Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 9, August 2021: 8019-8026

Adopting Edge Computing to Reduce Transmission Latency and Enhance Security in Medical Diagnosis Service

G Vamsi Krishna #1, M Vijayalakshmi #2, Nidamanuri Srinu #3
#1,#3 Associate. Professor, #2 M.Tech., Scholar
Dept of Computer Science & Engineering, Qis College of Engineering and Technology

Abstract:
Any user may submit the symptoms at any time, anywhere for medical diagnosis as machine learning becomes more prevalent. One reason to use edge computing is to decrease transmission time and latency while providing real-time diagnostic service. However, medical data is exposed when using data-driven machine learning, which must be built over a massive quantity of medical data. Privacy has to be preserved. We propose a lightweight privacy-preserving medical diagnostic method termed LPME, which may help to address the problems described above. The LPME redesigns the Extreme Gradient Building (XGBoost) model based on the edge-cloud model which incorporates encrypted model parameters to reduce quantities of ciphertext calculations to plaintext calculation. Additionally, LPME may offer discreet and fast diagnosis with edge security and privacy protection. Our research shows that LPME's security, efficiency, and efficacy are maximized.

1. Introduction:

It is now widely used in medical diagnosis and in mobile diagnostic allowing users to provide symptoms at any time, with diagnosis findings sent to them instantaneously. A machine learning-based diagnosis offers significant benefits, including increasing the quality of healthcare service and reducing the cost of costly diagnostic services. Machine learning-based medical diagnostics has therefore drawn significant interest from researchers and businesses alike. Telemedicine applications have been more popular and have led to many new healthcare, clinical, and mobile telemedicine needs. On the other hand, the bloom is also fraught with difficulties, for example, limited training data, security issues, and privacy concerns. In the medical field, there is a significant problem with the cost and duration of collecting medical data. Medical data, which is often stored in a single medical source, is difficult to use in machine learning when it's incomplete. It is essential to exchange training data with many medical institutions in order to properly develop an accurate diagnostic model. The capabilities of cloud computing have allowed for an expansive amount of information to be stored, while simultaneously providing an almost limitless ability to process that data. Yet, in the ever-increasing connections between mobile users and the cloud, there is unwanted Delayed reaction to diagnosis is directly connected to the lives and health of patients, as well as medical safety, particularly for those with acute illness (e.g., pneumonia, heart disease). To solve this problem, a new computing paradigm called edge computing has been suggested to lower latency and improve computational efficiency by utilising edge nodes, which are located near mobile users. Machine learning methods that use edge computing (see [14], [15], [16]) have received considerable attention in recent years and are important in improving diagnostic performance using edge computing. The edge network in Fig. 1, as opposed to the one in Fig. 2, has a small number of edge nodes (i.e., medical organisations) with constrained storage capability and processing capacity. To better understand medical diagnosis's weaknesses, the focus should be placed on a robust edge model, which is fast and trustworthy. The most up-to-date machine learning model, XGBoost, which offers great prediction performance in the distributed environment, is also proving its prowess in Kaggle contests. Plus, the tree-based structure offers the advantages of easier understandability and explainability. A lot of projects have been conducted using the XGBoost model for medical diagnostics [17], [18], [19], however they neglect the critical privacy problem during training. In fact, people with private illnesses (e.g., HIV, Hepatitis B virus) The condition is worsened by its presence. In order to ensure privacy, these individuals must be shielded. In addition, the medical data include a significant quantity of sensitive information, which has recently been constrained by privacy rules (e.g., GDPR [20] and HIPPA [21]), leading to the following: a greater restriction on data release, especially when in unencrypted form. In the edge computing context, the need to preserve the privacy of medical diagnostics necessitates.

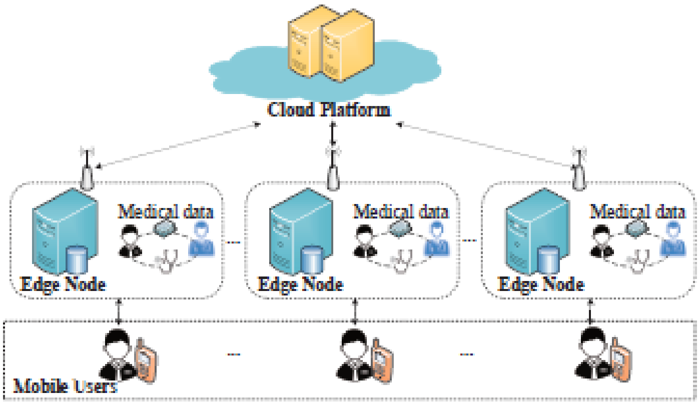

Fig. 1: The framework of edge computing for medical diagnosis.

A potential answer to worries about privacy leaks is Homomorphic Encryption (HE), which allows for the encryption of data while preserving data confidentiality. Privacy-preserving machine learning techniques based on cloud-based frameworks have their reach extended to edge computing through the addition of a single-cloud model or dual-cloud mode [2]. Even while the single cloud model [8] has better security than the dual cloud model, the former still has a greater risk of exposing secret keys since they are kept on the single cloud. Once a cloud is infiltrated, confidential information becomes public. However, the belief that two semihonest Additionally, machine learning's training phase entails safe computing over encrypted data. Given the rise of outsourced encrypted data, computationally expensive algorithms are needed, particularly for the resource-stressed edge nodes, which is the first significant challenge. The need of taking lightweight into account with privacy-preserving machine learning in edge computing cannot be ignored. To overcome the aforementioned difficulties, we propose a lightweight privacy-preserving XGBoost over encrypted model parameters We outline the use of edge computing to safely carry out medical diagnosis in this article, which we call lightweight privacy-preserving medical diagnosis, or LPME.

2. Related Work:

Earlier work on privacy-preserving machine learning [14], [15] suggested privacy preservation

After then, several proposals were suggested to ensure privacy protection. Fu et al. [16] described a privacy-preserving non-negative matrix factorization technique that relies on addition HE, but since the parameters of the matrix factorization are acquired by another party during the computing process, this may lead to a possible privacy leak.

Ma et al. [19] presented a random tree architecture for the protection of privacy using Paillier cryptosystem that performed precise, safe training on encrypted data. Wang et al. [32] designed a privacy-preserving collaborative neural network that use encryption to build a model that keeps private information confidential. Mohassel et al. [13] created a privacy-preserving neural network training method for faster learning. These methods [19], [20], and [23] all use the HE-based privacy preserving technique, which is viable for machine learning.

In order to address the problem previously mentioned, a model sharing-based privacy-preserving machine learning framework was developed. This framework sends encrypted model parameters to a separate service rather than storing them locally. It can not only ensure machine learning training, but also shift outsourced computation Yu et al. [18] proposed a framework for models created from data owned by multiple owners but which do not reveal local data.

While this system does not utilise encryption, it does make use of random numbers, which are less vulnerable to inference attacks, therefore resulting in less privacy leaks [18]. Cheng et al. [19] then offered a safe XGBoost model with encrypted parameter values. This information, however, can be cracked by someone else. Local data is also at risk due to settings containing sensitive information. In their paper, Li et al. [20] recommended a secure classification service using SVM models that they said could not be used for privacy-preserving model training. Aono et al. [21] designed a privacy-preserving deep learning system that used end-to-end encryption on datasets so that servers could not access participant local data, which significantly decreased execution times.

In the single-cloud scenario, Wang et al. [22] discovered that the previously described methods [20], [21] leak privacy. Because of the training models' anonymity, it is simple for them to be compromised when the cloud is.

To get over the privacy issue in the single-cloud approach, the dual-cloud concept is used. Liu et al. [21], [20] showed that dual-cloud server architecture, which uses secure computing, is both secure and accurate. In addition, the dual-cloud non-collusion model was shown to provide a better degree of security than the single-cloud approach, as shown by Hu et al. [21]. Even one server being breached cannot result in privacy being lost, as long as the other server remains secure.

Liu et al. [21] proposed an edge computing extension of the safe computation paradigm based on a dual-cloud architecture. Transmitting encrypted data between two cloud servers to ensure secure computation, which incurs the communication burden and heavy computational overhead, is unfortunately unavoidable.

It is impracticable for each resource-limited edge node to complete safe computation, which requires five modular exponentiation operations, two modular addition operations, and six modular multiplication operations. Zhang et al. [23] suggested a privacy-preserving feature transform that is lightweight, yet The literature we know of doesn't account for the trade-off between privacy and light weight in edge computing, as in [22]. In addition to the benefits of speed and real-time training, we developed a privacy-preserving machine learning algorithm that offers excellent privacy protections for edge nodes.

3. System Model

Within our system architecture, the four core components that are included are the Key Generation Center (KGC), Cloud Platform (CP), Edge Nodes (ENs), and Management Units (MUs), which are shown in Figure 1. If NEN is part of the system, These organisations communicate via secure channels like SSL and Secure Sockets Layer (TLS). The following is an illustration of the particular role of each organisation: Key in the middle. KGC has absolute faith in the development, administration, and distribution of our system's secret keys, and the secret shares are sent to other businesses for safe computation in the future node on the edge A medical institution (EN) with storage space and computational limits is a healthcare facility with restricted medical data. During the training, an EN is taught to cooperate with other ENs to create a global model that is securely shared once it is encrypted infrastructure CP is completely flexible for storing and processing data. Then it uses the worldwide best model parameters for global model construction, receiving encrypted ENs model parameters first user on a cell phone A MU may send an encrypted request for a diagnostic to a nearby EN, which will return the diagnosis with the results. To protect privacy, the EN and MU share a private diagnostic phase computation.

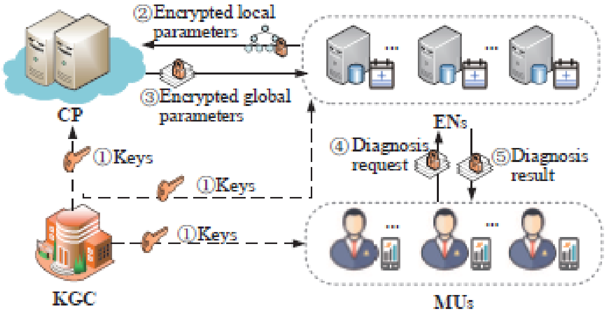

Fig. 2: System model.

Design goals:

Our approach is to develop a machine-learning framework for privacy-reservation with safe training, precise diagnostics and low weight adverse computing. These design goals are displayed Security. Local EN models include private information not accessible for privacy reasons, according to (a). It is impossible to leak parameters and intermediate computation outputs during the building of the global model. All MU requests sent to the ENs and all diagnostic Efficiency. The LPME system must ensure the accuracy of a global model that has been trained for medical diagnosis and must maintain a moderate demand on ENs and MUs. A reliable and accurate diagnostic service must be available for MUs to get correct diagnostic findings.

4. Results

Instead of the typical healthcare cloud computing, we utilised the iFogSim toolkit to mimic the fog network by creating an area where the fog might exist. The iFogSim provides you the simulated setup results. It's easy to see the outcomes of a process if you don't have any technology accessible. The simulator offers you a little bonus on top of your usual shift. We performed a lot of trials for 5 different monitoring devices used in this simulation. The following chart shows the mean delay and network utilisation for the five options. Our fog-based architecture's latency doesn't seem to be much affected by the connected device configurations, according to the table. The network implementation of the fog computing architecture is much less than the fog-only approach. The iFogSim toolbox replicates the five parameters included in the original iFogSim application. There are several types of monitoring devices for each of the five combinations. The four monitoring devices in Config 1, eight in Config 2, 16 in Config 3, 32 in Config 4, and 64 in Config 5 are, accordingly, As a consequence, each configuration generates a different outcome when simulated. The monitoring devices employed in the settings have an average network interval duration of 20,000 bytes and 5 ms, while 1000 million instructions are loaded into CPUs.

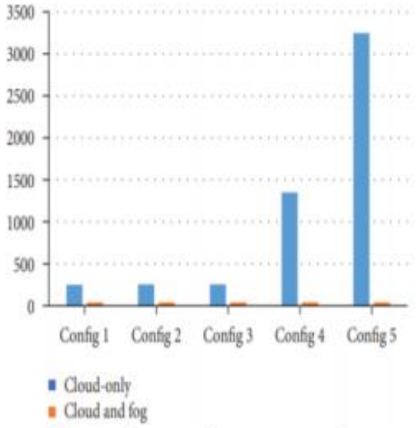

Figure: Average latency comparison

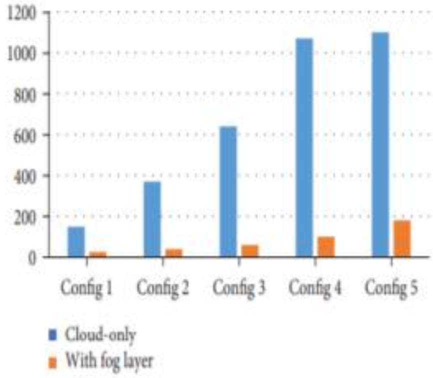

Figure: Network usage comparison

CloudSim is used for the cloud layer with iFogSim in tandem for the cloud plus fog simulation. From the diagrams, we can see that the fog layer complexity does not impact network latency or use. Fog computing is really very beneficial for the The results from the table clearly show that the fog layer The image shows the energy usage of fog computing and cloud use separately. It can be ascertained from the illustration that the majority of the energy used for fog computing happens at the system's edge. While cloud computing does consume a little amount of energy, most of it is spent in data centres or in the cloud the analysis. In addition, we identified a problem with latency. The fog computing network in our health informatics software will transmit data between different layers. Depending on the conditions, the quantity of data and the length of time involved will vary. Thus, the latency is variable. When data being analysed must be This is the time it takes for data from the fog layer to be returned to the IoT sensor; ef is the time it takes to evaluate edge devices and ee is the time spent in the cloud for the assessment. We will use these two equations to evaluate the time it takes to complete each task. Fog computing applications often depend on real-time network processing, and thus this is of great importance. These computations should be real-time and responsive to latency. Calculations are made difficult by a number of methods. By keeping packets in a fog node cache for a while, data from a previous transmission is avoided. Data packets in this category may be recharged with fresh data packets according to a few different renewal mechanisms. It is important to make data packets reach their most efficient number of edge devices in a smart way. Analysis of the Security: A fog layer in the cloud computing architecture may reduce security risk for patient data by decreasing the chance of a data centre failure resulting in data loss. But, at the same time, the data is kept on the cloud. The dangers to patient privacy are heightened with this. By utilising a secret key to encrypt patient data, we are able to guarantee patient privacy in this procedure.

5. Conclusion

The article outlined a lightweight XGBoost privacy architecture that can provide edge nodes with strong confidentiality and edge node privacy, as well as medical diagnostics with up-to-date information. XGBoost may be securely built using the LPME system to offer medical diagnostics quickly and with no risk of privacy breach. Data set tests conducted in the real world have shown that the LPME system is both safe and efficient in edge computing.

REFERENCES

[1] X. Wang, J. Ma, Y. Miao, X. Liu, and R. Yang, “Privacy-preserving diverse keyword search and online pre-diagnosis in cloud computing,” IEEE Transactions on Services Computing, 2019.

[2] Y. Zhang, C. Xu, H. Li, K. Yang, J. Zhou, and X. Lin, “HealthDep: An efficient and secure deduplication scheme for cloud-assisted ehealth systems,” IEEE Trans. Industrial Informatics, vol. 14, no. 9, pp. 4101–4112, 2018.

[3] B. Fu, P. Liu, J. Lin, L. Deng, K. Hu, and H. Zheng, “Predicting invasive disease-free survival for early-stage breast cancer patients using follow-up clinical data,” IEEE Transactions on Biomedical Engineering, 2018.

[4] A. Galletta, L. Carnevale, A. Bramanti, and M. Fazio, “An innovative methodology for big data visualization for telemedicine,” IEEE Transactions on Industrial Informatics, vol. 15, no. 1, pp. 490–497, 2018.

[5] I. Kononenko, “Machine learning for medical diagnosis: history, state of the art and perspective,” Artificial Intelligence in medicine, vol. 23, no. 1, pp. 89–109, 2001.

[6] C. P. Friedman, A. K. Wong, and D. Blumenthal, “Achieving a nationwide learning health system,” Science Translational Medicine, vol. 2, no. 57, pp. 57cm29–57cm29, 2010.

[7] Y. Miao, X. Liu, K.-K. R. Choo, R. H. Deng, H.Wu, and H. Li, “Fair and dynamic data sharing framework in cloud-assisted internet of everything,” IEEE Internet of Things Journal, 2019.

[8] Y. Miao, Q. Tong, K.-K. R. Choo, X. Liu, R. H. Deng, and H. Li, “Secure online/offline data sharing framework for cloud-assisted industrial internet of things,” IEEE Internet of Things Journal, 2019.

[9] Y. Miao, J.Weng, X. Liu, K.-K. R. Choo, Z. Liu, and H. Li, “Enabling verifiable multiple keywords search over encrypted cloud data,” Information Sciences, vol. 465, pp. 21–37, 2018.

[10] T. Ouyang, R. Li, X. Chen, Z. Zhou, and X. Tang, “Adaptive user-managed service placement for mobile edge computing: An online learning approach,” in Proc. IEEE Conference on Computer Communications (INFOCOM’19). IEEE, 2019, pp. 1468–1476.

[11] P. Dai, K. Liu, X. Wu, H. Xing, Z. Yu, and V. C. Lee, “A learning algorithm for real-time service in vehicular networks with mobileedge computing,” in Proc. IEEE International Conference on Communications (ICC’19). IEEE, 2019, pp. 1–6.

[12] F.Wang, C. Zhang, J. Liu, Y. Zhu, H. Pang, L. Sun et al., “Intelligent edge-assisted crowdcast with deep reinforcement learning for personalized qoe,” in Proc. IEEE Conference on Computer Communications (INFOCOM’19). IEEE, 2019, pp. 910–918.

[13] C.-C. Lin, D.-J. Deng, Y.-L. Chih, and H.-T. Chiu, “Smart manufacturing scheduling with edge computing using multi-class deep q network,” IEEE Transactions on Industrial Informatics, 2019.

[14] G. S. Aujla, R. Chaudhary, K. Kaur, S. Garg, N. Kumar, and R. Ranjan, “Safe: Sdn-assisted framework for edge–cloud interplay in secure healthcare ecosystem,” IEEE Transactions on Industrial Informatics, vol. 15, no. 1, pp. 469–480, 2018.

[15] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, “eseiz: An edge-device for accurate seizure detection for smart healthcare,” IEEE Transactions on Consumer Electronics, 2019.

[16] A. M. Rahmani, T. N. Gia, B. Negash, A. Anzanpour, I. Azimi, M. Jiang, and P. Liljeberg, “Exploiting smart e-health gateways at the edge of healthcare internet-of-things: A fog computing approach,” Future Generation Computer Systems, vol. 78, pp. 641– 658, 2018.

[17] A. A. Ogunleye and W. Qing-Guo, “Xgboost model for chronic kidney disease diagnosis,” IEEE Transactions on Computational Biology and Bioinformatics, 2019.

[18] M. Nishio, M. Nishizawa, O. Sugiyama, R. Kojima, M. Yakami, T. Kuroda, and K. Togashi, “Computer-aided diagnosis of lung nodule using gradient tree boosting and bayesian optimization,” PloS one, vol. 13, no. 4, p. e0195875, 2018.

[19] A. R. Rao and D. Clarke, “A fully integrated open-source toolkit for mining healthcare big-data: architecture and applications,” in Proc. IEEE International Conference on Healthcare Informatics (ICHI’16). IEEE, 2016, pp. 255–261.

[20] E. Union, “General data protection regulation,” 2018, https:// gdpr-info.eu/.

[21] U. S. D. of Health and H. Services, “Health insurance portability and accountability act1996,” 1996, https://www.hippa.com/.

[22] D. Liu, Z. Yan,W. Ding, and M. Atiquzzaman, “A survey on secure data analytics in edge computing,” IEEE Internet of Things Journal, 2019.

[23] Y. Miao, X. Liu, R. H. Deng, H.Wu, H. Li, J. Li, and D.Wu, “Hybrid keyword-field search with efficient key management for industrial internet of things,” IEEE Transactions on Industrial Informatics, 2018.

[24] Y. Miao, J. Ma, X. Liu, X. Li, Z. Liu, and H. Li, “Practical attributebased multi-keyword search scheme in mobile crowdsourcing,” IEEE Internet of Things Journal, vol. 5, no. 4, pp. 3008–3018, 2017.

[25] Y. Miao, X. Liu, K.-K. R. Choo, R. H. Deng, J. Li, H. Li, and J. Ma, “Privacy-preserving attribute-based keyword search in shared multi-owner setting,” IEEE Transactions on Dependable and Secure Computing, 2019.

[26] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” in Proc. International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT ’99). Springer, 1999, pp. 223–238.

[27] X. Liu, R. H. Deng, Y. Yang, H. N. Tran, and S. Zhong, “Hybrid privacy-preserving clinical decision support system in fog–cloud computing,” Future Generation Computer Systems, vol. 78, pp. 825– 837, 2018.