Articles for Oral Presentation

COMPSCI 725, S2 2015
Version 1.01 of 2015-07-21: corrected Rizwan's name
Clark Thomborson, Rizwan Asghar

Instructions:

- Please read through the abstracts of these articles.
- Select three which you'd be willing to present orally.
- Send an email to Clark on cthombor@cs.auckland.ac.nz, listing your three choices by surname of the last author and the year of publication e.g. "Komanduri 2014, Silver 2014, Juels 2013".
- Be careful to include your UPI in all course-related email. Your instructor may not be able to guess your University identity, if you're using an external email account.
- You should list your highest preference first.
- Clark will assign up to three students to an article, in a first-come-first-served fashion.
- You may delay sending your email until the end of the enrolment period (on Friday of the second week of lectures). However any delay will lessen your chances of being assigned your first preference.

Passwords

1. [Komanduri 2014] Saranga Komanduri, Richard Shay, Lorrie Faith Cranor, Cormac Herley, and Stuart Schechter. "Telepathwords: Preventing weak passwords by reading users' minds." In 23rd USENIX Security Symposium (USENIX Security 14). San Diego, CA: USENIX Association, pp. 591-606. 2014. [Download]

   Abstract: To discourage the creation of predictable passwords, vulnerable to guessing attacks, we present Telepathwords. As a user creates a password, Telepathwords makes real-time predictions for the next character that user will type. While the concept is simple, making accurate predictions requires efficient algorithms to model users' behavior and to employ already-typed characters to predict subsequent ones. We first made the Telepathwords technology available to the public in late 2013 and have since served hundreds of thousands of user sessions.

   We ran a human-subjects experiment to compare password policies that use Telepathwords to those that rely on composition rules, comparing participants' passwords using two different password-evaluation algorithms. We found that participants create far fewer weak passwords using the Telepathwords-based policies than policies based only on character composition. Participants using Telepathwords were also more likely to report that the password feedback was helpful.

2. [Silver 2014] David Silver, Suman Jana, Eric Chen, Collin Jackson, and Dan Boneh. "Password managers: Attacks and defenses." In Proceedings of the 23rd Usenix Security Symposium. 2014. [Download]

Abstract: We study the security of popular password managers and their policies on automatically filling in Web passwords. We examine browser built-in password managers, mobile password managers, and 3rd party managers. We observe significant differences in autofill policies among password managers. Several autofill policies can lead to disastrous consequences where a remote network attacker can extract multiple passwords from the user's password manager without any interaction with the user. We experiment with these attacks and with techniques to enhance the security of password managers. We show that our enhancements can be adopted by existing managers.

3. [Juels 2013] Ari Juels, and Ronald L. Rivest. "Honeywords: Making password-cracking detectable." In Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS), pp. 145-160. ACM, 2013. [Download]

   Abstract: We suggest a simple method for improving the security of hashed passwords: the maintenance of additional "honeywords" (false passwords) associated with each user's account. An adversary who steals a file of hashed passwords and inverts the hash function cannot tell if he has found the password or a honeyword. The attempted use of a honeyword for login sets off an alarm. An auxiliary server (the "honeychecker") can distinguish the user password from honeywords for the login routine, and will set off an alarm if a honeyword is submitted.

4. [Kelley 2012] Patrick Gage Kelley, Saranga Komanduri, Michelle L. Mazurek, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Julio Lopez. "Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms." In 2012 IEEE Symposium on Security and Privacy (SP), pp. 523-537. IEEE, 2012. [Download]

   Abstract: Text-based passwords remain the dominant authentication method in computer systems, despite significant advancement in attackers' capabilities to perform password cracking. In response to this threat, password composition policies have grown increasingly complex. However, there is insufficient research defining metrics to characterize password strength and using them to evaluate password-composition policies. In this paper, we analyze 12,000 passwords collected under seven composition policies via an online study. We develop an efficient distributed method for calculating how effectively several heuristic password-guessing algorithms guess passwords. Leveraging this method, we investigate (a) the resistance of passwords created under different conditions to guessing; (b) the performance of guessing algorithms under different training sets; (c) the relationship between passwords explicitly created under a given composition policy and other passwords that happen to meet the same requirements; and (d) the relationship between guessability, as measured with password-cracking algorithms, and entropy estimates. Our findings advance understanding of both password-composition policies and metrics for quantifying password security.

5. [Wright 2012] Nicholas Wright, Andrew S. Patrick, and Robert Biddle. "Do you see your password? Applying recognition to textual passwords." In Proceedings of the

Eighth Symposium on Usable Privacy and Security (SOUPS), p. 8. ACM, 2012. [Download]

> Abstract: Text-based password systems are the authentication mechanism most commonly used on computer systems. Graphical passwords have recently been proposed because the pictorial-superiority effect suggests that people have better memory for images. The most widely advocated graphical password systems are based on recognition rather than recall. This approach is favored because recognition is a more effective manner of retrieval than recall, exhibiting greater accuracy and longevity of material. However, schemes such as these combine both the use of graphical images and the use of recognition as a retrieval mechanism. This paper reports on a study that sought to address this confound by exploring the recognition of text as a novel means of authentication. We hypothesized that there would be significant differences between text recognition and text recall conditions. Our study, however, showed that the conditions were comparable; we found no significant difference in memorability. Furthermore, text recognition required more time to authenticate successfully.

Identification and Authentication

6. [Polakis 2014] Iasonas Polakis, Panagiotis Ilia, Federico Maggi, Marco Lancini, Georgios Kontaxis, Stefano Zanero, Sotiris Ioannidis, and Angelos D. Keromytis. “Faces in the distorting mirror: Revisiting photo-based social authentication.” In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS), pp. 501-512. ACM, 2014. [Download]

   > Abstract: In an effort to hinder attackers from compromising user accounts, Facebook launched a form of two-factor authentication called social authentication (SA), where users are required to identify photos of their friends to complete a log-in attempt. Recent research, however, demonstrated that attackers can bypass the mechanism by employing face recognition software. Here we demonstrate an alternative attack that employs image comparison techniques to identify the SA photos within an online collection of the users' photos.
   >
   > In this paper, we revisit the concept of SA and design a system with a novel photo selection and transformation process, which generates challenges that are robust against these attacks. The intuition behind our photo selection is to use photos that fail software-based face recognition, while remaining recognizable to humans who are familiar with the depicted people. The photo transformation process creates challenges in the form of photo collages, where faces are transformed so as to render image matching techniques ineffective. We experimentally confirm the robustness of our approach against three template matching algorithms that solve 0.4% of the challenges, while requiring four orders of magnitude more processing effort. Furthermore, when the transformations are applied, face detection software fails to detect even a single face. Our user studies confirm that users are able to identify their friends in over 99% of the photos with faces unrecognizable by software, and can solve over 94% of the challenges with transformed photos.

7. [Thomas 2014] Kurt Thomas, Dmytro Iatskiv, Elie Bursztein, Tadek Pietraszek, Chris Grier, and Damon McCoy. “Dialing back abuse on phone verified accounts.” In

Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS), pp. 465-476. ACM, 2014. [Download]

> Abstract: In the past decade the increase of for-profit cybercrime has given rise to an entire underground ecosystem supporting large-scale abuse, a facet of which encompasses the bulk registration of fraudulent accounts. In this paper, we present a 10 month longitudinal study of the underlying technical and financial capabilities of criminals who register phone verified accounts (PVA). To carry out our study, we purchase 4,695 Google PVA as well as pull a random sample of 300,000 Google PVA that Google disabled for abuse. We find that miscreants rampantly abuse free VOIP services to circumvent the intended cost of acquiring phone numbers, in effect undermining phone verification. Combined with short lived phone numbers from India and Indonesia that we suspect are tied to human verification farms, this confluence of factors correlates with a market-wide price drop of 30-40% for Google PVA until Google penalized verifications from frequently abused carriers. We distill our findings into a set of recommendations for any services performing phone verification as well as highlight open challenges related to PVA abuse moving forward.

8. [Xie 2012] Yinglian Xie, Fang Yu, Qifa Ke, Martín Abadi, Eliot Gillum, Krish Vitaldevaria, Jason Walter, Junxian Huang, and Zhuoqing Morley Mao. “Innocent by association: Early recognition of legitimate users.” In Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS), pp. 353-364. ACM, 2012. [Download]

> Abstract: This paper presents the design and implementation of Souche, a system that recognizes legitimate users early in online services. This early recognition contributes to both usability and security. Souche leverages social connections established over time. Legitimate users help identify other legitimate users through an implicit vouching process, strategically controlled within vouching trees. Souche is lightweight and fully transparent to users. In our evaluation on a real dataset of several hundred million users, Souche can efficiently identify 85% of legitimate users early, while reducing the percentage of falsely admitted malicious users from 44% to 2.4%. Our evaluation further indicates that Souche is robust in the presence of compromised accounts. It is generally applicable to enhance usability and security for a wide class of online services.

9. [Perito 2011] Daniele Perito, Claude Castelluccia, Mohamed Ali Kaafar, and Pere Manils. “How unique and traceable are usernames?” In Privacy Enhancing Technologies, LNCS 6794, pp. 1-17. Springer Berlin Heidelberg, 2011. [Download]

> Abstract. Usernames are ubiquitously used for identification and authentication purposes on web services and the Internet at large, ranging from the local-part of email addresses to identifiers in social networks. Usernames are generally alphanumerical strings chosen by the users and, by design, are unique within the scope of a single organization or web service. In this paper we investigate the feasibility of using usernames to trace or link multiple profiles across services that belong to the same individual. The intuition is that the probability that two usernames refer to the same physical person strongly depends on the “entropy” of the username string itself. Our experiments, based on usernames gathered from

real web services, show that a significant portion of the users' profiles can be linked using their usernames. In collecting the data needed for our study, we also show that users tend to choose a small number of related usernames and use them across many services. To the best of our knowledge, this is the first time that usernames are considered as a source of information when profiling users on the Internet.

10. [Sun 2011] San-Tsai Sun, Eric Pospisil, Ildar Muslukhov, Nuray Dindar, Kirstie Hawkey, and Konstantin Beznosov. “What makes users refuse web single sign-on?: an empirical investigation of OpenID.” In Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS), p. 4. ACM, 2011. [Download]

    Abstract: OpenID is an open and promising Web single sign-on (SSO) solution. This work investigates the challenges and concerns web users face when using OpenID for authentication, and identifies what changes in the login flow could improve the users' experience and adoption incentives. We found our participants had several behaviors, concerns, and misconceptions that hinder the OpenID adoption process: (1) their ex- isting password management strategies reduce the perceived usefulness of SSO; (2) many (26%) expressed concerns with single-point-of-failure related issues; (3) most (71%) held the incorrect belief that the OpenID credentials are being given to the content providers; (4) half exhibited an inability to distinguish a fake Google login form, even when prompted; (5) many (40%) were hesitant to consent to the release of their personal profile information; and (6) many (36%) ex- pressed concern with the use of SSO on websites that contain valuable personal information or, conversely, are not trust- worthy. We also found that with an improved affordance and privacy control, more than 60% of study participants would use Web SSO solutions on the websites they trust.

Access Control

11. [Birgisson 2014] Amar Birgisson, Joe Gibbs Politz, Úlfar Erlingsson, Ankur Taly, Michael Vrable, and Mark Lentczner. “Macaroons: Cookies with contextual caveats for decentralized authorization in the cloud.” In Network and Distributed System Security (NDSS) Symposium. 2014. [Download]

    Abstract: Controlled sharing is fundamental to distributed systems; yet, on the Web, and in the Cloud, sharing is still based on rudimentary mechanisms. More flexible, decentralized cryptographic authorization credentials have not been adopted, largely because their mechanisms have not been incrementally deployable, simple enough, or efficient enough to implement across the relevant systems and devices.

    This paper introduces macaroons: flexible authorization credentials for Cloud services that support decentralized delegation between principals. Macaroons are based on a construction that uses nested, chained MACs (e.g., HMACs) in a manner that is highly efficient, easy to deploy, and widely applicable.

    Although macaroons are bearer credentials, like Web cookies, macaroons embed caveats that attenuate and contextually confine when, where, by who, and for what purpose a target service should authorize requests. This paper describes macaroons and motivates their design, compares them to other credential systems, such as cookies and SPKI/SDSI, evaluates and measures a prototype implementation, and discusses practical security and application considerations. In particular, it is considered how macaroons can enable more fine-grained authorization in the Cloud, e.g., by strengthening mechanisms like OAuth2, and a formalization of macaroons is given in authorization logic.

12. [Mondal 2014] Mainack Mondal, Yabing Liu, Bimal Viswanath, Krishna P. Gummadi, and Alan Mislove. "Understanding and specifying social access control lists." In Symposium on Usable Privacy and Security (SOUPS). 2014. [Download]

    Abstract: Online social network (OSN) users upload millions of pieces of content to share with others every day. While a significant portion of this content is benign (and is typically shared with all friends or all OSN users), there are certain pieces of content that are highly privacy sensitive. Sharing such sensitive content raises significant privacy concerns for users, and it becomes important for the user to protect this content from being exposed to the wrong audience. Today, most OSN services provide fine-grained mechanisms for specifying social access control lists (social ACLs, or SACLs), allowing users to restrict their sensitive content to a select subset of their friends. However, it remains unclear how these SACL mechanisms are used today. To design better privacy management tools for users, we need to first understand the usage and complexity of SACLs specified by users.

    In this paper, we present the first large-scale study of fine-grained privacy preferences of over 1,000 users on Facebook, providing us with the first ground-truth information on how users specify SACLs on a social networking service. Overall, we find that a surprisingly large fraction (17.6%) of content is shared with SACLs. However, we also find that the SACL membership shows little correlation with either profile information or social network links; as a result, it is difficult to predict the subset of a user's friends likely to appear in a SACL. On the flip side, we find that SACLs are often reused, suggesting that simply making recent SACLs available to users is likely to significantly reduce the burden of privacy management on users.

13. [Chronopoulos 2013] Konstantinos Chronopoulos, Maria Gouseti, and Aggelos Kiayias. "Resource access control in the Facebook model." In Cryptology and Network Security (CANS), pp. 179-198. Springer International Publishing, 2013. [Download]

    Abstract: We study the fundamental security properties of *resource access control* as suggested by the operation of current social networks including Facebook. The "facebook model", which treats the server as a trusted party, suggests two fundamental properties, "owner privacy" and "server consistency", and two different modes of revocation, implicit and explicit. Through black-box experimentation, we determine Facebook's implementation for resource access control and we analyze its security properties within our formal model. We demonstrate, by the construction of explicit attacks, that the current implementation is not secure: specifically, we attack privacy with implicit revocation and server consistency. We evaluate the implications of the attacks and

we propose amendments that can align the current implementation with all its intended security properties. To the best of our knowledge this is the first time that a security analysis of the Facebook resource access control mechanism is performed within a proper security model.

Anonymisation and De-anonymisation

14. [Caliskan-Islam 2015] Aylin Caliskan-Islam, Richard Harang, Andrew Liu, Arvind Narayanan, Clare Voss, Fabian Yamaguchi, and Rachel Greenstadt. "De-anonymizing Programmers via Code Stylometry." Web manuscript, 500 kB, 17 pp., last modified "15/02/2015 9:56:55 p.m." Available: https://www.cs.drexel.edu/~ac993/papers/caliskan_deanonymizing.pdf, 21 July 2015. Note: a version of this article has been accepted for publication at the 24th USENIX Security Symposium, August 2015.

Abstract: Source code authorship attribution could provide proof of authorship in court, automate the process of finding a cyber criminal from the source code left in an infected system, or aid in resolving copyright, copyleft and plagiarism issues in the programming fields. In this work, we investigate methods to deanonymize source code authors of C++ using coding style. We cast source code authorship attribution as a machine learning problem using natural language processing techniques to extract the necessary features. The Code Stylometry Feature Set is a novel representation of coding style found in source code that reflects coding style from properties derived from abstract syntax trees. Such a unique representation of coding style has not been used before in code attribution.

Our random forest and abstract syntax tree-based approach attributes more authors (250) with significantly higher accuracy (95%) on a larger data set (Google Code Jam) than has been previously attempted. Furthermore these novel features are more robust than previous approaches, and are still able to attribute authors even when code is run through commercial obfuscation with no significant change in accuracy. This analysis also produces interesting insights relevant to software engineering. We find that (i) the code resulting from difficult programming tasks is easier to attribute than easier tasks and (ii) skilled programmers (who can complete the more difficult tasks) are easier to attribute than less skilled programmers.

15. [Humbert 2015] Mathias Humbert, Kévin Huguenin, Joachim Hugonot, Erman Ayday, and Jean-Pierre Hubaux. "De-anonymizing Genomic Databases Using Phenotypic Traits." In 15th Privacy Enhancing Technologies Symposium (PETS), 2015. [Download]

Abstract: People increasingly have their genomes sequenced and some of them share their genomic data online. They do so for various purposes, including to find relatives and to help advance genomic research. An individual's genome carries very sensitive, private information such as its owner's susceptibility to diseases, which could be used for discrimination. Therefore, genomic databases are often anonymized. However, an individual's genotype is also linked to visible phenotypic traits, such as eye or hair color, which can be used to re-identify users in anonymized public genomic databases, thus raising severe privacy issues. For

instance, an adversary can identify a target's genome using known her phenotypic traits and subsequently infer her susceptibility to Alzheimer's disease. In this paper, we quantify, based on various phenotypic traits, the extent of this threat in several scenarios by implementing de-anonymization attacks on a genomic database of OpenSNP users sequenced by 23andMe. Our experimental results show that the proportion of correct matches reaches 23% with a supervised approach in a database of 50 participants. Our approach outperforms the baseline by a factor of four, in terms of the proportion of correct matches, in most scenarios. We also evaluate the adversary's ability to predict individuals' predisposition to Alzheimer's disease, and we observe that the inference error can be halved compared to the baseline. We also analyze the effect of the number of known phenotypic traits on the success rate of the attack. As progress is made in genomic research, especially for genotype-phenotype associations, the threat presented in this paper will become more serious.

16. [McDonald 2012] Andrew WE McDonald, Sadia Afroz, Aylin Caliskan, Ariel Stolerman, and Rachel Greenstadt. "Use fewer instances of the letter 'i': Toward writing style anonymization." In Privacy Enhancing Technologies, LNCS 7384, pp. 299-318. Springer Berlin Heidelberg, 2012. [Download]

    Abstract: This paper presents Anonymouth, a novel framework for anonymizing writing style. Without accounting for style, anonymous authors risk identification. This framework is necessary to provide a tool for testing the consistency of anonymized writing style and a mechanism for adaptive attacks against stylometry techniques. Our framework defines the steps necessary to anonymize documents and implements them. A key contribution of this work is this framework, including novel methods for identifying which features of documents need to change and how they must be changed to accomplish document anonymization. In our experiment, 80% of the user study participants were able to anonymize their documents in terms of a fixed corpus and limited feature set used. However, modifying pre-written documents were found to be difficult and the anonymization did not hold up to more extensive feature sets. It is important to note that Anonymouth is only the first step toward a tool to acheive stylometric anonymity with respect to state-of-the-art authorship attribution techniques. The topic needs further exploration in order to accomplish significant anonymity.