A Cloud Security Model Based On Machine Learning and Neuron Network

Masroor Khan
masroor.khaaan@gmail.com
University of Bedfordshire, Luton

DOI: 10.29322/IJSRP.9.02.2019.p8659
http://dx.doi.org/10.29322/IJSRP.9.02.2019.p8659

Abstract:
The aim of the paper is to enhance the cloud security and improve the data efficiency in cloud outsourcing. As the security of the cloud data is an important issue in cloud outsourcing, proper algorithm is necessary. Here in this paper a mechanism for the cloud security has been proposed. The mechanism propose two algorithm, one is based on machine learning and another is based on neural network. The machine learning base algorithm is based on the KNN algorithm and the neural network technique is based on data fragmentation and hashing technology. Both of these algorithms are efficient to enhance the cloud security through encryption for cloud data in the cloud server.

I. Introduction:

The advancement of technology has taken the computing to a whole new level and one of the latest development in this context is the introduction of cloud computing. It has revolutionized the concept of distributed computing and thanks to this advanced technology, high performance computing has become affordable and accessible to many (Zhao, Li and Liu 2014). Cloud computing refers to distribution of computing including hardware, software to the consumer through internet. It is distributed in three main category such as platform as a service or PaaS, software as a service or SaaS, infrastructure as a service or IaaS. Although depending on the model chosen it is decided that what is delivered as a service, the core concept remains the same which is distribution of computing source through outsourcing.

Although cloud computing has already influenced so many industries including technology, business, management, logistics and numerous other industry as well (Almorsy, Grundy and Müller 2016). However the latest trend is the machine learning where cloud computing has huge potential and integration of cloud computing for machine learning is an interesting area of research as the mainstream application of machine learning is yet to be commercialized at full extent. Hence research around this topic is not only an interesting but important research as well.

Another important invention that has attracted attention of developer and investor in the field of machine learning is the concept of neural network (Xie *et al.* 2014). It refers to an intelligent computing system where the nodes areinterconnected in the similar fashion which is seen in the neuron system. Not only that, the way information is processed resembles the way human brain process information. It is one of the significant inventions in recent era in the field of machine learning and advanced computing

II. Related work:

Almorsy, Grundy and Müller (2016) have discussed how machine learning actually works. In machine learning the computers are trained with dataset so that the machines become more intelligent, intuitive and powerful. The motive behind machine learning is to help the machines take decision on behalf of human based on the input data. This where computing system in machine learning differs than the traditional computing system. Here the purpose of the computers are not only to process the data that is fed to the system but to decide what actions to be taken based on the nature of the data sets. In order to achieve such complex and advanced skills for the computers, rigorous training of the computers with huge data set is necessary. In fact the machines are trained with millions of data set so that it becomes easier for the machines to process various data and take decision with reference to the data set with increased accuracy and efficiency. Here the aim is to fully automate the process without intervention of human intelligence.

The author has suggested for K-NN (K Nearest Neighbour) technique for cloud security. This method is applied for classification of data. The classification helps to determine which data to be kept secured, hence private and which data is kept public this kind of classification according to the author plays and vital role in enhancing data security (Chang and Ramachandran 2016). The data which are less sensitive is kept public and data which is private needs to be encrypted with proper encryption. However without proper classification this is not possible and hence the technique according to the author is an important one for cloud data security. However the paper has only considered RSA algorithm for the encryption of the private data and this might be drawback when the data need advanced level of encryption for the data security in the cloud server.

The author has discussed the Impact of neural network for machine learning. According to although Machine learning is an excellent option for securing the cloud computing as it provides better security but it is more powerful when it is integrated with neural network. The reason cited by the author here is that although machine learning algorithms are capable of handling a huge amount of data the performance of the algorithms so not improve over time and the amount of data feed (Samarati *et al.* 2016). However, in the neural network the algorithms are designed such that the performance of the algorithm not only improves over time, it becomes more efficient with increase in the amount of data it deals with. The more data the algorithms process, it becomes more efficient which is not the case for most of the traditional machine learning algorithm.

The authors have proposed for an approach that combines data encryption with data fragmentation. The idea for the integration as specified by the author is to facilitate the data distribution (Bhangotra and Puri 2015). Data distribution is important to properly manage the cloud data when more than one service provider is involved for the cloud outsourcing. Fragmentation is necessary to efficiently distribute data among different service providers and encryption provides the necessary support for the data security.

A multi key encryption has been proposed to preserve the privacy of the sensitive cloud data. This method search for data in the cloud server and then the search result is ranked according to the nature of the data and level of encryption attached (Zibouh, Dalli and Drissi 2016). One benefit of the method is that it secures the information without outflow of the privacy information of the user data.

Li *et al.* (2017) discusses the reason why outsourcing of cloud is a good option for machine learning. According to author as machine learning involves processing of large amount of data source, it requires complex and powerful computer system setup. In order to design such advanced and complex computing system powerful computing resource with large storage is needed. Hence outsourcing is an excellent option for machine learning. Cloud computing provides all the necessary resource such as powerful computing architecture needed for large data processing and it also offers support for large storage. Additionally the cloud platform is scalable which makes it excellent choice for machine learning because the platform cam be extended as per the requirement and data processing speed. Although it is possible to develop the platform in-house, it is not an effective choice as it is costly, requires more human resource for development and maintenance. However with outsourcing these issues are resolved with enhanced productivity and cost effectiveness.

The author describes Challenges of outsourcing for cloud computing. According to author, outsourcing does provide benefits especially in terms of cost, it has certain drawbacks as well. One of the primary concern for outsourcing is to compromise with the security (Bost *et al.* 2015). As computer servers and networks stores data exclusive to the organization, it is sensitive in nature. Hence the security of the data is a major concern for outsourcing cloud service and deploy the machine learning architecture in the cloud server. If the data is hacked or somehow extracted without illegal means, then the company might reveal a lot of sensitive and organization exclusive data to the hackers without the intention to do so. It only means that the company will be at a position where they have to address issues like financial crisis, reduced organisation value and less customer trust. Hence the according to the author, data security remains the most important issues in outsourcing data in cloud computing

While discussing about cloud security and how data integrity should be addressed in cloud outsourcing, the authors have suggested that the primary strategy is to ensure that the vendor is proper and is a trusted one as majority of the core network is maintained by the vendor and vendor plays an important role in this context in securing the data outsourced by the company or the owner (Bost *et al.* 2015). If the vendor is not proper and reputed one, chances are that the platform, software or the infrastructure provided as service is not secured with latest hardware and software, thus compromising with the security. Hence choice of vendor is very important for data security, especially when outsourcing important and valuable data exclusive to the organizations.

As the data security is a major drawback and it need careful strategy to obtain the benefits data outsourcing provides in cloud computing context such as less expensive computing platform and better productivity (Gilad-Bachrach *et al.* 2016). However the author thinks that only trusted vendor is not sufficient for securing the data outsourced by the data owner. Advance technology is need in this context. The author suggest machine learning for enhancing the security of cloud computing. According to the author machine learning is an excellent choice for securing the data in the cloud computing when the server is deployed and managed by the vendor.

(Xia *et al.* 2016) have suggested a data fragmentation for securing data in the cloud sever. According to the author it has several benefits such as manipulation in cloud data, optimized cloud storage, data distribution and most importantly enhanced data security. In this approach all the fragmented data is considered and the aim is to secure all the fragmented data in the cloud server.

Yuan and Yu (2014) have suggested a flexible and scalable distribution protocol for verification of cloud security and the cloud data. Here special focus has been provided for ensuring cloud storage security as according to the author this is an effective technique for ensuring quality of service or QOS which is an essential feature for any cloud security model. The protocol that has developed and analysed by the authors in the research paper depends on the erasure code which is responsible for data access and data consistency.

Jiang *et al.* (2018) specify that one of the effective and efficient techniques of cloud data security is the hashing algorithm which provides better data security and data authentication. The author also specifies that hashing is effective in identifying the exact location of data do that it is easier to track where the data is data is stored on a storage among the fragmented data that is stored on the device. Here the author has suggested for dynamic hashing where data is possible to store and erase as per the demand and the requirement which according to the author is an excellent feature to ensure data security, especially when the data is stored in the cloud server.

III. Proposed work

The previous proposed solutions in machine learning algorithm for enhancing cloud security has their own benefits and drawbacks as well. Hence in this paper two mechanism has been proposed. One based on the machine learning algorithm and another based on neural network model.

IV. Algorithm for machine learning

In this research paper the machine learning algorithm that has been proposed is known as modified ensemble learning technique. The algorithm has been proposed for enhancing the existing KNN technique as it has some limitations in terms of execution and security features.

In this method the evaluation of the classifier is done through cross validation with the K-fold technique. In the K-fold technique each layers need to be trained in the following method:

- First the data set is classified in two different category: one set is classified as training set and another set is classified as testing set
- Then the classifier as each base level layer is produced with the training set defined in the previous step
- The base level generates predictions and those predictions are then assembled. The assembled predictions are then transferred to the Meta level. The prediction are then validated at the Meta level. After the verifications and validations are completed the final result of the predictions is then provided in details
- Then the dataset associated with complete layer which is different than the simple K-fold dataset is considered to train each different layer associated with the training set
- Now to provide enhance security for data that is confidential and needs strong authentication, the traditional RSA cryptography algorithm has been combined with HMAC. HMAC provides enhanced authentication when combined with RSA encryption. HMAC refers to hashed message authentication code which is a cryptographic checksum. This is not only stored in the local machine but transferred to the cloud as well with chipper text. In order to access the data stored in the cloud, hackers need to match the MAC+ cipher text stored in the cloud with the cryptographic checksum that is stored in the machine at the local server. Hence it becomes difficult for the hackers to hack the server and access the data
- HMAC is obtained with the following formula:
- HMAC (Key, msg) =HMAC (KeyXORout) ||H ((Key XORin) ||msg))
  Here Key refers to the original key, m referees to the message which needs to be secured through encryption.

V. Algorithm for neural network

In this paper a cloud security model has been proposed for improving the security of the cloud computing and cloud server. The model is based on neural cloud computing known as Neural Data Security Model or NDSM. This model is highly efficient for the cloud security as it enhances authentication of the cloud server with advanced encryption. The encryption model that has been considered for this model is the Dynamic Hashing Fragmentation. The concept of Dynamic Hashing Fragmentation is an important concept for fragmentation of sensitive data so that the data more secured in the cloud server. This cryptographic model based on the neural network not only encrypt the data, but also increase the level of confidentiality based on the data type. Data that needs to be treated with increased confidentiality should be encrypted with this neural based security model. This model or neural security algorithm is effective for increasing security as well as security for sensitive data.

Two algorithms are associated with this model and each of this two algorithm has to be executed properly in order to implement the model for the enhancement of cloud data security.

In order to perform the data fragmentation, data has to be first normalized. Out of the two algorithms associated with the model, each of the algorithms is designed for the data fragmentation, in a different way though.

The first algorithm for the data fragmentation works in the following way:

a) Algorithm 1

The procedure that needs to be followed for the data Fragmentation has been discussed in details in the following section:

Input – set of Cloud data obtained from the cloud database
Output – Encrypted and sensitive data obtained in the fragmented from
Step 1: Start.
Step 2: Access data from the cloud database where sensitive data is stored and read it after the data is obtained from the server.
Step 3: Sensitive data that is obtained from the cloud database are then fragmented. The fragmentation styles include vertical, horizontal as well as hybrid
Step 4: After the data is fragmented with the appropriate fragmentation style, it is then stored through dynamic hashing. After the data is fragmented with dynamic hashing, it is then read through hash function.
Step 5: Then dynamic hashing is applied for the encryption of sensitive data. It includes two steps:
5.1 The with the help of the cryptography algorithm, information is secured through encryption mechanism that is designed for the neural network security model
5.2 When request for data access is received from the users, then the data is encrypted with public key. However when the data is decrypted private key has to be used for the decryption. This is done to ensure that the sensitive data is secured properly for increased data security.
Step 6: The data which is encrypted with the neural model encryption mechanism, it is then stored with the help of hashing structure.
Step 7: Stop.

b) Algorithm 2:

The mechanism for executing the dynamic hashing fragment model
Input – indexing value is set for the Cloud data
Output –Sensitive data is obtained in the encrypted form
Step 1.

Various data centres in the cloud server are analysed for the access to the cloud database so that sensitive data from the cloud database is obtained.

Step 2.

The key and non-key attributes are fragmented by the horizontal and vertical fragmented mechanism. Attributes are identified and classified into two categories including key attributes and key attributes. Based on the classification, different types of fragmented mechanism is applied. In order to fragment the key and non-key attributes, horizontal and vertical fragmented mechanism is applied

Step 3.

After the data sets are properly fragmented, data sets are then arranged with dynamic hashing.

Step 4.

After the data sets are fragmented with hashing, those hashing functions are then considered to retrieve the data that are sensitive in nature

Step 5.

The fragmented data which is obtained as output is then provided to the cloud data security model.

Step 6:

Stop.

VI. Conclusion

The paper concludes that although cloud computing is revolutionizing the file of information and technology with affordable advance computing and outsourcing, it still has some security issues that needs to be addressed to improve the adoption of cloud computing in enterprise and business application where security of the data plays an important role. Although traditional encryption algorithm helps to provide security to cloud data, it still needs improvement. Hence machine learning and neural network algorithm is considered in this context. Machine learning algorithm helps computers system to recognize data based on data priority and data sensitivity. The ability of machine learning algorithm to recognize data as per the priority and sensitivity makes machine learning based algorithm more secured and efficient in securing cloud based data. Hence machine learning based cryptographic algorithm is more secured than the traditional cryptographic algorithms.

The modified ensemble learning technique which has been discussed in this context is based on machine learning technique and it is more efficient than the traditional KNN technique and improves various limitations of the technique.

Another important technique in cloud security is neural network algorithm. Two neural network algorithms that have been analysed here applies data fragmentation and hashing technique which increases the security of the data in the cloud server to a great extent.

VII. Further research:

Cloud computing and cloud security is an emerging topic in the field of computing and information technology. Hence as newer methods are being developed to enhance the security of the cloud data, the technique to violate the security mechanism is also evolving. Hence a mechanism which work perfectly in data security might not be that much effective provided improved and advance security threat is emerged in the field of cloud security. Hence a there is always an opportunity and scope for further research to enhance the cloud security. Hence further research work should be conducted in this context to bring improvement in the existing algorithm or bring something completely new that is more efficient, more powerful than the previous algorithms.

References:

Almorsy, M., Grundy, J. and Müller, I., 2016. An analysis of the cloud computing security problem. *arXiv preprint arXiv:1609.01107*.

Aslett, L.J., Esperança, P.M. and Holmes, C.C., 2015. A review of homomorphic encryption and software tools for encrypted statistical machine learning. *arXiv preprint arXiv:1508.06574*.

Bhangotra, V. and Puri, A., 2015. Enhancing cloud security by using hybrid encryption scheme. *Int J Adv Engg Tech/Vol. VI/Issue IV/Oct.-Dec*, *34*, p.40.

Bost, R., Popa, R.A., Tu, S. and Goldwasser, S., 2015, February. Machine learning classification over encrypted data. In *NDSS*.

Chang, V. and Ramachandran, M., 2016. Towards achieving data security with the cloud computing adoption framework. *IEEE Trans. Services Computing*, *9*(1), pp.138-151.

Gilad-Bachrach, R., Dowlin, N., Laine, K., Lauter, K., Naehrig, M. and Wernsing, J., 2016, June. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In International Conference on Machine Learning (pp. 201-210).

Gilad-Bachrach, R., Dowlin, N., Laine, K., Lauter, K., Naehrig, M. and Wernsing, J., 2016, June. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In *International Conference on Machine Learning*(pp. 201-210).

Jiang, X., Kim, M., Lauter, K. and Song, Y., 2018, October. Secure Outsourced Matrix Computation and Application to Neural Networks. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1209-1222). ACM.

Li, J., Li, J., Chen, X., Jia, C. and Lou, W., 2015. Identity-based encryption with outsourced revocation in cloud computing. *Ieee Transactions on computers*, *64*(2), pp.425-437.

Li, P., Li, J., Huang, Z., Gao, C.Z., Chen, W.B. and Chen, K., 2017. Privacy-preserving outsourced classification in cloud computing. *Cluster Computing*, pp.1-10.

Li, P., Li, J., Huang, Z., Li, T., Gao, C.Z., Yiu, S.M. and Chen, K., 2017. Multi-key privacy-preserving deep learning in cloud computing. *Future Generation Computer Systems*, *74*, pp.76-85.

Samarati, P., di Vimercati, S.D.C., Murugesan, S. and Bojanova, I., 2016. Cloud security: Issues and concerns. *Encyclopedia on cloud computing*, pp.207-219.

Xia, Z., Wang, X., Sun, X. and Wang, Q., 2016. A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data. *IEEE Trans. Parallel Distrib. Syst.*, *27*(2), pp.340-352.

Xie, P., Bilenko, M., Finley, T., Gilad-Bachrach, R., Lauter, K. and Naehrig, M., 2014. Crypto-nets: Neural networks over encrypted data. *arXiv preprint arXiv:1412.6181*.

Yuan, J. and Yu, S., 2014. Privacy preserving back-propagation neural network learning made practical with cloud computing. *IEEE Transactions on Parallel and Distributed Systems*, *25*(1), pp.212-221.

Zhao, F., Li, C. and Liu, C.F., 2014, February. A cloud computing security solution based on fully homomorphic encryption. In *Advanced Communication Technology (ICACT), 2014 16th International Conference on* (pp. 485-488). IEEE.

Zibouh, O., Dalli, A. and Drissi, H., 2016. CLOUD COMPUTING SECURITY THROUGH PARALLELIZING FULLY HOMOMORPHIC ENCRYPTION APPLIED TO MULTI-CLOUD APPROACH. *Journal of Theoretical & Applied Information Technology*, *87*(2).