Hints & Guides

- Some useful materials related to this project
  - PHP
  - MySQL
  - phpMyAdmin
  - Base64 encoding
  - Frequency Analysis
  - XOR encryption
  - Hash function and Hash Collision
  - robots.txt
  - Temporary files & Git

PHP

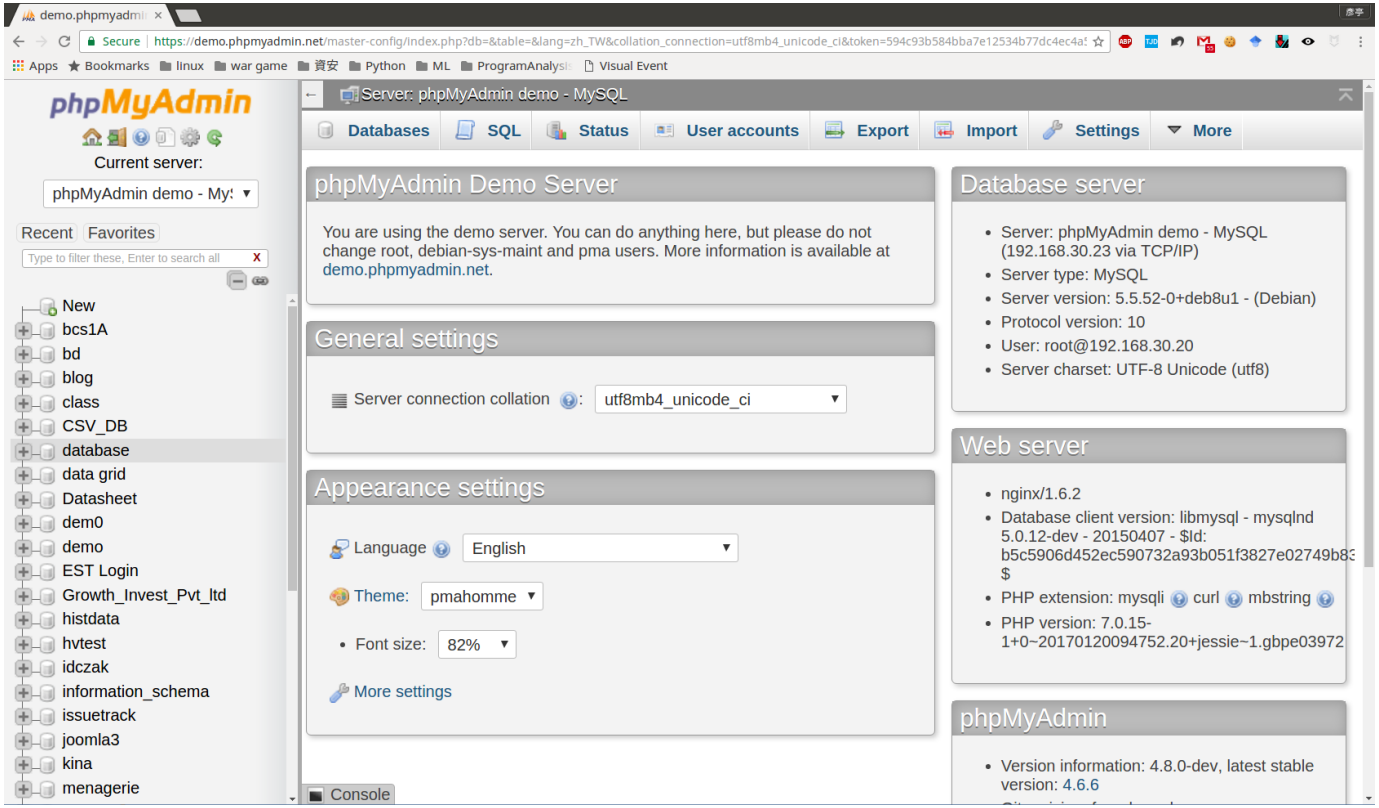
- *“PHP is a popular general-purpose scripting language that is especially suited to web development. Fast, flexible and pragmatic, PHP powers everything from your blog to the most popular websites in the world.”* --- PHP Official Website
- Bob builds his website with PHP.

MySQL

- MySQL is a very famous relational database management system.
- Bob uses MySQL to store his blog data.

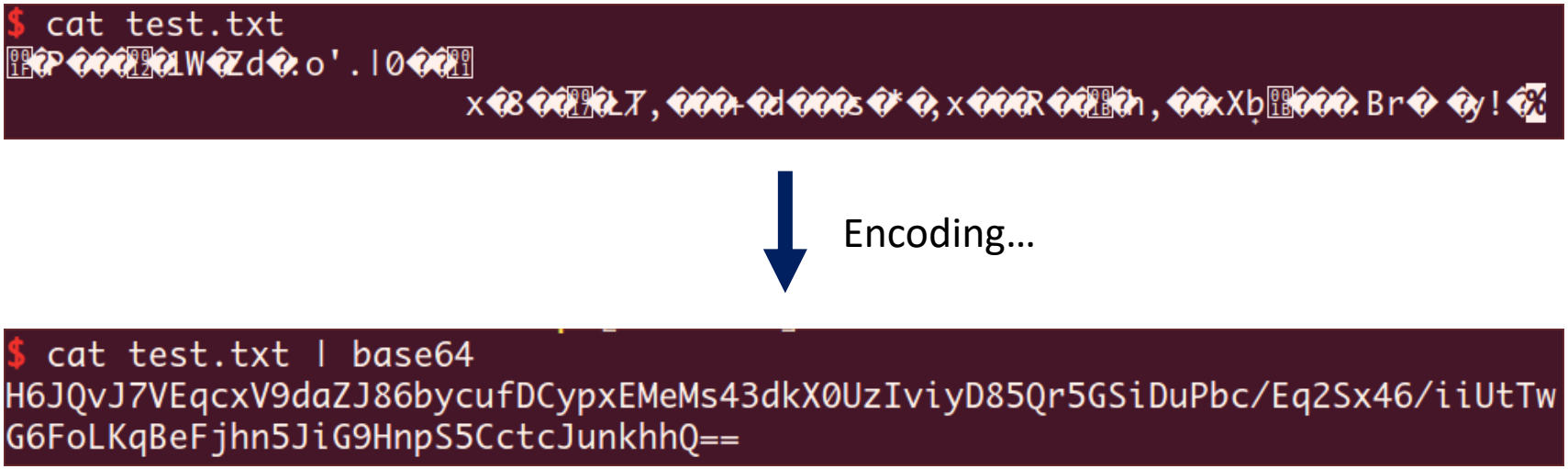
phpMyAdmin

- phpMyAdmin is a free software tool written in PHP, intended to handle the administration of MySQL over the web.



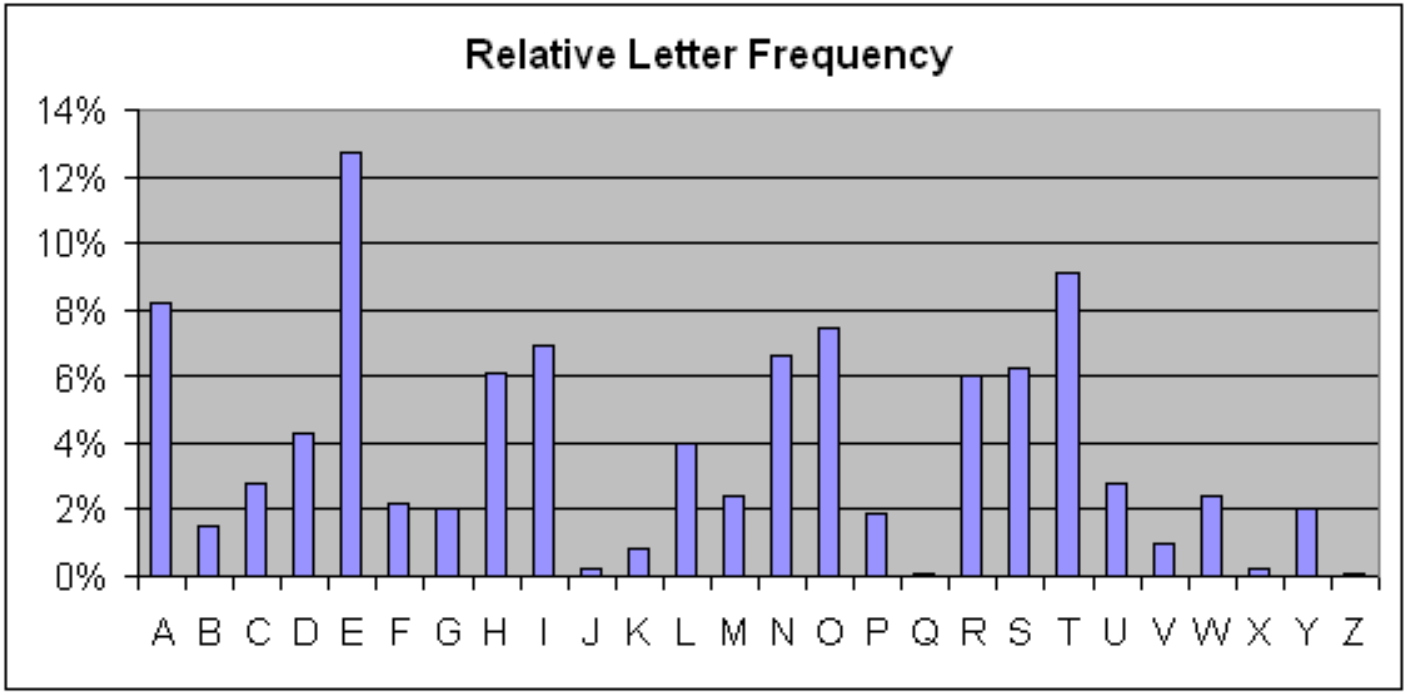
Base64 Encoding

- Base64 is a group of similar binary-to-text encoding schemes.



Frequency Analysis

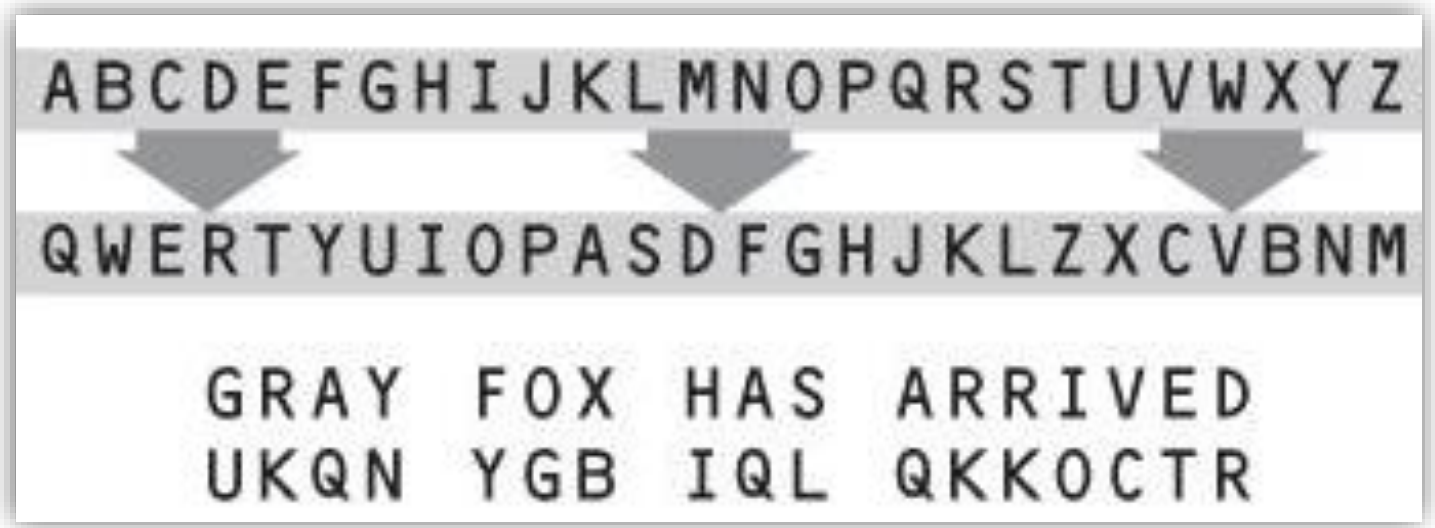
- Frequency analysis is an useful method to break some simple encryption method over certain plaintext (such as english language).



[Image from]: https://en.wikibooks.org/wiki/Cryptography/Frequency_analysis

Frequency Analysis(cont.)

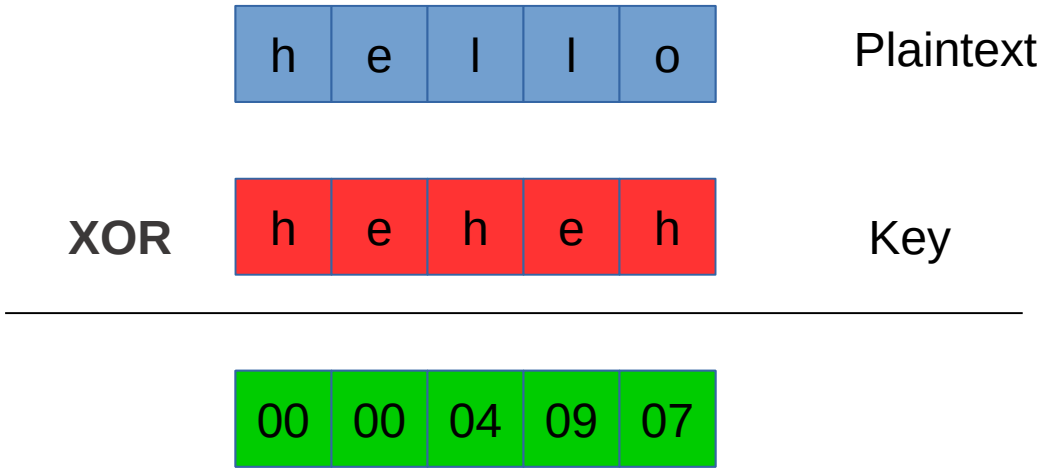
- Example:
  - Substitution cipher



[Image from]: http://www.infosectoday.com/Articles/Intro_to_Cryptography/Introduction_Encryption_Algorithms.htm

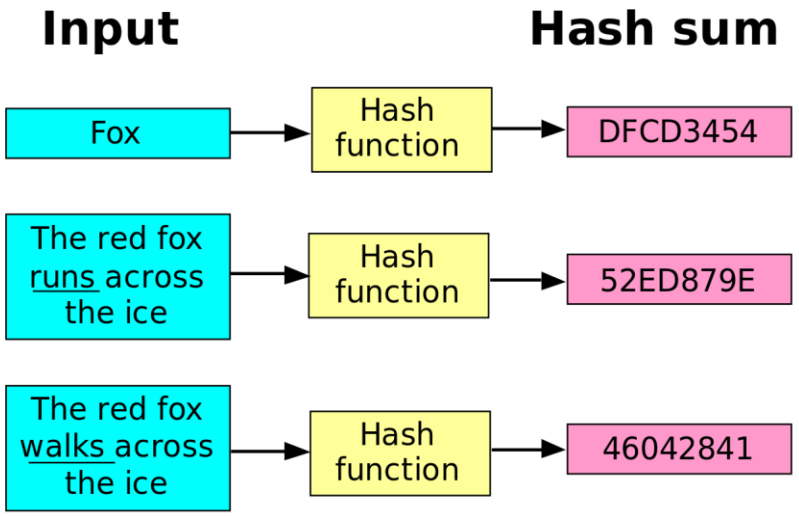
XOR Encryption

- XOR the key and the plaintext to get ciphertext.
- Example:
  - Plaintext: hello
  - Key: he



Hash Function

- A hash function is any function that can be used to map data of arbitrary size to data of fixed size(digest).
- Example: md5, sha1, sha256...etc.



[Image from]: https://zh.wikipedia.org/wiki/%E6%95%A3%E5%88%97%E5%87%BD%E6%95%B8

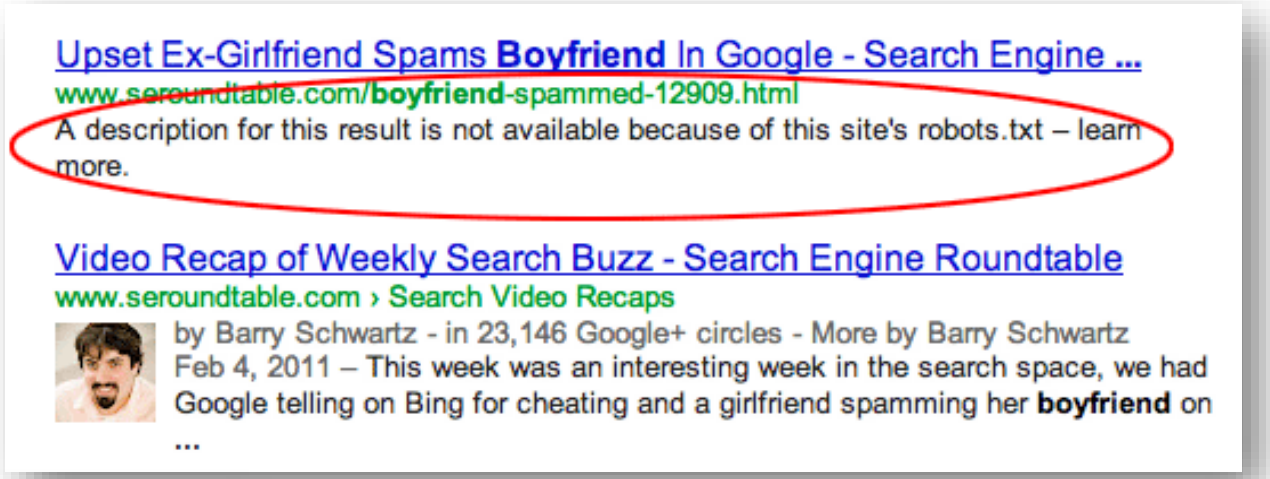
Hash Collision

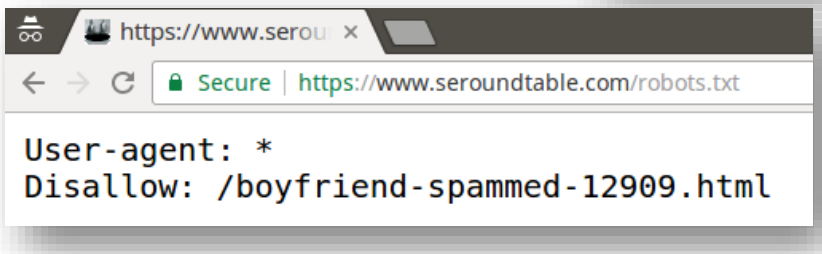
- If we can find an input ***x*** for a given hash function ***h*** and hash digest ***d***, such that ***h***(***x***) = ***d***, it's called **hash collision**.
- Simply, h(x1) = h(x2)
- Example:

```
1  def my_hash(s):
2      return sum(bytearray(s, 'utf-8')) % 25
3
4  if my_hash('test') == my_hash('heln'):
5      print('Hash collision')
```

robots.txt

- We use robots.txt to inform the search engine crawlers or robots about which files or path of the website should not be scanned or accessed.





[Image from]: https://www.seroundtable.com/boyfriend-spammed-12909.html

Temporary Files & Git

- Some sensitive files may leak a lot of information if you put them on the public web server.
- How to prevent?
  - Remove useless temporary files
  - Prevent sensitive data from being accessed
- Examples:
  - Git repo folder
  - Backup files (xxx.bak, xxx.old...etc.)
  - Vim editor temporary files
  - Other temporary files (xxx.tmp...etc.)