May Honey Encryption Ruin The Hacker’s Honeymoon?

Course: ITM 323.SF2 Security, Privacy and Compliance

Instructor: Dr. Bhanu M Viswanadha

Student: Sadegh Bamohabbat Chafjiri

Student ID: 0588462

Date: Feb 26 2019

Agenda Page

I. Executive summary

In this project, I introduce a data breach that occurred for Adidas in 2018. The hacker claimed that he gained access to a vast amount of customers' confidential data stored on the database of Adidas. Afterwards, Adidas confirmed that a few million of its clients might be impacted by that data breach. The catastrophe came into focus when Adidas had no clear answer about how and when this breach happened. …. In this project, firstly, I focus on the problem statement in section II. I show how the lack of a monitoring system in Adidas led to the data breach. In section III, I show how in the absence of a monitoring system a company has to spend extra money on protection, and there is no assurance that its course of action will add value to system protection. In section IV, I introduce the honeypot and honey encryption schemes. In section V, I combine both schemes to meet the 3 basis of CIA triad: confidentiality, integrity and availability of personal information stored in the database and finally, I recommend a new scheme based on collision-free scheme for "near match inputs", which NSA introduced for Hashing systems recently. This scheme can reduce the adversary's chance to identify the real master key among all the fake passwords in the honey encryption scheme.

II. Problem Statement

In early 2018, Adidas became aware of a security breach by an "unauthorized party." The unauthorized party claimed that he gained access to customer data stored on the database of Adidas. On June 26th 2018, Adidas announced the security breach on its US website. According to Humphries (2018), Adidas spokeswoman confirmed that a few million of its clients might be impacted by the data breach. The stolen data included customers' addresses, contact details and email addresses, login information including usernames and encrypted passwords. Hopefully, the data breach did not include financial information. The stolen passwords had been encrypted which means the hacker would have to decrypt the passwords before he could misuse them. However, the confidentiality of the customers' accounts was impacted by the security breach. According to White (2018) the investigation is still in progress, and Adidas is working with leading cybersecurity firms and law enforcement agencies to determine the scope of the incident. More information will be declared when the investigation is done.

III. Key requirements to address the problem

According to Wycislik-Wilson (2018), Adidas suffered from the lack of a monitoring system to detect and control the threat in a timely manner. Consequently, Adidas could not confirm or reject without reservation if the data breach claimed by the hacker had occurred or not. Another issue is that taking the unnecessary action to protect the data privacy, Adidas will spend its resources on unverifiable threats created by fake issues. To address this issue, the IT infrastructure of Adidas should meet some key requirements. Firstly, Adidas should be equipped with a powerful monitoring system. This system enables Adidas to monitor malicious activities.

Secondly, Adidas should deploy an advanced design of encryption algorithm to protect the passwords so that hackers are not able to decipher encrypted passwords.

Just as math enables cryptographers to protect data from password-cracking, it also helps hackers to discover the secret key. According to Mazerik (2014), it is a never ending “cat and mouse game.” For example, encryption algorithms should resist hackers who apply the clustering method to classify encrypted messages. By using the clustering method, they try to find the near-match inputs which causes a collision in output and then formulates some bits of the key. In this approach, search algorithms can reduce the search key space to less than brute force, and it empowers hackers to remove invalid keys with a time complexity less than exhaustive search. By finding a correlation or logical formula between input and output, computational complexity is reduced and ultimately the hacker can identify the valid key. So, a well-designed encryption algorithm should show a stronger resistance against cryptanalysis. Moreover, it seems that memorizing a long password in the One-Time-Pad scheme is tough for users. According to Juels (2018), one percent of passwords is “123456”. Fifty percent of people use weak passwords that can be easily cracked by intruders. Therefore, it seems the information leakage in input can help hackers to find the master key. This is true for the data breach that occurred for Adidas. There is this risk that the hacker finds the correlation between the master key and encrypted passwords. Therefore, a well-designed encryption algorithm should have a good statistical characteristic to resist cryptanalysis.

IV. Key research findings

An applicable monitoring system for a retail company such as Adidas is the honeypot system. The honeypot system is a mechanism to detect and control an adversary who wants to get an unauthorized access to databases. Honeypot stores false data to detect the intruder’s activity. False data seems to be an actual part of the website while it is only a trap for the adversary. Honeypot is an isolated and monitored part of the website to detect malicious activity (Honeypot computing, 2018). This system has a graphical dashboard and provides a high level of interaction and monitoring for admin. Network admin can use different features including email traps, malware honeypot, database honeypot against unauthorized access (Honeypot computing, 2018). Honeypot can give login warning to admin against malicious activities and black hats (Dargin, 2017). This monitoring system can be deployed with firewalls and includes false information to confuse intruders and waste their time. Honeypot can protect real data against intruders. When an unauthorized party tries to access the data base, the honeypot system will direct him /her to a pool of fake information. A well-configured honeypot system confuses intruders with fake passwords and wastes their time on stealing bogus data. Afterwards, this monitoring system posts alarm on malicious activities and gives reports to admin about incidents including the IP address of the intruder, time and date of the incident and the hacker’s technical skills. The adversary will not be aware of the monitoring on which the admin is able to configure his/her activities. The hacker does not realize the bogus data is a predetermined monitoring trap for hunting him.

In addition, Adidas IT system can deploy a honey encryption scheme to provide a stronger data encryption against brute force cracking on cloud computing systems (Honey Encryption, 2015). The main advantage of the Honey encryption scheme is a deflection mechanism proposed by Ari Juels and Thomas Ristenpart at Eurocrypt, a cryptography conference in 2014 (Honey Encryption, 2018). It has a high resistance against brute-force attack. Its benefit is a high level of encryption standard and follows the concept behind the honeypot to detect and to mislead the adversary when he/she attempts to decrypt the master key. However, this method is technically a little different from honeypot (Honey Encryption, 2015). The idea behind the encryption characteristic is similar to one-time-pad but by using an efficient operation. It provides the advantage of a distributed security. In the one-time-pad encryption scheme, the master key should have a length equal to or longer than plain text. In honey encryption, the sender can encrypt the plain text with a weaker key while it is still hard for the adversary to crack the key. In other words, honey encryption will add extra system complexity to the key entropy. The system complexity is created by a "fake password vault generator." This generator is fed by different resources of leaked/sample passwords. The generator puts bogus passwords in a master vault where all confidential information such as passwords and other confidential parameters are stored. So, an adversary is not able to have easy access to confidential information such as original passwords. In honey encryption, all passwords except original ones are bogus and called "Honey words or decoys" and the set of all passwords including the original password is called "sweet words" (Mazerik, 2014). The sweet words can resemble real passwords. The process of verification for a real password is that honey encryption has a "honey checker" which stores an index of real passwords. When a legal user uses a correct index of the password, authentication will go in a correct path, and it grants the user to access confidential data.

The strengthen of the Honey encryption scheme is that when an adversary attempts to get in the system by guessing the master key located in the vault, attacker will be supplied with honey words and will receive the wrong data. These bogus data canbe wrong pins, passwords or credit card numbers or CVVs which confuse the intruder.

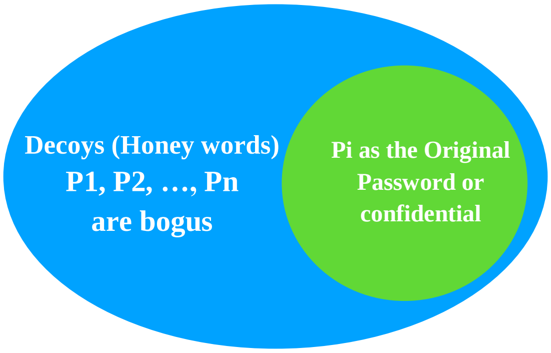

Fig 1. A large encryption passwords vault; Intruder has to distinguish Pi as the original password among all honey words (Mazerik, 2014)

In contrast, there is no scenario designed for malicious activities in modern encryption. For example, vulnerability occurred in the Lastpass breach only because the user chose a weak and meaningful master key from dictionary not from non-dictionary words and the intruder could guess the whole master key only by decrypting a part of the master key.

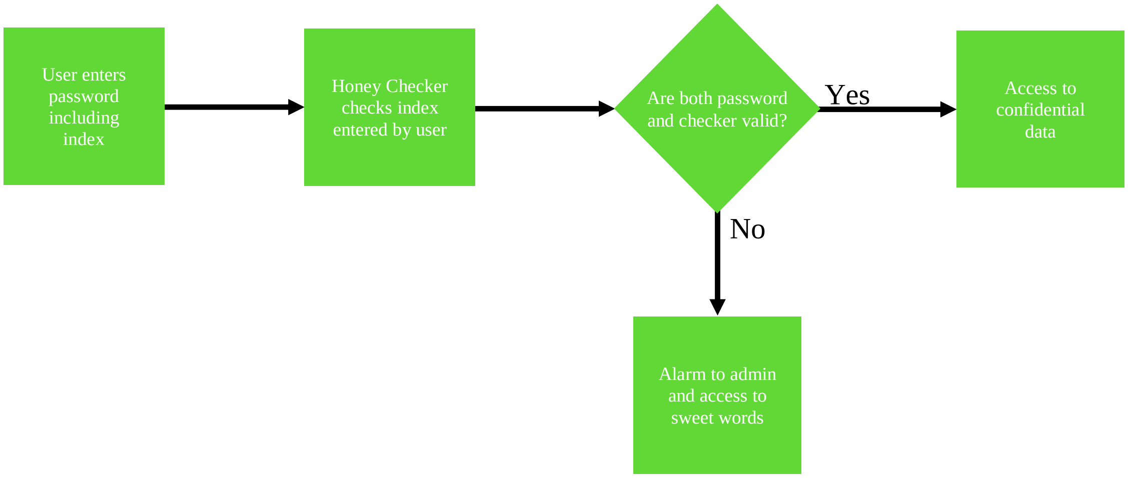
User enters password including index
Honey Checker checks index entered by user
Are both password and checker valid?
Yes
Access to confidential data
No
Alarm to admin and access to sweet words

Fig 2. Process of the honey encryption (Mazerik, 2014)

V. Solution Options/Conclusion/Recommendations

Lack of a monitoring system created a gap between what Adidas employed to protect customers' personal information and what is essential to employ. It is clear that the absence of a monitoring system negatively impacted Adidas IT service management, and lack of logs and reports led to inadequate course of action. Of the available countermeasures proposed in section IV, I will put a short term plan into action by deploying the honey encryption scheme in the Adidas database. The plan protects the confidentiality, integrity and availability of confidential data including users' personal information stored in the application layer. This scheme can confuse the adversary with many stored decoys and will add value to system protection. To achieve the purpose of the recommended plan, I will train Adidas's research and development team. I will employ a cryptographer to design a new algorithm based on honey encryption or extend the previous algorithm with the honey encryption scheme. This algorithm must have a good statistical characteristic against typical cryptanalysis. In addition, it should have an additional feature to provide a pool of fake passwords. So, in the worst scenario, if the intruder finds a vulnerability in the encryption algorithm, he/she still needs to find a master key from a pool of bogus passwords stored in the vault, which reduces risk of data breaches.

As a long term plan, I will deploy a combination of honeypot and honey encryption schemes. It can protect data confidentiality by providing a higher level of monitoring as well as adding system complexity against intruders, which means if an unauthorized party wants to get into the system and steal data, the system will detect attack and warn admin. The best scheme for Adidas is the high-interaction honeypot. It enables the system to host several services by employing multiple honeypots and virtual machines (VMware); otherwise, Adidas should use one physical computer for each honeypot.

VI. Next Step

By deploying the combination of honeypot and honey encryption schemes, the risk still remains that the adversary might find the “near match inputs” generating same output (see fig 3). “Near match inputs” of encryption elements/hash functions means there are 2 inputs having only a few bits in difference to each other, and they make an output-collision. “Near match inputs” can help intruders filter search space and remove decoys and finally distinguish the master key among a pool of honey words. “Near match inputs” enable intruders to search for the master key by measuring the shortest “Hamming distance.” Hamming distance is a good measure to check near match words in input with less complexity. So, the intruder can misuse the property of near match passwords to distinguish the master key which has the shortest “Hamming distance” from reached sweet words. In the NSA Patent Portfolio V5.1 (2019), NSA recommended the idea of Collision-free hashing for near match input which is applicable for the honey encryption scheme (See appendix). Therefore, the next step it is essential to modify the characteristic of honey encryption algorithm and make honey encryption scheme free from “near match inputs” causing collision.

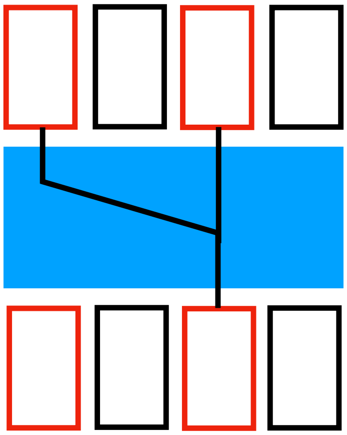

Fig 3: Collision in the near match inputs

VII. Conclusion

In this project, I introduced a data breach that happened to Adidas in 2018. The hacker accessed a vast amount of customers' confidential data stored in the database. This project focused on lack of control and monitoring system in Adidas and introduced two countermeasures to address this issue: honeypot and honey encryption schemes. Afterwards, I combined both schemes to meet 3 basis of CIA triad: confidentiality, integrity and availability of personal information stored in database, and finally, as the next step, I recommended a new technique presented by NSA called Collision-Free Hashing for Near-match inputs. This technique can reduce the adversary's chance to identify the real master key among all fake passwords in the honey encryption scheme.

References

Humphries, M. (2018). Adidas Website Hacked, Millions of US Customer Details Stolen. Retrieved from https://www.pcmag.com/news/362173/adidas-website-hacked-millions-of-us-customer-details-stole

White, W. (2018). Adidas Data Breach 2018: What U.S. Customers Should Know. Retrieved from https://investorplace.com/2018/06/adidas-data-breach-2018/

Wycislik-Wilson, M. (2018). Adidas data breach may have exposed personal data of American customers. Retrieved from https://betanews.com/2018/06/29/adidas-data-breach/

Mazerik, R. (2014). Honey Encryption. Retrieved from https://resources.infosecinstitute.com/honey-encryption/#gref

Juels, A. (2018). An Introduction to Honey Encryption. Retrieved from https://www.skyhighnetworks.com/cloud-security-blog/cryptographic-parlor-tricks-for-passwords-an-introduction-to-honey-encryption/

Honeypot computing. (2018). Retrieved from https://en.wikipedia.org/wiki/Honeypot_(computing)

Dargin, M. (2017). Increase your network security: Deploy a honeypot. Retrieved from https://www.networkworld.com/article/3234692/increase-your-network-security-deploy-a-honeypot.html

Honey Encryption. (2018). Retrieved from https://en.wikipedia.org/wiki/Honey_encryption

NSA PATENT PORTFOLIO V5.1 (2019) Retrieved from https://www.nsa.gov/Portals/70/documents/what-we-do/research/technology-transfer/nsa-technology-transfer-program.pdf

Appendix

Collision free for the near match inputs: in this scheme intruder cannot perform a Hamming distance technique to filter bogus passwords because it does not guarantee that the near match inputs will generate a collision.

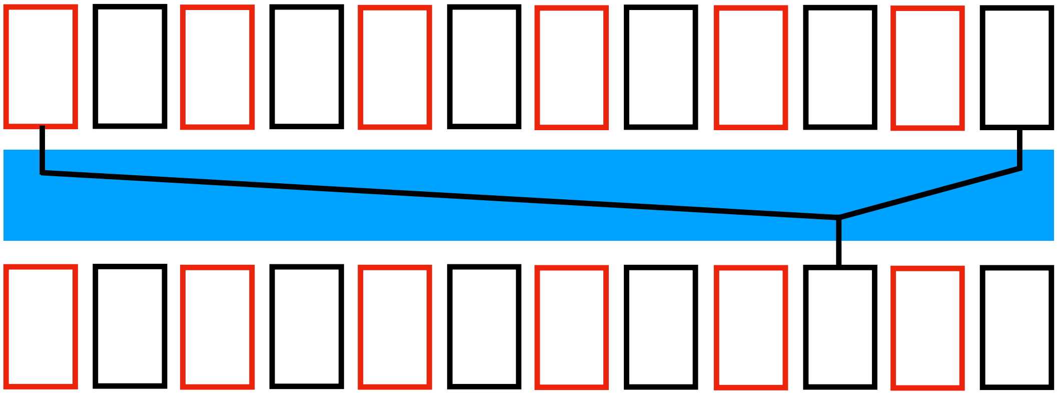

Fig 4: collision free technique for the near match inputs