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

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 24th Feb 2018. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-02

Title: SECURING DATA AND ACHIEVING EFFICIENCY FOR CONSUMERS DATA

Volume 08, Issue 02, Pages: 98–102.

Paper Authors

MR.K.SRINIVASA RAO, B.ANIL KUMAR

Vignan's Lara Institute of Technology & Science

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

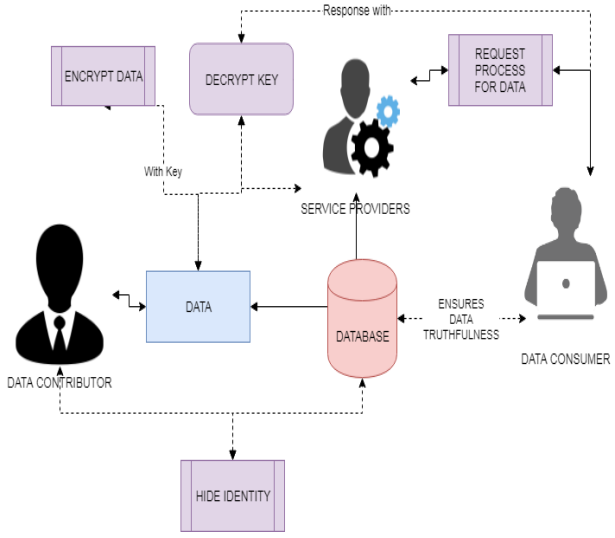
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SECURING DATA AND ACHIEVING EFFICIENCY FOR CONSUMERS DATA

MR.K.SRINIVASA RAO[1], B.ANIL KUMAR[2]
Assistant Professor[1], Department of M.C.A ,Vignan's Lara Institute of Technology & Science
M.C.A Student[2],Department of M.C.A ,Vignan's Lara Institute of Technology & Science

Abstract:
As a significant business paradigm, many online information platforms have emerged to satisfy society's needs for person-specific data, where a service provider collects raw data from data contributors, and then offers value-added data services to data consumers. However, in the data trading layer, the data consumers face a pressing problem, i.e., how to verify whether the service provider has truthfully collected and processed data? Furthermore, the data contributors are usually unwilling to reveal their sensitive personal data and real identities to the data consumers. In this paper, we propose TPDM, which efficiently integrates Truthfulness and Privacy preservation in Data Markets. TPDM is structured internally in an Encrypt-then-Sign fashion, using partially holomorphic encryption and identity-based signature. It simultaneously facilitates batch verification, data processing, and outcome verification, while maintaining identity preservation and data confidentiality. We also instantiate TPDM with a profile matching service and a data distribution service, and extensively evaluate their performances on Yahoo! Music ratings dataset and 2009 RECS dataset, respectively. Our analysis and evaluation results reveal that TPDM achieves several desirable properties, while incurring low computation and communication overheads when supporting large-scale data markets.

Architecture



Introduction:
In the era of big data, society has developed an insatiable appetite for sharing personal data. Realizing the potential of personal data's economic value in decision making and user experience enhancement, several open information platforms have emerged to enable person-specific data to be exchanged on the Internet [1], [2], [3], [4], [5]. For example, Gnip, which is Twitter's enterprise API platform, collects social media data from Twitter users, mines deep insights into customized audiences, and provides data

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

analysis solutions to more than 95% of the Fortune 500 [2].However, there exists a critical security problem in these market-based platforms, i.e., it is difficult to guarantee the truthfulness in terms of data collection and data processing,especially when privacies of the data contributors are needed to be preserved. Let's examine the role of a pollster in the presidential election as follows. As a reliable source of intelligence, the Gallup Poll [6] uses impeccable data toassist presidential candidates in identifying and monitoring economic and behavioral indicators. In this scenario, simultaneously ensuring truthfulness and preserving privacy require the Gallup Poll to convince the presidential candidates that those indicators are derived from live interviews without leaking any interviewer's real identity (e.g., social security number) or the content of her interview. If raw data sets for drawing these indicators are mixed with even a small number of bogus or synthetic samples, it will exert bad influence on the final election result. Ensuring truthfulness and protecting the privacies of data contributors are both important to the long term healthy development of data markets. On one hand, the ultimate goal of the service provider in a data market is to maximize her profit. Therefore, in order to minimize the expenditure for data acquisition, an opportunistic way for the service provider is to mingle some bogus or synthetic data into the raw data sets.

Existing system:

To integrate truthfulness and privacy preservation in a practical data market, there are four major challenges. The first and the thorniest design challenge is that verifying the truthfulness of data collection and preserving the privacy seem to be contradictory objectives. Ensuring the truthfulness of data collection allows the data consumers to verify the validities of data contributors' identities and the content of raw data, whereas privacy preservation tends to prevent them from learning these confidential contents. Specifically, the property of non-repudiation in classical digital signature schemes implies that the signature is unforgeable, and any third party is able to verify the authenticity of a data submitter using her public key and the corresponding digital certificate, i.e., the truthfulness of data collection in our model. However, the verification in digital signature schemes requires the knowledge of raw data, and can easily leak a data contributor's real identity. Regarding a message authentication code (MAC), the data contributors and the data consumers need to agree on a shared secret key, which is unpractical in data markets. Yet, another challenge comes from data processing, which makes verifying the truthfulness of data collection even harder. Nowadays, more and more data markets provide data services rather than directly offering raw data. The following three reasons account for such a trend: 1) For the data contributors, they have several privacy concerns.Nevertheless, the service-based trading mode, which has hidden the

sensitive raw data, alleviates their concerns; 2) for the service provider, semantically rich and insightful data services can bring in more profits; 3) for the data consumers, data copyright infringement and datasets resale are serious. However, such a data trading mode differs from most of conventional data sharing scenarios, e.g., data publishing. Besides, the result of data processing may no longer be semantically consistent with the raw data, which makes the data consumer hard to believe the truthfulness of data collection. In addition, the digital signatures on raw data become invalid for the data processing result, which discourages the data consumer from doing verification as mentioned above. Moreover, although data provenance helps to determine the derivation history of a data processing result, it cannot guarantee the truthfulness of data collection.

Proposed system:

In this Project, by jointly considering above four challenges, we propose TPDM, which achieves both Truthfulness and Privacy preservation in Data Markets. TPDM first exploits partially holomorphic encryption to construct a cipher text space, which enables the service provider to launch data services and the data consumers to verify the correctness and completeness of data processing results, while maintaining data confidentiality. In contrast to classical digital signature schemes, which are operated over plaintexts, our new identity-based signature scheme is conducted in the cipher text space. Furthermore, each data contributor's signature is derived from her real identity, and is unforgivable against the service provider or other external attackers. This appealing property can convince data consumers that the service provider has truthfullycollected data. To reduce the latency caused by verifying a bulk of signatures, we propose a two-layer batch verification scheme, which is built on the linearity of admissible pairing. At last, TPDM realizes identity preservation and revocability by carefully adopting ElGamal encryption and introducing a semi-honest registration center. We summarize our key contributions as follows. To the best of our knowledge, TPDM is the first secure mechanism for data markets achieving both data truthfulness and privacy preservation. TPDM is structured internally in a way of Encrypt then- Sign using partially holomorphic encryption and identity-based signature. It enforces the service provider to truthfully collect and to process real data. Besides, TPDM incorporates a two-layer batch verification scheme with an efficient outcome verification scheme, which can drastically reduce computation overhead.

Modules:

1. **DATA COLLECTION**

   Data contributors are providing the data to share the details with service providers who take care of the data and secrecy of the data. Data contributor uploads the data with key and which is encrypted. In order to maintain secrecy of the sensitive information about data contributors, the id is maintained rather than their details.

2. **CONSUMING DATA**

Data consumers are requesting the data to consume from service provider. If service providers accept the request means they are offering the data with decryption key and the id of the data contributor, in order to maintain the truthfulness of the data and also preserve the privacy of the data contributors.

3. **DATA EXTRACTION**

Data can be extracting from the encrypted data. With encrypted data and key, data consumers can decrypt the data from the signature (i.e., id) we can ensures the truthfulness of the data from the data owner's signature.

4. **GRAPHICAL NOTATIONS**

The collected data are representing as graphical form which is help to identify the best way ofanalyzing the performance of proposed system. The graphs are different like pie chart, bar chart and column chart. The better way to understand the data in which it helps to find the best method among available.

Conclusion:

In this paper, we have proposed the first efficient secure scheme TPDM for data markets, which simultaneously guarantees data truthfulness and privacy preservation. In TPDM, the data contributors have to truthfully submit their own data, but cannot impersonate others. Besides,the service provider is enforced to truthfully collect and process data. Furthermore, both the personally identifiable information and the sensitive raw data of data contributors are well protected. In addition, we have instantiated TPDM with two different data services, and extensively evaluated their performances on two real-world datasets. Evaluation results have demonstrated the scalability of TPDM in the context of large user base, especially from computation and communication overheads. At last, we have shown the feasibility of introducing the semi-honest registration center with detailed theoretical analysis and substantial evaluations.

References

[1] "Microsoft Azure Marketplace," https://datamarket:azure:com/home/.
[2] "Gnip," https://gnip:com/.
[3] "DataSift," http://datasift:com/.
[4] "Datacoup," https://datacoup:com/.
[5] "Citizenme," https://www:citizenme:com/.
[6] "Gallup Poll," http://www:gallup:com/.
[7] M. Barbaro, T. Zeller, and S. Hansell, "A face is exposed for AOL searcher no. 4417749," New York Times, Aug. 2006.
[8] "2016 TRUSTe/NCSA Consumer Privacy Infographic – US Edition," https://www:truste:com/resources/privacy-research/ncsa-consumer-privacy-index-us/.
[9] K. Ren,W. Lou, K. Kim, and R. Deng, "A novel privacy preserving authentication and access control scheme for pervasive computing environments," IEEE Transactions on Vehicular Technology, vol. 55, no. 4, pp. 1373–1384, 2006.
[10] M. Balazinska, B. Howe, and D. Suciu, "Data markets in the cloud: An opportunity

for the database community," PVLDB, vol. 4, no. 12, pp. 1482–1485, 2011.
[11] P. Upadhyaya, M. Balazinska, and D. Suciu, "Automatic enforcement of data use policies with datalawyer," in SIGMOD, 2015.
[12] T. Jung, X.-Y. Li, W. Huang, J. Qian, L. Chen, J. Han, J. Hou, andC. Su, "AccountTrade: accountable protocols for big data trading against dishonest consumers," in INFOCOM, 2017.
[13] G. Ghinita, P. Kalnis, and Y. Tao, "Anonymous publication of sensitive transactional data," IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 2, pp. 161–174, 2011.
[14] B. C. M. Fung, K.Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing: A survey of recent developments," ACM Computing Surveys, vol. 42, no. 4, pp. 1–53, Jun. 2010.
[15] R. Ikeda, A. D. Sarma, and J.Widom, "Logical provenance in dataoriented workflows?" in ICDE, 2013.
[16] M. Raya and J. Hubaux, "Securing vehicular ad hoc networks," Journal of Computer Security, vol. 15, no. 1, pp. 39–68, 2007.
[17] T. W. Chim, S. Yiu, L. C. K. Hui, and V. O. K. Li, "SPECS: secure and privacy enhancing communications schemes for VANETs," Ad Hoc Networks, vol. 9, no. 2, pp. 189 – 203, 2011.
[18] D. Boneh, E. Goh, and K. Nissim, "Evaluating 2-dnf formulas on ciphertexts," in TCC, 2005.
[19] R. A. Popa, A. J. Blumberg, H. Balakrishnan, and F. H. Li, "Privacy and accountability for location-based aggregate statistics," in CCS, 2011.
[20] J. H. An, Y. Dodis, and T. Rabin, "On the security of joint signature and encryption," in EUROCRYPT, 2002.