IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 4, April 2015.
www.ijiset.com
ISSN 2348 – 7968

Privacy Preserving Data Aggregation and Proficient in Mobile Sensing

S.Deepika[1], Bharathi Sri[2]

1 PG Scholar, Department Of Computer Science And Engineering

2 Assistant Professor, Department Of Computer Science And Engineering,

ABSTRACT

Mobile sensing exploits information contributed by mobile users to create subtle inferences regarding folks and their close and therefore is often applied to environmental observation, traffic observation and attention. However, the large-scale preparation of mobile sensing applications is hindered by the shortage of incentives for users to participate and therefore the issues on doable privacy outflow. Also, they are doing not contemplate the Min combination, that is kind of helpful in mobile sensing. to deal with these issues, we have a tendency to propose Associate in Nursing economical protocol to get the total combination, that employs Associate in Nursing additive homomorphic encoding and a unique key management technique to support massive plaintext area.

Key words: Mobile sensing, Data aggregation, security, Wireless sensor network.

1. INTRODUCTION

Most smart phones are equipped with a rich set of embedded sensors such as camera, microphone, GPS, accelerometer, ambient light sensor, gyroscope, and so on. The data generated by these sensors provide opportunities to make sophisticated inferences about not only people but also their surrounding and thus can help improve people's health as well as life. 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. For example, the average amounts of daily exercise that people do can be used to infer public health conditions. The average or maximum level of air pollution and pollen concentration in an area may be useful for people to plan their outdoor activities.

Other statistics of interests include the lowest gasoline price in a city, the highest

ISSN 2348 – 7968

moving speed of road traffic during rush hour, and so on. Although aggregation statistics computed from time series data are very useful, in many scenarios, the data from users are privacy-sensitive, and users do not trust any single third-party aggregator to see their data values. For instance, to monitor the propagation of a new flu, the aggregator will count the number of users infected by this flu. However, a user may not want to directly provide her true status if she is not sure whether the information will be abused by the aggregator. Accordingly, systems that collect users' true data values and compute aggregate statistics over them may not meet users' privacy requirement. Thus, an important challenge is how to protect the users' privacy in mobile sensing, especially when the aggregator is untrusted.

One possible way for achieving privacy is to perturb the answers to such queries independently of one another, thereby ensuring that even revealing a few true answers does not help infer anything about the perturbation of other answers. However pointed out that if the time-series exhibit certain patterns, and then independent perturbation of query answers can be distinguished from the original answers and filtered out. Authors consider perturbing time series data to defend against several privacy attacks, but they do not provide any formal privacy guarantee, without which data owners may not publish sensitive data in the fear of unforeseen privacy attacks. On the other hand, formal privacy guarantees like differential privacy that work well for relational data, seem too hard to achieve for time series data. For instance, standard differentially private techniques can result in a noise of to each query answer, where n is the number of queries to answer, making the query answers practically useless if a long sequence of queries is to be answered.

2. RELATED WORKS

In this paper [4] S. B. Eisenman, E. Miluzzo,N. D. Lane,R. A. Peterson was represents : an outline and example implementation of the system design, associate degree analysis of sensing associate degreed abstract thought that quantifies wheeler performance and also the wheeler setting; a report on networking performance in an environment characterized by bicycle quality and human unpredictability; and an outline of motorbike web system user interfaces.

In this paper [9] T-H. Hubert Chan1, Elaine Shi2, and Dawn Song was represent however associate degree world

ISSN 2348 – 7968

organization trustworthy person will gather info and learn mixture statistics from a population while not harming individual privacy. For instance, think about a sensible grid operator UN agency needs to trace the overall electricity consumption of vicinity each quarter-hour, for programming and optimization functions. Since such power consumption information will reveal sensitive info regarding individual's presence and activities, we tend to would like to perform such aggregation in an exceedingly privacy-preserving manner.

In this paper [3]Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts presents VTrack, a system for period estimation victimisation this device information that addresses 2 key challenges: energy consumption and device undependability. Whereas GPS provides extremely correct location estimates, it's many limitations: some phones don't have GPS in any respect, the GPS device doesn't add "urban canyons" tall buildings and tunnels or once the phone is within a pocket, and also the GPS on several phones is power-hungry and drains the battery quickly.

In this paper [1] Zekeriya Erkin1 and Gene Tsudik explores some straightforward and comparatively economical cryptanalytic privacy techniques that permit spatial group-wide aggregation of good meter measurements. We tend to additionally think about temporal aggregation of multiple measurements for one good meter. whereas our work is on no account the primary to tackle this subject, we tend to believe that planned techniques square measure appealing as a result of their simplicity, few assumptions and peer based mostly nature, i.e., no would like for any on-line aggregators or trustworthy third parties.

3. DESCRIPTION OF THE PROPOSED SCHEME:

The 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 use 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

ISSN 2348 – 7968

loads. Another nice property of our protocol is that it only requires a single round of user-to-aggregator communication.

Based on the sum aggregation protocol, we propose a protocol to obtain the Min aggregate. To our best knowledge, this is the first privacy-preserving 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.

DATA AGGREGATION

Data aggregation is any process in which information is gathered and expressed in a summary form, for purposes such as statistical analysis. A common aggregation purpose is to get more information about particular groups based on specific variables such as age, profession, or income. The information about such groups can then be used for Web site personalization to choose content and advertising likely to appeal to an individual belonging to one or more groups for which data has been collected. For example, a site that sells music CDs might advertise certain CDs based on the age of the user and the data aggregate for their age group. Online analytic processing is a simple type of data aggregation in which the marketer uses an online reporting mechanism to process the information.

ENCRYPTION KEY GENERATION

$$k_i = \left(\sum_{s' \in S_i} h(f_{s'}(t)) \right) \bmod M$$

DECRYPTION KEY GENERATION

$$k_0 = \left(\sum_{s' \in S} h(f_{s'}(t)) \right) \bmod M$$

Thus, the encryption keys satisfy the security requirement of the underlying cryptosystem.

$$\sum_{i=1}^{n} k_i = \left(\sum_{i=1}^{n} \sum_{s \in S_i} h(f_s(t)) \right) \bmod M$$

$$= \left(\sum_{s \in S} h(f_s(t)) \right) \bmod M$$

$$= k_0 \bmod M$$

SECURITY LEVEL

If the aggregator knows the c secrets used by a user, it can obtain the encryption key of the user. We can derive the probability that the aggregator finds the c secrets used by a

ISSN 2348 – 7968

user. Let pb denote the probability that in a single trial the aggregator can successfully guess the secrets assigned to the user. Recall that is the maximal fraction of users that collude with the aggregator. A fraction of users may collude against the aggregator to reveal the aggregate.

$$P_b = \frac{1}{\binom{(1-\gamma)nc}{c}}$$

To achieve this goal, they need to obtain all the secrets that the aggregator has. However, each user only knows a subset of the secrets. So long as not all users collude, they cannot obtain all the secrets.

4. EXPERIMENTAL RESULTS

There are four performance metrics considered in the simulations:

1) Packet delivery ratio (PDR) is the ratio of the amount of information packets received by a destination node and the number of data packets generated by a source node; 2) Throughput is the total size of data packets properly received by a destination node each second; 3) Average end-to-end delay is the mean of end-to-end delay among a source node and a destination node with CBR traffic; 4) Communication Overhead is the size of Type Length Value (TLV) blocks in total messages, which are used to bring trust values; 5) Routing load is the ratio of the number of control packets transmitted by nodes to the number of data packets received successfully by destinations during the simulation.

5. CONCLUSION

An honest-but-curious key dealer correctly follows our protocol steps, but wants to get users' data values from the transcript of messages in our protocol. To provide privacy under this model, our protocol adds one more encryption and decryption to the data that each user submits to the aggregator. 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. The computation cost of our scheme is increased when dealing with dynamic joins and leaves. To evaluate the computation cost, we measured the running time when there is redundancy in security. One scheme can obtain the accurate Min, while the other

ISSN 2348 – 7968

one can obtain an approximate Min with provable error guarantee at much lower cost. To deal with dynamic joins and leaves, here proposed a scheme that utilizes the redundancy in security to reduce the communication cost for each join and leave.

REFERENCES

[1] Z. Erkin and G. Tsudik, "Private Computation of Spatial and Temporal Power Consumption with Smart Meters," Proc. Int'l Conf. Applied Cryptography and Network Security (ACNS '12),
pp. 561-577, 2012.

[2] 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.

[3] 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.

[4] 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.

[5] 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 SensorSystems (SenSys '09), pp. 85-98, 2009.

[6] 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.

[7] 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.