International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

PATA: A Protocol for Mobile Sensing Privacy-Aware and Trustworthy Data Aggregation

Devendra Hapase[1], M. D. Ingle[2]

[1]ME Computer (Engineering), Jayawantrao Sawant College of Engineering, Hadapsar Pune-28, Savitribai Phule Pune University, Pune, India

[2]Assistant Professor, (Computer Engineering), Jayawantrao Sawant College of Engineering, Hadapsar Pune-28, Savitribai Phule Pune University, Pune, India

Abstract: ***Now days essentiality of the mobile device such as smartphones and tablets, in which having various number of mobile sensing applications are used in different area such as for air pollution monitoring and for healthcare survey. With the increasing capabilities of mobile devices such as smartphones and tablets, there are more and more mobile sensing applications such as air pollution monitoring and healthcare. These applications generally combine the information given by mobile users to show the information about human's activities or surroundings. When the data supplied by mobile users is decent and trustworthy then the mobile sensing works properly. System may be damage if the mobile users act like eavesdropper due to malicious impact and submit forge data. Sometimes, mobile sensing users unable to submit data because of security issue. To overcome on these issues, a new privacy-aware and trustworthy data aggregation protocol for mobile sensing proposed. PATA protocol permits the server to collect the data which is add by mobile users without knowing the data of individual user. At the same time, if malicious users try to submit wrong data, system will be found or the denied aggregation result by the server. In this way, the malicious users' effect on the aggregation result is effectively limited. The knowing of wrong data works even if multiple malicious users collude. Security measurement determined that this scheme can give the trustworthy and privacy preserving goals, and implementation of scheme shows that it take low computation cost and low power consumption.***

Keywords: Wireless Sensor Network, Data Aggregation, Trusted Authority, Homomorphic encryption, Blind Signature.

1. Introduction

Mobile nodes, especially smart devices, are plays the very important role in our daily routines. With increased the sensing abilities such as camera, microphone, accelerometer, GPS, etc. and communication capabilities for the data services such as 4G, WiFi and Bluetooth, mobile devices can supply helpful information to infer of our daily life such as location, health, activity, etc. through the GPS or social network as well as the weather study like air pollution, temperature, noise, traffic, etc. and hence help to enhanced the punctuation of life.

Mobile sensing applications are very much popular in some field such as traffic monitoring [1], pollution monitoring [2], health monitoring [3]. For these useful applications, information aggregation are uploaded by mobile users are useful for inferring about the environment and people's life or identifying essential phenomena in a particular area.

For example, the average time people spend on social media would useful us think about people's social life pattern. We can also study about flu who the affected by particular disses and we can easily provide the vaccines to that patients or provide the particular facilities for that affected area. The maximum speed of the moving bike or vehicle on road traffic in rush time can put up an essential information about traffic jam and help people to set schedules for various task and choose path accordingly.

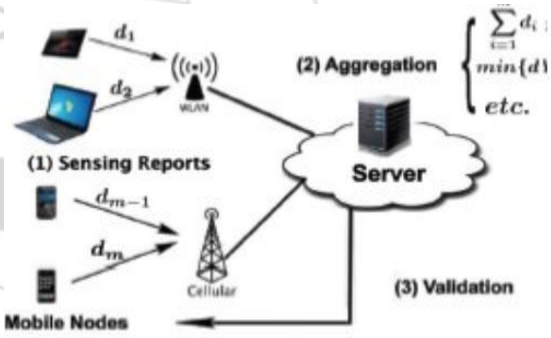

Figure 1: SYSTEM MODEL

Although these mobile applications are very helpful, in many cases, but the selected data may increase the privacy concerns. User may not trust the aggregation server or does not allow another person to check their data in plaintext. For example, to examine the flu trend in a particular area, the servers require gathering information about the infected users with flu.

However, the infection status of a user is sensitive and if we do not provide the security for their data they are no more interested to participate in any sensing task. Another essential problem is whether the data offer by mobile users should be trusted or not. It is realistic that a user's device malfunctions and offer wrong sensing results to the server, or even worse, the mobile user is malicious and intentionally forges data to mislead the server.

In either case, the server finished with wrong aggregation component and hence creates wrong inferences. There are many more existing tasks on carrying user security in data aggregation [4], [5]. However, they take up that mobile users

are trustworthy. As per study, a single forged data can make the aggregation result significantly deviate from the true value and become useless. Another way, there are also many more existing task on trustworthy data aggregation in wireless sensor networks [6], [7]. However, they do not assume about user data privacy, and we are unfamiliar about of any data aggregation protocol that focused on both privacy and trustworthiness problems.

For various mobile sensing applications, some recent research [8] addresses trust and privacy simultaneously. Their root of trust is that multiple users can examined the same data at the same time and location, and thus a measurement that deviates from the majority can be deemed as untrusted. However, this is wrong for a wide range of aggregation applications, where all users creates their own data individually and the validity of their data cannot be verified from other users' data. For such applications, their methods may erroneously sanction honest users with various data. Moreover, since they focused on general mobile sensing applications rather than data aggregation, only anonymity is offer for security. However, anonymity-based approaches are known to be vulnerable to tracking attacks [9].

In [1], privacy and trustworthiness problems is implemented together for data aggregation in mobile sensing. It also offer data trustworthiness by ensuring that every user must apply a true data value predefined by the aggregation server.

In case, if a server wants to count the number of users affected by a flu, it can determined two values 1 (mention as affected person) and 0 (mention as unaffected person) for each and every user to select from given data. The proposed method guarantee that a malicious user who submits values other than 1 and 0 will be detected by the server, which conclude that the malicious user cannot submit 1, 000 to win over the server that many people have caught this flu. As a comparison of this system to existing system is that preserving the privacy by submitting data in an anonymous way, proposed system provide security by hiding the user data from the server, i.e., user's data is not displayed to the server in clear text.

2. Technical Keywords

1) **Blind signature:** Blind signature is a form of digital signature. It can be implemented with common public key signing schemes, e.g., RSA and DSA. To get a blind signature on message x, the user combines it with a random "blinding factor". The blinded message x0 is sent to the signer and the signer signs it using its private key. The resulting signature s0 is sent back to the user. The user then removes the blinding factor from s0 to get a signature s on x. s is identical to the signature generated by a normal signing protocol and thus can be verified with the signer's public key. This blind signature technique can ensure both unlinkability and unforgeability.
2) **Data Value Vector:** The data value vector may be linear or non-linear. For example, to monitor the flu trend by collecting how many users are infected, the server can set the data value vector as [0; 1], where 0 means that a user is fine and 1 means that a user has caught this flu. The data value vector is a novel concept to provide data trustworthiness and restrict the behavior of malicious users. It is a vector of all data values that are considered to be valid by the server, sorting in increasing order
3) **Trusted Authority:** Assume that a trusted authority is responsible for generating secrets for the server and MNs. The assumption can be relaxed to an honest-but-curious key dealer which does not collude with the aggregation server. The key dealer follows proposed protocol to generate secrets for all parties but it may try to infer MNs' data value by eavesdropping communications between the server and MNs. Under this relaxed assumption, we only need to add an encryption and decryption step to the Data Aggregation phase: each MN encrypts the encrypted report ci with a pre-shared key with the server and sends the final result to the server. To get the aggregate statistics, the server first decrypts each MN's cipher text using the pre-shared key and then does the Aggregation and Decryption.

3. Literature Survey

This part talks about the researches that have done about secure data aggregation in mobile sensing in detailed manner.

The approach described in [1] proposed a PATA protocol. PATA protocol permits the server to collect the data which is add by mobile users without knowing the data of individual user. At the same time, if malicious users try to submit wrong data, system will be found or the denied aggregation result by the server. Inthis way, the malicious users' effect on the aggregation result is effectively limited

In [2] Authors developed a new system called as VTrack, which is used for predicting travel time estimation depending on the data sensor data which will help to solve the two main issues 1) Sensor Unreliability and 2) Energy Utilization. GPS provides the precise location estimation, but still it has many drawbacks such as most phones doesn't support GPS, GPS don't work in urban areas such as tall buildings and tunnels also the GPS on many phones are power consuming in such cases VTrack can work on alternative, less power consuming but noisier sensors like WiFi to predict users trajectory as well as the travel time in the specified route. VTrack make use of Hidden Markov Model (HMM) based map matching method and estimation of travel time technique which can add scattered data to get the most probable road segments on which user has driven and to attribute travel time to those segments. By test results authors demonstrated that VTrack can handle noise and outages in location estimates and provides proper delays for delay-aware routing algorithms.

Author in paper [3] designed, developed, evaluated and also shared user experiences ofhis application named as NoiseSpy. This system is a sound sensing system which runs on mobile and turns the mobile phones in to data logger for monitoring environmental noise which is also low cost. It can explore city area while collaboratively visualizing noise levels in real-time. NoiseSpy software consolidates the

sound levels with GPS data for creating a map of sound levels that were encountered during a journey.

In [4] authors developed a Heart-to-Heart (H2H) system, for system is used to authenticate external medical device controllers and also for programmers to Implantable Medical Devices (IMDs), which haves pacemakers and cardiac defibrillators these are the therapeutic medical devices which can be partially or fully embedded in human body. For providing non-invasive reprogramming and also for data readout they have built-in radio communication. Due to improper authentication protocols some of IMDs can expose patients to the-air attack and physical harm. Proposed system uses ECG (Heartbeat data) for an authentication technique, which makes sure to allow access only for medical instruments in physical contact with an IMD-bearing patient. They also developed a novel method for retrieving time-varying randomness from ECG signals for use in H2H. Authors also developed a new cryptographic device pairing protocol which makes use of this randomness to prevent the attacks by active adversaries, while meeting the practical challenges of lightweight implementation and noise tolerance in ECG readings. Lastly, authors elaborate an end-to-end development in an ARM-Cortex M-3 microcontroller which demonstrates the practicality of H2H in present IMD hardware.

In [5] authors presented the design and evaluation of PriSense. It's a new way to preserving privacy data aggregation in people-centric urban sensing system. The basic concept of PriSense is data slicing as well as mixing and support a numbers of statistical additive and non-additive aggregation function including Sum, Average, Variance, Count, Max/Min, Median, Histogram, and Percentile with exact aggregation results. This introduced technique can support strong user privacy against a tunable threshold number of colluding users and aggregation servers.

In [6] authors presented PASTE, which is the first differentially private aggregation algorithm for distributed time-series data that gives good practical utility without any trusted server. PASTE solves two important challenges in participatory data-mining applications those are (i) individual users collect temporally correlated time-series data, and (ii) an untrusted third-party aggregator wishes to run aggregate queries on the data. To ensure differential privacy for time-series data despite the presence of temporal correlation, to address this, PASTE incorporates two new algorithms. PASTE uses the Fourier Perturbation Algorithm (FPAk). To deal with the absence of a trusted central server, PASTE uses the Distributed Laplace Perturbation Algorithm (DLPA) that maximizes noise in a distributed way in order to guarantee differential privacy.

4. Proposed Work

4.1 Problem Statement

Now days essentiality of the mobile device such as smartphones and tablets, in which having various number of mobile sensing applications are used in different area such as for air pollution monitoring and for healthcare survey. With the increasing capabilities of mobile devices such as smartphones and tablets, there are more and more mobile sensing applications such as air pollution monitoring and healthcare. These applications generally combine the information given by mobile users to show the information about human's activities or surroundings. When the data supplied by mobile users is decent and trustworthy then the mobile sensing works properly. System may be damage if the mobile users act like eavesdropper due to malicious impact and submit forge data. Sometimes, mobile sensing users unable to submit data because of security issue

4.2 Problem Definition

A new privacy-aware and trustworthy data aggregation protocol for mobile sensing is proposed. PATA protocol permits the server to collect the data which is added by mobile users without knowing the data of individual user. At the same time, if malicious users try to submit wrong data, system will be found or the denied aggregation result by the server. In this way, the malicious users' effect on the aggregation result is effectively limited.

4.3 Proposed Solution

The overview of PATA protocol is shown in the Figure 2

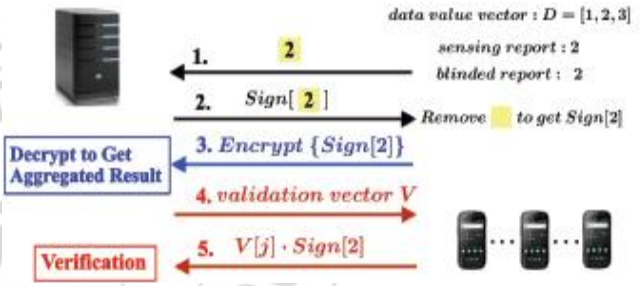

Figure 2: Overview of PATA protocol

This approach is working under four phases: Setup, Data Preprocessing, Data Aggregation, and Data Validation.

1. Setup: In this phase, a set of secrets are generated and assigned to the server and MNs, which will be used to generate encryption and decryption keys during the Data Aggregation phase. The secret distribution and key generation process ensure that the server can only decrypt the aggregated result but cannot decrypt any individual's report. Also, the server will generate system parameters, RSA keys and the data value vector and then broadcast them to all MNs.

2. Data Preprocessing: Each MN picks a value D[j] from the data value vector D based on its sensing result and uses D[j] in its sensing report. For example, if the sensing result is 2:1 and the data value vector is [1; 2; 3], the MN will use 2 as its data value in the report. Then each MN gets its sensing report blindly signed by the server so that the server cannot see the data value. This signed report will be used in the following two phases by MNs. Since an MN cannot generate a signature itself, it will not be able to change the data value.

3. Data Aggregation: In this phase, each MN encrypts the signed report it obtained during the Data Preprocessing phase using its encryption key and then sends it to the server. On receiving all ciphertexts, the server aggregates them together and then decrypts it using its decryption key to get the aggregated result. The server will not be able to decrypt any individual MN's report because their keys satisfy certain conditions.

4. Data Validation: In this phase, the server validates if each MN has submitted a report with valid data from the data value vector and has followed our protocol honestly. Firstly, the server generates a validation vector V based on the data value vector and the task ID, and sends V to each MN. On receiving this validation vector, the MN multiplies its own signed report with the corresponding element in the validation vector. The result is sent to the server who can verify if the MN has cheated or not based on this result. The whole process is done with encryption, and thus the server is able to validate each MN's report without knowing the data value.

5. Conclusion

A new privacy-aware and trustworthy data aggregation protocol PATA has been studied in this paper. This protocol utilizes blind signature, holomorphic encryption and a novel encrypted vector-based data validation technique. Without knowing any individual's data, server might be aggregate user's data. This protocol also actives the server to validate users' trustworthiness in a privacy-preserving manner based on whether they submit valid data report following protocol. Security measurement proves that method used can secure the privacy of mobile users from a curious server and simultaneously it can protect the system from malicious users. Theoretical conclusion of method proves that protocol runs very fast even when the plaintext space is large and many mobile users exist in the system. Very low power consumed by the smartphones and computational devices. These conclusions show that the given protocol is suitable for a broad range of mobile sensing applications with typical plaintext spaces and resource component.

References

[1] Jingyao Fan, Qinghua Li and Guohong Cao, "Privacy-Aware and Trustworthy Data ggregation in Mobile Sensing", in IEEE Conference on Communications and Network Security (CNS), 2015
[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, " in Proc. of ACM SenSys, 2009
[3] E. Kanjo, "NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping, " MONET, vol. 15, no. 4, 2010.
[4] M. Rostami, A. Juels, and F. Koushanfar, "Heart-to-Heart (H2H):Authentication for Implanted Medical Devices, " in ACM SIGSAC, 2013.
[5] J. Shi, Y. Zhang, and Y. Liu, "Prisense: Privacy-Preserving Data Aggre-gation in People-Centric Urban Sensing Systems, " in INFOCOM, 2010.
[6] V. Rastogi and S. Nath, "Differentially Private Aggregation of Distributed Time- Series with Transformation and Encryption, " in Proc. of ACM SIGMOD, 2010.
[7] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure Information Aggregation in Sensor Networks, " in Proc. of ACM SenSys, 2003.
[8] W. Zhang, S. K. Das, and Y. Liu, "A Trust Based Framework for Secure Data Aggregation in Wireless Sensor Networks, " in Proc. of IEEE SECON, 2006.
[9] L. Kazemi and C. Shahabi, "TAPAS: Trustworthy Privacy-Aware Participatory Sensing, " KAIS, vol. 37, no. 1, 2013.
[10] H. Zang and J. Bolot, "Anonymization of Location Data Does Not Work: A Large- Scale Measurement Study, " in Proc. of ACM MobiCom, 2011.
[11] R. Zhang, J. Shi, Y. Zhang, and C. Zhang, "Verifiable Privacy-Preserving Aggregation in People-Centric Urban Sensing Systems, " Selected Areas in Communications, IEEE Journal on, vol. 31, no. 9, 2013
[12] M. M. Groat, W. He, and S. Forrest, "KIPDA: k-Indistinguishable Privacy-preserving Data Aggregation inWireless Sensor Networks, " in Proc. of IEEE INFOCOM, 2011.
[13] Q. Li and G. Cao, "Efficient Privacy-Preserving Stream Aggregation in Mobile Sensing with Low Aggregation Error, " in PETS, 2013.
[14] X. Xu, Q. Wang, J. Cao, P.-J. Wan, K. Ren, and Y. Chen, "Locating Malicious Nodes for Data Aggregation in Wireless Networks, " in Proc. of IEEE INFOCOM, 2012.
[15] W. Hu and G. Cao, "Energy-Aware Video Streaming on Smartphones, "in Proc. Of INFOCOM, 2015.
[16] K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz, "Sensing Techniques for Mobile Interaction, " in Proc. of ACM UIST, 2000.

Author Profile

Mr. Devendra S. Hapase, is currently pursuing M.E (Computer) from Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India -411007. He received his B.E (Computer) Degree from SKNCOE, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India -411007. His area of interest is network security and web& data .mining

Prof. M.D Ingle, received his M Tech. (Computer) Degree from Dr. Babasaheb Ambedkar Technological University, Lonere, Dist. Raigad-402 103, Maharashtra, India. He received his B.E (Computer) Degree from Govt college of Engineering, Aurangabad, Maharashtra, India. He is currently working as M.E coordinator and Asst Prof (Computer) at Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India - 411007. His area of interest is network security and web & data mining.