CIS-TC Newsletter

Communications and Information Security Technical Committee

http://cis.committees.comsoc.org/

Issue 11 – December 2023

Contents

Message from the Chair 1

CISTC Technical Recognition Award 2023.....2

CISTC Early Career Award 2023 2

Featured Topics 3

CIS-TC Organized Symposia 5

CIS-TC Affiliate Conferences 5

Forthcoming Meeting 5

MESSAGE FROM THE CHAIR

Dear CISTC Members,

Many advances have been made as we are striding towards the next generation of communication and information security systems. IEEE ComSoC CISTC continues to serve as the platform for security researchers and practitioners to exchange information, share ideas and best practices, resources, and knowledge. Here, we are very pleased to welcome you to the present issue of CISTC Newsletter.

In Globecom 2023, we were very excited to have our first in-person TC meeting in the post-pandemic era. Different from previous virtual TC meetings, we not only joyfully engaged in conversations with old friends but also forged connections with new friends and new TC members. During the TC meeting, we discussed a lot, including CISTC Technical Recognition Award, Early Career Award, Webinars, and others. In particular, we agreed on the importance of our CISTC Newsletter, as it has indeed served as one of the essential ways to reshape relationships with our communities. All TC members really appreciate our Newsletter Chair, Professor Francesco Chiti, for his great effort in organizing our CISTC Newsletter, making it become an open and inclusive platform for our TC members. Hence, I would like to encourage you to send us any scientifical, technical contributions or even news that you believe could be beneficial to our community. Meanwhile, I also do hope you will enjoy in reading this Issue.

Finally, let me conclude this message wishing you and your loved ones, on behalf of all the CISTC Officers, peaceful and healthy season holidays together with a renewed and prosperous 2024.

Sincerely,
Rongxing Lu
Chair of CISTC

CIS-TC Newsletter
Communications and Information Security Technical Committee
http://cis.committees.comsoc.org/

CISTC Technical Recognition Award 2023

The Award Committee chaired by Professor Francesco Chiti, with unanimous consent decided to give CIS-TC Technical Recognition Award 2023 to Professor Prof. Yi Qian from University of Nebraska-Lincoln, (NE, USA) for "contributions in the advancement of communications and information security".

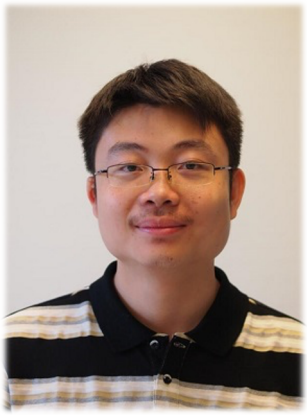
The Committee congratulate with the Awarded and he will be awarded during the CIS-TC meeting held at IEEE Globecom 2023 in Kuala Lumpur (Malaysia).

Yi Qian *(Fellow, IEEE) received the Ph.D. degree in electrical engineering from Clemson University, Clemson, SC, USA. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Nebraska–Lincoln (UNL). Prior to joining UNL, he worked in the telecommunications industry, academia, and government. Some of his previous professional positions include serving as a Senior Member of scientific staff and a Technical Advisor at Nortel Networks, a senior systems engineer and a technical advisor at several startup companies, an Assistant Professor at the University of Puerto Rico at Mayaguez, and a Senior Researcher at the National Institute of Standards and Technology. He has research and industry experience in information and communication network security, wireless communications and networks, wireless sensor networks, vehicular communication networks, smart grid communications, broadband satellite communications, optical communications, high-speed communications and networks, and the Internet of Things. His research interests include cyber security and communication network security, and computer and communication networks. He was the Chair of the IEEE Technical Committee for Communications and Information Security. He was the Technical Program Chair of the IEEE International Conference on Communications 2018. He serves or has served on the Editorial Boards of several international journals and magazines, including as the Editor-in-Chief for IEEE Wireless Communications from July 2018 to June 2022. He was a Distinguished Lecturer of the IEEE Vehicular Technology Society and a Distinguished Lecturer of the IEEE Communications Society.*

CISTC Early Career Award 2023

The Award Committee chaired by Professor Francesco Chiti, with unanimous consent decided to give CIS-TC Early Career Award 2023 to Dr. Weizhi Meng from Technical University of Denmark (Denmark) for "his contributions in the advancement of communications and information security".

The Committee congratulate with the Awarded and he will be awarded during the CIS-TC meeting held at IEEE Globecom 2023 in Kuala Lumpur (Malaysia).



Weizhi Meng *(Senior Member, IEEE) received the Ph.D. degree in computer science from the City University of Hong Kong (CityU), Hong Kong. He was a Research Scientist with the Department of Infocomm Security (ICS), Institute for Infocomm Research, A*STAR, Singapore, and a Senior Research Associate with the Department of Computer Science, CityU. He is currently an Associate Professor with the Department of Applied Mathematics and Computer Science, Cybersecurity Section, Technical University of Denmark, Denmark. His primary research interests include cybersecurity and intelligent technology in security, including intrusion detection, smartphone security, biometric authentication, HCI security, trust computing, blockchain in security, and malware analysis.*

CIS-TC Newsletter
Communications and Information Security Technical Committee
http://cis.committees.comsoc.org/

FEATURED TOPICS

"Enhancing Secure Keyword Search for Cloud Storage: The Charm of Decentralized Trust"

Cheng Huang[1], and Rongxing Lu[2]
[1]Department of ECE, University of Waterloo, Canada
[2]Faculty of Computer Science, University of New Brunswick, Canada

While secure keyword search for cloud storage has been extensively studied, this topic continues to draw considerable attention from both academia and industry in recent years. This interest stems from the growing convenience users find in storing and managing data in remote clouds. However, they concurrently harbor serious concerns about data security when their data is out of their control. As most commercial cloud services merely provide plain file storage services to support rich data management functionalities, a fundamental trust conflict emerges. Users must rely on cloud service providers (CSPs), such as Dropbox and OneDrive, to protect their data and assume that these CSPs will not be compromised by adversaries in the future.

In this context, the study of secure cloud storage solutions continues to be a vibrant field[1-4]. Among many areas of focus, a significant research interest is the enablement of keyword search on encrypted outsourced data. Specifically, to maintain data security, users encrypt their files before uploading them to the cloud, which creates the challenge of searching encrypted data. To address this issue, a cryptographic primitive, searchable encryption, was proposed[4]. By utilizing searchable encryption, users can perform keyword searches to locate the files they need while keeping their files secure. But most of existing searchable encryption-based keyword search schemes normally suffer from search pattern leakages and access pattern leakages. A well-known attack leveraging access patterns is the file-injection attack, where an attacker meticulously injects specific files into a user's outsourced database[6]. By doing so, the attacker may have a great chance to successfully guess the user's queries merely by observing the patterns in which the user accesses these files. Furthermore, there exist additional attacks aimed at exploiting the vulnerabilities caused by search pattern leakages. An attacker, without needing detailed knowledge of a user's outsourced database and merely armed with its statistical information, can accurately identify the keywords included in the user's queries with a remarkably high probability[7].

To this end, private information retrieval (PIR) techniques are introduced into secure keyword search schemes recently, which can be broadly categorized into two groups:

- Secure Keyword Search based on Single-Server PIR[8]: The scheme allows a user to locate the files matching the queried keywords from a database stored on a single server, without revealing to the server which keyword is being searched and which files are matched.
- Secure Keyword Search based on Multi-Server PIR[9]: The scheme enables a user to identify files that match the queried keywords from replicated databases from multiple servers (two servers in most cases), without disclosing to the server which keyword is being searched and which files are matched.

In the first category, partially or fully homomorphic encryption (HE) is often utilized to fulfill security requirements[10]. As a result, the efficiency of HE-based secure keyword search primarily hinges on the performance of HE itself, as well as the innovative techniques developed to leverage it. In the second category, security is built upon a more relaxed assumption: it requires that at least one server is not compromised by adversaries and does not collude with them, thus establishing a decentralized trust model. Despite this model not adopting the most stringent security assumption, it can lead to remarkably efficient keyword constructions, thereby drawing substantial research attention.

Given these considerations, we propose an efficient and secure keyword search scheme[1] based on the decentralized trust model, leveraging the cryptographic primitive known as Distributed Point Function (DPF)[11]. The choice of DPF is motivated by its exceptional communication efficiency when used to construct PIR. DPF possesses a fascinating property: it enables two servers to jointly evaluate a private function using only one non-zero value, while maintaining the confidentiality of that value. This property allows DPF-based PIR to maintain logarithmic communication costs. Building upon DPF-based PIR, we can naturally design secure keyword search schemes using bitmap-based indexes, where each row represents a bitmap of words for a document. These bitmap-based indexes are encrypted before being outsourced, and users can search for a keyword by privately retrieving the corresponding column. However, despite their simplicity, bitmap-based indexes may become excessively large to accommodate words from all stored documents.

To address this limitation, bloom-filter-encoded indexes are intrduced into the secure DPF-based keyword search schemes[9]. As depicted in Figure 1, the keywords of a document are compressed into a Bloom Filter (BF), which is then inserted as a single row in the keyword indexes. With a fixed size based on the maximum number of keywords associated with a document, BF-encoded indexes offer more efficient storage. Similar to bitmap-based indexes, BF-encoded indexes can be encrypted and searched. During a keyword search, the keyword is mapped to a Bloom Filter using k hash functions to determine the positions, and the corresponding k columns are privately retrieved using DPF. The retrieved columns are then combined to generate a k-column query

result, and users can scan each row of the result. If a row consists entirely of ones, it indicates that the corresponding document contains the searched keyword.

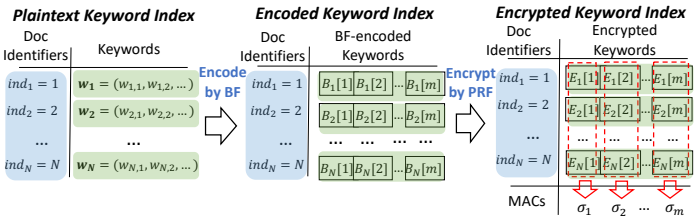


Figure 1. High-level flows of constructing an encoded and encrypted keyword index using BF and pseudorandom functions (PRF)

While the aforementioned schemes perform well, they may not always be appropriate due to the inherent characteristics of Bloom Filters (BF). Generally, a BF's size (m) and the number of required hash functions (k) are largely determined by the number of inserted elements (n) and the false-positive rate (ε, where $0 < \varepsilon < 1$). Notably, for BF-encoded indexes, a document may have a large number of associated keywords, i.e., n is large, amd a negligible false positive rate is needed for keyword search, ε is small. These conditions can lead to a situation where $k * m$ (the number of DPF evaluations performed at server side to search for a single keyword) grows largely. This increased computational load at the server-side results in longer search latency, significantly degrading user experience. For instance, our experiments show that with a database of 4,000 documents and 1,000 keywords per document, searching for a single keyword with a false positive rate of 10^{-5} requires over 17 seconds.

To reduce search latency, our idea is to empower cloud servers to obliviously "aggregate" all columns related to query keywords. This will allow servers to perform DPF evaluations m times rather than $k * m$ times. To attain this, we propose new encoding methods based on garbled bloom filter and cuckoo filter to generate keyword indexes. We also adapt multi-point DPF evaluation based on cuckoo hashing to reduce the complexity of DPF evaluation at the server-side to $O(3m)$. Furthermore, we expand DPF-based keyword search with a segmentation method that divides a search query into M segments (M ranges from 1 to m), reducing the server-side complexity to $O(3M)$.

Certainly, these improvements do not come without costs: system initialization time and updating delay are prolonged, and the communication overheads of searching one keyword increase marginally (e.g., an extra 14%). Nevertheless, these extra costs are balanced by a significant search latency reduction of more than 100 times, thus making this trade-off worthwhile. In addition, our scheme also offers verifiability to identify malicious servers that return incorrect query results. We attain this feature by applying Wegman-Carter message authentication codes (WCMAC) that support homomorphic XOR and integrity checking, since the basic MAC methods cannot adapt to oblivious aggregation operations. To ensure the confidentiality of encoded indexes, we devise a double encryption method leveraging set-constrained pseudorandom functions (SCPRF). This method enables users to encrypt the encoded index prior to outsourcing, thereby enhancing the overall security of the scheme.

Particularly, we give two dfferent constructions of the proposed scheme, dubbed "OC" and "VC", and compared them with a baseline ("BL") in our experiments[1,9]. We represent the experiments under the semi-honest model with "SEMI" and those under the malicious model with "MAL". Solutions that consider collusion attacks, are denoted as "CA". In such as case, The experiment settings are denoted by: BL-SEMI, BL-MAL, OC-SEMI, OC-MAL, OC-MAL-CA, VC-SEMI, VC-MAL, and VC-MAL-CA. In the experiment, we adopt a 128-bit security level and select two datasets for comprehensive experimentation: one being a subset of a real Wikipedia dataset comprising 4,000 documents (https://dumps.wikimedia.org) and the other being a synthetic dataset. Utilizing Python, we preprocess each Wikipedia document to extract keywords, with a maximum of 1,000 keywords extracted from a single document. As for the synthetic dataset, we choose 50,000 files to simulate a variety of cloud storage scenarios.

The experimental results show that, when the false positive rate is smaller, the improvement in search efficiency becomes more pronounced. This is due to the fact that both OC and VC reduce the complexities of DPF evaluation on the server side from $O(km)$ to $O(3M)$ and $O(2M)$ respectively, and a smaller false positive rate leads to a larger k. The effect is also more significant when M is smaller. Figure 2 and Figure 3 demonstrate the performance of single-keyword search with varying numbers of keywords and documents. It is evident that our scheme surpasses the baseline in terms of search latency. It is worth noting that, the number of keywords appears to have a more significant impact on search latency than the number of documents. This is because the cloud servers only need to perform relatively lightweight XOR operations with an increase in the number of documents, whereas an increase in the number of keywords results in a heavier load on DPF evaluation.

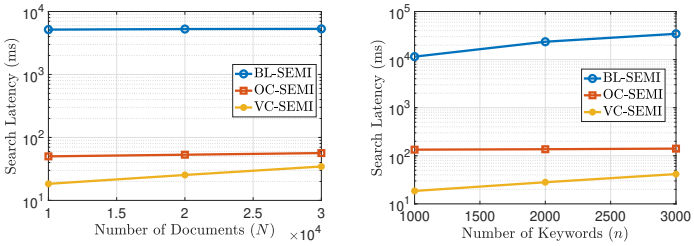


Figure 2. The search latency of the baseline and our scheme with $\varepsilon = 10^{-4}$, $n = 500$, and $N = 1{,}000$ (Semihonest model)

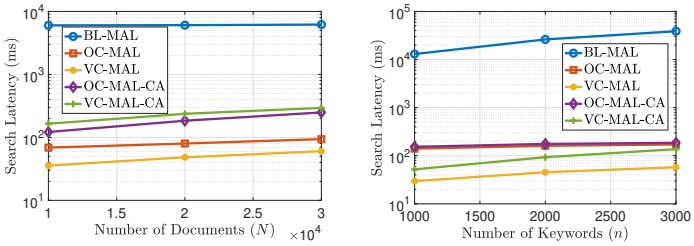


Figure 3. The search latency of the baseline and our scheme with $\varepsilon = 10^{-4}$, $n = 500$, and $N = 1{,}000$ (Malicious model)

In summary, the proposed multi-client keyword search scheme for cloud storage services successfully fulfills the requirements of security and efficiency within the decentralized trust model. The practicality of the

scheme has been validated through the development of a proof-of-concept prototype, accessible at https://github.com/EnderCheng/KeywordSearch. Furthermore, we give two possible future research directions. Firstly, search functionality can be extended by by incorporating support for expressive features, including Boolean queries. Secondly, pre-processing techniques can be investiagated to further enhance search efficiency.

Reference

[1] C. Huang, D. Liu, A. Yang, R. Lu, and X. Shen, "Multi-client Secure and Efficient DPF-based Keyword Search for Cloud Storage", IEEE Transactions on Dependable and Secure Computing, to appear.
[2] X. Zhang, C. Huang, Y. Su, J. Qin, and X. Shen, "Divertible Searchable Symmetric Encryption for Secure Cloud Storage", Proc. IEEE Globecom'22.
[3] R. A.Mahdavi, and F. Kerschbaum, "Constant-weight PIR: Single-round Keyword PIR via Constant-weight Equality Operators". Proc. USENIX Security'22.
[4] S. Zhang, S. Ray, R. Lu, Y. Guan, Y. Zheng, and J. Shao, "Efficient and Privacy-Preserving Spatial Keyword Similarity Query over Encrypted Data", IEEE Transactions on Dependable and Secure Computing, to appear.
[5] Y. Zheng, R. Lu, F. Yin, J. Shao, and H. Zhu, "Achieving Practical Symmetric Searchable Encryption with Search Pattern Privacy over Cloud", IEEE Transactions on Services Computing, Vol. 15, No. 3, pp. 1358-1370, 2022.
[6] Y.Zhang, J.Katz, and C.Papamanthou, "All your queries are belong to us: the power of file-injection attacks on searchable encryption", Proc. USENIX Security'16.
[7] E.M. Kornaropoulos, C.Papamanthou, and R.Tamassia, "The state of the uniform: Attacks on encrypted databases beyond the uniform query Distribution", Proc. IEEE S&P'20.
[8] H. Corrigan-Gibbs, A. Henzinger, and D. Kogan, "Single-server private information retrieval with sublinear amortized time", Proc. EUROCRYPT'22.
[9] E.Dauterman, E.Feng, E.Luo, R.A. Popa, and I.Stoica, "DORY: An encrypted search system with distributed trust", Proc. of USENIX OSDI'20.
[10] A. Viand,P. Jattke, and A. Hithnawi, "SoK: Fully homomorphic encryption compilers", Proc. IEEE S&P'21.
[11] E. Boyle, N. Gilboa, and Y. Ishai, "Function secret sharing: Improvements and extensions" Proc. ACM CCS'16.

CIS-TC Organized Symposia

- *Communications and Information System Security Symposium* – **Globecom2023,** 4-8 December 2023, Kuala Lumpur, Malaysia *"Intelligent Communications for Shared Prosperity"* (https://globecom2023.ieee-globecom.org/)
- *Communications and Information System Security Symposium* – **ICC2024**, 9–13 June 2024, Denver, CO, USA *"Scaling the Peaks of Global Communications"* (https://icc2024.ieee-icc.org/)

CIS-TC Affiliate Conferences

- **Conference: ICACT 2024**
  - 26th International Conference on Advanced Communications Technology
  - February 4-7, 2024, Phoenix Pyeongchang, Korea
  - https://www.icact.org/
- **Conference: WTS 2024**
  - Wireless Telecommunications Symposium of WTS 2024
  - April 10-12, 2024, San Francisco Bay Area (Oakland), California, USA
  - http://www.wtsconference.org/
- **Conference: SmartNets 2024**
  - International Conference on Smart Applications, Communications and Networking 2024
  - May 28-30 2024, Harrisonburg, Washington DC, USA (In-person / Virtual conference)
  - https://smartnets.ieee.tn/

Forthcoming Meeting

The next IEEE ComSoc's Communication & Information Security TC (CISTC) meeting will be held at IEEE ICC 2024 "Scaling the Peaks of Global Communications" on 9–13 June 2024 in Denver, CO (USA).

Every Member is more than welcome to join it and to provide your valuable contribution.