ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Keeping the privacy safe for Back- Propagation Neural Network Learning Made Practical with Cloud Computing

C.Sampath kumar
M.Tech Student,
Department of CSE,
Sree Rama institute of Technology and Science,
Kuppenakuntla, Penuballi, Khammam,TS India.

P.Nageswara Rao
Assistant Professor,
Department of CSE,
Sree Rama institute of Technology and Science,
Kuppenakuntla, Penuballi, Khammam,TS India.

ABSTRACT:

To improve the accuracy of learning result, in practice multiple parties may collaborate through conducting joint Back-Propagation neural network learning on the union of their respective data sets. During this process no party wants to disclose her/his private data to others. Existing schemes supporting this kind of collaborative learning are either limited in the way of datapartition or just consider two parties. There lacks a solution that allows two or more parties, each with an arbitrarily partitioned data set, to collaboratively conduct the learning. This paper solves this open problem by utilizing the power of cloud computing. In our proposed scheme, each party encrypts his/her private data locally and uploads the ciphertexts into the cloud.

The cloud then executes most of the operations pertaining to the learning algorithms over ciphertexts without knowing the original private data. By securely offloading the expensive operations to the cloud, we keep the computation and communication costs on each partyminimal and independent to the number of participants. To support flexible operations over ciphertexts, we adopt and tailor the BGN "doubly homomorphic" encryption algorithm for the multiparty setting. Numerical analysis and experiments on commodity cloud show that our scheme is secure, efficient, and accurate.

Index Terms:

Privacy reserving, learning, neural network, back-propagation, cloud computing, computation outsource.

INTRODUCTION:

BACK-PROPAGATION [18] is an effective method for learningneural networks and has been widely used in various applications. The accuracy of the learning result, despite other facts, is highly affected by the volume of high-quality data used for learning. As compared to learning with only local data set, collaborative learning improves the learning accuracy by incorporating more data sets into the learning process [21], [11], [20]: the participating parties carry out learning not only on their own data sets, but also on others' data sets. With the recent remarkable growth of new computing infrastructures such as cloud computing, it has been more convenient than ever for users across theInternet, who may not even know each other, to conduct joint/collaborative learning through the shared infrastructure [13], [14].Despite the potential benefits, one crucial issue pertaining to the Internet-wide collaborative neural network learning is the protection of data privacy for each participant. In particular, the participants from different trust domains may not want to disclose their private data sets, which may contain privacy or proprietary information, to anybody else. In applications such as healthcare, disclosure of sensitive data, for example, protected health information (PHI) [2], is not only a privacy issue but of legal concerns according to the privacy rules such as Health Insurance Probability and Accountability Act (HIPAA) [1]. To embrace the Internet-wide collaborative learning, it is imperative to provide a solution that allows the participants, who lack mutual trust, to conduct neural network learning jointly without disclosing their respective private data sets. Preferably, the solution shall be efficient and scalable enough to support an arbitrary number of participants, each possessing arbitrarily partitioned data sets.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Existing System:

To improve the accuracy of learning result, in practice multiple parties may collaborate through conducting joint Back-Propagation neural network learning on the union of their respective data sets. During this process no party wants to disclose her/ his private data to others. Existing schemes supporting this kind of collaborative learning are either limited in the way of data partition or just consider two parties. There lacks a solution that allows two or more parties, each with an arbitrarily partitioned data set, to collaboratively conduct the learning.

Proposed System:

In our proposed scheme, each party encrypts his/her private data locally and uploads the cipher texts into the cloud. The cloud then executes most of the operations pertaining to the learning algorithms over cipher texts without knowing the original private data. By securely offloading the expensive operations to the cloud, we keep the computation and communication costs on each party minimal and independent to the number of participants. To support flexible operations over cipher texts, we adopt and tailor the BGN "doubly homomorphic" encryption algorithm for the multiparty setting. Numerical analysis and experiments on commodity cloud show that our scheme is secure, efficient, and accurate.

RELATED WORK:
Several privacy preserving BPN network:

learning schemes have been proposed recently. Schlitter [19] introduces a privacy preserving BPN network learning scheme that enables two or more parties to jointly perform BPN network learning without disclosing their respective private data sets. But the solution is proposed only for horizontal partitioned data. Moreover, this scheme cannot protect the intermediate results, which may also contain sensitive data, during the learning process. Chen and Zhong [6] propose a privacy preserving BPN network learning algorithm for two-party scenarios. This scheme provides strong protection for data sets including intermediate results. However, it just supports vertically partitioned data. To overcome this limitation, Bansal et. al. [4] enhanced this scheme and proposed a solution for arbitrarily partitioned data. Nevertheless, this enhanced scheme, just like [6], was proposed for the two-party scenario.

Directly extending them to the multiparty setting will introduce a computation/communication complexity quadratic in the number of participants. In practical implementation, such a complexity represents a tremendous cost on each party considering the already expensive operations on the underlying groups such as elliptic curves. However, [4] just considers the two-party scenario though it supports arbitrarily partitioned data set. To our best knowledge, none of existing schemes have solved all these challenges at the same time. There still lacks an efficient and scalable solution that supports collaborative BPN network learning with privacy preservation in the multiparty setting and allows arbitrarily partitioned data sets. Our Contribution. In this work, we address this open problem by incorporating the computing power of the cloud. The main idea of our scheme can be summarized as follows: each participant first encrypts her/his private data with the system public key and then uploads the ciphertexts to the cloud; cloud servers then execute mostof the operations pertaining to the learning process over the ciphertexts and return the encrypted results to the participants; the participants jointly decrypt the results with which they update their respective weights for the BPN network.

During this process, cloud servers learn no privacy data of a participant even if they collude with all the rest participants. Through offloading the computation tasks to the resource-abundant cloud, our scheme makes the computation and communication complexity on each participant independent to the number of participants and is, thus, highly scalable. For privacy preservation, we decompose most of the subalgorithms of BPN network into simple operations such as addition, multiplication, and scalar product. To support these operations over ciphertexts, we adopt the Boneh, Goh, and Nissim (BGN) [5] "doubly homomorphic" encryption algorithm and tailor it to split the decryption capability among multiple participants for collusion-resistance decryption. As decryption of [5] is limited to small numbers, we introduce a novel design in our scheme such that arbitrarily large numbers can be efficiently decrypted. To protect the intermediate data during the learning process, we introduce a novel random sharing algorithm to randomly split the data without decrypting the actual value. Thorough security analysis shows that our proposed scheme is secure. Experiments conducted on Amazon Elastic Compute Cloud (Amazon EC2) [15], over real data sets from UCI machine learning repository [12], show that our scheme significantly outperforms existing ones in computation/communication cost and accuracy loss.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Our contribution can be summarized as follows:
To our best knowledge, this paper is the first that provides privacy preservation for multiparty (morethan two parties) collaborative BPN network learning over arbitrarily partitioned data.. Thorough analysis investigating privacy and efficiency guarantees of the proposed scheme is presented; real experiments on Amazon cloud further show our scheme's several magnitudeslower computation/communicational costs than the existing ones. We tailor [5] to support multiparty secure scalar product and introduce designs that allow decryption of arbitrary large messages. These improvements can be used as independent general solutions for other related applications.

MODELS AND ASSUMPTIONS:

System Model:

We consider a system composed of three major parties: a trusted authority (TA), the participating parties (data owner), and the cloud servers (or cloud). TA is the party only responsible for generating and issuing encryption/decryption keys for all the other parties. It will not participate in any computation other than key generation and issuing. Each participating party s, denoted as Ps, owns a private data set and wants to perform collaborative BPN network learning with all other participating parties. That is, they will collaboratively conduct learning over the arbitrarily partitioned data set, which is private and cannot be disclosed during the whole learning process. We assume that each participating party stays online with broadband access to the cloud and is equipped with one or several contemporary computers, which can work in parallel if there are more than one.

Security Model:

Our scheme assumes the existence of a trusted authority, who is trusted by all the parties, TA has the knowledge of system secret key and will not participate in any computation besides the key generation and issuing. TA is allowed to learn about each party's private data whenever necessary.

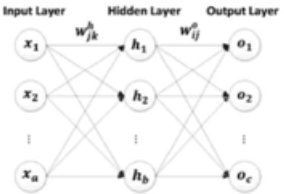

Fig. 1. Configuration of BP network.

We claim that the existence of TA is useful when investigation is needed in case some malicious party intentionally interrupts the system, say using bogus data sets. In real life, parties such as the government agents or organization alliances can be the TA. Although the existence of TA is helpful, we leave the completely distributed solution as a future work. The participating parties do not fully trust each other. Therefore, they do not want to disclose their respective private data (except for the final weights learned by the network) to any other parties than TA. The cloud is not fully trusted by the participating parties either, i.e., the cloud is not allowed to learn about the sensitive information, such as original data sets and intermediate data. In this paper, we follow the curious-but-honest model [8]. That is, all the parties (i.e., all the participating parties and the cloud) will honestly follow our protocol but try to discover others' private data as much as possible. A number malicious of participating parties may collude among themselves and/ or with the cloud.

OUR PROPOSED SCHEME:

Problem Statement and Scheme Overview Problem Statement.In this paper, we aim at enabling multiple parties to jointly conduct BPN network learning without revealing their private data. The input data sets owned by the parties can be arbitrarily partitioned. The computational and communicational costs on each party shall be practically efficient and the system shall be scalable.

Scheme Overview.

To achieve the above goals, the main idea of our proposed scheme is to implement a privacy preserving equivalence for each step of the original BPN network learning algorithm described in Algorithm 1. Different from the original BPN network learning algorithm, our proposed scheme lets each party encrypt her/his input data set and upload the encrypted data to the cloud, allowing the cloud servers to perform most of the operations, i.e., additions and scalar products. To support these operations over ciphertexts, we adopt and tailor the BGN "doubly homomorphic" encryption [5] for data encryption. Nevertheless, as the BGN algorithm just supports one step multiplication over ciphertext, the intermediate results, for example, the intermediate products or scalar products, shall be first securely decrypted and then encrypted to support consecutive multiplication operations as described in Algorithm 1.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

For privacy preservation, however, the decrypted results known to each party cannot be the actual intermediate values, for example, the values of the hidden layer For this purpose, we design a secret sharing algorithm that allows the parties to decrypt only the random shares of the intermediate values. The random shares allow the parties to collaboratively execute the following steps without knowing the actual intermediate values. Data privacy is thus well protected. The overall algorithm is described in Algorithm 2, which is the privacy preserving equivalence of Algorithm 1. To support the operations defined in Algorithm 2, we propose three other cloud-based algorithms: Algorithm 3 for secure scalar product and addition, Algorithm 4 for secure random sharing, and Algorithm 5 for sigmoid function approximation. After the entire process of the privacy preserving learning, all the parties jointly establish a neural network representing the whole data set without disclosing any private data to each other.

CONCLUSIONS:

In this work, we proposed the first secure and practical multiparty BPN network learning scheme over arbitrarily partitioned data. In our proposed approach, the parties encrypt their arbitrarily partitioned data and upload the ciphertexts to the cloud. The cloud can execute most operations pertaining to the BPN network learning algorithm without knowing any private information. The cost of each party in our scheme is independent to the number of parties. This work tailors the BGN homomorphic encryption algorithm to support the multiparty scenario, which can be used as an independent solution for other related applications. Complexity and security analysis shows our proposed scheme is scalable, efficient, and secure. One interesting future work is to enable multiparty collaborative learning without the help of TA.

REFERENCES:

[1] "The Health Insurance Portability and Accountability Act of Privacy and Security Rules," http://www.hhs.gov/ocr/privacy , 2013.
[2] "National Standards to Protect the Privacy of Personal HealthInformation," http://www.hhs.gov/ocr/hipaa/finalreg.html, 2013.
[3] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematics Tables, Dover- Books on Mathematics. Dover, 1964.
[4] A. Bansal, T. Chen, and S. Zhong, "Privacy Preserving Back- Propagation Neural Network Learning over Arbitrarily Partitioned Data," Neural Computing Applications, vol. 20, no. 1, pp. 143-150, Feb. 2011.
[5] D. Boneh, E.-J. Goh, and K. Nissim, "Evaluating 2-DNF Formulas on Ciphertexts," Proc. Second Int'l Conf. Theory of Cryptography (TCC '05), pp. 325-341, 2005.
[6] T. Chen and S. Zhong, "Privacy-Preserving Backpropagation Neural Network Learning," IEEE Trans. Neural Network, vol. 20, no. 10, pp. 1554-1564, Oct. 2009.
[7] L. Cun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network," Proc. Advances in Neural Information Processing Systems, pp. 396-404, 1990.
[8] S.D.C. di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati, "Over-Encryption: Management of Access Control Evolution on Outsourced Data," Proc. 33rd Int'l Conf. Very LargeData Bases (VLDB '07), pp. 123-134, 2007.
[9] T. El Gamal, "A Public Key Cryptosystem and a Signature SchemeBased on Discrete Logarithms," Proc. Advances in Cryptology (CRYPTO '85), pp. 10-18, 1985.
[10] S.E. Fahlman, Faster-Learning Variations on Back-Propagation: AnEmpirical Study, pp. 38-51. Morgan Kaufmann, 1988.
[11] K. Flouri, B. Beferull-lozano, and P. Tsakalides, "Training a SVMBased Classifier in Distributed Sensor Networks," Proc. 14th European Signal Processing Conf., pp. 1-5, 2006.
[12] A. Frank and A. Asuncion, UCI Machine Learning Repository, 2010.
[13] R. Grossman and Y. Gu, "Data Mining Using High Performance Data Clouds: Experimental Studies Using Sector and Sphere," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD '08), pp. 920-927, 2008.
[14] R.L. Grossman, "The Case for Cloud Computing," IT Professional, vol. 11, no. 2, pp. 23-27, Mar. 2009.
[15] A. Inc., Amazon Elastic Compute Cloud (Amazon EC2), Amazon Inc.,http://aws.amazon.com/ec2/#pricing, 2008.
[16] R. Law, "Back-Propagation Learning in Improving the Accuracy of Neural Network-Based Tourism Demand Forecasting," Tourism Management, vol. 21, no. 4, pp. 331-340, 2000.
[17] A.J. Menezes, P.C.V. Oorschot, S.A. Vanstone, and R.L. Rivest, Handbook of Applied Cryptography. CRC Press, 1997.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[18] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representations by Error Propagation," Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, pp. 318-362, MIT Press, 1986.
[19] N. Schlitter, "A Protocol for Privacy Preserving Neural Network Learning on Horizontal Partitioned Data," Proc. Privacy Statistics in Databases (PSD '08), Sept. 2008.
[20] S. Stolfo, A.L.P.S. Tselepis, A.L. Prodromidis, S. Tselepis, W. Lee, D.W. Fan, and P.K. Chan, "JAM: Java Agents for Meta-Learning over Distributed Databases," Proc. Third Int'l Conf. Knowledge Discovery and Data Mining, pp. 74-81, 1997.
[21] B. Yang, Y.-d. Wang, and X.-h. Su, "Research and Design of Distributed Neural Networks with Chip Training Algorithm," Proc. First Int'l Conf. Advances in Natural Computation (ICNC '05)- Vol. Part I, pp. 213-216, 2005.
[22] A.C. Yao, "Protocols for Secure Computations," Proc. 23rd Ann. Symp. Foundations of Computer Science (SFCS '82), pp. 160-164, 1982.
[23] J. Yuan and S. Yu, "Privacy Preserving Back-Propagation LearningMade Practical with Cloud Computing," Proc. Eighth Int'l ICST Conf. Security and Privacy in Comm. Networks (SecureComm '12), Sept. 2012.
[24] S. Zang and S. Zhong, "A Privacy-Preserving Algorithm for Distributed Training of Neural Network Ensembles," Neural Computing and Applications, vol. 22, no.1, pp. 269-282, 2013.

Author's:

C.Sampath Kumar is a student of Sree Rama Institute of Technology & Science, Kuppenakuntla,Penuballi, Khammam, TS,India.Presently he is Pursuing his M.Tech (CSE) from this collegeHis area of interests includes Information Security, Cloud Computing, Data Communication & Networks.

Mr. P.Nageswa Rao is an efficient teacher, received M.Tech from JNTU Hyderabad is working as an Assistant Professor in Department of C.S.E, Sree Rama Institute of Technology & Science, Kuppenakuntla, Penuballi, Khammam, AP,India. He has published many papers in both National & International Journals. His area of Interest includes Data Communications & Networks, Database Management Systems, Computer Organization, C Programming and other advances in Computer Applications.