Poster: Toward a Theoretical Privacy Framework for Electronic Locks in Context of Home Security Monitoring System for Clouds of Things

Depeng Li
Department of Information and
Computer Sciences
University of Hawaii at Manoa
1680 East West Road, Honolulu, HI,
USA
depengli@hawaii.edu

ABSTRACT

Current popular schemes e.g. homomorphic cryptography are extensively deployed to preserve privacy in a limited level but without a formal privacy model, we can neither offer privacy guarantee nor quantify the privacy loss. In this paper, we raise a few privacy-related questions, one after another, with the e-Lock state changes in a smart home as an example. In a novel privacy framework we proposed, the questions are partially addressed with the utilization of a set of theoretical models e.g. hidden markov model, differential privacy and information flow with belief. Since our paper is still at its start phase, we plan to accomplish the framework and wish can inspire colleagues' interests in this area.

Categories and Subject Descriptors

K.4.1 [**Computer and Society**]: Public Policy Issues – *Privacy*;

General Terms

Measurement, Documentation, Theory.

Keywords

Electronic Lock, Formal Privacy Model, Privacy Preservation.

1. INTRODUCTION

In smart homes, proliferation of sensors and actuators equipped with communication capability enable occupants to remotely access or control an array of automated home electronic devices by entering a single command or a PIN number [1]. A current trend is that smart home, with occupancies' consent, could cooperate with third parties such as a home security monitoring company to get 24/7 security protection service. However, the participatory sensing application could potentially reveal the true data of occupants - it is highly possible that they are misused without well-designed privacy preservation.

The privacy violation for smart home [2] - even for some simple appliance such as electronic locks (e-Lock) [3], - is a pressing challenge today and increasingly affects all occupancies given the fact that captured true data can be misused to infer personal activities. The insight is based on the observation that some intermittent activities such as e-Lock switched on/off could possibly infer personal absence/presence at the smart home.

A power replay attack [3] is explored against most popular and commercially endorsed electronic lock. However, privacy preservation and privacy analyses (e.g. [4]) for e-Lock in context of smart home security monitoring system has not been presented. More importantly, it lacks a formal model of privacy analysis – privacy guarantee and quantitative evaluation are desirable. The pertinent privacy related questions should be addressed regarding time-series e-Lock state data: (1) How much privacy is lost and to what extend if all e-Lock states in smart home are open for access? (2) If perturbation methods which introduce uncertain noise to true personal data is deployed, could aggregator still queries leak privacy? (3) If the aggregator (e.g. server or the cloud of the security company) is not trusted, could we protect the privacy by access control scheme based on information flow?

Our contributions: we supply a theoretical privacy framework to analyze the privacy leakage through accommodating fundamental functionality e.g. sharing / hiding, perturbation and access control for aggregated time-series e-Lock state dataset for smart home system stored in cloud of things. We not only carefully study potential privacy inference but also try to address corresponding concerns about privacy loss in case of true dataset, of distorted dataset and of dataset under information flow protection:

(1) When the switching on/off operations of e-Locks are open to access, they are treated as the input and, in turn, could be modeled as real-valued correlated Gaussian random variable. Based on that, hidden Markov Chain model is provided to measure the occupancies' absence in correlation with the e-Lock's switching on/off operations

(2) To offer privacy protection, perturbing noisy into real data is assumed. We then invoke differential privacy method to analyze significant distinguish for specific occupancies' absence of the smart home.

(3) Turning on/off operations of e-Locks are transmitted to the security companies' server or even their cloud. They may be borrowed by third party for investigation in future. It is possible that aggregator is untrusted. An attempt to utilize hyberproperty information flow is taken to shield privacy.

2. ARCHITECTURE OF E-LOCK IN SMART HOME

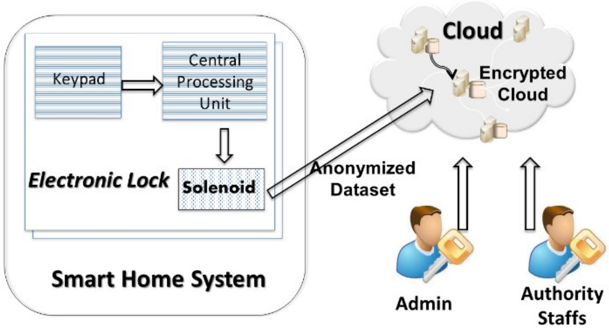

Figure 1 – Overview of e-Lock [3] in Smart Home of Cloud Things

Figure 1 demonstrates the overview of the e-Lock in context of smart home with the cloud-things services. E-Locks in a smart home includes three components – keypad, central processing unit and solenoid (actuator). If the credential inputted at keypad is valid, a signal is sent to solenoid to changes the e-Lock's state from off to on. The application generates time-series categorical data which is aggregated to the server/cloud of the security company through secure communication channel. The admin or even authority staffs (e.g. policeman) may be able to query or even access the anonymized dataset from now on.

3. PRIVACY LOSS AND THREAT MODEL

3.1 Privacy Loss Scenario

Privacy for residence occupancy: An e-Lock state change C_i can let an adversary infer that the resident is presence or absence with the support of a temporal correlation of participatory sensing data.

Example I: Alice is the only one at home and then e-Lock's state is changed. Eve can probably infer that Alice may open the door and leave or Alice has accompany. If it is the former, Eve can take the risk to break in.

Untrusted third party aggregator that peek privacy: An adversary queries the collected data-set to steal the privacy by taking advantage of the strong correlation among successive values in the series.

Example II: Eve observe that Alice left the community. Eve can query the number of e-Lock state changes in the community at two successive time slots to guess which house Alice left.

Untrusted third party: the collected dataset could be borrowed by third party to accomplish research duties such as optimization or investigation tasks such as criminal inquiry. Unveiling time-series true/raw data may violate householder's privacy.

3.2 Threat Models

Like other researches [5] in areas of privacy preservations, we assume that smart devices (e.g. e-Lock, etc.) in smart home and the cloud/server obey network communication schemes. However, both users and the aggregators could be untruthful since they can lie and they also have the intension to combine the information if possible. However we need at least a fraction of them (e.g. a majority) are honest.

4. PROPOSED PRIVACY FRAMEWORK

4.1 Markov Chain

We assume that the state of an e-Lock L_i, (where $L_i \in L = \{L_1, L_2, \dots, L_n\}$, n is the number of e-Locks in households) is sampled when there is an unlock/lock action. $L_i \in \{0,1\}$ where 0 denotes unlocked and 1 locked. Let array L_t^n denote the state of all e-Lock at time t. There are 2^n possible state of all e-Locks.

Assume there are m family memebers. The presence of each person in the household is also monitored as $P_i \in \{0,1\}$ where 0 denotes absence and 1 presence. At the time instant t, the presence of all members $\{P_1, P_2, \dots, P_m\}$ is denoted as an array P_t^m. There are 2^m possible state of the presence of all m family members.

Thus, we model the joint probability distribution of the e-Lock states and the presence state over x time instants:

$$P(L_t^n, P_t^m) = \prod_{t=1}^{x} P(L_t^n | L_{t-1}^n) P(P_t^m | L_t^n) \qquad (1)$$

Based on (1), we can deduce a hidden Markov model for presence of persons, which can be characterized by three parameters: (a) the initial presence, (b) a state distribution and (c) a conditional distributions. After defining the 3 inputs with concrete details, our hidden Markov model should assess the interrelated association between the pair (L, P) in which, array P with all elements being 0 is what both the burglar and the security company are interested in.

4.2 Differential Privacy

Let I_i denote all e-Lock state change data related with one smart home or even a community. Denote $I = \sum_i^n I_i$ which is the collected dataset related with n persons $\{ I_1, I_2, \dots, I_n\}$. We demand the following holds

$$\Pr[A(I) = x] \leq e^{\epsilon} \Pr[A(I') = x] \qquad (2)$$

where Pr is a probability distribution over randomness of algorithm $A(I)$ where I is the input, I' is the addition or removing of one single user, and x is an any value output.

Let $\boldsymbol{Q} = \{Q_1, Q_2, \dots Q_n\}$be any query sequence, we demand the following holds:

$$|Q(I) - Q(I')|_p \leq \Delta_p(Q) \qquad (3)$$

where $p \in \{1,2\}$, $Q(I)$ and $Q(I')$ are each vectors, $\Delta_1(Q)$ measures Manhattan distance $\sum_i |Q_i(I) - Q_i(I')|$ and $\Delta_2(Q)$ Euclidean distance ($\sqrt{\sum_i (Q_i(I) - Q_i(I'))^2}$).

4.3 Hyberproperty

The malicious third-party may query the true data and revise its belief from the keep going interaction thereafter [6]. An experiment $£ = < S, b_H, \sigma_H, \sigma_L >$ is processed where S is the query system, b_H denotes prebelief about high state, σ_H denotes high state and σ_L denotes low state. The third-party / agent predicts the output distribution p'_A and S produces a state $\sigma' \in \mathrm{p}' = [\![S]\!](\dot{\sigma_L} \oplus \dot{b_H})$. The agent can infer a postbelief: $b'_H = (\mathrm{p}'_A | o) \upharpoonleft H$ where o is the low projection of the output state. With £, we, instantiating Bayes' rule on these probabilities, get Bayesian inference:

$$BI(£, o) = \frac{b_H(\sigma_H) \bullet ([\![S]\!] \dot{\sigma_L} \oplus \dot{\sigma_H} \upharpoonleft H))(o)}{(\sum \sigma'_H : b_H(\sigma'_H) \bullet ([\![S]\!] \dot{\sigma_L} \oplus \dot{\sigma_H} \upharpoonleft H))(o))} \qquad (4)$$

5. DISCUSSION AND FUTURE WORKS

Our future works will centrally focus on privacy preservation via perturbing distributed noisy information to time-series e-Lock state change data to minimize the privacy loss with lower utility-privacy tradeoff. In addition, how to extend hidden Markov Chain method to precisely quantify privacy loss and corresponding counter-measures via differentially private protection will be studied.

6. REFERENCES

[1] D. J. Cook, "How smart is your home?" *Science,* vol. 335, no. 6076, pp. 1579-1581, 2012.

[2] T. Kim, L. Bauer, J. Newsome, A. Perrig, and J. Walker. "Access right assignment mechanisms for secure home networks." *Journal of Communications and Networks*, vol. 13, no. 2, pp. 175-186, April 2011.

[3] S. Oh, J. Yang, A. Bianchi, and H. Kim, "Poster: power replay attack in electronic door locks", extended abstract, IEEE Sym. Security and Privacy, pp. 1-2, 2014, San Jose, CA.

[4] C. Dwork, "Differential privacy: A survey of results." *In Theory and Applications of Models of Computation*, pp. 1-19, 2008.

[5] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption." In *2010 ACM SIGMOD* pp. 735-746. ACM, 2010.

[6] M. Clarkson, A. Myers, and F. Schneider. "Quantifying information flow with beliefs." *Journal of Computer Security* vol. 17, no. 5: pp. 655-701,2009.