Semaine 8 : Post-incident
IPL I317B Sécurité : labo

Olivier Choquet & Thibault Vanwersch

10 novembre 2020

Partie 1 : les hash de mot de passe

J’ai une terrible nouvelle à vous apprendre ! La Bank de fer a été piraté ! On sait pas trop comment mais nos informateurs ont trouvé un dump de la base de donnée des utilisateurs sur le dark-moodle-web !

Pour ce labo, nous allons utiliser l’outil John The Ripper qui est « un outil Open Source d'audit de la sécurité des mots de passe et de récupération des mots de passe ». En moins politiquement correct, on pourrait dire que c’est un cracker de mot de passe.

Télécharger *John The Ripper* en version 1.9 (ou plus) : https://www.openwall.com/john/

- Pour Windows, vous pouvez récupérez des binaires :
  - Téléchargez l’archive *1.9.0-jumbo-1 64-bit* en zip ou 7z
  - Extrayez l’archive et lancer un shell (shell WSL, cmd.exe ou même powershell) dans le dossier « run » où vous pouvez executer : *.\john.exe --help*
- Sur Ubuntu, vous allez devoir le compiler :
  - `sudo apt-get install build-essential libssl-dev`
  - `sudo apt-get install yasm libgmp-dev libpcap-dev libnss3-dev libkrb5-dev pkg-config libbz2-dev zlib1g-dev`
  - `wget https://www.openwall.com/john/k/john-1.9.0-jumbo-1.tar.gz`
  - `tar xf john-1.9.0-jumbo-1.tar.gz`
  - `cd john-1.9.0-jumbo-1/src/`
  - `./configure`
  - `make -s clean && make -sj4`
  - `cd ../run/`
  - `./john --help`

Exercices :

1. Dans un premier temps, essayons de voir quels sont les mots de passe les plus triviaux.
    1. Téléchargez une liste des 1050 mots de passe les plus courant sur le github suivant : https://github.com/danielmiessler/SecLists/tree/master/Passwords/Common-Credentials
    2. Lancez une bruteforce sur les mots de passe qui ont fuité :
       $ ./john --format=bcrypt --wordlist=best1050.txt les_hash.txt
2. Pour récupérer quelques mots de passes supplémentaire, il n'est pas délirant d'essayer les logins comme mot de passe (certains utilisateurs négligeant pourraient avoir utilisé la même valeur pour les deux champs). Créez votre propre wordlist contenant la liste des logins utilisateurs et testez là sur les hashs.
3. Il est temps maintenant de sortir l'artilerie (plus) lourd : la célèbre liste de mot de passe *RockYou* que vous trouverez en tar.gz dans le dossier *Passwords/Leaked-Databases/* du même github.
    - Relancer john avec celle-ci.
    - Qu'elle est l'ETA de john ? (si vous tapez sur retour à la ligne pendant l'execution de john, il vous affiche des informations sur l'état courant.)
    - Avez-vous l'espoir de finir le test de cette liste avant la fin du cours ?
4. Tous les hashs de ce dump commencent par le même « truc ». À quoi correspond ce « truc » et quel problème cette répétition pose-t-elle ?
5. Comme leurs formes le laisse suggérer, les mots de passe sont hashé avec bcrypt-blowfish qui est volontairement lent. Que ce serait-il passé si on avait utilisé une autre fonction de hash comme par exemple md5 ou sha256 ?
    - Vous pouvez utiliser la fonction de test de john pour avoir des valeurs empiriques a comparer :
      .\john.exe --test --format=Raw-MD5
      .\john.exe --test --format=Raw-SHA256

Partie 2 : les logs serveurs

Pour cette deuxième partie du labo, nous allons investiguer des logs de serveur web. Lors du premier TP, vous avez eu un petite aperçu du trafic sur les sites internets du cours hors des périodes scolaire … place maintenant à des fichiers de logs un peu plus gros !

Pour ces exercices, le choix des outils est libre, vous pouvez utiliser du python pour parcourir ces fichiers mais il est également possible d'utiliser les outils linux/bash plus classique pour ce genre de tâche : grep, wc, cut, awk, count, sort, uniq, ...

Exercice

1. Plusieurs période de trafique anormal ont eu lieu pendant la période de temps reprise dans ce fichier, comptez les requêtes pour chaque jour et affichez un classement des jours les plus actif (ayant reçu le plus de requête)
2. Un rapide coup d'œil au contenu du fichier semble montrer que beaucoup de requêtes venaient d'un nombre restraint d'adresse IP. Affichez un deuxième classement mais celui-ci pour les adresses ip ayant effectué le plus de requêtes.
3. Tant qu'à classer des trucs, pouvez-vous essayer d'identifier la page qui a reçu le plus de requêtse ? C'est peut-être à cause d'elle que le site reçoit autant de trafique ...
4. (Bonus) Si vous avez fait ces exercices en python, recommencer en bash.