Abstract—The Internet of things(IoT) makes our life more intelligent. Its combination with the cloud server can solve big data processing problem to meet users' needs and bring us great convenience. However, there are two challenges we need to face, data sharing and Key-Leakage. To solve above challenges, attribute-based encryption is used to achieve data sharing combined with searchable encryption. Most of existing attribute-based searchable encryption schemes are inefficient and not suitable for Internet of things devices because of the large amount of attributes and keys. And the key-leakage problem is serious in practice which very little literature focused on it. In order to address both problems, in this paper, we propose a key aggregation searchable encryption scheme based on blockchain with auxiliary input, which is capable of achieving secure data sharing on encrypted data. Our scheme is presented through a novel CPA secure(Chosen Plaintext Attack) scheme. We prove our scheme is CCA secure(Chosen Ciphertext Attack) against Key-Leakage under the Decisional Diffie-Hellman assumption and Goldreich-Leivin Theorem. Moreover, we adopt the proposed scheme to establish a data sharing system based on blockchain, which improves search efficiency and connects the global ecology. In addition, extensive performance evaluations are conducted, and the results indicate our scheme is really efficient in cloud computing-enhanced IoT. Index Terms—Cloud computing, Outsourcing, Data sharing, Key-Leakage, Blockchain, CCA secure. I. INTRODUCTION WITH the increasing convergence of the Internet of things (IoT) with our daily life, it has gradually realized the transformation from interconnected everything to intelligent-connected everything, which brings us convenience. Data sharing realized by it can break the ”information island” problem among various domains [1]–[3]. However, after entering the era of big data, the huge amount of data has limited the application of IoT. Meanwhile, the cloud appears in our field of vision with its powerful data-processing capabilities, for enhancing IoT. The data sharing realized by combination of IoT and cloud also makes our life more intelligent and convenient [4]. For example, in Internet of Vehicles(IoV), vehicle-to-vehicle interaction can be realized through IoT and cloud, and road section information can be obtained by sharing data so as to realize intelligent transportation. Jie Niu is with the School of Mathematics and Staticstics, Xidian University, Xian, China. e-mail: Niu Jie1754@163.com. Xuelian Li is with the School of Mathematics and Staticstics, Xidian University, Xian, China. e-mail: xuelian202@163.com. Juntao Gao is with the School of Telecommunication and Engineering, Xidian University, Xian, China. e-mail: jtgao@mail.xidian.edu.cn. Yue Han is with the School of Mathematics and Staticstics, Xidian University, Xian, China. e-mail: Han Yue0526@163.com. Manuscript received May 14, 2019; revised November 21, 2019. Fig. 1. The challenges in IoT. Clearly, IoT will be of great benefit to our daily lives. However, in order to fully take advantage of combination of IoT and cloud, we have to address some challenges lying ahead in IoT. Firstly, data in different domains cannot be shared. A large amount of data needs network transmission, as well as data analysis and other functions. Moreover, these shared resources are limited. If shared resources in IoT cannot be distributed in a balanced way, which is likely to result in poor utilization rate of shared resources. Because there are different clouds in different domains, the situation limits the expansion and application of ecosystem architecture in IoT. Secondly, in the process of data sharing, it may suffer from malicious attacks, i.e., the leakage of user's keys. Key-Leakage is a serious security problem. No matter how secure cryptosystem is designed, as long as key is leaked, all cryptographic operations about the entire key are no longer secure. And it is easier for an adversary to physically

acquire key than to break actual difficulty assumption. For instance, in an identity-based cryptosystem, the disclosure of user's key means that corresponding identity information (such as ID number) needs to be modified, which is very serious in real life. The situation of cyber security is not optimistic, which has seriously threatened people's normal life and even national security. How to use cryptography to reduce the possibility of Key-Leakage and achieve data sharing is of great significance. As shown in Fig.1, we can see two challenges in IoT: Data sharing and KeyLeakage. For example, in the data sharing scenario in Fig.1, the medical institution needs patients' information stored in the hospital for medical research. At the same time, the hospital also needs to collect all relevant information of patients in order to provide them accurate services. This situation requires data sharing to better serve humanity. However, in the process of data sharing, some external adversaries may obtain secret JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015 2 TABLE I COMPARISONS OF RELATED WORK. Scheme Searchable Encryption Attributed-based Encryption Key Aggregation Cryptosystem Data-Sharing Anti-Key-Leakage Anti-Internal Keyword Guessing Attacks [6] √ × × × × - [5, 7 − 10] √ × × √ × - [11 − 17] √ √ × √ × - [18] × × √ √ × - [19 − 21] × × √ √ √ - [22 − 23] × √ √ √ × - [24] √ × √ √ × × [26 − 27] √ × √ × × × [28] √ × × √ × √ key through some means (such as side channel attack), which seriously damages the interests of data users. In order to solve the above problems, this paper proposes a data sharing system by key aggregation searchable encryption based on blockchain with the auxiliary input model(BAIKASE). We introduced blockchain technology to establish the data sharing system. Blockchain, with the characteristics of distributed storage, highly transparent information, nontampering and effective credit sharing, is increasingly loved by various industries. It can integrate data of various domains on a unified platform, which makes data sharing more convenient and saves costs. However, in the application of blockchain, there are corresponding deficiencies in some existing schemes to achieve data sharing , i.e., they don't consider key-Leakage problem. Our scheme can not only solve the above two challenges, but also extend to all domains of IoT. Moreover, we can connect the global ecology through blockchain. If possible, we can establish a blockchain in a continent or a region to facilitate more open and transparent face-to-face communication around the world and to better serve mankind. In addition, the characteristics of blockchain that is always safe and effective can also provide us with a variety of micro services, making life more convenient and better. A. Related Work 1) key aggregation searchable encryption (KASE): Boneh [5] first proposed a public key encryption with keyword search scheme to achieve efficient data retrieval from encrypted data. Subsequently, various keyword search algorithms are proposed to provide diverse search functions in [6]–[11]. Furthermore, many keyword search encryption schemes(SE) require cloud server to store encrypted files and combine attribute-based encryption(ABE) to achieve data sharing. Attribute-based keyword search (ABKS) encryption scheme are proposed in [12]–[17]. All these schemes can achieve fine-grained access control of encrypted data. Inspired by ABE, Sun et al.[15] presented the first attribute-based keyword search scheme with efficient user revocation (ABKS-UR) providing scalable finegrained search authorization. Taking into account the more realistic scenario, Liu et al.[17] proposed an online/offline attribute-based signature scheme for sharing of Mobile Health Records. However, the mentioned above

schemes incur high computational costs and are not suitable for mobile devices with power consumption constraints(such as mobile phones). To solve the problem, the related works introduced the key aggregation system (KAC). Chu et al. [18] first proposed a Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage. Meanwhile, the author put forwards a public challenge: How to design a leakage-resilient cryptosystem. Based on [18], Patranabis et al. proposed a provably secure key-aggregate cryptosystems with broadcast aggregate keys for online data sharing on the cloud [19]. But their scheme isn't a leakage-resilient cryptosystem. In 2019, Wang [20] came up with a CPA secure leakage-resilient Key-Aggregate cryptosystem with Auxiliary input based on [19]. Later, many schemes improved KAC and made it suitable for realistic scenarios in [21]–[23].However, Cui combined the KAC system with SE scheme and proposed the first KASE scheme [24]. In [24], they addressed a practical problem, by proposing the novel concept of key-aggregate searchable encryption and instantiating the concept through a concrete KASE scheme, in which a data owner only needs to distribute a single key to a user for sharing a large number of documents, and the user only needs to submit a single trapdoor to the cloud for querying the shared documents. Later, many works [25]–[27] based on [24] have been proposed. In [27], Liu et al. came up with a verifiable searchable encryption with aggregate keys for data sharing system, where a data owner need only distribute a single aggregate key to others. However, the authors [28] argued that Cui's scheme is vulnerable to cross pairing attacks(or dictionary or guessing). Cross-pairing attack is a special internal keyword guessing attack(IKGA). This kind of attack can make keyword divulge, causing privacy to divulge thereby. Therefore, we just need to make the keyword and ciphertext have no direct connection to prevent the keyword guessing attack. We will discuss the specific security in this paper. The notion of leakage resilient cryptography has been proposed in recent years, and a large number of efforts have been made in this topic. In general, there are mainly three leakage models. (1) Bounded retrieval model [29], [30]. Under this model, the adversary can obtain bounded leakage information of the internal state of the whole scheme; (2) Continual leakage model [31], [32]. The adversary can obtain a limited amount of leaked information in each time period and the total amount is unlimited; (3) Auxiliary input model [33]–[35](AI). Under this model, the adversary can not recover secret key and other information through the leaked information. That is to say, even such a function information-theoretically reveals the entire secret key SK, it still computationally infeasible to recover SK from f(SK). Based on the above brief analysis, the auxiliary input model is relatively safe, so we will combine AI with KASE to construct the anti-leakage scheme. This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2956322, IEEE Internet of Things Journal TABLE II THE SYMBOL DESCRIPTIONS IN OUR SCHEME. Notation Meaning Notation Meaning Notation Meaning λ System parameters kagg The aggregate key ECDSA Elliptic Curve Digital Signature Algorithm G, GT Cyclic group with order p Tr The aggregate trapdoor CPA Chosen Plaintext Attack g A generator of group G Tri The single trapdoor CCA Chosen Ciphertext Attack e A bilinear pairing map M0 The message needed decrypt AI-CPA Chosen Plaintext Attack with Auxiliary Input Hi Cryptographic hash function Sign The signature of encrypted files AI-CCA Chosen Ciphertext Attack with Auxiliary Input n the maximum possible number of documents KSIG The secret key of signing IKGA Internal Keyword Guessing Attacks

Ppub System public key VSIG The secret key of verifying SEAI Strong Extraetor with Auxiliary Input m0 A point of G used for user registration α, u0, r1, r11 The random numbers PPT Probability Polynomial Adversary ID Identity ~s, ~u,~t, K~ s The random vectors DDH Decisional DiffieHellman Assumption (sk, pk) User's secret keys gi, X0, U0, ~αs, d(m0), dID R1, R11, V, as, bs, h2, h3 The computing component CBDH Computational Bilinear Diffie-Hellman Assumption CT = (C1, C2, C3) The ciphertext of files DO The Data Owner PBFT Practical Byzantine Fault Tolerance CW The ciphertext of keywords DU The Data User PoW Proof of Work h, h1 The hash value of keywords' ciphertext CS The cloud server In Table I, we summarize the resistance of the mentioned above schemes to Key-Leakage Attacks and other security. 2) Blockchain in IoT: IoT is experiencing exponential growth, but it still suffers from privacy and security vulnerabilities. Conventional security and privacy approaches tend to be inapplicable for IoT, mainly due to its decentralized topology and the resource-constraints of the majority of its devices. Therefore, the blockchain technology is introduced in IoT to solve the problems [36]–[41]. In [37], they proposed a blockchain-based secure and privacy-preserving PHI sharing scheme for diagnosis improvements in e-Health systems. And in [39], the proposed architecture is hierarchical, and consists of smart homes, an overlay network and cloud storages coordinating data transactions with blockchain to provide privacy and security. In [41], the authors proposed a lightweight blockchain-based architecture for IoT that virtually eliminates the overheads of classic blockchain, while maintaining most of its security and privacy benefts. B. Our Contributions In this paper, in view of above challenges, which only support data sharing through ABE and cannot resist KeyLeakage attack, we propose a new anti-key-leakage attack data searching scheme, called Blockchain-based Key Aggregation Searchable Encryption Scheme with auxiliary input (BAI-KASE). What's more, we established a data sharing system based on blockchain with corresponding consensus mechanism, for cloud computing-enhanced IoT. Specifically, the contributions of this paper are threefold as follows. 1) First, a BAI-KASE scheme is proposed, which enables users to not only aggregate different encryption keys into one for saving scarce bandwidth, but also introduce Auxiliary Input model for secret keys' security enhancement. We give the detailed analysis to show that our scheme is AI-CCA secure(Chosen Ciphertext Attack with Auxiliary Input) under DDH assumption and Goldreich-Leivin Theorem. The BAI-KASE scheme is based on an AI-CPA secure(Chosen Plaintext Attack with Auxiliary Input)scheme. Both schemes can resist Key-Leakage attacks with auxiliary input model. Moreover, BAI-KASE can also resist Internal Keyword Guessing Attacks(IKGA). 2) Secondly, a data sharing system is established, which is based on private blockchain and public blockchain with Practical Byzantine Fault Tolerance Algorithm (PBFT) and Proof of Work (PoW) consensus mechanism. It can effectively solve the problem of access control and data sharing to improve efficiency. Our scheme is capable of realizing not only intra-blockchain sharing but also inter-blockchain sharing. In the private blockchain, as long as you satisfy the access control attribute of private blockchain and register, you can freely access the shared data stored in the cloud server in the private blockchain. What's more, if you want to access shared data in other private blockchains, you can get the data you are interested in through some legal data transactions. Besides, users can verify data integrity and non-tampering feature through blockchain. Our blockchain system is secure through detailed security analysis. 3) Thirdly, the efficiency and performance analysis of the proposed scheme is justified by concrete implementations. We simulate our scheme on mobile phone and computer. In the next place, we compare the performance on BAI-KASE with others in terms of computation costs to show the advantage of our scheme. Moreover, we illustrate the storage overhead and communication overhead on blockchain, and

the results indicate our scheme is really efficient and the payload of blockchain is small, which suits in cloud computingenhanced IoT. This rest of paper is organized as follows. Section II gives the necessary background knowledge. Later, we establish a data sharing system based on blockchain and BAI-KASE scheme in detail in Section III. Finally, in Section IV and V we conduct