Chapter 10
Password Recovery Research and its Future Direction

Vrizlynn L. L. Thing
Institute for Infocomm Research, Singapore

Hwei-Ming Ying
Institute for Infocomm Research, Singapore

ABSTRACT

As users become increasingly aware of the need to adopt strong password, it brings challenges to digital forensics investigators due to the password protection of potential evidentiary data. On the other hand, due to human nature and their tendency to select memorable passwords, which compromises security for convenience, users may select strong passwords by considering a permutation of dictionary words. In this chapter, the authors discuss the existing password recovery methods and identify promising password recovery approaches. They also present their previous work on the design of a time-memory tradeoff pre-computed table coupled with a new sorting algorithm, and its two new storage mechanisms. The results on the evaluation of its password recovery performance are also presented. In this chapter, the authors propose the design of a new password recovery table by integrating the construction of common passwords within the enhanced rainbow table to incorporate the two promising password recovery approaches. They then present the theoretical proof of the feasibility of this technique.

INTRODUCTION

Being the most common authentication method, passwords are widely used to protect valuable data and to ensure a secure access to systems/machines. However, the use of password protection presents a challenge for investigators while conducting digital forensics examinations.

In some cases, compelling a suspect to surrender his password would force him to produce evidence that could be used to incriminate him, thereby violating his right against self-incrim-

DOI: 10.4018/978-1-4666-2190-9.ch010

ination. Therefore, this presents a problem for the authorities. It is then necessary to have the capability to access a suspect's data without expecting his assistance.

While there exist methods to decode hashes to reveal passwords used to protect potential evidence, lengthier passwords with larger characters sets have been encouraged to thwart password recovery. Awareness of the need to use stronger passwords and active adoption have also rendered many existing password recovery tools inefficient or even ineffective.

The more common methods of password recovery techniques are based on brute force, precomputed tables, dictionary attack, breaking hashing algorithms and more recently, using rainbow tables.

In the brute force attack, every possible combination of the password characters in the password space is attempted to perform a match comparison. It is an extremely time consuming process and recovering a strong password could take weeks or even months. However, due to its exhaustive generation and search, the password will be recovered eventually if sufficient time is given. Cain and Abel (Cain and Abel, 2012), John the Ripper (John The Ripper, 2012) and LCP (LCPSoft, 2012) are some popular tools which support brute force attacks.

The precomputed tables, on the other hand, hold all the possible passwords in the password space and their corresponding hash values, which were generated beforehand. The tables were then sorted and stored. Given a password hash, the corresponding password can be easily retrieved from the tables. The disadvantage is the need for an extremely large storage space. For example, to generate precomputed tables for 10-character passwords whereby the characters can be any printable ASCII characters, and the hash output for each password is 16 bytes (e.g. MD5 hash), the precomputed tables will be 1.317ZB (Zetta Bytes). Therefore, this method is not feasible for strong password (e.g. within a large password space).

The dictionary attack method composes of loading a file of dictionary words into a password cracking tool to search for a match of their hash values with the stored one. It is a subset of the precomputed tables as only commonly used passwords were stored in the file. If the password is not a dictionary word, the recovery would fail.

Research attempting to discover and identify the weaknesses of hashing algorithms have also been useful in the application of passwords or encryption keys recovery. This method is based mainly on the collision of hashes in specific hashing algorithms (Contini, 2006; Fouque, 2007; Sasaki, 2007; Sasaki, 2008). However, they are too complex and extremely time-consuming to be used for performing password recovery during forensics investigations. The methods are also applicable to specific hashing algorithms with inherent weakness only.

In the time-memory tradeoff method (Hellman, 1980), a hybrid of brute force attack and precomputed tables is used. Large number of different passwords are repeatedly hashed and reduced to forms password chains. Only the head and tail of these chains are stored in tables. During a recovery, the password hash goes through a series of reduction and hashing until a match with one of the stored tails is found. The chain which generates this password hash can be identified and the password can be recovered. The storage space and recovery speed can be adjusted flexibly by setting the rounds of hashing and reduction, and the available storage space, accordingly. This method can be applied to retrieve Windows login passwords encrypted into LM or NTLM hashes (Todorov, 2007). Passwords encrypted with hashing algorithms such as MD5 (Rivest, 1992), SHA-2 (NIST, 2002) and RIPEMD-160 (Dobbertin, 1996) are also susceptible to this recovery method.