ENGINEERING IN ADVANCED RESEARCH SCIENCE AND TECHNOLOGY

ISSN 2350-0174
Vol.02, Issue.11
November -2020
Pages: -241-247

PRESERVING PRIVACY IN SMGS FOR INFERENTIAL ATTACKS

[1] YADALA YAMINI [2] KAMBHAM SALIVAHANA REDDY, M. tech (Assistant professor)
[1,2] Global college of engineering and technology, Dept. of CSE

ABSTRACT: In developing countries, reliable electricity access is often undermined by the absence of supply from the national power grid and/or load shedding. To alleviate this problem, smart micro-grid (SMG) networks that are small scale distributed electricity provision networks composed of individual electricity providers and consumers, are being increasingly deployed. To ensure the reliable operation of SMGs, monitoring is necessary for data collection and state estimation processes. However, highly calibrated and trustworthy smart meters that are ideally suited to perform such monitoring tasks are often costly and non-ideally suited to SMGs which operate under unreliable communication network infrastructures. As a result, SMGs are an easy target to an adversary who can very easily gain access to private information by monitoring transmission between nodes in the SMG network, and launch inference-based privacy attacks. These attacks lead to electricity theft and grid instability problems in the SMG. The widely popular differential privacy (DP) technique (a rigorous technique in the family of privacy-preserving data publishing (PPDP) techniques to mathematically guarantee the preservation of data privacy) does not address multi-attribute correlations, that are inherently exploited by an adversary in inference attacks. In this paper we propose HIDE, an oblivious computationally efficient, and rigorous information-theoretic privacy engineering framework for datasets/databases arising in the SMG environments that robustly accounts for multi-attribute correlations while preserving data privacy in a provably optimal fashion. A salient and powerful advantage of HIDE is its ability to generate optimal utility-privacy tradeoffs (computationally efficiently) when the privacy preserving entity in the worst case might have no prior statistical information that links a user's private data with his public data.
Keywords: micro-grid, privacy, database, correlation, computationally efficient

1 INTRODUCTION: A significant percentage of the world's population (approximately about 20%) are without reliable access to electricity, and live mostly in rural and isolated regions of developing countries. In such regions, electricity access is primarily undermined by the absence or intermittence of supply from the national power grid [1]. One way of providing power in rural and remote communities is through smart micro-grids (SMGs) that are networks based on a distributed renewable power generation and a low-cost communication infrastructures. To ensure the reliable operation of SMGs, monitoring is necessary for data collection and state estimation processes. However, highly calibrated and trustworthy smart meters that characterize Advanced Metering Infrastrcutures (AMIs) and are ideally suited to perform such monitoring tasks, are often costly and non-ideally suited to SMGs which operate under unreliable communication network infrastructures. As a result, SMGs are an easy target to an adversary who can very easily gain access to private information by monitoring (eavesdropping) transmission (e.g., time series data on power consumption) between nodes in the SMG network, and launching inferencebased privacy attacks, which subsequently affects the stability of smart micro-grid operation [2][3] (see Section 2). One might argue that cryptographically encrypting transmission data on a communication channel is a safe option to prevent intercepting adversaries launching privacy attacks.

However, in the context of SMGs, there are two main issues with this approach: (i) recent encryption based solutions for the urban smart grid are computationally expensive [4][5], leave alone their practicality in SMGs, and (ii) collecting nonanonymized and un-sanitized high-frequency energy usage data by the utilities is the status quo of smart grid data management in the industry - primarily because household energy consumption data may be used by the utilities and third-party industries to profile energy consumption patterns for intelligent management services (e.g., load balancing, dynamic billing, demand response) and meeting business interests respectively [6]. Therefore, our goal is to prevent adversaries and legal stakeholders from obtaining sensitive smart metering data in a computationally efficient manner, while still enabling the stakeholders to perform their respective functions. In this paper The main challenge to ensuring that private consumer information in an SMG remains private is the fact that public information is often correlated with private information, and this correlation can lead to an inference attack whereby an adversary can have access to the private data of users in a dataset (see Section 2 for details). A popular solution approach to solving the inference attack problem is to distort the release of public information in a manner so as to prevent the leaking of private user data from public data. However, the distortion should be done in a manner to provide significant utility to the party, e.g., third party, receiving the distorted public data, at the same time should be resource amenable to operate in unreliable network infrastructure environments characteristic of SMGs. In this section, we motivate our research based on the principles adopted in existing research on appropriate ways to distort publicly released information. Non-Rigorous Anonymization Approaches - Various nonrigorous anonymization approaches, as part of the PPDP technique have been proposed in existing literature (see [7] for a detailed survey) to preserve data privacy. Well known approaches include t-closeness, k-anonymity, l-diversity, (c, t) - isolation, β-likeness, and their variants. The basic principle behind these approaches is that a data publisher anonymizes or seeks to hide the identity and/or sensitive private attributes of record owners in a publicly released database. However, such approaches suffer from three main drawbacks: (i) they are primarily heuristics and do not mathematically guarantee the preservation of privacy [7], (ii) do not account for the statistical correlations between attributes of a database while preserving privacy, (iii) are domain specific, i.e., non-oblivious, and (iv) computationally intractable if one needs to first find the optimal value of parameters t, k, etc., for a given dataset. To this end, one of our goals is to design a computationally efficient oblivious PPDP mechanism for publicly released data that explicitly accounts for the correlations between the private and public data of users and mathematically guarantees the preservation of privacy.

2. RELATED WORK:

They provide a mechanism for defining naming conventions, certificate constraints, and certificate policies, but they do not specify how these should be used.

Trust Anchor Security One major component of a secure PKI enabled system is the requirement that each RP must have secure methods to load and store the root of trust or trust anchor (TA). The TA is typically a CA at the top of a CA hierarchy. RPs trust certificate holders because they trust the TA which trusts a CA which trusts the end certificate holders. Trust is evidenced by a chain of certificates rooted at the trust anchor. If an adversary could change the root of trust for any RP, that RP could be easily compromised. Certificate Attributes Smart grid to continue to function and other portions of the grid infrastructure are unreachable it will be essential for smart grid devices to be able to authenticate and determine the authorization status for each other without the need to reach a back-end security server. To do this two additional capabilities would be required. First, smart grid certificates will require policy attributes to indicate the applicability of the certificate to a given application. Second a local source of performing certificate status will be required. Smart Grid PKI Tools Standard smart grid operators would have to familiarize themselves with PKI concepts terminology and risks. Standards alone may not

necessarily provide a cost effective solution. Given set of standards it would be possible for vendors to develop smart grid PKI tools which are based on these standards. Tools would greatly ease the process of managing the PKI components needed to support the smart grid application. These tools will be knowledgeable of the appropriate smart grid certificate policy and certificate format standards and are used to programmatically enforce compliance to those standards. Such tools will enhance interoperability reduce the burden of running the PKI and ensure that appropriate security requirements are adhered it is reasonable to expect that the cumulative vulnerability of the system may also be vast. Virtually all parties agree that the consequences of a smart grid cybersecurity breach can be enormous. New functions such as demand response introduce significant new attack vectors such as a malware that initiates a massive coordinated and instantaneous drop in demand potentially causing substantial damage to distribution, transmission, and even generation facilities. Considering the incredible size of the threat and wide-ranging potential consequences from cyberattacks the smart grid cybersecurity protection requirements must be extreme. The grid will require a comprehensive security plan that encompasses virtually all aspects of grid operations. One component of such a plan includes trusted computing platforms. Basic trusted computing model. platforms and associated mechanisms are used to ensure that malware is not introduced into software processing devices. There are two categories of devices for which the malware protection problems should be considered: embedded computer systems and general purpose computer systems. Embedded systems are computer systems that are designed to perform a specific task or set of tasks. They are intended to run only software that is supplied by the manufacture. General purpose systems are intended to support third party software purchased by the specific consumer who purchased the system. A PC is an excellent example of a general purpose system. A microwave oven, or cable television set-top box are examples of embedded systems. Thus problem of malware protection should be considered separately for each category. For embedded systems the problem of protecting the system against the installation of malware can be solved with high degrees of assurance. First manufacturer must implement secure software development processes many standard models for such processes are defined in. Second if the device is intend to be field upgradable the manufacturer must provide a secure software upgrade solution. The predominant method of doing this is to manufacture the embedded systems hardware with secure storage containing keying material for a software validation. The hardware isconfigured with the public key of a secure signing server operated by the manufacturer. The device can validate any newly downloaded software prior to running it. The proactive approach can provide higher levels of assurance than can be obtained with a reactive approach such as a virus checker. Additional security can be obtained by validating the software each time the device boots up. Such techniques are referred to as high assurance boot. HAB techniques typically rely on core software in secure hardware to validate boot-block code. The boot-block code then validates the operating system and the OS in turn validates the higher level applications. Each validation step is performed with public key or keys preinstalled in the secure hardware. For devices which are intended to run for long periods of time without booting it is useful to have a method of performing secure software validation on running code. It is possible to have background tasks that can periodically perform such functions without disrupting the operations of the device. It is further possible to couple such background validation steps with other operational aspects of the device, such that if the device is found to be compromised, secure hardware on the device needed to bring up and maintain security associations with remote entities will prevent the local device from establishing and maintaining security associations with the remote entities. Device attestation is needed to ascertain the devices on the network, true identities, ahead of any manual or automated provisioning at the site. Device attestation techniques accredited manufacturers can factory install device attestation certificates in each smart grid device. These device attestation certificates are used only to assert the device manufacturer, model, serial number, and that the device has not been tampered with. These certificates

coupled with the appropriate authentication protocol can be used by the energy service provider to ensure that the device is exactly what it claims to be. In order to support device attestation the device

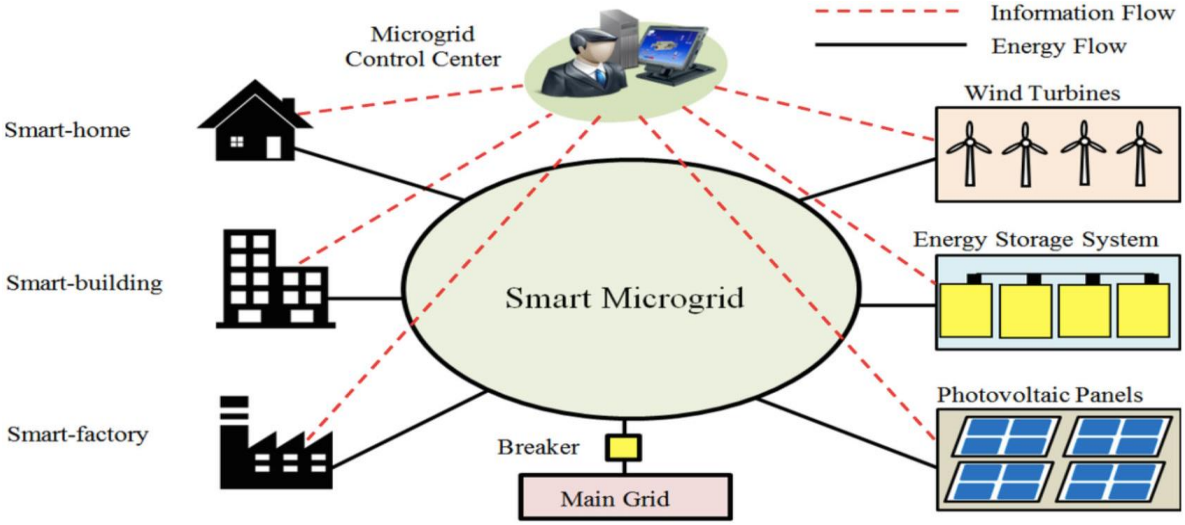
3. IMPLEMENTATION:

We propose an oblivious information-theoretic privacy-utility tradeoff model through which optimal privacy-utility tradeoffs are obtained via solutions to optimizations problems. Being oblivious, the model output is independent of the database conditioned on the value of the true query output (thus also making the model domain and application independent). Our model takes into account the fact that the true joint distribution that links users' (record owners) privat data and data to be released might be unknown in the worst case, and this practical assumption poses a significant challenge to obtaining optimal privacy utility tradeoffs.

Advantages:

a computationally efficient mechanism, HIDE, that leads to optimal privacy utility tradeoffs in SMGs via the solution of optimization problems that model the tradeoffs. As the integral component of HIDE, we first propose a general solution methodology to solve the information theory driven privacy- utility tradeoff optimization problems when the true joint statistical distribution between SMG users' private and public data (see Section 2.2.1) is unknown in the worst case to the optimizer (utility/consumer),

Architecture:



An SMG network architecture model is typically comprised of three components: a power network, a communications network, and the SMG users (see Fig. 1). Here, SMG users comprise energy consumers (households), energy producers (utilities and households), and a super user or a utility company. Producers could assume the role of consumers, and vice versa when it comes to renewable power supply/demand. The power network includes a group of distributed generators supported by renewable energy generation sources such as solar power. In this network, micro generation emanating from a subset or all of the users is connected via a distribution network. However, due to the volatility of renewable energy, the amount of energy generated is not guaranteed to match the power demand on the grid at all times. In order to alleviate mismatch issues, there is the presence of a communication network that enables power consumption/production data collection, monitoring, and message interactions.

The Privacy Setting in HIDE:

We consider a general single-user (consumer) tabled database setting2 where each record showcases energy consumption data for a particular consumer at a given sampled time instant multiple records denoting consumption information for the same user in different time instants. Each record has a private consumption data component, represented by the discrete random variable (usually a vector) S

that is correlated with the consumer's non-private (or public) data, represented via a discrete random variable (usually a vector) X.

Threat Model and the Privacy Metric in HIDE:

Threat Model. We consider threats that are inferential in nature, where an adversary with potentially unbounded computational resources can infer S from Y .3 We assume that the DBM in the worst case might only have partial knowledge on the exact nature of PS;X (e.g., due to noisy or insufficient data). However, the adversary might in his best case have complete knowledge of the exact nature of PS;X in terms of side information available from various sources. We also assume that the adversary in the worst case has complete knowledge of the privacy mapping, i.e., PY jX, i.e., he knows the entire distribution.

4. CONCLUSION:

we proposed HIDE for the smart micro-grid (SMG), a computationally efficient information theory based privacy preserving mechanism for consumer energy consumption datasets having multiple correlated attributes. HIDE leads to optimal utility-privacy tradeoffs, when the privacy preserving entity in the worst case might have no prior information that links a user's private data and his public data to be released. In this regard, as the integral component of HIDE, we first proposed an optimization theory based general solution methodology to solve information theory driven privacy-utility tradeoff optimization problems when the true joint statistical distribution between users' private and public data is unknown in the worst case to the optimizer (utility/consumer), and the correlations between public and private data attributes are being accounted for. Second, we then showed that these optimization problems are convex for a certain general class of objective functions and constraints. For such problems, we provided details of our solution methodology that uses properties of Hilbert spaces in functional analysis theory, and guarantees an optimal solution for a given optimization problem, in polynomial time. In addition, our proposed solution methodology accurately estimated true non-linear correlations between the private and public data attributes even in the presence of less number of data samples, something of a reality in SMGs which might not have the infrastructure capability to handle frequent data sampling.

REFERENCES [1] Anthony R. Metke and Randy L. Ekl —Security Technology for Smart Grid Networks‖ [2] Binod Vaidya, Dimitrios Makrakis, and Hussein T. Mouftah, University of Ottawa —Authentication and Authorization Mechanisms for Substation Automation in Smart Grid Network‖ [3] Ting Liu, Member, IEEE, YangLiu, YashanMao, Yao Sun, XiaohongGuan, Fellow, IEEE, Weibo Gong, Fellow, IEEE, and Sheng Xiao‖ A Dynamic SecretBased Encryption Scheme for Smart Grid Wireless Communication‖ IEEE TRANSACTIONS ON SMART GRID‖ [4] Fadi Aloula*, A. R. Al-Alia , Rami Al-Dalkya, Mamoun Al-Mardinia,Wassim El-Hajjb‖ Smart Grid Security: Threats, Vulnerabilities and Solutions‖ [5] Aaron St. Leger, Member, IEEE, John James, Senior Member, IEEE, Dean Frederick, Senior Member,IEEE‖ Smart Grid Modeling Approach for Wide Area Control Applications‖ [6] G. N. Ericsson, —Cyber Security and Power System Communication—Essential Parts of A Smart Grid Infrastructure,‖ IEEE Trans. Power Delivery, vol. 25, issue 3, July 2010, pp. 1501–07. [7] S. Fries et al., —Security for the Smart Grid – Enhancing IEC 62351 to Improve Security in Energy Automation Control,‖ Int'l. J. Advances in Security, vol 3, no. 3–4, 2010, pp. 169 83. [8] B. Vaidya, D. Makrakis, and H. T. Mouftah —Provisioning Substation- Level Authentication in the Smart Grid Networks,‖ Proc. IEEE MILCOM 2011, Nov. 2011, pp. 1189–94.

[9] Y. Zhu and R. Bettati, "Anonymity vs. information leakage in anonymity systems," in 25th IEEE International Conference on Distributed Computing Systems (ICDCS'05), pp. 514–524, IEEE, 2005. [10] G. Smith, "On the foundations of quantitative information flow," in International Conference on Foundations of Software Science and Computational Structures, pp. 288–302, Springer, 2009. [11] L. Yang and F. Li, "Detecting false data injection in smart grid innetwork aggregation," in Smart Grid

Communications (SmartGridComm), 2013 IEEE International Conference on, pp. 408–413, IEEE, 2013. [12] P. L. Ambassa, A. V. Kayem, S. D. Wolthusen, and C. Meinel, "Privacy violations in constrained micro-grids: Adversarial cases," in Advanced Information Networking and Applications Workshops (WAINA), 2016 30th International Conference on, pp. 601–606, IEEE, 2016. [13] H. Li, S. Gong, L. Lai, Z. Han, R. C. Qiu, and D. Yang, "Efficient and secure wireless communications for advanced metering infrastructure in smart grids," IEEE Transactions on Smart Grid, vol. 3, no. 3, pp. 1540–1551, 2012. [14] P. L. Ambassa, S. D. Wolthusen, A. V. Kayem, and C. Meinel, "Robust snapshot algorithm for power consumption monitoring in computationally constrained micro-grids," in Innovative Smart Grid Technologies-Asia (ISGT ASIA), 2015 IEEE, pp. 1–6, IEEE, 2015. [15] J. Lines, A. Bagnall, P. Caiger-Smith, and S. Anderson, "Classification of household devices by electricity usage profiles," in International Conference on Intelligent Data Engineering and Automated Learning, pp. 403–412, Springer, 2011. [16] F. du Pin Calmon and N. Fawaz, "Privacy against statistical inference," in Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on, pp. 1401–1408, IEEE, 2012. [17] N. Merhav and M. Feder, "Universal prediction," Information Theory, IEEE Transactions on, vol. 44, no. 6, pp. 2124–2147, 1998. [18] J. W. Miller, R. Goodman, and P. Smyth, "On loss functions which minimize to conditional expected values and posterior probabilities," Information Theory, IEEE Transactions on, vol. 39, no. 4, pp. 1404–1408, 1993. [19] T. M. Cover and J. A. Thomas, Elements of information theory. John Wiley & Sons, 2012. [20] A. Ghosh, T. Roughgarden, and M. Sundararajan, "Universally utility-maximizing privacy mechanisms," SIAM Journal on Computing, vol. 41, no. 6, pp. 1673–1693, 2012. [21] S. Boyd and L. Vandenberghe, Convex optimization. Cambridge university press, 2004. [22] R. Ahlswede and P. Gacs, "Spreading of sets in product spaces and ´ hypercontraction of the markov operator," The annals of probability, pp. 925–939, 1976. [23] V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On maximal correlation, hypercontractivity, and the data processing inequality studied by erkip and cover," arXiv preprint arXiv:1304.6133, 2013. [24] H. O. Hirschfeld, "A connection between correlation and contingency," in Mathematical Proceedings of the Cambridge Philosophical Society, vol. 31, pp. 520–524, Cambridge Univ Press, 1935. [25] A. Renyi, "On measures of dependence," ´ Acta mathematica hungarica, vol. 10, no. 3-4, pp. 441–451, 1959. [26] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," SIAM Journal on Applied Mathematics, vol. 28, no. 1, pp. 100–113, 1975.