Research on Security of Business Scenario in Computing Force Network

Yiran Zhang[1], Sixu Guo[1(✉)], Huizheng Geng[1(✉)], Li Su[1], and Fei Liu[2]

[1] China Mobile Research Institute, Beijing, China
{zhangyiran,guosixu,genghuizheng,suli}@chinamobile.com
[2] China Mobile Tianjin, Tianjin, China
liufei@tj.chinamobile.com

Abstract. With the development of the digital economy era, emerging technologies and applications have brought about the continuous growth of data volume, and all walks of life have put forward more urgent needs for computing power and network. Facing the in creasing demand for computing power and network, computing force network came into being. However, the integration of multi-source computing power and network has broken the traditional security domain, and the computing force network will face more serious security and privacy problems. According to the type of computing power, this paper divides the business scenarios of computing force network into three categories: storage task, computing task and training task, and studies the security work of computing force network respectively.

Keywords: computing force network · data security · privacy protection · computing and network convergence

1 Introduction

Computing force network (CFN, Computing Force Network) is a new information infrastructure with computing as the center and network as the foundation [1–3]. It integrates the Network, cloud, data, intelligence, security, edge, end and chain (ABCDNETS) deeply. The purpose of computing force network is to improve the utilization rate of computing power resources, realize the optimization and efficient use of network and computing power, so that computing power resources can be "one-point access, instant use". The arrangement management system of the computing force network is the core center of the computing force network. By flexibly combining the atomic capabilities of the computing force network, it realizes the unified management, unified arrangement, intelligent scheduling and global optimization of the computing force network resources downward, improves the efficiency of the computing force network, provides the scheduling capability interface of the computing force network upward, and supports the diversified services of the computing force network.

However, the computing force network absorbs a large number of ubiquitous distributed computing power nodes. The operators of the computing force network may not

N. Radojević et al. (Eds.): ICAID 2022, AHIS 7, pp. 1201–1209, 2023.
https://doi.org/10.2991/978-94-6463-010-7_124

have complete control over the nodes and cannot guarantee the security and credibility of the computing nodes themselves. Moreover, there are great differences in the security capabilities provided by computing nodes in the network, which greatly improves the breadth and difficulty of security protection. The security arrangement management system of computing force network is an important part of computing force network arrangement management system, which can configure corresponding security measures according to the security requirements of different computing power tasks of users to ensure the safe operation of services.

According to the types of computing power, this paper divides the business scenarios of computing force network into three types: storage task, computing task, training task, and expounds the work of computing force network security arrangement management system in each scenario.

2 Storage Task

Computing force network can provide storage services based on magnetic needle and distributed storage. Its typical scenarios include warm, cold data storage, backup in medical, cultural tourism, public services, urban management, finance, industry and Commerce and other industries. However, computing force network has no complete control over computing power resources, and there are security problems such as data privacy disclosure in practical applications. Encryption is undoubtedly one of the important ways to protect the security and privacy of data stored in computing force network.

Traditional encrypted storage solutions require the data owner to perform complex data encryption tasks locally, which burdens users and improves the quality of user service experience.

Based on the cloud storage encryption architecture proposed by Microsoft Research [4], this paper proposes a computing force network encryption storage system, as shown in Fig. 1. In the figure, certificate generation, data processing, data verification and token generation are the core modules of the security arrangement management system of computing force network, which are located in the trusted region of computing force network. Certificate generation module is responsible for generating access certificates for authorized users; data processing module is responsible for data partitioning, encryption, coding and other operations before data storage to computing power resources; data

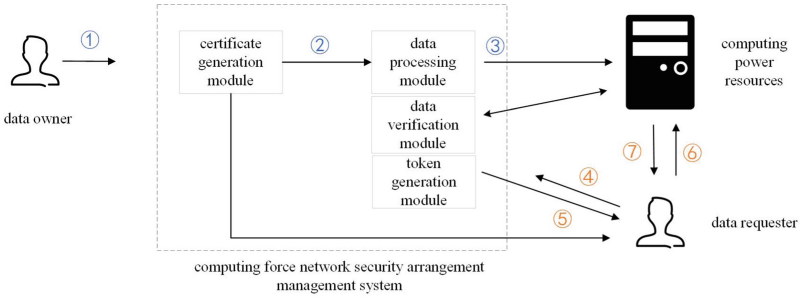

Fig. 1. The data encryption algorithm in computing force network

verification module is responsible for verifying the data integrity of computing power resources; token generation module is responsible for data access tokens.

Computing force network encryption storage system has two phases: the data uploading stage of the data owner and the query stage of the data requester. In the data uploading stage of the data owner, ① the data owner accesses the computing force network and uploads the data to be stored; ② the certificate generation module generates the certificate; ③ data processing module processes data. In the query stage of data requesters, ④ data requesters send access requests, upload keywords; ⑤ computing force network verifies the identity of users, sends data tokens and vouchers to data requesters; ⑥ data requesters use tokens to extract ciphertext from computing power resources; ⑦ data requesters use credentials to decrypt data.

The data encryption algorithm of the data processing module in Fig. 1 can be divided into two types: data storage technology based on encryption technology; secure search technology based on encrypted data. The comparison of these two methods is shown in Table 1.

In data storage technology based on encryption technology, the traditional encryption and decryption technology is usually combined to ensure data security. Proxy re-encryption is a common form of data encryption. The proxy server is set up in the network to convert the ciphertext of the data owner into the ciphertext that can be decrypted by the authorized data visitor, which can achieve better data sharing.

Encrypting data before uploading it to computing power resources can ensure data privacy and security, but it will affect data availability and make data query tasks difficult. To address the above problems, Searchable encryption was proposed and used to provide secure search services based on encrypted data. The user submits the query keyword or query condition to the computing force network, and the computing resource server searches the keyword index to find the data that meets the condition, and then returns the query result to the user.

Table 1. Comparison of data security storage technologies in computing force networks

Technical	Technical characteristics	Major problems
data storage technology based on encryption technology	Adopt traditional encryption technology	The security mechanism is complex and has security risks, and the time and space cost is too large
secure search technology based on encrypted data	Encryption algorithms support ciphertext query	Basic operations such as addition, subtraction, multiplication and division are not supported

3 Computing Task

Computing force network can provide computing services by relying on chips such as CPU. Its typical scenarios include Internet behavior analysis, location capability analysis in urban management, public services, finance, industry and Commerce and other industries, as well as home wide log synthesis calculation of operators. However, the information collected by these services and applications often contains a lot of sensitive information, including medical history, income, identity, interest and location. The sharing, publishing, analysis, calculation and other operations of these information will directly or indirectly disclose the privacy of users. The security arrangement management system of computing force network will analyze the security requirements of computing power task and design the corresponding privacy calculation algorithm to ensure the security of data calculation process.

Before researching the privacy computing algorithm design/selection scheme for the computing force network security scheduling system, this paper classifies the current mainstream privacy computing technology, which is mainly divided into three directions: the first category is privacy computing technology based on cryptography, including differential privacy (DP), secure multi-party computing (SMC), homomorphic encryption (HE), differential privacy (DP), zero knowledge proof (ZKP); the second category is the fusion of artificial intelligence represented by federated learning (FL) and privacy protection technology; the third category is privacy computing technology based on trusted hardware represented by trusted execution environment (TEE). The comparison of different privacy computing methods is shown in Table 2.

Based on the privacy computing framework [5], this paper proposes a privacy computing algorithm design/selection scheme for the computing force network security arrangement management system, as shown in Fig. 2.

1. Semantic extraction. The privacy information I is extracted from the plaintext information M which is used to represent the indivisible and disjoint privacy (semantic) features in the information.
2. Scene extraction. According to the types and semantics of each privacy component in privacy information I, the application scenarios are abstracted, including the types of computing tasks, constraints, etc.
3. Privacy computing algorithm selection/design. Select/design appropriate privacy protection methods. If there is a scheme that is available and meets the conditions, it is directly selected; if there is no one, it is redesigned.

   (1) Pretreatment. Preprocess the privacy information I to determine the data distribution characteristics, value range, etc.
   (2) Algorithm framework. According to the application scenario and application category, the privacy protection algorithm is determined, and the algorithm steps and the combination relationship between steps are given in detail.

Table 2. Comparison of privacy computing methods

Tec	implementation	advantage	Major problems
SMC	Data systems and structuring	Theoretically, no third party is required	The computation force is large, the time is long
HE	Data occlusion	No data information loss	High computational power consumption; With the increase of data, the computing speed slows down significantly
DP	Data denoising	The magnitude of "noise" can be added as required	After adding "noise", the accuracy of data decreases
ZKP	Data system and structure	Achieve a specific purpose and provide only minimal information	Theoretical safety is not fully proven and widely accepted; Some types of computation are less efficient; The degree of standardization is not high
FL	Data model	Support more AI algorithms; Distributed architecture reduces the cost of computing power	The quality of data model is uneven; High communication complexity
TEE	Data isolation	No loss of data information; No algorithm limitation in the domain; High versatility and low development difficulty	May face channel test attack

(3) Algorithm parameter. Determine the specific value of the corresponding parameters of the privacy preserving algorithm.
(4) Algorithm combination. According to the application scenario, different steps are combined within the algorithm, or between different algorithms, so as to improve the security or performance.

4. Algorithm performance analysis. According to relevant evaluation criteria, the privacy protection effect, time complexity and space complexity of the selected privacy protection scheme are evaluated. When the algorithm performance does not reach the expectation, the feedback mechanism is implemented, including two situations: ① If the scene abstract is not properly, the scene will be abstracted and iterated again; ② If the scene abstract is correct, the privacy protection scheme is adjusted and improved to achieve a satisfactory algorithm effect.

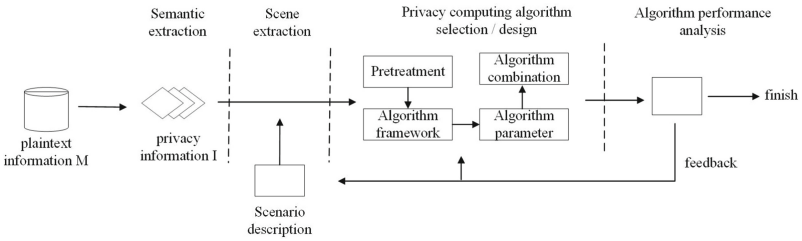

Fig. 2. The privacy computing algorithm design/selection scheme of computing force network security arrangement management system

4 Training Task

Computing force network relies on GPU, NPU, DPU and other computing power to realize model training of personnel certification, document audit and intelligent recommendation in intelligent manufacturing, meteorology, finance, transportation and other industries. In the training task of the computing force network, due to the large amount of data and the need to protect users' privacy, the data of the terminal node cannot leave the local centralized to the central cloud server. Therefore, it is necessary to add edge nodes between users and the cloud to form a training model of cooperation between edge servers and cloud servers.

Figure 3 shows the face recognition system supporting privacy protection in the computing force network, which can flexibly schedule the deployment location of tasks according to the resource requirements of businesses. Tasks are deployed in the edge room during business idle hours such as weekends, and tasks will be dispatched to the room during business busy hours such as rush hours, and the on-board system will be called to share the resource pressure during super busy hours such as large-scale hot activities. It includes model training and inference prediction. The model training part should be deployed in a centralized computing resource pool, such as in a large cloud computing center to complete complex model training tasks offline. Since the transmission of facial information to the cloud computing center far away will bring a large network delay, the inference prediction part should be deployed in the edge room closer to the user, so as to quickly realize the task of face recognition. The cloud will publish the constructed model to the edge node, which is responsible for processing the data uploaded by the terminal node and generating the processing scheme. Similar architecture can also be used in the Internet of vehicles [6, 7], traffic video surveillance [8] and other occasions.

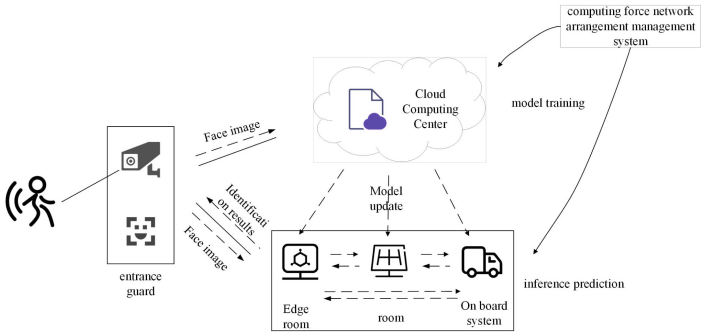

Fig. 3. The face recognition system supporting privacy protection in the computing force network

Table 3. Comparative experiment of several secure storage technologies in computing force network

Scheme	Symmetry	Security	Computational complexity
SWP [9]	symmetric	CPA	O(n)
BCO [10]	Asymmetric	CKA2	O(m)
PKEET [11]	Asymmetric	OW-CCA	$O(\log^3 n)$

5 Experiment and Analysis

For storage tasks, according to different application scenarios and security requirements, experiments are carried out on the data encryption algorithm in computing force network proposed in Sect. 2, in which the data processing module selects several typical encrypted storage schemes: Song first proposed symmetric searchable encryption scheme [9], Dan Boneh first proposed asymmetric searchable encryption scheme using bilinear pairing [10], Yang first proposed public key encryption with equality test [11], which can judge whether the ciphertext encrypted by different public keys corresponds to the same plaintext, and compare the security and time complexity of these schemes, as shown in Table 3. Chosen Plaintext Attack (CPA) indicating that ciphertext and index will not disclose Plaintext information without query. Adaptive Chosen Keyword Attack (CKA2) allows the attacker to send query requests based on searched tokens and search results, ciphertext and index will not disclose plaintext. One way-chosen Ciphertext Attack (OW-CCA) indicates that the attacker can also make or select some Ciphertext at will and obtain the plaintext after decryption. In Table 3, n represents the total number of plaintext space words and m represents the total number of data records.

SWP [9] is a typical symmetric searchable encryption method, which can only resist CPA attacks. BCO [10] can resist CPK2, and its efficiency is lower than that of symmetric searchable encryption algorithm [9] due to the use of bilinear pair operation. PKEET [11] satisfies OW-CCA and has low computational complexity. The specific technology used in the storage scheme of computing force network depends on different application scenarios and security requirements.

For computing tasks, many factors such as hardware performance, transmission cost and time constraints of user devices need to be taken into account when privacy protection technology is actually used. For example, when organizations with a large amount of sensitive data, such as hospitals or banks, act as users, they need to use homomorphic encryption technology to ensure the security of the model. When individuals with weak computing power are users, differential privacy technology needs to be used to ensure the efficiency of the model. In a distributed environment, multi-party cooperative training of a machine learning model may use secure multi-party computing technology to ensure the privacy of all parties. More and more research is devoted to the combination of SMC, HE, DP and other methods to achieve a reasonable trade-off between data privacy and utility.

For the training task, it includes two stages: the training stage of the machine learning model and the prediction stage of the model. In different stages, it faces different privacy threats. In addition, due to the technology of machine learning itself, the protection methods adopted are also different. For example, homomorphic encryption technology is mostly used in the prediction stage of neural networks, but rarely used in the training stage The reason is that machine learning itself is a computing intensive task, which costs a lot of computation and communication. Even without encryption, high-throughput computing units are required, and homomorphic encryption also costs a lot of computation and communication.

6 Conclusions

In this paper, the business scenarios of computing force network are divided into three categories: storage task, computing task, training task, and the work of computing force network security arrangement management system is analyzed respectively. In order to ensure the safe operation of computing force network, there are still many technical challenges: on the one hand, computing force is the basic unit of measurement of computing force network, and the classification of security classification and sensitivity level of computing force has not reached a consensus in the industry, resulting in difficult unified computing force billing and scheduling; On the other hand, the security capability of computing power resources changes in real time in the network, so it is of practical significance to study the dynamic real-time perception and prediction of computing power security capability and the more intelligent security arrangement management method.

References

1. Yang Y (2019) Multi-tier computing networks for intelligent IoT. Nat Electron 2(1):4–5
2. Tang X, Cao C, Wang Y et al (2021) Computing power network: the architecture of convergence of computing and networking towards 6G requirement. China Commun 18(2):11
3. Liu ZN, Li K, Wu LT et al (2020) CATS: cost aware task schedulinging multi-tier computing networks. J Comput Res Dev 57(9):1810
4. Kamara S, Lauter K (2010) Cryptographic cloud storage. In: Proceedings of the International Conference on Financial Cryptography and Data Security. FC 2010. LNCS, vol 6054. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-14992-4_13

5. Li F, Li H, Niu B, Chen J Privacy computing: concept, computing framework, and future development trends. Engineering 5(6):1179–1192
6. Bao W, Wu C, Guleng S et al (2021) Edge computing-based joint client selection and networking scheme for federated learning in vehicular IoT. China Commun 18(6):39–52
7. Shinde SS, Bozorgchenani A, Tarchi D et al (2021) On the design of federated learning in latency and energy constrained computation offloading operations in vehicular edge computing systems. IEEE Trans Veh Technol
8. Sada AB, Bouras MA, Ma J et al (2019) A distributed video analytics architecture based on edge-computing and federated learning. In: 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), pp 215–220. IEEE
9. Song DX, Wagner D, Perrig A (2000) Practical techniques for searches on encrypted data. In: IEEE Symposium on Security & Privacy, pp 44–55. Berkeley
10. . Boneh D, Crescenzo GD, Ostrovsky R et al (2004) Public Key Encryption with Keyword Search. In: Cachin C, Camenisch JL (eds) Advances in Cryptology - EUROCRYPT 2004. EUROCRYPT 2004. LNCS, vol 3027. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-24676-3_30
11. Yang G, Tan CH, Huang Q et al (2010) Probabilistic public key encryption with equality test. In: Pieprzyk J (eds) Topics in Cryptology - CT-RSA 2010. CT-RSA 2010. LNCS, vol 5985. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-11925-5_9