International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

PRIVACY-PRESERVING PROTOCOL TO OBTAIN THE SUM AGGREGATE OF TIME-SERIES DATA

BODA MAHESH[1] , LEELASRAVANTHI[2]
1. PG Scholar, Dept of CSE, Balaji Institute of Technology, Warangal, TS, India.
2Assistant Professor, Dept of CSE, Balaji Institute of Technology, Warangal, TS, India.

Abstract: The works on sensor data aggregation assume a trusted aggregator, and hence cannot protect user privacy against an untrusted aggregator in mobile sensing applications. Several recent works consider the aggregation of time series data in the presence of an untrusted aggregator. To protect user privacy, they design encryption schemes in which the aggregator can only decrypt the sum of all users' data but nothing else. Rastogi and Nath use threshold Paillier cryptosystem to build such an encryption scheme. To decrypt the sum, their scheme needs an extra round of interaction between the aggregator and all users in every aggregation period, which means high communication cost and long delay the proliferation and ever-increasing capabilities of mobile devices such as smart phones give rise to a variety of mobile sensing applications. This paper studies how an untrusted aggregator in mobile sensing can periodically obtain desired statistics over the data contributed by multiple mobile users, without compromising the privacy of each user. Although there are some existing works in this area, they either require bidirectional communications between the aggregator and mobile users in every aggregation period, or have high-computation overhead and cannot support large plaintext spaces. Also, they do not consider the Min aggregate, which is quite useful in mobile sensing. To address these problems, we propose an efficient protocol to obtain the Sum aggregate, which employs an additive homomorphic encryption and a novel key management technique to support large plaintext space. We also extend the sum aggregation protocol to obtain the Min aggregate of time-series data. To deal with dynamic joins and leaves of mobile users, we propose a scheme that utilizes the redundancy in security to reduce the communication cost for each join and leave. Evaluations show that our protocols are orders of magnitude faster than existing solutions, and it has much lower communication overhead.

Keywords: Threshold Paillier Cryptosystem.

I. INTRODUCTION

A sensor network consists of large number of sensors deployed in a region for the purpose of event monitoring or detection. The sensors are preprogrammed to listen for specific events. For example, a sensor network deployed in a high security region might be programmed to detect infrared heat signals to indicate an intruder. Figure 1 shows a typical sensor network deployment. Each node in a sensor network is responsible for observing and reporting various dynamic properties of their surroundings in a time critical manner. These mobile and miniaturized information devices are equipped with embedded processors, wireless communication circuitry, information storage capability, smart sensors and actuators. These sensor nodes networked in an ad hoc way, with little or no fixed network support, to provide the surveillance and targeting information for dynamic control. Sensor devices are mobile, subject to failure, deployed spontaneously and repositioned for more accurate surveillance. Despite these dynamic changes in configuration of the sensor network, critical real-time information must still be disseminated dynamically from mobile sensor data sources through the self-organizing network infrastructure to the components that control dynamic re-planning and reoptimization of the theatre of operation based on newly available information. With large number of sensor devicesbeing quickly and flexibly deployed in most impromptu networks, each sensor device must be autonomous and capable of organizing itself in the overall community of sensors to perform coordinated activities with global objectives.

When spontaneously placed together in an environment, these sensor nodes should immediately know about the capabilities and functions of other sensor Nodes and work together as a community system to perform cooperative tasks and networking functionalities. Sensor networks need to be self-organizing since they are often formed spontaneously from large number of mixed types of nodes and may undergo frequent configuration changes. Some sensor nodes may provide networking and system services and resources to other sensor nodes. Others may detect the presence of these nodes and request services from them. The characteristics of sensor nodes necessary for creating self-organizing sensor networks are agility, self awareness, self configurability and autonomy. Sensor nodes with these features will have capabilities for self assembling impromptu networks that are incrementally extensible and dynamically adaptable to device failure and degradation, mobility of sensor nodes and changes in task and network requirements. Nodes are aware of their own capabilities and those of other nodes aroundthem which may provide the networking and system services or resources that they need. Although nodes are autonomous, they may cooperate with one another to disseminate information or assist each other in adapting to changes in the network configuration. An impromptu community of these nodes may cooperate to provide continual coordinated services while some nodes may be newly deployed or removed from the spontaneous community. Nodes will act in response to environmental events and relay collected and possibly aggregated information through the multi-hop wireless network in accordance with desired system functionality.

The inherently dynamic and distributed behavior of these networks, coupled with inherent physical limitations such as small instruction and data memory, constrained energy resources, short communication radii and a low bandwidth medium in which to communicate, make developing communication protocols difficult. Using these sensors as a basis for development, the software architecture and communication stack residing on these devices are built taking into consideration the prolific research in the areas of ad-hoc networking, data aggregation, cluster formation, distributed services, group formation, channel contention and power conservation. An event is an abstraction, identifying anything from a set of sensor readings, to the nodes processing capabilities. For the purpose of the simulation studies in this project, events are assumed to be localized phenomenon, occurring in a fixed region of space. This assumption will hold for a wide variety of sensor-net applications, since many external events are localized themselves.

II. RELATED WORK

Many works have addressed various security and privacy issues in mobile sensing networks and systems (e.g., [10], [2]), but they do not consider data aggregation. There are a lot of existing works on security and privacy-preserving data aggregation, but most of them assume a trusted aggregator and cannot protect user privacy against untrusted aggregators. Yang et al. proposed an encryption scheme that allows an untrusted aggregator to obtain the sum of multiple users's data without knowing any specific user's data. Their scheme requires expensive rekeying operations to support multiple time steps and thus may not work for time-series data. Shi et al. proposed a privacy-preserving data aggregation scheme based on data slicing and mixing techniques. Their scheme is not designed for time-series data. It may not work well for time-series data, since each user may need to select a new set of peers in each aggregation interval due to mobility. Besides, their scheme for non additive aggregates requires multiple rounds of bidirectional communications between the aggregator and mobile users which means long delays. In contrast, our scheme obtains those aggregates with just one round of unidirectional communication from users to the aggregator.multiplies them together and sends the aggregate ciphertext to all users. Each user decrypts a share of the sum aggregate. The aggregator collects all the shares and gets the final sum. Their scheme requires an extra round of interaction between the aggregator and users in every aggregation period. Erkin and Tsudik [12] also proposed an aggregation scheme based on Paillier cryptosystem, but it requires communications between every pair of users in every aggregation period. Based on an efficient additive homomorphic encryption scheme, Rieffel et al. [9] proposed a construction that does not require an extra round of interaction between the aggregator and the users. In their scheme, the computation and storage cost is roughly equal to the number of colluding users that the system can tolerate. Thus, their scheme has high overhead to achieve good resistance to collusion, especially when the system is large and a large number of users collude. In contrast, our scheme tolerates a high fraction of colluding users with very small cost even when the system is large.

Acs and Castelluccia [13] also proposed a scheme based on additive homomorphic encryption, but in their scheme each node shares a pairwise key with any other node. Shi et al. [7] proposed a construction for sum aggregation based on the assumption that the Decisional Diffie-Hellman problem is hard over finite cyclic groups. In their construction, each user sends her ciphertext to the aggregator and no communication is needed from the aggregator to the users. To decrypt the sum, their construction needs to traverse the possible plaintext space of sum and thus, it is not efficient for a large system with large plaintext spaces. Chan et al. [8] extended the construction with a binary interval tree technique, but their scheme still has the limitation in plaintext spaces. Jawurek and Kerschbaum [14] proposed a scheme that provides differential privacy for sum.Our aggregation protocol for sum can be used as a building block of their scheme to improve the computational efficiency. Also, existing works [3] do not consider the Min of time series data.

III. MODULES

A. Node Sensing

An important challenge is how to protect the users' privacy in mobile sensing, especially when the aggregator is untrusted. Most previous works on sensor data aggregation assume a trusted aggregator, and hence cannot protect user privacy against an untrusted aggregator in mobile sensing applications. Moreover, none of these existing schemes considers the Min aggregate (i.e., the minimum value) of time-series data, which is also important in many mobile sensing applications.

B. Mobile sensing

To sense the all connected nodes in server and Mobile sensing applications such as environmental monitoring, traffic monitoring, healthcare, and so on. In many scenarios, aggregation statistics need to be periodically computed from a stream of data contributed by mobile users, to identify some phenomena or track some important patterns.

C. Privacy

Thus, an important challenge is how to protect the users' privacy in mobile sensing, especially when the aggregator is untrusted .Most previous works on sensor data aggregation assume a trusted aggregator, and hence cannot protect user privacy against an untrusted aggregator in mobile sensing applications. To our best knowledge, this is the first privacypreserving solution to obtain the Min of time-series data in mobile sensing with just one round of user-to-aggregator communication. Our protocols for Sum and Min can be easily adapted to derive many other aggregate statistics such as Count, Average, and Max. Since users may frequently join and leave in mobile sensing, we also propose a scheme that employs the redundancy in security to reduce the communication cost of dealing with dynamic joins and leaves.

D. Data aggregation

The aggregator collects the ciphertexts of users, multiplies them together, and sends the aggregate ciphertext to all users. Each user decrypts a share of the sum aggregate. The aggregator collects all the shares and gets the final sum. However, their scheme requires an extra round of interaction between the aggregator and users in every aggregation period. Erkin and Tsudik also proposed an aggregation scheme based on Paillier cryptosystem, but it requires communications between every pair of users in every aggregation period. Based on an efficient additive homomorphic encryption scheme, Rieffel et al. Proposed a construction that does not require an extra round of interaction between the aggregator and the users. In their scheme, the computation and storage cost is roughly equal to the number of colluding users that the system can tolerate. Thus, their scheme has high overhead to achieve good resistance to collusion, especially when the system is large and a large number of users collude. In contrast, our scheme tolerates a high fraction of colluding users (e.g., 30 ercent) with very small cost even when the system is large. Acs and Castelluccia also proposed a scheme based on additive homomorphic encryption, but in their scheme each node shares a pairwise key with any other node. Shi et al. proposed a construction for sum aggregation based on the assumption that the Decisional DiffieHellman problem is hard over finite cyclic groups. In their construction, each user sends her ciphertext to the aggregator and no communication is needed from the aggregator to the users.

IV. PRIVACY PRESERVED DATA AGGREGATION FOR MOBILE SENSING

A new protocol for mobile sensing to obtain the sum aggregate of time-series data in the presence of an untrusted aggregator. Our protocol employs an additive homomorphic encryption and a novel key management scheme based on efficient HMAC to ensure that the aggregator can only obtain the sum of all users' data, without knowing individual user's data or intermediate result. In our protocol, each user only needs to compute a very small number of HMACs to encrypt her data. Hence, the computation cost is very low and the protocol can scale to large systems with large plaintext spaces, resource constrained devices and high aggregation loads. Another nice property of our protocol is that it only requires a single round of user-to-aggregator communicationBased on the sum aggregation protocol, we propose a protocol to obtain the Min aggregate. To our best knowledge, this is the first privacypreserving solution to obtain the Min of time-series data in mobile sensing with just one round of user-to aggregator communication. Our protocols for Sum and Min can be easily adapted to derive many other aggregate statistics such as Count, Average and Max. Since users may frequently join and leave in mobile sensing, we also propose a scheme that employs the redundancy in security to reduce the communication cost of dealing with dynamic joins and leaves.

V. CONCLUSION

To facilitate the collection of useful aggregate statistics in mobile sensing without leaking mobile users' privacy, we proposed a new privacy-preserving protocol to obtain the Sum aggregate of time-series data. The protocol utilizes additive homomorphic encryption and a novel, HMAC based key management technique to perform extremely efficient aggregation. Implementation-based measurements show that operations at user and aggregator in our protocol are orders of magnitude faster than existing work. Thus, our protocol can be applied to a wide range of mobile sensing systems with various scales, plaintext spaces, aggregation loads, and resource constraints. Based on the Sum aggregation protocol, we also proposed two schemes to derive the Min aggregate of time-series data. One scheme can obtain the accurate Min, while the other one can obtain an approximate Min with provable error guarantee at much lower cost. To deal with dynamic joins and leaves, we proposed a scheme that utilizes the redundancy in security to reduce the communication cost for each join and leave. Simulation results show that our scheme has much lower communication overhead than existing work.

VI. REFERENCES

[1] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M.Hansen, E. Howard, R. West, and P. Boda, "Peir, the Personal Environmental Impact Report, As a Platform for Participatory Sensing Systems

Research," Proc. ACM/USENIX Int'l Conf. Mobile Systems, Applications, and Services (MobiSys '09), pp. 55-68, 2009.

[2] A. Thiagarajan, L. Ravindranath, K. LaCurts, S. Madden, H.Balakrishnan, S. Toledo, and J. Eriksson, "VTrack: Accurate, Energy-Aware Road Traffic Delay Estimation Using Mobile Phones," Proc. ACM Seventh Conf. Embedded Networked Sensor Systems (SenSys '09), pp. 85-98, 2009.

[3] S. Consolvo, D.W. McDonald, T. Toscos, M.Y. Chen, J. Froehlich,B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J.A. Landay, "Activity Sensing in the Wild: A Field Trial of Ubifit Garden," Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI '08), pp. 1797- 1806, 2008.

[4] J. Hicks, N. Ramanathan, D. Kim, M. Monibi, J. Selsky, M. Hansen, and D. Estrin, "AndWellness: An Open Mobile System for Activity and Experience Sampling," Proc. Wireless Health, pp. 34-43, 2010.

[5] N.D. Lane, M. Mohammod, M. Lin, X. Yang, H. Lu, S. Ali, A.Doryab, E. Berke, T. Choudhury, and A. Campbell, "Bewell: A Smartphone Application to Monitor, Model and Promote Wellbeing,"Proc. Fifth Int'l ICST Conf. Pervasive Computing Technologies for Healthcare, 2011.

[6] V. Rastogi and S. Nath, "Differentially Private Aggregation of Distributed Time-Series with Transformation and Encryption,"Proc. ACM SIGMOD Int'l Conf. Management of Data, 2010.

[7] E. Shi, T.-H.H. Chan, E. Rieffel, R. Chow, and D. Song, "Privacy-Preserving Aggregation of Time-Series Data," Proc. Network and Distributed System Security Symp. (NDSS '11), 2011.

[8] T.-H.H. Chan, E. Shi, and D. Song, "Privacy-Preserving Stream Aggregation with Fault Tolerance," Proc. Sixth Int'l Conf. Financial Cryptography and Data Security (FC '12), 2012.

[9] E.G. Rieffel, J. Biehl, W. van Melle, and A.J. Lee, "Secured Histories: Computing Group Statistics on Encrypted Data While Preserving Individual Privacy," http://arxiv.org/abs/1012.2152,2010.

[10] P.-A. Fouque, G. Poupard, and J. Stern, "Sharing Decryption in the Context of Voting or Lotteries," Proc. Fourth Int'l Conf. Financial Cryptography (FC '00), pp. 90- 104, 2000.

[11] MNDOLI, "Mnosha Permissible Exposure Limits," http://www.dli.mn.gov/OSHA/PDF/pels.pdf, 2013.

[12] S.B. Eisenman, E. Miluzzo, N.D. Lane, R.A. Peterson, G.-S.Ahn,and A.T. Campbell, "The Bikenet Mobile Sensing System for Cyclist Experience Mapping," Proc. ACM Fifth Int'l Conf. Embedded Networked Sensor Systems (SenSys '07), pp. 87-101, 2007.

[13] M.G. Apte, W.J. Fisk, and J.M. Daisey, "Indoor Carbon Dioxide Concentrations and SBS in Office Workers," Proc. Healthy Buildings Conf., pp. 133-138, 2000.

[14] Z. Zhu and G. Cao, "APPLAUS: A Privacy-Preserving Location Proof Updating System for Location-Based Services," Proc. IEEE INFOCOM, 2011.

[15] Q. Li and G. Cao, “Mitigating Routing Misbehavior in Disruption Tolerant Networks,” IEEE Trans. Information Forensics and Security, vol. 7, no. 2, pp. 664-675, Apr. 2012.

[16] E.D. Cristofaro and C. Soriente, “Short Paper: Pepsi— Privacy-Enhanced Participatory Sensing Infrastructure,” Proc. Fourth ACM Conf. Wireless Network Security (WiSec ’11), pp. 23-28, 2011.

[17] Q. Li, S. Zhu, and G. Cao, “Routing in Socially Selfish Delay Tolerant Networks,” Proc. IEEE INFOCOM, pp. 1-9, 2010.

[18] Q. Li, W. Gao, S. Zhu, and G. Cao, “A Routing Protocol for Socially Selfish Delay Tolerant Networks,” Ad Hoc Networks, vol. 10, no. 8, pp. 664-675, 2012.

[19] D. Bonet, E.-J. Goh, and K. Nissim, “Evaluating 2-DNF Formulas on Ciphertexts,” Proc.Second Int’l Conf. Theory of Cryptography (TCC ’05), 2005.

[20] C. Gentry, “Fully Homomorphic Encryption Using Ideal Lattices,”Proc. ACM Symp. Theory of Computing (STOC ’09), pp. 169-178,2009.

[21] C. Castelluccia, A.C.-F. Chan, E. Mykletun, and G. Tsudik, “Efficient and Provably Secure Aggregation of Encrypted Data in Wireless Sensor Networks,” ACM Trans. Sensor Networks, vol. 5,no. 3, pp. 20:1-20:36, 2009.

[22] Y. Yang, X. Wang, S. Zhu, and G. Cao, “SDAP: A Secure Hopby-Hop Data Aggregation Protocol for Sensor Networks,” ACM Trans. Information and System Security, vol. 11, no. 4, article 18, 2008