Résumé de l'article de recherche *Telepathwords*

Nicola Spanti

Janvier 2017

Table des matières

1 Article de référence

Ce document est un résumé en français d'un article de recherche en anglais. L'article en question est "Telepathwords : Preventing Weak Passwords by Reading Users' Minds". Il est issu de la collaboration de *Carnegie Mellon University* (Saranga Komanduri, Richard Shay, and Lorrie Faith Cranor) et *Microsoft Research* (Cormac Herley and Stuart Schechter). Il a été présenté durant le 23e *Security Symposium* de USENIX.

— ISBN : 978-1-931971-15-7
— Date : 20-22 août 2014
— Page web : `https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/komanduri`
— Article (en PDF) : `https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-komanduri.pdf`
— Résumé audio en anglais (MP3) : `https://2459d6dc103cb5933875-c0245c5c937c5dedcca3f1764ec`
`ssl.cf2.rackcdn.com/sec14/komanduri.mp3`
— Résumé vidéo en anglais (MP4) : `https://2459d6dc103cb5933875-c0245c5c937c5dedcca3f1764ec`
`ssl.cf2.rackcdn.com/sec14/komanduri.mp4`

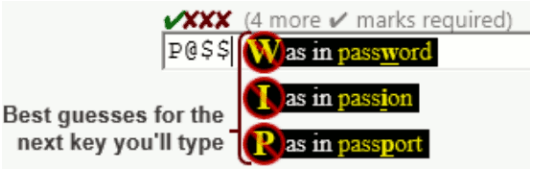

FIGURE 1 – L'interface graphique proposée

2 Résumé

2.1 Introduction

Les mots de passe sont une méthode d'authentification courante. Malheureusement les individus choisissent souvent des mots de passe prédictibles (mots du dictionnaire, remplacement de la lettre "o" par le nombre zéro, etc). Pour tenter de remédier à ce problème, des chercheurs proposent *Telepathwords*. Ce système tentent de deviner quel mot de passe un individu va taper, et l'informe de ses prédictions. Cela permet d'inciter ou obliger à s'enregistrer avec un mot de passe jugé imprévisible par le système. En d'autres termes, celui-ci permet de définir des bonnes conduites ou des régles pour un mot de passe. Il a été comparé avec d'autres pour évaluer son efficacité en fonction de différents critères.

2.2 Interface utilisateur

À côté de la zone de saisie du mot de passe, l'individu est informé des 3 lettres que le système juge les plus probables qu'il tape, celles-ci sont barrés pour mettre en avant qu'il est conseillé de les éviter. Il est également affiché la raison qui a déterminé à chaque prédiction de lettre. De plus, au dessus de chaque lettre tapée, il est indiqué si c'était ou pas une des lettres anticipées par le système (ou un de ses substituts courants comme l'arobase "@" pour la lettre "a"). Au dessus de la zone de saisie, il peut également être notifié un nombre minimum de lettres non anticipées à ajouter pour que le mot de passe soit accepté. Ainsi, l'utilisateur ou l'utilisatrice est informé·e en temps réel de la robustesse de son mon mot de passe, tout en sachant pourquoi, comme l'illustre la figure 1.

2.3 Architecture

Telepathwords utilise une architecture client-serveur. Cela permet d'avoir beaucoup de données du côté du serveur pour faire la meilleure prédiction possible. En effet, les tests ont été faits avec 1,5 Go de données. De plus, dans le cadre du Web (dans lequel l'usage des mots de passe est courant), une telle quantité de données est actuellement inacceptable.

Le mot de passe final sera envoyé en clair au serveur, lui envoyer les versions intermédiaires n'est donc pas problématique. Il faut néanmoins chiffrer le transport. On peut noter qu'il y a un risque de latence avec cette architecture, on peut l'amoindrir avec un cache.

2.4 Algorithmes de prédiction

Plusieurs algorithmes de prédiction sont utilisés. Il donne des prédictions et des raisons qui sont notées. Ainsi uniquement les mieux notées sont présentées à la personne tapant un mot de passe.

Ils utilisent un modèle de données optimisé pour prendre peu de place et être rapide à parcourir vers l'avant (c'est-à-dire les prochaines lettres). De plus, les majuscules et les espaces ont été enlevés.

Les algorithmes mis en place permettent de détecter :

— *les séquences de caractères courantes*, que ce soit à l'intérieur d'un mot (par exemple "prédi" pourrait être suivi de "re" ou "iction") ou à la suite d'un mot (par exemple "vie" pourrait être suivi par "privée" ou "intime"), avec la gestion des séparations par des nombres ou des caractères non alphanumériques (pour gérer des cas comme "pa1234ssword" et "12x34y678z9") et les substituts courants (comme le dollar "$" pour un "s")
— *la proximité des caractères sur un clavier*, ce qui permet d'éviter "123456" ou "azerty", mais cela suppose de connaitre la disposition du clavier
— *les répétitions*, par exemple "xyabcabcabc" ou "abcdefabc"
— *les mots écrits un caractère sur 2*, comme "p*a*s*s*w*o*r*d" ou "ppaasswwoorrdd"

2.5 Limitations

— La base de données pour faire les prédictions est un élément central du système. Mais il n'est pas aisé de trouver un corpus à jour des mots les plus utilisés dans les mots de passe, ainsi que de gérer toutes les langues et dispositions de clavier.
— L'architecture client-serveur peut poser des problèmes de latence. De plus, elle nécessite d'utiliser une puissance de calcul bien plus grande que d'autres systèmes visant à pousser à définir un mot de passe robuste, il en est de même pour la mémoire (stockage de la base de données, mise en RAM d'une partie, et le potentiel cache des résultats).
— Il n'y a pas de gestion des séquences de caractères inversés (comme "gfedcba" au lieu de "abcdefg").
— Le système ne garde pas en mémoire les mots de passe définis. Cela évite une potentielle fuite, mais ne permet pas au système d'apprendre de lui-même de nouveaux comportements.

2.6 Comparaisons avec d'autres systèmes

Des individus ont participé à tester *Telepathwords* et d'autres systèmes visant à définir un mot de passe robuste. Pour chaque système, des informations en temps réel étaient données à l'utilisateur ou l'utilisatrice. Après la saisie d'un mot de passe correct, des questions étaient posées pour évaluer le système auquel l'individu a été confronté, ses habitudes vis-à-vis des mots de passe et sa localisation géographique. Ensuite il doit resaisir le mot de passe, il a 5 essais. Il lui est demandé de retenir le mot de passe, puisqu'il lui est demandé 2 jours plus tard (avec 5 essais). Pour finir, l'individu est questionné sur sa façon de retenir le mot de passe et s'il stocke ou non ses mots de passe.

L'unique élément qui variait (aléatoirement) selon l'individu était la règle qui validait un mot de passe. 2 règles (telepath et telepath-v) étaient issues de *Telepathwords*, avec elles le mot de passe devait contenir au moins 6 caractères non prédis (y compris le premier caractère qui ne peut pas être prédit), la différence entre les 2 est qu'une affiche le mot de passe tandis que l'autre le cache (avec des ronds ou des carrés). La règle basic8 nécessite au minimum 8 caractères pour être satisfaite. 3class8 exige au moins 8 caractères et l'usage de 3 classes de caractères parmi 4 (minuscules, majuscules, chiffres, symboles). 3class12 est comme 3class8 mais pour au moins 12 caractères. 3class8-d n'admet de mot d'un dictionnaire en plus des conditions de 3class8.

Il a été constaté que le temps d'enrollement moyen est plus long avec telepath et telepath-v qu'avec les autres règles (jusqu'à 3 fois plus). La difficulté perçue pour trouver un mot de passe est similaire entre 3class8-d, telepath, telepath-v, et plus grande qu'avec les autres. La perception de sécurité des mots de passe est moindre avec basic8 et 3class8, les règles de *Telepathwords* ne l'améliorent par rapport à 3class12 et 3class8-d. L'aide visuelle de telepath et telepath-v est jugée meilleure que les autres pour la compréhension. D'après les testeurs et testeuses, 3 règles permettent d'une manière identique ou presque d'aider à définir un mot de passe robuste, il s'agit de 3class12, telepath et telepath-v.

En sécurité, on évalue souvent la sécurité d'un système en fonction de son maillon le plus faible. Avec ce postulat, on peut évaluer la sécurité d'une règle de mot de passe en fonction du mot de passe le plus faible qu'elle engendre. 2 algorithmes de mesure de la sécurité de mots de passe ont été utilisés. Avec les 2, le plus faible de ceux faits avec telepath et telepath-v est meilleur qu'avec les autres règles.

2.7 Conclusion

Telepathwords informe mieux l'individu sur la qualité. De plus, les mots de passe faits avec ses règles sont significativement meilleures que les règles sans dictionnaire. Malgré que la définition du mot de passe soit perçue plus difficile, ce n'est pas le cas de sa mémorisation. Il est efficace pour renforcer le maillon le plus faible pour les mots de passe.