Using Side Channel Attacks in the Human Computational Model

Orr Dunkelman[1]*

[1]ESAT/SCD-COSIC, The Pope-Abiding University (of Old) Leuven
Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium
`orr.dunkelman@esat.kuleuven.be`

Abstract. Recent works have looked at how to use human knowledge or computability in order to define or to achieve security. We take this approach one step further, and suggest to use cryptanalytic attacks in the human computational model. We have identified several instances where side channel attacks on human beings may be useful. We show for each of these cases the appropriate use of side channel attacks, and estimate the time complexity of the involved attacks.

1 Introduction

Standard cryptographic assumptions are based on some complexity theory assumptions (e.g., the existence of some one-way function, the hardness of computing discrete logarithm). In recent years the interaction between human beings and the computational process has led to the introduction of a more precise definitions. For example, in [1], CAPTCHAs[1] are used as a distinguishing mechanism between real people and machines, resulting in a secure stenographic covert channel.

Recently, this approach was taken one step forward in order to define security notions for hash functions. In [8], the basic problem of defining a secure collision resistant hash function is tackled. Even though there is a security notion for such functions [9], this notion has to assume the existence of a family of compression functions. Otherwise, let $h(\cdot)$ be some hash function. As $h(\cdot)$ compress inputs of arbitrary length into a fixed-size strings, there must exist two values a, b, such that $h(a) = h(b)$. Thus, there is a simple, $O(1)$ algorithm that finds collisions in $h(\cdot)$, namely

```
print(a,b).
```

Thus, for any hash function $h(\cdot)$ there exists such an algorithm. Thus, resulting to the standard approach of "let's assume that there is no efficient algorithm

* The research done is this paper was generously supported by the FAO-Schwarz scholarship.

[1] CAPTCHAs are *Completely Automated Public Turing test to tell Computers and Humans Apart.* Most of them are distorted pictures of text, that only human beings can read, while no OCR software can read the text embedded in the CAPTCHA.

that can break our scheme", fails in the case of hash functions. This led [8] to discuss the notion of human knowledge, as even though there exists such an algorithm, we have no idea which of the possible print algorithm is the one that truly produces a collision.

In this paper we continue to pursue the study of interaction between cryptography and human beings. We show that by applying side channel attacks to humans, we can improve the security of the cryptographic schemes used (by finding and eliminating the weaknesses found using our attacks). We then show that it is possible to use side channel attacks to improve the everyday life of many people, even when no cryptographic protocols are present. This is the first time side channel attacks are suggested to be deployed for a large number of users simultaneously.

The paper is organized as follows: Section 2 covers the previous essential results in the field of human computability. Section 3 deals with the cryptographic uses of side channel attacks on humans. We explore the non-cryptographic uses of side channel attacks in Section 4. Finally, we do not summarize the paper in Section 5.

2 Human Computability

The first thought that comes to mind when dealing with human computation devices is the fact that they are very slow, prune to err, and use very old technology. It is well known that all human computation units today use base 10 computation, making them incompatible with other computational devices.[2]

However, human computability has been a very important in the history of computation. It is a well known fact that the first code breakers had used only human computation devices. Moreover, due to the fact that many such devices are available, the human computational model is well studied.

Human computation had several uses in the world of cryptography and information security. For example, CAPTCHAs are used a tool to distinguish between software that tries to send spam and human beings that try to send spam. One of the techniques that was developed by the spammers is to delegate the problem of solving these CAPTCHAs to other human computing devices. One of the means used by them, is to require human computing devices to prove their year of manufacturing using the same CAPTCHAs as the CAPTCHAs the spammer's software tries to solve.

The fact that CAPTCHAs can (till today) effectively distinguish between a software running on a modern CPU and a software running over a brain has led researchers to suggest CAPTCHA based stenographic covert channel [1]. The main idea behind the encryption scheme is the fact that the generation of CAPTCHAs actually present a strong one way function which is hard to invert.

Another recent approach to human computability was given in [8]. It is a well known problem that defining the security of hash functions in the standard

[2] We note that several models have used base 60 for their computations. This only shows that no one should let humans do computers work.

cryptographic model is impossible. This follows the problem that the standard definition assumes that there is no efficient algorithm which breaks the scheme. As in the case of hash functions, efficient algorithms which produce collisions exist for each and every hash function, the standard approach fails.

This problem was addressed in [9], where the solution was to define families of hash functions. If for each such family there are many possible hash functions, no efficient algorithm would be able to produce collisions without knowing in advance the chosen hash function. The main problem of this approach toward security notions is the fact that most common hash functions are not keyed families, but rather a specific function. Moreover, there is still a research done on the topic of what way the keying of the family of hash functions should be approached.

Rogaway has noticed that by considering the human computational model, one can find consistent and meaningful security notions for hash functions [8]. This is done by pointing out the fact that even if there is an efficient algorithm which produces a collision for the hash function, we have no idea which algorithm it is.

The main idea is that if a hash function is used in an higher protocol (which is also the case), then the existence of an efficient attacker against the protocol would imply the existence of an efficient collision finding algorithm. Thus, in case no one knows how to attack the protocol, then there is no known efficient collision finding algorithm.

The human knowledge in cryptography is obviously recursive enumerable (as the number of research papers in cryptography is finite[3]).Thus, it is very easy to prove that a given hash function is collision resistant given enough translators from Chinese into English.

The above examples only show that there is a very tight relation between security and human computability. We research this relation in the following sections.

3 Cryptographic Uses of Human Side Channel Attacks

When considering a cryptographic environment, one has to take into consideration the delicate relation between the various components in order to assess its security. This is problem has led to the introduction of the concept of universally composable primitives [4], i.e., primitives whose security is not affected from interaction with other primitives. This is especially important for cases where there are many instances of related primitives running. For example, a simple zero knowledge protocol can become inherently insecure once the same statement is proved simultaneously to two (colluding) adversaries.

As much as this approach leads to a much stronger cryptography, it neglects one of the most essential and important component of the system — the human operating the protocol (or the software that operates the protocol). For example,

[3] Unless you are a reviewer.

any protocol based on passwords becomes insecure once the human is willing to exchange the password for a candy [2].

This led [3] to propose a new method for cryptanalysis of cryptographic schemes. The main approach taken in [3] is the fact that human beings in possession of secret keys or passwords are likely to disclose the information under physical side channel attacks. For example, it was observed that under power analysis (where the human is used as the measurement tool to the amount of power consumed by the closed circuit), passwords can be easily retrieved. This is a much faster and more accurate approach than trying to fiddle with the silicon. It is easy to see that while power analysis on smart cards may need several thousands of samples, power analysis on humans is expected to require significantly lower number of samples (even though the process of generating a sample may be longer).

Another side channel attack that recently affected the cryptographic world is the cache attacks [5–7]. We claim that in the process of history, the cash attack against human beings was much more successful, and usually required a smaller number of samples (again, obtaining one sample might take longer). Many human computing devices have been corrupted using cash attacks, and besides disclosing secret information under this attack, there are several well known cases of interfering with a voting protocols due to this attack.

There are also several side channel attacks which are specific only to human computing devices. The blackmail attack (as was observed in [3]), is a very powerful attack that can be used to extract information from such computing devices. In many of these attacks, the computing device under attack becomes more and more cooperative as the attack continues, leading to an increased efficiency.

Another side channel attack that can be applied only to humans is the seduction attack. This approach may vary according to the specific implementation. It is well known that during the cold war, the Russian intelligence agencies have used this approach to find out classified information. Even though the Russian model for attack was based on blond hair, there are known cases in history where the used model had a different hair color. Even though the specific model to use for the attack is not necessarily known in advance, using several prototypes would cover most of the manufactured human computing devices.

We summarize this section by noting that there are probably many more side channel attacks that can be applied to human computing devices. We are sure that this field would yield a great deal of research, as it can be useful for deciphering many of the secret communications used by terrorists.

4 Medical Uses of Human Side Channel Attacks

While the previous section dealt with cryptanalytic aspect of side channel attacks, we show that side channel attacks can also be used for improving the medical treatment of various illnesses. We shall investigate mostly schizophrenia, even though the ideas can be easily applied to other illnesses as well, including (but not restricted to) multiple personality disorder, generalized anxiety disorder, and bipolar disorder.

Schizophrenia is a mental disease. Some of those who suffer from it are diluted to hear voices speaking in their heads,[4] telling them what to do. Usually, humans who suffer from Schizophrenia are likely to become paranoid, and in some cases they might become violent due to these voices.

There are several chemical solutions to Schizophrenia. Most are based on supplying the patient with sedatives, which reduces the amount of voices, as well as the tendency to act according these voices. However, this approach is troublesome. First, it usually requires the good will of the patient to accept such drugs. Second, the patient may suffer some side effects, mostly related to the fact that he (or she) is being sedated.

We note that voices inside the head of a person are eventually represented as electrical currents. Thus, by performing a side channel attack that investigate these currents, can lead to a better understanding of the patient. For example, by examining the electrical activity in the patient's brain, it is possible to detect when he (or she) hears these voices using the standard techniques of identifying computational operations based on the power consumption.

After the phase of collecting the patient's data, and especially the power consumption of the voices, we can easily identify when the voices become violent. This can allow for a much better classification of patients according to the violence in their heard voice, which in turn would allow to use smaller amounts of sedatives on the less violent patients.

Another possible side channel approach might be fault analysis. Once we identify the electrical currents associated with the voices, we can try and counter them (by injecting the opposite currents). This fault in the computation would result in the patient not hearing the voices, and thus being free of this symptom.

5 C'est Ne Pas Un Summary

Even though the authors are sure that the results of this paper are true, we find that experimental data is to be obtained. However, as the process of physically obtaining the data require us to leave the chair, and file a request with the Helsinki committee, we find it very impractical.

References

1. Luis von Ahn, Manuel Blum, Nicholas J. Hopper, John Langford, *CAPTCHA: Using Hard AI Problems for Security*, Advances in Cryptology, proceedings of Eurocrypt 2003, Lecture Notes in Computer Science, vol. 2656, pp. 294–311, Springer, 2003.
2. BBC, *Passwords revealed by sweet deal*, available online at *http://news.bbc.co.uk/2/hi/technology/3639679.stm*, 2004.

[4] Referred in medical jargon as *auditory hallucinations.*

3. David Beynon, *Practical Key Recovery*, Journal of Craptology, Volume 0, No. 1, 1999.
4. Yehuda Lindell, *General Composition and Universal Composability in Secure Multi-Party Computation*, proceedings of 44th Symposium on Foundations of Computer Science (FOCS 2003), pp. 394–403, IEEE Computer Society, 2003.
5. Dag Arne Osvik, Adi Shamir, Eran Tromer, *Cache Attacks and Countermeasures: The Case of AES*, proceedings of CT-RSA 2006, Lecture Notes in Computer Science, vol. 3860, pp. 1–20, Springer, 2006.
6. Daniel Page, *Theoretical use of cache memory as a cryptanalytic side-channel*, technical report CSTR-02-003, Dept. of Computer Science, University of Bristol, available online at *http://www.cs.bris.ac.uk/Publications/pub_info.jsp?id=1000625*, 2002.
7. Collin Percival, *Cache missings for fun and profit*, BSDCan 2005, available online at *http://www.daemonology.net/hyperthreading-considered-harmful/*, 2005.
8. Phillip Rogaway, *Formalizing Human Ignorance: Collision-Resistant Hashing without the Keys*, proceedings of Vietcrypt 2006, Lecture Notes in Computer Science, vol. 4341, pp. 221–228, Springer, 2006.
9. Phillip Rogaway, Thomas Shrimpton, *Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance*, proceedings of FSE 2004, Lecture Notes in Computer Science, vol. 3017, pp. 371-388, Springer, 2004.