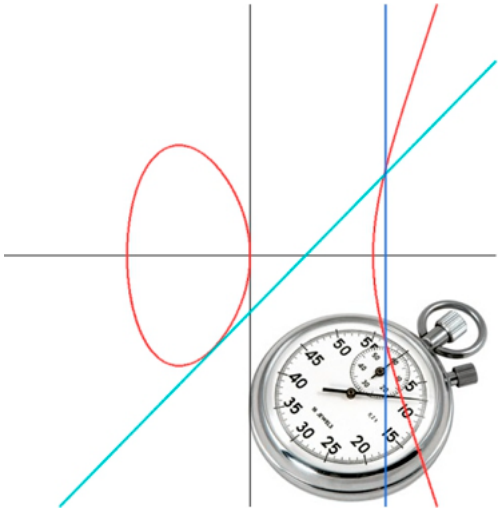
THE SECURITY
NEWSLETTER
#19

Together with a claimed success rate of 60% for breaking Microsoft's visual CAPTCHA system [6] and looking at current advances in signal processing and computer science, especially artificial intelligence, it seems to be only a matter of time, that the only difference between humans and computer in solving CAPTCHAs is that it is more cumbersome for humans.

P. BAUM

OpenSSL: A Timing Attack Vulnerability on the ECDSA Implementation

In a previous security newsletter, we reported an attack against OpenSSL [7]. Researchers at the University of Michigan injected faults during RSA signing algorithm by timely varying the power supply. This fault attack permitted to recover the 1024-bit private key. The exploited vulnerability affected the exponentiation implementation. Recently, researchers at Finnish University discovered a new vulnerability [8]. They focused on the OpenSSL's elliptic curve digital signature algorithm (ECDSA). Using a timing attack, they were able to recover the 160-bit private key. The vulnerability affects the Montgomery ladder implementation. Montgomery Ladder is an exponentiation algorithm that successively adds a point along an elliptic curve.



A timing attack is a side channel attack that attempts to break a cryptosystem by analyzing the time taken by the secret key operations. Every operation in a computer takes time to execute, and the time can differ based on the key value; with precise time measurements of each operation, an attacker can retrieve the input key. The OpenSSL's RSA implementation was an early target for attacks based on timing [9]. Meanwhile suitable protections like constant run-time implementations have been designed to thwart timing attacks.

In a nutshell, the attack targets a TLS server that authenticates using ECDSA signatures. Using the timing of the TLS handshake execution and the exchanged messages, the authors mounted an attack that recovers the private key. This attack is remotely feasible and successful when occurring within a Local Area Network (LAN). Interestingly, the attacker is passive; he only spies the TLS communication without exchanging any message with the server. Collecting a total of 16,384 signatures is sufficient for recovering the private key. However, the authors did not prove the feasibility of the attack across the Internet. The attack is sensitive to the network latency. When the network noise is high, the timings might not be meaningful.

This attack is applicable if:

- The attacker is within the same LAN as the TLS server,
- The TLS server holds a certificate embedding an ECDSA public key,
- The TLS server uses OpenSSL's ECDSA implementation for the authentication.

At the time of writing this report, the last version of OpenSSL does not fix the vulnerability. In the meantime, one should be cautious with the usage of the Elliptic Curve algorithms implemented in OpenSSL. Here are some practical mitigations for the attack:

- Implement the patch provided in the paper,
- Use another library that has a constant run-time implementation for the Montgomery ladder,
- Use a TLS certificate embedding another type of key, like an RSA or a DSA key.

Simple and generic solutions do not exist. Identifying timing attack vulnerabilities is a difficult task. This is not the first timing attack against OpenSSL implementation, and it will not be the last. OpenSSL is certainly the most widely scrutinized cryptographic library. Identifying such vulnerabilities leads to more secure software. This is what makes OpenSSL a trusted reference. This attack targeted both OpenSSL 0.9.8o and OpenSSL 1.0.0a

M. KARROUMI

6 Jeff Yan and Ahmad Salah El Ahmad, "A Low-cost Attack on a Microsoft CAPTCHA", April 2008, http://homepages.cs.ncl.ac.uk/jeff.yan/msn.htm.
7 Patrice Auffret and Mohamed Karroumi, "Attacking OpenSSL," Technicolor Security Newsletter 5, no. 16 (2010).
8 B. B Brumley and N. Tuveri, "Remote Timing Attacks are Still Practical ," Cryptology ePrint Archive, no. Report 2011/232 (2011), eprint.iacr.org/2011/232.pdf.
9 David Brumley and Dan Boneh, "Remote timing attacks are practical," Computer Networks 48, no. 5 (August 5, 2005): 701-716, http://www.sciencedirect.com/science/article/pii/S1389128605000125.

THE SECURITY
NEWSLETTER
#19

SONY AND THE LEAKING KEY

The design of cryptographic systems is delicate and subtle as it is prone to errors, which may have dramatic consequences. The methodology usually adopted by cryptographers consists of three steps. First, one defines what needs to be protected and against what types of attacks. Second, one defines the attacker's resources, that is, what the attacker is allowed to do. Third, one proves (or at least provides very strong arguments) that the proposed cryptographic system meets the so-defined security notion. For example, for digital signature schemes, the goal of the attacker is to produce a valid digital signature on a message for which the corresponding signature is unknown. Moreover, it is customary to ask that this should hold if the attacker can obtain signatures on messages of her choice. The corresponding security notion is called existential unforgeability against chosen-message attacks (EUF-CMA).

Does it mean that a cryptographic system built using the above methodology will not be attacked? No! The security can be undermined by an attack not covered by the model. Moreover, not only the system but also its implementation needs to be secure. In what follows, we will see how a digital signature scheme recommended by the cryptographic community - or more exactly a slightly tweaked version thereof, can be attacked. This is another lesson to learn. The specifications of a cryptographic algorithm must be strictly followed. Any deviation may be fatal.

Last December, a team of hackers called fail0verflow reported that they had discovered the ECDSA private key[10] that was used to sign authorized Sony's PlayStation 3 firmware. ECDSA is the elliptic curve analogue of the Digital Signature Algorithm (DSA). It appears in many cryptographic standards, including ANSI X9.62, IEEE P1363, FIPS 186-2, and SECG.

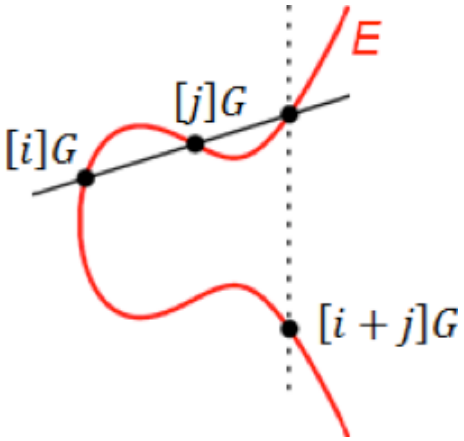


Figure 1: Addition on an Elliptic Curve

Here is how the signing process works. The domain parameters are an elliptic curve E over a finite field and a base point G on E of prime order n. Note that the points on an elliptic curve form an algebraic group under the chord-and-tangent law. The group operation is noted additively with + and we write $[k]G$ for $G+G+...+G$ (k times) and $x(P)$ for the x-coordinate of a point P. ECDSA requires a cryptographic hash function H used to produce a digest of the message m being signed. The private signing key is d. Signing process is as follows:

1. Randomly choose $k \in \{1, \ldots, n-1\}$.
2. Compute $r = x([k]G) \mod n$; if $r = 0$ go to Step 1.
3. Compute $s = k^{-1}(H(m)+dr) \mod n$; if $s = 0$ go to Step 1.
4. Return signature $\sigma = (r,s)$.

It is worth noting here that the ECDSA specifications require k to be random. In particular, since k is chosen at random in the set $\{1,...,n-1\}$ where n is (at least) a 160-bit value, this implies that the chances for a same k to be involved in two different signatures are very small. If k is repeated to generate signatures, the private key d can be recovered. To see it, suppose that $\sigma_1=(r, s_1)$ and $\sigma_2=(r, s_2)$ are two ECDSA signatures sharing the same k on two different messages m_1 and m_2. Then, since $s_1 = k^{-1}(H(m_1) + dr) \mod n$ and $s_2 = k^{-1}(H(m_2) + dr) \mod n$, it follows that $ks_1 - H(m_1) \equiv dr \equiv ks_2 - H(m_2) \pmod n$. Hence, we get $k(s_1 - s_2) \equiv H(m_1) - H(m_2) \pmod n$ and therefore $k = \frac{H(m_1)-H(m_2)}{s_1-s_2} \mod n$. The value of k then yields the value of private key d as $d = \frac{ks_1-H(m_1)}{r} \mod n$.

The attack by the `fail0verflow` team for recovering the PS3 signing key was not an attack against ECDSA but rather an attack against a poor implementation. The pitfall was the reuse of 'random' values. This illustrates once more that a slight implementation error may have huge implications.

Although Sony's signature private key has most probably never left its Hardware Secure Module, its value is available on the Net. It is now possible to sign any arbitrary firmware that will be accepted by the deployed Sony PS3. Sony deploys a new generation of hardware with enhanced security and of course using a new signing key.

M. JOYE

10 http://www.ps3blog.net/2010/12/30/ps3-private-keys-discovered-homebrewpiracy/

THE SECURITY
NEWSLETTER
#19

IDENTIFYING PHRASES IN ENCRYPTED VoIP

Traditional forms of telephony are being more and more frequently replaced by Voice over Internet Protocol (VoIP). These systems use audio codecs to encode the speech, enabling the transmission of the digital audio as an audio stream over an IP network. The confidentiality of the communication may be ensured by encrypting the VoIP packets if they are transmitted over unsecure networks. Nevertheless it was shown [11] that the language of the encrypted conversations can be identified if the VoIP packets are first compressed with variable bit rate (VBR) codecs and then encrypted with a length preserving stream cipher. Since many widely available VoIP products (including Skype) make use of VBR this vulnerability is not limited to one product. In 2010 Wright et al. [12] go one big step further in attacking encrypted VoIP calls by identifying phrases when the audio is encoded using variable bit rate (VBR) codecs.

Principles of Speech Coding

Voice data is transmitted as a Real-time Transport Protocol stream over UDP, which carries the audio data, compressed using a special speech codec such as GSM or G.728. At some fixed interval, the codec takes *n* samples from the stream and compresses them into a packet for efficient transmission across the network. Many voice codecs are based on a technique called Code-Excited Linear Prediction (CELP). The CELP encoder performs a brute-force search over the entries in a codebook of audio vectors and outputs the one that most closely reproduces the original audio. Therefore the quality of the compressed sound is determined by the number of entries in the codebook. The larger the codebook the higher the quality of the encodings and the more bits are used in indexing, resulting in higher bit rates and therefore larger packets. In CELP variants the encoder adaptively chooses the bit rate for each packet in order to achieve a good balance of audio quality and network bandwidth.

Basic Observation

Phonetic models of speech break the sounds into several different categories, including vowels and fricatives as well as stops and affricatives. Each of these canonical sounds is called a phoneme and the

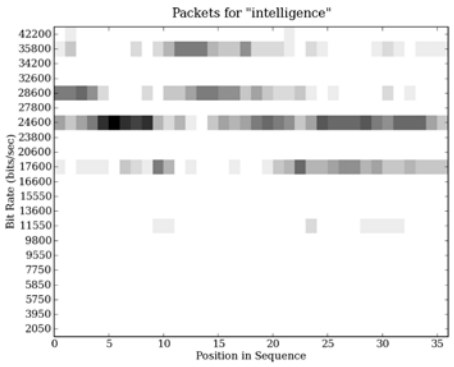


Figure 2: Bit rates for "intelligence" [13]

pronunciation for each word in a language can then be given as a sequence of phonemes. In turn the difference in bit rate for the phonemes enables the recognition of words and phrases according to the sizes of the CELP packets. It is the correlation between building blocks of speech (phonemes) and the length of the packets that a VoIP codec outputs which is exploited by the attack.

Spotting Sentences

The objective to the attack is to spot given spoken phrases in encrypted VoIP communications with minimal knowledge. That means without having information about the speaker or access to spoken content. The approach to crack the encrypted VoIP communication uses machine learning techniques and can be coarsely outlined as follows:

1. Build a model based on profile Hidden Markov Model (HMM) techniques.
2. Transform the profile HMM into a model suitable for performing searches on packet sequences.
3. Search the encrypted VoIP stream with the profile HMM.

Building the model is a complex task, since the human speech has a high degree of variability and the adaptive compression performed by the codec may contribute additional variance to the stream of packet sizes. To tackle this problem matching algorithms from the speech recognition and bioinformatics communities are applied, which use techniques based on hidden Markov models [14]. The problem of searching a protein database of fragments of known protein families is similar to the problem of searching a stream of

11 Charles V Wright et al., "Language identification of encrypted VoIP traffic: Alejandra y Roberto or Alice and Bob?," in Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium (Boston, MA: USENIX Association, 2007), 4:1–4:12, http://portal.acm.org/citation.cfm?id=1362903.1362907.

12 Charles V Wright et al., "Uncovering Spoken Phrases in Encrypted Voice over IP Conversations," ACM Trans. Inf. Syst. Secur. 13, no. 4 (December 2010): 35:1–35:30.

13 Charles V Wright et al., "Uncovering Spoken Phrases in Encrypted Voice over IP Conversations," ACM Trans. Inf. Syst. Secur. 13, no. 4 (December 2010): 35:1–35:30.

14 L. R Rabiner, "A tutorial on hidden Markov models and selected applications inspeech recognition," Proceedings of the IEEE 77, no. 2 (February 1989): 257-286.

packets sizes for instances of a word or phrase. Therefore the profile HMM [15] techniques developed to perform multiple sequence alignment of protein families in protein databases are adapted to the problem of finding words and sentences in encrypted VoIP traffic.

One of the challenges in this attack is to be independent from the speaker and without having access to the spoken content. To prepare the profile HMM for searching an appropriate synthetic training set has to be constructed. This construction is based on the word models developed by the speech recognition community. The basic idea is that all spoken words in a language are formed by concatenating phonemes, like words are strings of letters. In this paper a phonetic pronunciation dictionary is used together with a library of examples of packet sequences that corresponds to each phoneme in order to generate a synthetic training set for the phrases. This synthetic set is used to train the profile HMM which is used to search VoIP conversations.

Since in an encrypted VoIP communication the sentence to be matched may be surrounded by other sentences or silence, additional states are introduced in the HMM. The search in the sequence of packets is performed by applying Viterbi decoding [16] in order to find subsequences of packets that match the profile.

Experimental Setup and Results

The authors use the audio recordings from the TIMIT continuous speech corpus [17]. Speakers in the dataset include males and females with eight distinct regional dialects from across the United States. The library of packet sequences corresponding to phoneme, diphone, and triphone are built by encoding the audio with Speex in wideband VBR mode. The training data is used to build the HMMs to search for 122 target sentences.

For testing the VoIP communication has to be simulated. The VoIP conversation is simulated via testing data, which consists of randomly concatenated sentences from the speakers in the TIMIT test set. Having the test set the simulated conversation is encoded with wideband Speex in VBR mode.

The testing is performed by using the profile HMM to search for instances of each phrase in the resulting stream of packets lengths. The alignment of the packet lengths to the phrase HMM provides subsequences of packets with corresponding scores for each. Subsequences with scores above a given threshold t_p are considered hits. A true positive is a hit that contains all of the words in the given phrase where a hit which does not contain all words in the phrase is considered a false positive. To assess the performance of the approach the standard metrics named precision and recall (used in the information retrieval community) are measured for a variety of thresholds t_p.

Analysis of Results

The authors investigated the impact of noise, dictionary size, gender, and audio quality on the performance of the presented technique. The main influences on the performances are found to be the length of the phrase and the speaker of the phrase. Nevertheless the performance of the phrase spotting methodology can be increased by requiring only partial matching allowing the loss of 1 or 2 words in a sentence. In this case the technique is able to detect the presence of some phrases in encrypted VoIP conversations with recall and precision exceeding 90% with an average of about 50% for a wide variety phrases spoken by different speakers.

Techniques for Mitigation

Since the technique is based on the relation between packet sizes and ordering the two obvious approaches to defeat the technique are altering the packets order or quantize the packets to reduce the correlation between phrases and packet sizes. Since packet reordering is not really an option due to the real-time of VoIP, padding the packets by encrypting to 128, 256 and 512 bits can be used to reduce the information about the underlying audio. Although padding introduces additional overhead in the bandwidth it is shown that it provides a significant improvement regarding the security of VoIP calls.

Conclusion

The results of the study demonstrate that an attacker can spot a variety of phrases in realistic VoIP calls when the default encryption algorithms of the SRTP protocol are used. Phrases within encrypted calls can be identified with an average accuracy of 50 % and an accuracy of greater than 90 % for some phrases. Since many existing VoIP software (including the well-known Skype) use VBR encoding to compress the speech signals the presented approach is a serious attack of the user's privacy. With the use of padding techniques in encryption transforms confidentiality can be regained in VoIP conversations without compromising bandwidth too much.

M. ARNOLD

15 S R Eddy, "Multiple alignment using hidden Markov models," Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology 3 (1995): 114-120, http://www.ncbi.nlm.nih.gov/pubmed/7584426.

16 A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," IEEE Transactions on Information Theory 13, no. 2 (April 1967): 260- 269.

17 J. S Garofolo et al., TIMIT Acoustic-Phonetic Continuous Speech Corpus, CD (Linguistic Data Consortium, 1993).

THE SECURITY
NEWSLETTER
#19

NEW THREATS FOR SET-TOP BOXES

Traditional Threats

Set Top Boxes (STB) and Pay TV decoders have been available for more than two decades. Pay TV services started in the early 1980s. In 1984, the French operator Canal+ launched its first subscription-based channel. In 1985, HBO used Videocipher II to protect its satellite programs. In both cases, pirate decoders were soon available. Piracy was organized as quasi-industrial organizations [18]. The cat and mouse game started. Conditional Access System (CAS) providers learned from the early mistakes and currently make a very decent job to keep piracy at an acceptable level.

The objective of the hackers was to get unauthorized access to clear digital content. This can be either by extracting the clear descrambled video and redistributing it, or recently by redistributing the keys needed to descramble the content rather than the content itself. Usually, the defense relies on mainly two mechanisms:

- Clear digital video and keys are never available outside of a dedicated component in the STB. When temporarily stored in external memories, they are encrypted.
  Usually, an external smart card handles the decryption keys and returns them, if authorized, via a secure channel, to the STB that will use them to descramble content.
- The STB executes only authenticated application. It is impossible to run an arbitrary software on such device.

This is often referred as a closed garden environment.

The use of dedicated secure System on Chips (SoC) helps to build these two protections. Video descrambling, decoding, and rendering is entirely done within the safe enclosure of the chip, preventing any easy path to clear video. This type of SoC provides, first, secure storage and second, secure boot loader, which enables building a root of trust for the software running on the STB.

18 Charles Platt, "Satellite Pirates," Wired, 2004.

New Threats

Recently, new trends appeared in the STB market. For instance,

- STBs not only handle scrambled broadcast video protected by CAS, but they may also receive VOD content protected by Digital Rights Management (DRM) systems.
- STBs are connected to the Internet and host an embedded web browser. Thus, consumers may enjoy Internet on their TV screen through the STB.
- STBs are able to execute downloadable applications such as for instance games.
- Media gateways appear. They host in the same device the features of a STB and of a traditional gateway [19].

Each of these trends opens a breach in the traditional closed garden model. Each of them brings a new set of threats and security issues that will be hereafter analyzed.

At first sight, the addition of a DRM may seem straightforward. Unfortunately, it is not the case. Ideally, DRM should use the same protected path as CAS. This would require CAS and DRM to share some parts of the SoC. Nevertheless, the chip would have to strongly isolate the key secrets of CAS from the key secrets of DRM. Currently, SoCs are strongly CAS-typed. Therefore, it is likely that the implementation of DRM will remain software-based, as it is for general purpose computers.

In the closed garden context, the security of the firmware was the responsibility of the CAS provider. In the case of an additional software-based DRM, the corresponding security does not necessarily rely on CAS provider. Liability will have to be shared.

The Web browser that gives access to the open Internet is of course a serious threat. It is notorious that the Internet is a strong vector of contamination through malware, and other malicious sites. As for general purpose computer, a malicious site or applet may take the control of a host that is not properly protected. Most probably, such attack would not impair the security of the CAS that is hardware-based. DRM may be more at risk as they often are software-based. Nevertheless, it may seriously impair the user experience and thus tarnish the reputation of the network or broadcast operator. Limiting the browser to access only a white list of known trusted sites would be a secure option, but it would not be acceptable to consumers.

The huge success of downloadable applications for iOS™ or Android is a clear indicator of consumers' appetite for such features. Unfortunately, downloadable apps, even through a strictly controlled apps store, bring a set of new security issues such as data leakage, privacy issues, or even control of the host. The news about Android malware is a perfect illustration of these new risks.

19 "MediaEncore (Media Gateway)," Technicolor, n.d., http://www.technicolor.com/en/hi/digital-home/mediaencore.

THE SECURITY
NEWSLETTER
#19

Downloadable applications for a dedicated platform attract a special category of attackers: the hobbyists. Hobbyists like to be able to run their own, designed applications on these platforms. If the execution is only allowed to signed applications, they will try to jailbreak the system to run arbitrary applications. The most famous examples are game consoles or iPad™. This may enlarge the population of attackers for the STB world.

Gateways have their own set of threats [20] [21]. In the usual trust model of gateways/modems, the attacker was mostly coming from the outside. With the advent of media gateways, once more, the trust model may have to evolve. The attacker may be an "insider". As we have seen above, through a malicious site, or a malevolent downloadable application, the attacker may take control of the host. From that point, she may take control of the gateway functions. For instance, she may eavesdrop the communications, modify the settings of the firewall, or modify the configured DNS server. This last attack enables extremely efficient, stealthy phishing operations. The browser would in good faith believe that it is connected to the genuine site although the forged DNS server would redirect it to a malicious phishing site.

Some Remedies

As usual, after a threat analysis, the world seems awfully aggressive and dangerous. Will the new generation of STB open the Pandora box of malware? Will the new generation of STB become as risky as general purpose computers? Luckily, the answer is negative.

Although the STB is leaving the safe environment of closed garden, it does not enter into a widely open environment. In fact, it is possible to place the STB in a reasonably fenced environment. The cornerstone is the secure SoC. It offers two important services: an extremely robust boot loader and a secure key store. This allows authenticating and checking the integrity of an initial kernel. Once verified, this kernel can itself check the authenticity and integrity of a native firmware. Thus, after the boot, the STB has a trusted environment of execution. The reader may have identified the similarity with the Trusted Platform Module [22].

Then, the trusted and checked firmware can control which applications are executed, and which access to resources they are granted. A trusted application may be granted access to critical resources whereas a non trusted application may get access to a very limited set of resources such as remote control and display. This is called Access Control.

Furthermore, applications may be isolated from other applications by sandboxing [23] or container isolation. The objective is that if the hacker exploits a vulnerability in the application (buffer overflow, cross site scripting...), she cannot escalate her privileges to gain control on other parts of the STB than the current application.

Many other techniques may help to mitigate these threats. All require a deep knowledge of security, of the platform, and carefully crafted implementations of the security mechanism. This is tough work.

Conclusion

As new generations of STB propose new features, their security scope will extend beyond usual content protection. They are leaving the relatively safe environment of closed garden to enter in a more open environment. This evolution raises many new threats, similar to the ones usually encountered with general purpose computers.

Fortunately, STBs have some advantages compared to general purpose computers. It is possible to build a root of trust. Nevertheless, facing these new threats, STB manufacturers will have to carefully design new protection schemes in order to offer a better user experience to their customers.

E. DIEHL

20 "OSVDB: The Open Source Vulnerability Database", n.d., http://osvdb.org/.

21 Dan Goodin, "Worm breeds botnet from home routers, modems," The Register, March 24, 2009, sec. Security, http://www.theregister.co.uk/2009/03/24/psyb0t_home_networking_worm/.

22 "Trusted Computing Group: TPM", n.d., https://www.trustedcomputinggroup.org/groups/tpm/.

23 Raphael Gelloz and Michel Morvan, "Sandboxing," Technicolor Security Newsletter 5, no. 18 (Spring 2011): 8-11.

THE SECURITY
NEWSLETTER
#19

WHERE WILL WE BE?

IEEE International Conference on Image Processing (ICIP 2011),
Brussels, Belgium, September 11-14, 2011

- Paper presentation: Virtual view invariant domain for 3D video blind watermarking, by Javier Franco-Contreras, Séverine Baudry, and Gwenaël Doërr

CodenomiCON Europe 2011-07-22, 10 year anniversary edition,
Oulu, Finland, October 7, 2011

- Invited talk by Olivier Heen: Wi-Fi Testing: 802.11 and its implementations

Workshop on Computer Vision Methods in Blind Image Forensics (CVBIF 2011),
Barcelona, Spain, November 13, 2011

- Invited talk by Gwenael Doerr: 'Forensics Techniques to Deter Movie Piracy'

TECHNICOLOR SPONSORED CONFERENCES

16th European Symposium on Research in Computer Security (ESORICS 2011),
Leuven, Belgium, September 12-14, 2011

HACK.LU 2011,
Luxembourg, Luxembourg, September 19-21, 2011

13th IACR Workshop on Cryptographic Hardware and Embedded Systems (CHES 2011),
Nara, Japan, September 25-28, 2011

11th ACM Workshop on Digital Rights Management (ACM-DRM 2011),
Chicago, IL, USA, October 17, 2011

EXTENSIVE WORLDWIDE PRESENCE

Vancouver	Palo Alto	Mexico	Rennes	Piaseczno	Beijing
Hollywood	New York	Manaus	London	Rome	Bangkok
Indianapolis	Guadalajara	Paris	Madrid	Bangalore	Sydney

technicolor

TECHNICOLOR WORLDWIDE HEADQUARTERS
1, rue Jeanne d'Arc
92443 Issy-les-Moulineaux France
Tel. : 33(0)1 41 86 50 00 - Fax : 33 (0) 1 41 86 58 59

www.technicolor.com