Privacy and E-Authentication: The Dangers of Self-disclosure in Social Networks

Divakaran Liginlal[1] and Lara Khansa[2]

[1] Carnegie Mellon University at Qatar, Information Systems Program,
PO Box 24866 Doha, Qatar
[2] Virginia Polytechnic and State University, Pamplin College of Business,
2062 Pamplin Hall (0235), Blacksburg, VA 24060, USA
liginlal@cmu.edu, larak@vt.edu

Abstract. We propose a Bayesian model of privacy in e-authentication and develop associated entropy-based metrics. A major contribution of this work is the application of weighted entropy to characterize the user's privacy preferences. Further, we model the effects of side information on privacy and relate it to self-disclosure on Internet web sites and social networks. Specifically, our empirical study of Internet users' information disclosure habits within social networks along with the theoretical results provide insights into building a regulatory framework to address privacy concerns in e-authentication.

Keywords: Privacy metrics, authentication, Bayesian, weighted entropy, social network.

1 Introduction

The act of establishing confidence in the claimed identity of a user, device, or another entity in an information or communication system is called authentication [3]. Knowledge-based authentication (KBA) is an economic method of authenticating users for on-line transactions that occur on an infrequent basis, with the potential of providing reasonably secure authentication [13, 17]. KBA consists of matching one or more pieces of information (also called factoids) provided by an individual (claimant) against information sources associated with the claimant for the purpose of verifying his/her identity. The participants of a KBA scheme include individuals or other organizations asserting identity, relying parties, and verifiers [3]. Relying parties make use of an identity credential or assertion of identity to decide what actions need to be taken in a given application context. Often they use verifiers or trusted verification services to gain confidence in the identity assertion.

The term KBA has conventionally been used [13] in the context of e-government to encompass those systems relying on proprietary sources of claimant information that do not require prior registration by the user. However, systems that use cognitive passwords [12], otherwise known as recognition-oriented KBA [7], also fall under the umbrella of KBA. With the advent of Federal Financial Institutions Examination Counsel's (FFIEC) regulations mandating strong authentication for financial transactions,

R. Sharman, T.S. Raghu, and H.R. Rao (Eds.): WEB 2009, LNBIP 52, pp. 166–176, 2010.

e-banking systems have widely incorporated KBA as a secondary authentication method, for applications such as password resets. KBA systems rely on an effective matching of a claimant's real time responses to challenge questions against the identity attribute values supplied by trusted sources. A suitable matching function would need to take into consideration the effect of an attacker correctly guessing a response (guessability metric) and the ability of a genuine user to recall and enter the correct response (memorability). Chen and Liginlal [5] provide a comprehensive literature review of KBA and a discussion of KBA metrics and model selection from a security perspective. Our objective in this paper is to address the issue of privacy in KBA.

Information privacy is defined by Westin [26] as "the ability of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated and the subsequent uses of such information." Privacy considerations in authentication have attracted researchers' attention only recently [24]. In the context of authentication, three types of privacy issues arise [20]: (1) Decisional privacy aims to prevent interference with decisions about self and family; (2) Information privacy is the individual's right to control access to and dissemination of personal information; and (3) Communication privacy deals with the confidentiality of the individual's communications. Often communication privacy is bundled with information privacy.

2 A Bayesian View of Privacy in E-Authentication

A recent definition of KBA metrics [7] follows a frequentist approach to estimating the guessability of a set of statistically independent factoids. A Bayesian approach to modeling KBA, on the other hand, confers the ability to incorporate subjective probabilities or beliefs about attacker strategies. Further, the use of conditional probabilities helps to model parameters such as guessability and memorability of factoids, while also considering factoid dependencies.

Fig. 1 shows a Bayesian network [21] model of KBA (BN-KBA). BN-KBA is a directed acyclic graph representing a probabilistic model, with each node representing a random variable that can take on two or more possible values. The arrows indicate the existence of direct effects between linked variables. The associated conditional probabilities embody the strength of these influences. The hypothesis variable 'Authentication' assumes a state from {true, false} and the evidence variables $\{e_1, e_2, \ldots, e_7\}$ take states from {correct, wrong}. The values of the evidence variable correspond to whether or not factoid values are entered correctly. Associated with each node is a probability distribution to be specified in advance. The conditional probabilities $P(e_i = \text{correct} \mid \text{Authentication} = \text{true})$ represent the likelihood of a genuine user recalling and entering the evidence correctly, or in other words the memorability associated with the factoid. Similarly, $P(e_i = \text{correct} \mid \text{Authentication} = \text{false})$ represents the likelihood of an imposter guessing the value of the evidence correctly or, in other words, the guessability of the associated factoid. The nodes representing external knowledge sources such as the Social Security Administration (SSA), the Credit Bureaus (CB), and the Internal Revenue Service (IRS) take the states {trusted, untrusted} with associated probabilities denoting their trustworthiness.

In Fig. 1, we also depict the actual factoid variables taking on states corresponding to all possible values for a factoid in a specific KBA domain. We treat these factoid variables as hidden nodes, which are actually needed only for parameter estimation and factoid selection. Once we have selected a set of factoids, we can absorb the effect of the states of the knowledge sources and the factoid variables and perform the authentication step with only the evidence and hypothesis variables. The major strength of BN-KBA is that probabilistic inference can be made directly from the conditional probabilities. Thus, the posterior P(Authentication = true | e), where e is the set of all evidence variables, computed by applying Bayes' rule, determines the level of assurance one may gain from the authentication process. We refer the reader to [5] for details about parameter estimation and factoid selection for BN-KBA.

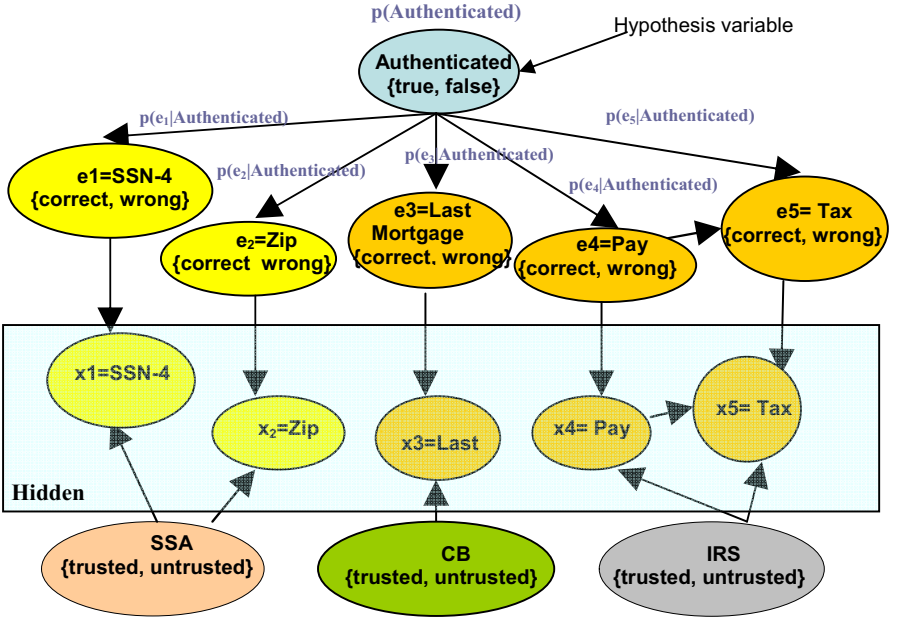

Fig. 1. A model of BN-KBA depicting the factoid variables and knowledge sources

2.1 Entropy Reduction in BN-KBA

The privacy characteristic of KBA is better understood by examining the nature of data embedded in the factoids. The factoids may be classified based on whether they are static or dynamic, and personal or not personal. Examples of static factoids are one's mother's maiden name and date of birth. Dynamic factoids, such as credit card balances, last paycheck amount, and last mortgage paid, exhibit the most dependencies. Another classification of factoids is based on whether they are personal or not personal. Personal factoids, such as favorite color or movie, are in general unique to a person. Examples of nonpersonal factoids are cell phone number and home address. For e-government applications, proprietary databases such as the social security

administration (SSA), credit bureau (CB), and the internal revenue service (IRS) serve as reference sources for verifying the response to user challenges. In general, data from such sources are either static or dynamic. One may also add data sourced directly from individuals, also referred to as personal entropy [10]. These require the user's prior registration with either the relying party or a verifier.

To better understand the privacy implications of KBA, one needs to also dissect the attacker strategies (See [3] for details). We consider, for the purpose of our discussion, a rational attacker, who knows all probable factoid values involved in a KBA domain deduced from public information about databases or through insider information or other attacks. The attacker may try out each value with equal probability, or the attacker knows the probability distribution of the factoid values and exploits that knowledge in the attack, or the attacker tries out the most probable values. If an attack is successful, the recovery from the disclosure of static information, whether personal or not personal, is often very hard to achieve. Also, if an imposter successfully authenticates as a genuine user, this would lead to access to transaction logs and other profile information of the user and ultimately identity theft.

Bayesian conceptualization of privacy. We consider an n-factoid vector **x**. Each univariate x_i in **x** follows a discrete probability distribution $p(x_i)$. Privacy metrics can be developed by considering the distribution of values in x_i, the attacker strategy, i.e., the information at the attacker's disposal, and the disutility [23] for the owner of the identity (id) for each x_i. Our assumption is that disutility can help factor in a user's assessment of the likelihood of a privacy breach that will be damaging to an individual in one form or another, i.e., reputation, financial, social, and legal.

Information entropy is a measure of the uncertainty associated with a random variable. The Bayesian Network in Fig. 1 may be considered as an information channel. When the factoid variables, denoted by the nodes marked x_1 to x_5, are instantiated with their corresponding values, information flows to remove the uncertainty associated with the hypothesis variable. This entropy reduction forms the basis for an authentication decision.

2.2 Privacy Metric for KBA

For our analysis, we adopt Shannon entropy that measures the average information content in a random variable [8].

Definition 1. Given a random variable x with an associated probability distribution $\{p_j\}_{j = 1, \ldots, k}$ satisfying $p_j \geq 0_{j = 1, \ldots, k}$ and $\sum p_j = 1$, Shannon entropy (or ordinary information entropy), is defined by,

$$H(x) = -c\sum_{x} p(x)\log[p(x)] \,, \tag{1}$$

where c is a positive constant, usually set equal to unity. The result is expressed in bits if the logarithm is to the base 2, and the notation H(p) is often used in place of H(x).

The joint entropy H(x, y) of a pair of discrete random variables with joint distribution p(x,y) is defined as,

$$H(x,y) = -\sum_{x}\sum_{y} p(x,y)\log p(x,y) \ , \tag{2}$$

Let $(x_1,x_2, \ldots, x_n)$ be drawn according to $p(x_1,x_2, \ldots, x_n)$. The Shannon entropy $H(x_1,x_2, \ldots, x_n)$ is defined as the sum of the individual Shannon entropies, if the x_i's are statistically independent. Otherwise, by applying the chain rule for entropy, one may express the joint entropy as a sum of the conditional entropies, i.e.,

$$H(x_1, x_2,..., x_n) = \sum_{i=1}^{n} H(x_i \mid x_{i-1},..., x_1) \ , \tag{3}$$

This leads to the following definitions of KBA privacy metrics $£_{KBA}$ at the domain level and $£_{KBA,\,id}$ at the identity level.

Definition 2 (Domain-level KBA privacy). In order to define domain-level KBA privacy ($£_{KBA}$), we consider the average amount of private information available about KBA users with respect to a selected set of factoids. We quantify $£_{KBA}$ by the Shannon entropy, of the selected n-factoid vector $\mathbf{x} = \{x_1 \ldots x_n\}$, i.e.,

$$£_{KBA} = H(p_x) \ , \tag{4}$$

If the factoids are independent, $£_{KBA}$ is the sum of the individual entropies of the factoids in **x**. It is important to note that dependencies only reduce the uncertainty associated with a factoid, thus leading to the notion of maximal privacy.

Given that privacy is a term more often associated with an individual's control over his/her personal information, it will be more appropriate to also define KBA privacy at the identity level. Each individual may have preferences about what factoids to be used in an authentication session, partly based on an estimate of the likelihood of a privacy breach. The concept of weighted Shannon entropy [1] provides a way of incorporating the disutilities, associated with user preferences.

Definition 3. Given a disutility vector $\{d_j\}_{j\,=\,1,\,\ldots,\,k}$ satisfying $\{d_j \geq 0\}_{j\,=\,1,\,\ldots,\,k}$ associated with each event indexed by j in the discrete probability distribution p_j, the weighted Shannon entropy [11] is defined by,

$$Hw(x) = -c\sum_{x} d_j p(x)\log\left[p(x)\right] \ , \tag{5}$$

Weighted entropy has been applied to a variety of contexts, including pattern recognition and econometric modeling [16]. For a factoid x_i, assume the j^{th} value is associated with an individual id. Let id assign a disutility d_j to the factoid. Then the weighted information associated with the factoid instance j is $I_d(d_j, p_{xij}) = -kw_j \log p_{xij}$ [1]. This weighted information measure also helps quantify the notion of personal entropy associated with an individual's factoid.

Consider a disutility vector $d_j = \{d_{1j} \ldots d_{nj}\}$, $\{d_{ij} \geq 0\}_{j\,=\,1\,\ldots\,n}$ associated with a factoid vector $x = \{x_1 \ldots x_n\}$ for an individual user identified by id. Let $\chi_j = \{\chi_{1j}, \ldots, \chi_{nj}\}$ represent the instance level values for id.

Definition 4 (Identity-level KBA privacy). In order to define identity-level KBA privacy ($£_{KBA,id}$), we consider the amount of private information available about a specific KBA user denoted by id, with respect to a selected set of factoids. $£_{KBA,id}$ under factoid independence (maximal privacy) can be quantified as the sum of the weighted information of the factoid instances as follows,

$$£_{KBA,id} = \sum_i £_{i,id} = -k \sum_i d_{ij} \log\left[p\left(x_i = \chi_{ij}\right)\right] , \tag{6}$$

Although we implicitly assume that the factoids are independent, the actual amount of privacy reduction due to the dependencies is determined by the mutual information across all the factoids.

We now show that the definition of KBA domain-level privacy is meaningful to our discussion of privacy and captures in essence the privacy expectation of KBA users considered as a whole.

Theorem 1. If users' privacy preferences are invariant across all identities and neutral across all factoids, $£_{KBA}$ is upper bounded by the expected value of $£_{KBA,\,id}$, aggregated over all identities.

Proof. If users' privacy preferences are invariant across all identities, their disutility vectors are the same, given a set of factoids. In addition, if users are neutral in their privacy preferences across all selected factoids, the disutility vector can be represented as di = {c, …, c}, where $c \geq 0$ is a constant. Under this condition, the expected value becomes,

$$\begin{aligned} E_{id}\left(£_{KBA,id}\right) &= E_{id}\left(\sum_i £_{i,id}\right) = -k_1 \sum_i \sum_j p\left(x_i = \chi_{ij}\right) \log\left[p\left(x_i = \chi_{ij}\right)\right] \\ &= -k_1 \sum_i H\left(p(x_i)\right) = -k_1 \sum_i p(x_i) \log\left(p(x_i)\right) \end{aligned} , \tag{7}$$

where $k_1 = kc$. In the definition of entropy it is customary to set the scaling value $k_1 = 1$. Thus, $£_{KBA} = H(p_x) \leq \sum_i H\left(p(x_i)\right) = -k_1 \sum_i p(x_i) \log\left(p(x_i)\right) = E_{id}\left(£_{KBA,id}\right)$. This concludes the proof.

3 Side Information and Its Influence on Privacy

The term side information is used in a variety of contexts such as in communication systems [22], financial markets [4, 15], and cryptography [2] to mean auxiliary information that is correlated to a significant variable of interest in a stochastic process.

3.1 Modeling Side Information in KBA

In the case of KBA, strategic knowledge about a KBA domain, or about specific identities gleaned from various sources, constitutes information that a potential adversary may exploit. Electronic sources of side information include home pages, web search

engines, email archives, social networking sites, and people search provider services. Inference attacks on identities, based on profile information publicly available about individuals, constitute a rich source of data given the proliferation of personal home pages and social networking sites [6, 14]. For the following discussion, we assume that side information is modeled as a random variable u_i correlated with a factoid x_i.

Definition 5. The expected privacy leak due to side information is equal to the amount of mutual information I(xi,ui) [8] in xi and ui.

Theorem 2. Side information reduces domain-level KBA privacy.

Proof. The proof follows from the definition of mutual information $I(x_i,u_i) = H(x_i) - H(x_i|u_i) \geq 0$.

Theorem 3. KBA privacy reduction in a factoid x_i due to side information u_i is upper bounded by $H(u_i)$.

Proof. The proof follows from a basic property of mutual information, i.e., $I(x_i,u_i) \leq H(u_i)$.

3.2 The Link to Information Disclosure

Information disclosure on a variety of Internet sites inflates the dangers of reduced privacy (and security) for knowledge-based authentication. Weighted entropy metrics allow the KBA designer to incorporate a user's preferences based on his/her estimates of the likelihood of a privacy breach at design time. The definition of these measures does not take into account the fact that a user is often unaware of auxiliary information that exists about a particular factoid due to self-disclosure or disclosure by others both online and in other social contexts. The uncertainty representing such auxiliary information is factored into the attacker strategies in the form of the guessability metric, which has relevance during a KBA challenge-response session. On the other hand, for factoid selection it will be useful to study online sources of side information that may be exploited to initiate identity-level attacks. Common examples are personal home pages, blogs, chat rooms, and social networks.

One may come up with three approaches to filtering identity-related information on the web. The first approach is one in which a user agent searches the web to find matching profiles and checks whether the selected factoids have already been disclosed. Involving the user interactively helps enhance the accuracy of the results. Such filtering methods could plausibly use ontology-based search tools such as Google's APIs. The second approach is targeted at social networks to harvest user profiles based on community interests and to infer side information about users. Chen and Liginlal [6] apply state-of-the-art statistical machine learning techniques to fit a classification model to textual data collected from MySpace and use the learned classifier to predict critical personal attributes, e.g., hobbies, salary, and occupation, of previously unseen user examples.

In this paper, we discuss an approach based on a user agent deployed to search small worlds [19] formed in popular social networks around communities conforming

to the specialized interests of a KBA user. We present results from our small world experiments on MySpace and Friendster to study the dangers of self-disclosure and ways of estimating the prior probabilities associated with side information.

3.3 The Dangers of Self Disclosure

Watts et al. [25] present a model that offers an explanation of social network searchability in terms of recognizable personal identities, i.e., sets of characteristics measured along a number of social dimensions. They demonstrate that their model can account for Milgram's [19] experimental findings that short paths of average size 6 exist between individuals in a large social network and that ordinary people are able to find these short paths.

We used the following methodology to build small worlds within a social network.

1. Repeat steps (i) through (iv) N times
   (i) Randomly (or through profile matching) select a user from the network, such that the selected user has about 150 immediate friends and has not been picked in an earlier step. The number 150 conforms to the Dunbar number which is a theorized cognitive limit to the number of individuals, with whom any one person can maintain stable social relationships [9]. Save the profile of the selected user.
   (ii) Randomly select the user's network of immediate friends from individuals whose settings are not private, who have a significant number (10% of the Dunbar number) of immediate friends, and who were not chosen in an earlier step. Save the profile of the selected friends.
   (iii) Repeat (recursively) step ii up to five hops for each selected friend, i.e. six interconnected friends from the root user in step i. This means that if the network cannot go to 6 levels we abandon our attempt and repeat the entire sequence i-iii.
   (iv) Parse the saved profiles of users constructed in steps i-iii. Create frequency counts corresponding to information disclosure trends for each attribute of interest to us from a KBA perspective.
2. Average the frequency counts of each attribute over all small worlds.
3. Repeat the entire experiment in steps 1 and 2, choosing new seeds for the random number generator used for sampling. Accumulate the information disclosure statistics and average over all experiments.

3.4 Results of Empirical Study

Table 1 shows the results of our experiments. We built in each experiment an average of 22 small worlds with N=100. The average number of users sampled in each experiment was 10,112 for MySpace and 9,871 for Friendster, out of which we found that nearly 31% users for MySpace and 29% users for Friendster had either restricted privacy settings or number of friends not displayed. The results from the publicly searchable space clearly point to the dangers of information revelation within social networks.

First, we test the hypothesis that self disclosure habits of individuals in small worlds are invariant across the two social networks. We instrumentalize the test by

Table 1. Average frequency count of factoids disclosed in small worlds in MySpace and Friendster

Factoid	MySpace Frequency Count %	Friendster Frequency Count %
Favorite Music	50.04	47.24
Favorite Movie	43.31	44.78
Favorite TV Show	38.73	40.53
Favorite Book	36.54	34.14
Favorite Sport	0.13	0.24
Favorite Color	0.41	0.72
Street grown up	0.00	0.01
Pet's name	0.01	0.03
Model of first car	0.04	0.02
Zipcode	0.00	0.01
High School Mascot	0.00	0.00

comparing the frequency count of factoid values disclosed in the two networks under consideration. The null hypothesis that the difference across both networks is insignificant could not be rejected (t-value = 0.014; p-value > 0.1).

The four attributes that are directly searchable are favorite music, favorite movie, favorite TV show, and favorite book. A large percentage of users revealed this information. Confidence in the trustworthiness of this information is high because we ensured in our sampling that each user had a significant number of friends in their immediate network and they were part of a small world that emanated from a user with at least 150 friends and extended five hops in diameter. We parsed the other fields such as "Interests" and "About Me" searching for information related to favorite sport, favorite color, street in which a user grew up, pet's name, model of first car, zipcode, and high school mascot. The results were not as significant as in the case of the directly searchable attributes. This could be partially attributed to the limitations of our search method although, we attempted to manually cross-validate our findings while using variations in the search tokens used. Regardless, the findings show that favorite sport, favorite color, pet's name, and model of first car were also disclosed by a small percentage of users.

3.5 Implications of the Study

The theoretical and empirical results from this research provide several useful guidelines for building a regulatory framework encompassing the three important stakeholders in an authentication process.

From the perspective of the relying party there is an implicit assumption in authentication that 'user consent' has been given for acquisition and storage of verifying data and its 'fair' usage. This assumption is fraught with dangers. Regulations mandating strong authentication need to specifically ensure that 'user consent' for acquiring and using personal data for the purposes of authentication doesn't necessarily entail loss of control by the user, resulting in privacy breaches with lifelong implications.

Regulations need to emphasize the need for proactive risk assessment in the context of authentication, selection of a risk minimization method possibly limiting the amount of

personal information the user is required to provide, and raising the level of awareness for administrators of authentication systems in the privacy implications. Verifiers and third party credential services need to be strictly regulated in terms of the safeguards for database protection. There is a need for federal privacy laws not only to complement state laws but also to provide stiffer sanctions for breach. The individual's control must necessarily extend to these services who may have acquired personal data from multiple sources, often without the knowledge of the owner. The individual must also have a say in what attributes can be provided to verifiers and must be able to revoke or withdraw the data. It is obvious that this is very challenging to implement practically.

Most KBA implementations require users to enter their personal information at registration time. There must, thus, be sufficient scrutiny in the manner in which factoids are acquired from personal entropy sources. Such acquisition must be orchestrated through a trusted and secure process totally within the control of the individual. In fact, instead of actual factoid data, weak keys generated from such data should be acquired and only such keys ought to be used in a challenge-response authentication [18]. Enhancing user awareness through user education must necessarily be a role played by relying parties and verifiers. While the concern of the KBA designer is to gain high levels of assurance in authentication, the user must be empowered to safeguard as well as track the dissemination of his/her private information. The KBA system may collude with the user to deploy privacy agents routinely on the web to mine websites and social networks and glean information about possible privacy breaches. Such agents must be deployed under user control so that the user has the option to trust the agent with sensitive personal information for the purpose of profile matching.

4 Conclusion and Future Research

In this paper, we considered a Bayesian view of privacy in KBA whereby privacy is quantified in terms of the information required to remove uncertainty about the hypothesis variable relative to an authentication session. This provides the basis for a definition of domain-level and identity-level KBA privacy metrics relying on information-theoretical concepts. A major contribution of our research work is the application of weighted entropy concepts that allow the incorporation of personal privacy preferences of a KBA user into the factoid selection process. Finally, we reported the study of social networking sites to show how user profiling and collaborative filtering methods can be applied to study the impact of side information on identity-level KBA security and privacy.

The empirical study reported in the paper is only meant to provide evidence of the drawbacks of side information. The important question we investigated relates to how much people really value privacy over security. This is the fundamental theme of this paper tackled from a rather theoretical perspective. An empirical study designed to differentiate between perceived and actual disclosure behavior will be useful to understand the implications of user participation in online services on KBA security and privacy. The models and metrics discussed in the context of KBA can be extended easily to multifactor authentication. Since Shannon entropy only facilitates the study of average-case effects of side information, it will be interesting to pursue alternate entropy formulations of KBA privacy. Finally, defining a unified framework for usability, privacy, and security in KBA is an important extension of this research.

References

1. Belis, M., Guiasu, S.: Quantitative-Qualitative Measure of Information in Cybernetic Systems. IEEE Transactions on Information Theory 14, 593–594 (1968)
2. Bennett, C.H., Brassard, G., Crépeau, C., Maurer, U.M.: Generalized Privacy Amplification. IEEE Transactions on Information Theory 41, 1915–1923 (1995)
3. Burr, W.E., Dodson, D.F., Polk, W.T.: Electronic Authentication Guideline: Recommendations of the National Institute of Standards and Technology. NIST Special Publication 800-63 Version 1.0.2, National Institute of Standards and Technology, NIST (2006)
4. Caroline, H.: Comparison of Insiders' Optimal Strategies Depending on the Type of Side-Information. Stochastic Processes and their Applications 115, 1603–1627 (2005)
5. Chen, Y., Liginlal, D.: Bayesian Networks for Knowledge-Based Authentication. IEEE Transactions on Knowledge and Data Engineering 19, 695–710 (2007)
6. Chen, Y., Liginlal, D.: Information Disclosure on the Internet-A Machine Learning Perspective. Working Paper, University of Wisconsin-Madison (2008)
7. Chokhani, S.: Knowledge Based Authentication (KBA) Metrics. In: KBA Symposium-Knowledge Based Authentication: Is It Quantifiable?, Gaithersburg, MD (2004)
8. Cover, T., Thomas, J.: Elements of Information Theory. Wiley, Chichester (2006)
9. Dunbar, R.I.M.: Neocortex Size as a Constraint on Group Size in Primates. Journal of Human Evolution 20, 469–493 (1992)
10. Ellison, C., Hall, C., Milbert, R., Schneier, B.: Protecting Secret Keys with Personal Entropy. Future Generation Computer Systems 16, 311–318 (2000)
11. Guiasu, S.: Weighted Entropy. Reports on Mathematical Physics 2, 165–179 (1971)
12. Haga, W.J., Zviran, M.: Cognitive passwords: from theory to practice. Data Processing and Communications Security 3, 19–23 (1989)
13. Hastings, N.E., Dodson, D.F.: Quantifying Assurance of Knowledge Based Authentication. In: 3rd European Conference on Information Warfare and Security, ECIW 2004 (2004)
14. He, J., Chu, W.W., Liu, Z.: Inferring Privacy Information from Social Networks. In: Mehrotra, S., Zeng, D.D., Chen, H., Thuraisingham, B., Wang, F.-Y. (eds.) ISI 2006. LNCS, vol. 3975, pp. 154–165. Springer, Heidelberg (2006)
15. Hillairet, C.: Comparison of Insiders' Optimal Strategies Depending on the Type of Side-Information. Stochastic Processes and their Applications 114, 1603–1627 (2005)
16. Karmeshu, J.: Entropy Measures, Maximum Entropy Principle and Emerging Applications. Springer, Heidelberg (2003)
17. Lawler, B.: Models of Knowledge Based Authentication (KBA). In: KBA Symposium-Knowledge Based Authentication: Is It Quantifiable?, Gaithersburg, MD (2004)
18. Lowry, S.: Challenge & Response within E-Authentication Framework. In: KBA Symposium-Knowledge Based Authentication: Is It Quantifiable?, Gaithersburg, MD (2004)
19. Milgram, S.: The Small-World Problem. Psychology Today 1, 61–67 (1967)
20. Millett, L.I., Holden, H.S.: Authentication and its Privacy Effects. IEEE Internet Computing 6, 54–58 (2003)
21. Pearl, J.: Probabilistic Reasoning in Intelligence Systems. Morgan, San Maleo (1988)
22. Shannon, C.E.: Channels with Side Information at the Transmitter. IBM J. Res. Develop, 289–293 (1958)
23. Theil, H.: Disutility as a Probability. Management Science 20, 109–116 (1980)
24. US National Research Council. Who Goes There? Authentication through the Lens of Privacy. Nat'l Academy Press, Washington (2003)
25. Watts, D.J., Dodds, P.S., Newman, M.E.J.: Identity and Search in Social Networks. Science 296, 1302–1305 (2002)
26. Westin, A.: Privacy and Freedom. Atheneum, New York (1967)