CS2421 Network and Server Security

Group 4

*The MD5 Cipher Text decoding based on dictionary attack*

**Division of Work**

LIU KA WO (DAVIS) 30215633 (report writing)

CHEUNG WAI CHUNG (CHRIS) 30208907 (testing)

PIN LING CHI (Victor) 30210831 (programming)

Timmy Yuen (search information/documentation)

**Investigation of the common password file protection techniques**

1.One-way function: The system stores only the value of a function based on

the user's password. When the user presents a password, the system transforms

that password and **compares** it with the stored value. In practice, the system

usually performs a one-way transformation (not reversible) in which the password is used to generate a key for the one-way function and in which a

fixed-length output is produced.

2.Access control: Access to the password file is limited to one or a very few accounts (only user who is “admin” can access into).

Moreover, It depends on what type of files you want to password protect. If you mean general documents such as Office files (Word, Excel, PPT), PDF etc, using

encryption method, then you can set a password with the internal feature equipped with. But the setback is you can't do it in batch. Otherwise, a good way is to put them in a Folder lock software, then you have one password to unlock all files in the volume.

Have lots of random symbols, capitals and numbers and no full words. For

example: S/!.7;]'@mc5{0}$)j€. It can much enhance the password strength and

then increase security of the password files.

**Investigation of the prevailing password cracking techniques**

1. **Dictionary attack**

Dictionary attack is a technique used by most of the regular hackers to determine the passphrase by trying their luck many times. Unlike its name, it works like a dictionary it is just a simple file which contains some unusual words that many people use as their password it is one of the easy ways to hack into anyone’s account.

1. **Phishing**

Phishing is the easiest method used by hackers. It does nothing, it simply ask users for their passwords but the process of asking password is unique and different, Hackers used to create the fake page, fake emails, fake apps etc. It simply asks you to log in with your Id and password and once you have entered the details, your details are transferred into hacker’s server.

1. **Bruce force attack**

It will try its level best and try every possible combination until the password is found. It takes Time.But You Can Easily Found if There Any Short Length Password. But nowadays, people are smarter, so the growing size of passwords is making Brute Force attack difficult to crack any password. It Quite similar to Dictionary attack You Can Say That the upgraded version of Dictionary attack.

1. **Rainbow table attack**

Rainbow Table is usually a large dictionary which contains loads of pre-calculated hashes and the passwords from which they were calculated. The major difference between Rainbow and other dictionary attacks are Rainbow table are specially optimised for hashes and passwords.

**Investigation of the common programming languages used in hackers tool and pros,cons**

Java has some very important features which make it very useful for hacking.

**Pros:**

1. definitely clearer, simpler, easier to write and hence more maintainable and understandable.

2.Comes with over 1,000 modules and many more are available in different other repositories.

3. some pre-built libraries which provide some great functionality.

4.outstanding performance, compile faster to hack.

**Cons:**

**1.** Bigger projects can be difficult to compile and build.

**2.Development is more expensive than in PHP or Python.**

PHP is most popular dynamic programming language, unlike JavaScript It is server-side programming language. PHP is strongly recommended to every beginner in Hacking and Penetration testing.

**Pros:**

1. open source, free to use,flexible
2. highly effective at accessing with various database type,perfect for web-based scripts

**Cons:**

1. Is a Interpreted Languages, and poor error handling also slower speed of execution.
2. Not suitable for making desktop applications

SQL is database programming language. Each and every data is stored in database so you should know about database programming and vulnerability .

**Pros:**

1. High Speed: SQL Queries can be used to retrieve large amounts of records from a database quickly and efficiently.
2. No Coding Required:  
   Using standard SQL it is easier to manage database systems without having to write substantial amount of code.

**Cons:**

1. some databases go for proprietary extensions to standard SQL to ensure vendor lock-in.

Part 5, 6, 7: Victor Pin

**5**, Well, for this program, I am using PHP to fulfill the requirement. I know that PHP is not a quiet good language for cracking, ( PHP is a script language which runs on server, C language and python are more efficient ) The reason for me to use PHP on this assignment is that I am familiar with PHP, and I am going to use different approach to access this assignment. Due to my understanding and the following reasons:

Well, the word “ crack” might be a wrong word here, because due to my research, the

encrypted md5 file is theoretically irreversible, which means it’s impossible for people to use reverse calculation to access the original file. The common approach to “crack” md5 file is exhaustive search. Which means use random algorithm to generate all the possible string combinations, encrypted them into md5 string. Because of the integrity of md5 based on hash algorithm, each string has only one corresponding md5 string, which means people can compare the md5 string with generated md5 string, if they are equal, then the original generated string is the target string, in the other word, the md5 collision. But this is such an inefficient way because we know that most of the password or user name strings are logical like “abc123”, “superman2012” but the computer generated random string has no logic because the computer can’t be think like human beings ( unless CS2421 is a machine learning or artificial intelligence course). And because of the resource constrains, the md5 collisions take a lot of resources and memories to do logical calculations, and a lot of resource is needed, it may take days to collision a 10 digit string for my poor laptop. So my approach for this assignment is, to find the “dictionary”, which is the collection of weak password and common password.

and user name, encrypted them to md5 and put them as related keys and values into the database, when the user is searching a md5 string with my program, I will just check if the target string are in the database, if the value matches, then return the user the related key of the md5 string, what is more, the dictionary may contains millions records, so the database optimization will be a serious part of this program.

1. **System Specifications**

Minimum configuration:  
Operating system: Windows XP (SP3) +, Red Hat Linux 5.0+.  
Processor: 2.2GHz Dual Core.

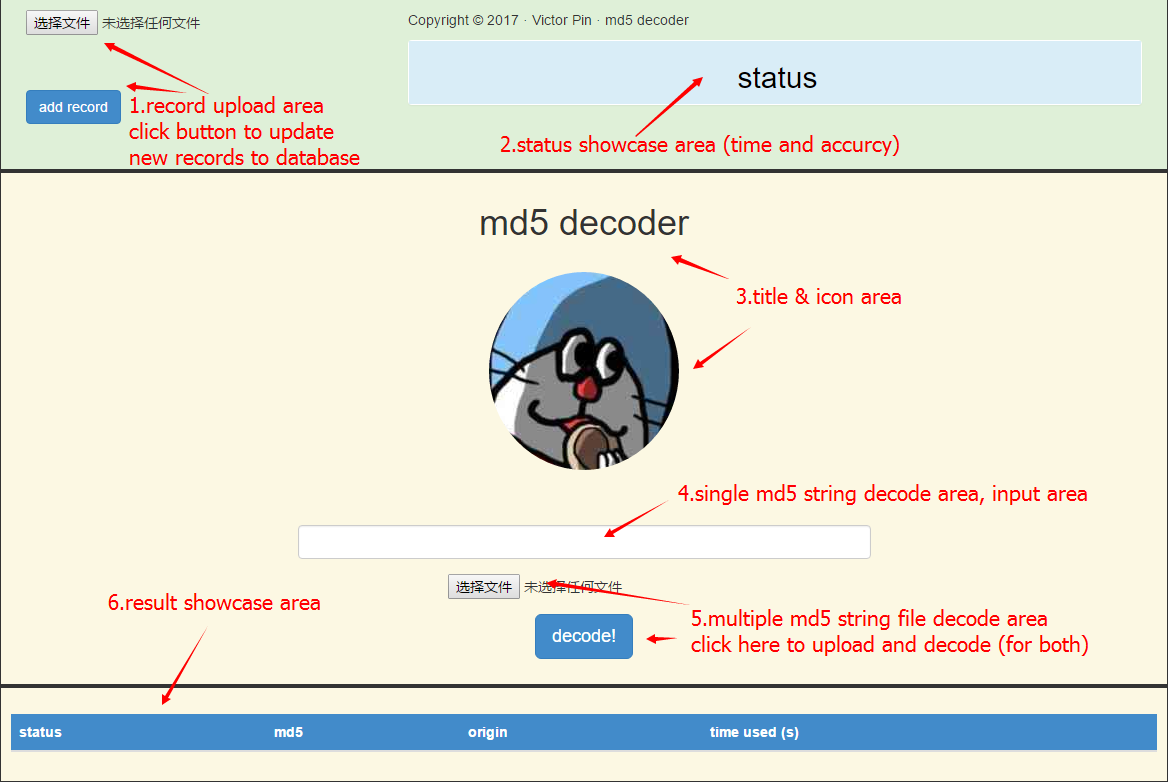
Environment: PHP 5.0 +.

Database: MYSQL DB.

Server: Apache 2.0 +.  
Memory: 2 GB RAM.  
Storage space: 500 MB of free space is required.

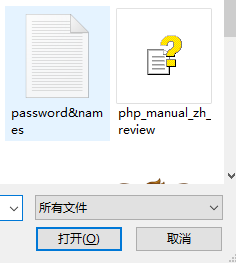
1. **Functional Specifications**

My decoder is a browser/client based PHP program that allow the user to access with webpage, what he needs is a local web server and the runnable PHP environment.

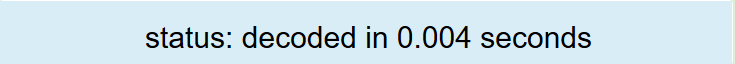


The system of the program has one panel, which is divided into 6 areas:

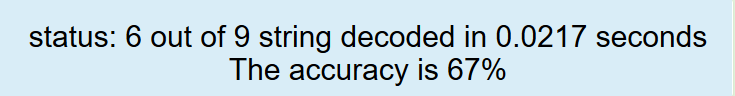
1. Record upload area: allow user to upload file contains plain text (password and username), than update it with corresponding md5 cipher text to the database.



1. Status Area: Once the cipher text been input, or uploaded, the status will show if the cipher text is decoded and the time taken.



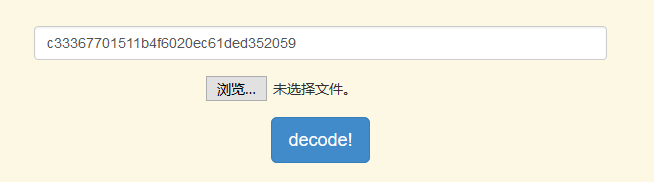
For multiple cipher texts, the accuracy of how many string is decoded will be shown.

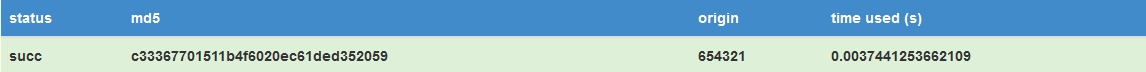


The newly added records will also be shown here.

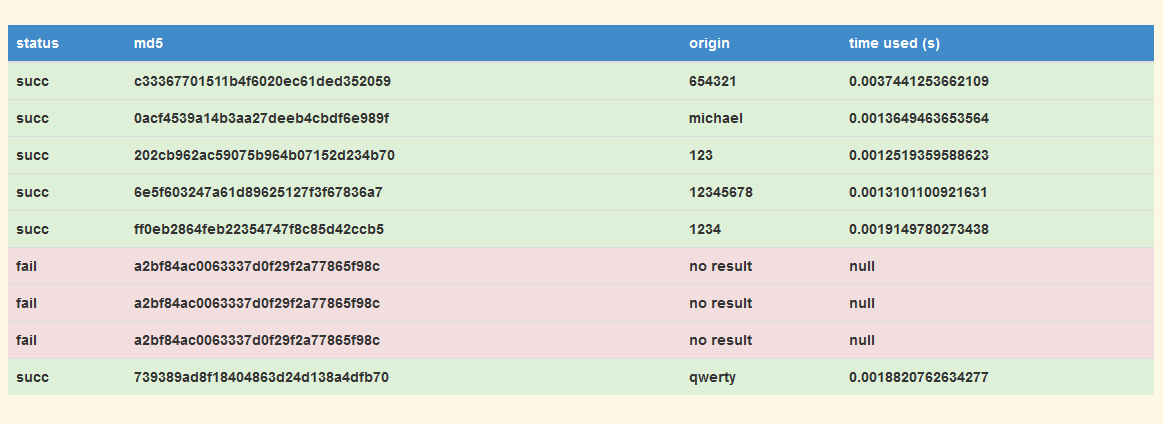


1. Title & logo: A cute logo, isn’t it?
2. Input Area: This area allows user to input single md5 cipher text, and click the “decode!” button to search for the corresponding plain text.





1. Cipher text upload Area: The user can upload a txt file contains multiple md5 cipher text string, once the user finish upload the file. The program will read though the file, put the md5 strings into database for comparing and return the result to the status panel and result area.
2. Result area: This area will display the decoded results, for the decoded string, the plain text and time taken will be shown, and nothing will be shown for the no decoded string.



1. Flow diagram

