Personalized Password Guessing

Phuong Cao, Hongyang Li, Adam Slagell, Klara Nahrstedt, Zbigniew Kalbarczyk• {pcao3,hli52,slagell, klara, kalbarcz}@illinois.edu

Coordinated Science Laboratory & National Center for Supercomputing Applications • University of Illinois at Urbana-Champaign

Background

Text-based passwords

Users often choose simple, dictionary-based passwords

Attacks targeting passwords

The number of password leaks are increasing recently, e.g., LinkedIn (6M), Adobe (250M).

Existing work on passwords

Existing works measure construction rules and strength of passwords without taking into account of user profiles and the service that requires passwords.

Threat Model

Offline targeted password cracking

1. An attacker obtained a hash password of a user and attempt to reveal the plain-text password from the hash.

2. The attacker collects user information from the Internet (e.g., social media) or social engineering:

For example: birth place, hobbies, occupation, favorite movies/song, etc.

3. The attacker generates a list of words related to the user profile, applies a number of transformational rules, and combines them to generate a personalized password.

4. The hashed value of the personalized passwords are compared with the obtained hash password.

Research Question

Question

Is personalized password cracking is a better approach compared to dictionary based or random guess approaches?

Goal

Evaluate the viability of cracking a user password faster using the user personal information obtained from the Internet or social engineering.

Hypothesis

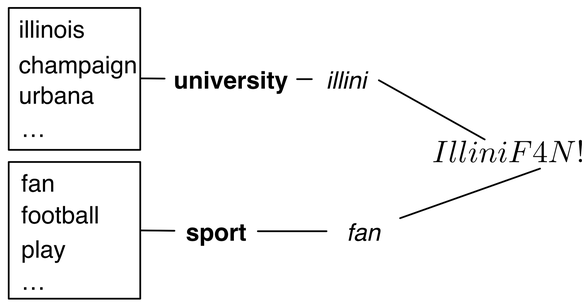
A user generates a password based on a list of words, each is drawn from a topic related to the user.

Approach

Using probabilistic graphical model to capture relationships between the user password and the user/service profile

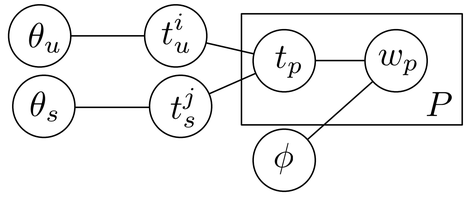
Running survey on real users to validate our hypothesis

System Workflow

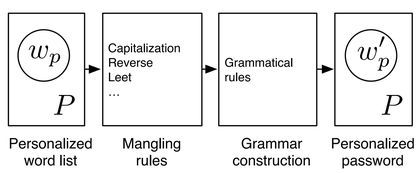


An example process of generating a password. A user who likes sport and studies at University of Illinois chooses two topics for the password: study and university. Two personalized words **illini** and **fan** are selected from the topics. The words are transformed using grammatical and mangling rules to the final password **IlliniF4N!**

Probabilistic Model



Generating P personalized words. A word wp is selected from a topic tp. The topic is based on a list of user profile topics tu and a list of service topics ts. A topic distribution is parameterized by θ and a word-topic distribution is parameterized by ϕ.



Generating a personalized password from P personalized words based on grammatical and mangling rules.

Applications

1. Suggest personalized, secure, and easy to memorize passwords to users.

2. Measure strength of a password using personalized metrics

Experiments

Personalized password guessing

Design: using questionnaire forms to collect user profiles and ask users to generate passwords for several types of services.

Goal: identify the relationships between the user/service profile and the generated passwords.

Personalized password generation

Design: using a sign-up or change password form to suggest personalized passwords to a user.

Goal: suggest personalized passwords and measure strength of a password provided by a user.

Conclusion

A probability based framework to automate generation of personalized passwords based on user and service profiles

Two experiments are designed to:

+ Evaluate the viability of cracking a user password faster using user/service profile.

+ Suggest personalized, secure, and easy to memorize passwords to users.

Acknowledgements

This work was supported in part by NSF grant CNS 10-18503 CISE and 1314891, by the Army Research Office under Award No. W911NF-13-1-0086, the National Security Agency (NSA) under Award No. H98230-14-C-0141, Air Force Research Laboratory and the Air Force Office of Scientific Research, under agreement No. FA8750-11-2-0084.

References

[1] Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wang, X. The Tangled Web of Password Reuse. In *NDSS* (2014).

[2] Koller, D., and Friedman, N. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[3] Yan, J. J., Blackwell, A. F., Anderson, R. J., and Grant, A. Password memorability and security: Empirical results. *IEEE Security & privacy 2*, 5 (2004), 25–31.