Volume 4 || Special Issue 8 || ICCEME 2019-2020 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

PRIVACY PRESERVING AND EFFICIENT INCENTIVE MECHANISM WITH COLLABORATIVE COMPUTING

Ms. Rokade Monika D.[1], Ms. Shelke Nutan S.[2]

Professor/ Student, Computer, SPCOE Dumbarwadi, Otur, Pune, India [1 2]

Monikarokade4@gmail.com[1], nutanshelke01@gmail.com[2]

Abstract- Collaborative computing utilizes different information servers to mutually finish information investigation, e.g., factual examination and surmising. One significant deterrent for it lies in privacy concern, which is straightforwardly connected with hubs' investment and the devotion of got information. Existing privacy-preserving ideal models for distributed computing and circulated information total just give hubs homogeneous privacy security without thought of hubs' various trust degrees to various information servers. We propose a two-phase structure that registers the normal worth while preserving heterogeneous privacy for hubs' private information. The new test is that in the reason of meeting privacy necessities, we ought to ensure the proposed system has a similar calculation precision with existing privacy-mindful arrangements. In this paper, hubs acquire heterogeneous privacy assurance despite various information servers through one-shot clamor bother. In light of the meaning of KL privacy, we determine the logical articulations of the privacy preserving degrees (PPDs) and measure the connection between various PPDs. At that point, we acquire the shut structure articulation of calculation exactness. Moreover, a proficient motivating force system is proposed to accomplish improved calculation exactness at the point when information servers have fixed financial plans. At last, broad reenactments are directed to confirm the got hypothetical outcomes.

Keywords: *Collaborative computing, average consensus, privacy preservation, incentive mechanism*

I INTRODUCTION

In the era of Internet of Things (IoT), more and more valuable data is generated in local smart devices, e.g., smartphones, smartwatches, even smart meters and refrigerators [2]. Drawing support from cloud computing framework, data servers can aggregate these data for in-depth analysis, such as statistical inference, personalized recommendation, and health monitoring. However, the amounts of data generated in IoT are unprecedented, which brings extremely large communication, computation and storage costs for data servers. Moreover, the data produced by smart devices are generally geo-distributed [3]. In particular, for cross-border data sources and servers, direct data communication may be prohibited by nations' policies. Hence, it is impractical for one single server to fulfill such data aggregation and analysis. A promising alternative computation framework is collaborative computing, which assigns the aggregation and computation tasks to multiple data servers and instructs servers to collaboratively complete the data analysis. Obviously, such a computation framework significantly mitigates the resource bottlenecks of single server. Edge computing [4] is a representative paradigm of collaborative computing. Also, some social networking service companies, such as Facebook and Twitter, have deployed multiple servers in different nations and utilized these servers to jointly underpin advanced tasks [5], [6]. Nevertheless, the data generated in IoT often contains users' sensitive information, such as location, medical data and energy consumption [7]–[9]. Moreover, privacy issues occur frequently, e.g., it is reported that the private information of about 50 million Facebook users has been disclosed [10]. Many countries formulate privacy policies to restrict data servers to make use of users' private data, and some privacy requirements specification methods have been proposed to map the privacy polices to a formal language in description logic to ensure consistency between these polices and data usage [11], [12]. Therefore, privacy has become an urgent concern for data aggregation and analysis [13]. In this paper, we view all users in IoT as generalized nodes which possess private data. Multiple geo-distributed data servers first aggregate data from different groups of nodes (corresponding to distinct locations), and then collaboratively complete statistical analysis. For a group of nodes, data servers are categorized into two types of data processors, which have different permissions to nodes' data. The server collecting their private data is the first type while other servers having no direct connection with them are viewed as the second type. The second type of data processors are less trustworthy since they have no permission to directly

access nodes' data. This brings a new critical concern for nodes when releasing private data, which is how to ensure data processors with distinct trust degrees receive data versions with different privacy protection. That is, heterogeneously privacy-preserving problem should be further investigated in collaborative computing.

II LITERATURE SURVEY

Security saving average consensus issue concerns ensuring the underlying conditions of members not to be unveiled while accomplishing average consensus. In existing written works, a large portion of works consider giving protection assurance to specialists when the various members are seen as a similar kind of security aggressors. Be that as it may, for a specialist, not the various members are deceitful. The specialist can set more vulnerable protection request against these believable members to get potential utility improvement. In this paper, we consider that network is made out of a few gatherings. Inside each gathering, specialists treat members inside and outside the gathering as two kinds of aggressors. At that point, a private average consensus algorithm (PACA) is proposed to give diverse security insurance against members inside and outside the gathering, by bothering operators' underlying states with irregular noises[1].

Universal detecting empowered by Wireless Sensor Network (WSN) advances cuts across numerous areas of cutting edge living. This offers the capacity to quantify, deduce and comprehend ecological pointers, from sensitive ecologies and regular assets to urban situations. The multiplication of these gadgets in an imparting impelling network makes the Internet of Things (IoT), wherein sensors and actuators mix flawlessly with nature around us, and the data is shared across stages so as to build up a commON OPERATING PICTURE (COP).[2]

Numerous huge associations gather enormous volumes of data every day in a geologically conveyed manner, at data bases on the globe. Regardless of their topographically assorted starting point the data must be prepared and dissected all in all to separate understanding. We call the issue of supporting huge scope geo-appropriated investigation WIDE-AREA BIG DATA (WABD).[3]

The expansion of Internet of Things (IoT) and the achievement of rich cloud administrations have pushed the skyline of another figuring worldview, edge registering, which calls for preparing the data at the edge of the network. Edge registering can possibly address the worries of reaction time necessity, battery life imperative, transfer speed cost sparing, just as DATA WELLBEING AND PRIVACY.[4]

Versatile examination on enormous data sets has been center to the elements of various groups at Facebook - both building and non engineering. Aside from impromptu investigation of data and formation of business knowledge dashboards by investigators over the organization, some of Facebook's site highlights are additionally founded on dissecting enormous data sets. These highlights run from straightforward announcing applications like Insights for the Facebook Advertisers, to further developed sorts, for example, companion recommendations[5]

As of late, there has been a generous measure of work for enormous scope data examination utilizing Hadoop-put together stages running with respect to huge bunches of item machines. A less investigated point is the manner by which those data, overwhelmed by application logs, are gathered and organized in any case. In this paper, we present Twitter's creation logging foundation and its advancement from application-explicit logging to a bound together "customer occasions" log position, where messages are caught in common, all around organized, adaptable Thrift messages. Since most investigation errands consider the client meeting as the fundamental unit of examination, we pre-emerge "meeting successions", which are reduced rundowns that can answer a huge class OF COMMON QUESTIONS QUICKLY[6]

In an Internet of Things network, different sensors send data to a combination community for it to gather an open speculation of intrigue. Notwithstanding, a similar sensor data might be utilized by the combination place to make deductions of a private sort that the sensors wish to secure. To display this, we receive a decentralized speculation testing structure with double open and private theories. Every sensor mentions a private objective fact and uses a nearby sensor choice guideline or protection mapping to sum up that perception freely of the other sensors.[7]

As of late, wireless sensor networks have been widely utilized in social insurance applications, for example, medical clinic and home patient checking. Wireless clinical sensor networks are progressively defenseless against spying, alteration, pantomime and replaying assaults than the wired networks. A great deal of work has been done to make sure about wireless clinical sensor networks.[8]

Deregulated power markets with time-differing power costs and open doors for buyer cost alleviation makes vitality stockpiling, for example, a battery, an appealing suggestion. Sharing an enormous limit battery over a gathering of homes in a network can lighten the financial hindrances as well as endeavor the way that clients' action designs don't really cover. Be that as it may, battery sharing incites rivalry for battery limit between the clients as a rule as they might need to boost their own cost reserve funds by involving more battery limit when the

power cost is low. Significantly, clients may have protection concerns when they speak with the common battery controller[9].

Sincerely charged pictorial materials are much of the time utilized in fear inquire about, however no current normalized picture database is devoted to the investigation of various fears. The current work portrays the consequences of two free examinations through which we looked to create and approve this sort of database—a Set of Fear Inducing Pictures (SFIP).[10].

III SYSTEM DESIGN

Our objectives are to i) analyze whether the two-phase average-computing framework can provide different privacy guarantee for nodes' private data against distinct privacy violators; ii) give in-depth analysis about the computation accuracy of the privacy-preserving scheme; iii) propose an incentive mechanism for data servers to assign appropriate compensation for nodes's privacy loss in order to obtain satisfied computation accuracy. We next discuss the details of the proposed framework

Perturbed Data Reporting: In Phase 1, each node first perturbs its private data with a Gaussian noise according to its privacy demand (against the first type of violators). Then, the node reports the noisy data to data server. Specifically, the perturbed data should provide privacy preservation with PPD

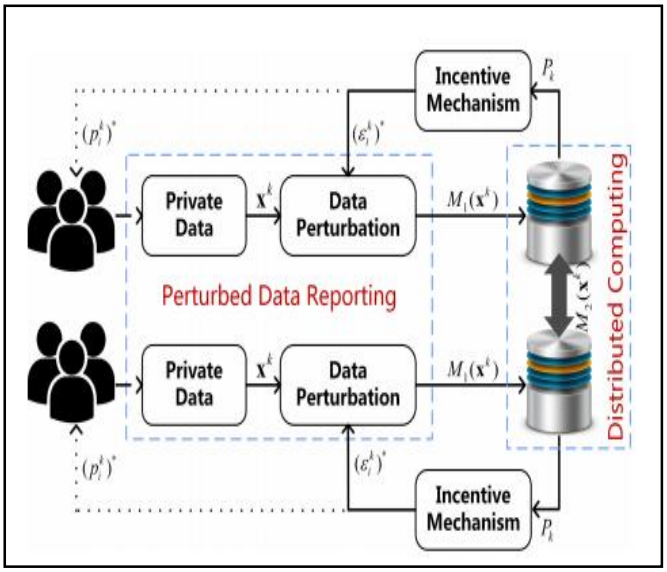

Fig.1 System Design

Distributed Average Computation: In Phase 2, different data servers collaboratively compute the average value using a general average consensus algorithm. When a server communicates its aggregated data with other servers, the released information about nodes' private data should preserve stronger privacy. Suppose the PPD in Phase 2. In addition, the computation result needs to meet the accuracy requirement.

Incentive Mechanism: There exists a conflict between computation accuracy and privacy protection. To provide an efficient solution to the conflict, we adopt an incentive mechanism to compensate nodes' privacy loss. After receiving the compensation, nodes will report perturbed data with PPDs corresponding to the rewards. Specifically, to minimize the computation deviation under a fixed incentive budget, each data server computes the compensation and PPD for each node by solving an optimization problem.

IV CONCLUSION

We proposed a two-phase computation framework to instruct data servers to aggregate nodes' private data and compute the average value collaboratively. More importantly, through one-shot noises perturbation, nodes obtain heterogeneous privacy guarantee against different types of privacy violators. We also derived the closed-form expressions of two PPDs and computation accuracy of the proposed framework. Then, in order to efficiently solve the conflict between privacy and accuracy, we devised an incentive mechanism for data servers, which provides optimal computation accuracy when servers have fixed incentive budget. Lastly, extensive simulations verified the obtained theoretical results.

REFERENCES

[1] X. Wang, J. He, P. Cheng, and J. Chen, "Privacy preserving average consensus with different privacy guarantee," in Proc. IEEE Annu. Amer. Control Conf., 2018, pp. 5189–5194.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (IoT): A vision, architectural elements, and future directions," Future Gener. Comput. Syst., vol. 29, no. 7, pp. 21645–1660, 2013.

[3] A. Vulimiri et al., "WANalytics: Geo-distributed analytics for a data intensive world," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2015, pp. 1087–1092.

[4] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," IEEE Internet Things J., vol. 3, no. 5, pp. 637–646, Oct. 2016.

[5] A. Thusoo et al., "Data warehousing and analytics infrastructure at facebook," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2010, pp. 1013– 1020.

[6] G. Lee, J. Lin, C. Liu, A. Lorek, and D. Ryaboy, "The unified logging infrastructure for data analytics at twitter," Proc. VLDB Endowment, vol. 5, no. 12, pp. 1771–1780, 2012.

[7] M. Sun, W. P. Tay, and X. He, "Toward information privacy for the internet of things: A nonparametric learning approach," IEEE Trans. Signal Process., vol. 66, no. 7, pp. 1734–1747, Apr. 2018.
[8] X. Yi, A. Bouguettaya, D. Georgakopoulos, A. Song, and J. Willemson, "Privacy protection for wireless medical sensor data," IEEE Trans. Dependable Secure Comput., vol. 13, no. 3, pp. 369–380, May/Jun. 2016.
[9] J. Yao and P. Venkitasubramaniam, "Privacy aware stochastic games for distributed end-user energy storage sharing," IEEE Trans. Signal Inf. Process. Netw., vol. 4, no. 1, pp. 82–95, Mar. 2018.
[10] N. Cohen, "Facebook isn't violating our privacy," 2018. [Online]. Available: https://www.nytimes.com/2018/03/29/opinion/facebookprivacy-zuckerberg- society
[11] T. D. Breaux, H. Hibshi, and A. Rao, "Eddy, a formal language for specifying and analyzing data flow specifications for conflicting privacy requirements," Requirements Eng., vol. 19, no. 3, pp. 281–307, 2014.
[12] D. Smullen and T. D. Breaux, "Modeling, analyzing, and consistency checking privacy requirements using Eddy," in Proc. ACM Symp. Bootcamp Sci. Secur., 2016, pp. 118–120.
[13] A. A. Cardenas, P. K. Manadhata, and S. P. Rajan, "Big data analytics for security," IEEE Secur. Privacy, vol. 11, no. 6, pp. 74–76, Nov./ Dec. 2013.
[14] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy preserving and copy-deterrence content-based image retrieval scheme in cloud computing," IEEE Trans. Inf. Forensics Secur., vol. 11, no. 11, pp. 2594–2608, Nov. 2016.
[15] Monika D. Rokade, Dr. Yogesh Kumar Sharma. "Identification of Malicious Activity for Network Packet using Deep Learning" IJAST, vol.29 No. (9s), 2324-2331, (2020)