iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** II **Month of publication:** February 2023

DOI: https://doi.org/10.22214/ijraset.2023.49139

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue II Feb 2023- Available at www.ijraset.com

Data Privacy using Multi-Keyword Search and Coordinate Matching over Encrypted Documents

Dr. Gireesh Agarwal[1], I Nikhila[2], K Rajavardhan[3], A Anuj Reddy[4]
[1, 2, 3, 4]*Department of CSE Vardhaman College of Engineering* Hyderabad, India

Abstract: ***With cloud computing, businesses have access to notjust data storage but also to networks and computers. Most cloud service providers operate on a pay-as-you-go basis, which means that their customers no longer have to invest much in infrastructure to take use of their services. As these services are not hosted on the company's local network, they are far simpler to administer and implement than more conventional infrastructure options. Cloud computing's appeal stems from these aspects, explaining its ever- increasing popularity. The benefits of storing information and programs on the cloud outweigh the disadvantages. As such, it raises issues that must be addressed head-on to guarantee a safe cloud computing setting. With more and more sensitive information being stored on the cloud at present, there are growing privacy and security issues about this information. The data is encrypted before being saved in the cloud to avoid this problem. An effective search methodis also required since the quantity of data typically kept is quite large. The two most important components of cloud computing that we cover here are encryption and searching. We propose a multi-keyword search method to search through encrypted clouddata in addition to a safe and efficient encryption technique for the data and queries stored in the cloud.***
Keywords: Data, privacy, private Key, public key, multi-keyword search, coordinate matching.

I. INTRODUCTION

In the information technology (IT) industry and the aca- demic community, cloud computing is becoming an in- creasingly popular paradigm due to its accessibility, cost- effectiveness, and other desirable qualities.Users are encour- aged to store their information on a cloud server and access it from a distant location. at any time, from any location, thanks to the cloud's convenient and cost-effective capabilities. Thus,it offers several advantages to IT companies, data owners, anddata consumers. To name a few:
Data owners may obtain storage space of varying sizes on the fly to suit their changing demands.
Users may view their data remotely at any moment.
Both data creators and its consumers are released from the burden of locally keeping the data. Because of not having to buy hardware and software, you may save money.While there are certainly advantages to using cloud storage, there are also serious concerns that must be addressed. Simply said, once data is uploaded to the cloud, its original owners no longer have access to it. Since the cloud is an open platform, it is possible for both external invaders and inside attackers to get access to the data stored there (malicious). Cloud computing has huge benefits in terms of both cost and ease of adoption. Benefiting from these features, both consumers and businesses are shifting toward cloud-based data storage [1] rather of investing heavily in on-premisessolutions. While cloud computing has many clear benefits,it also raises some valid concerns. Concerns about privacy arise when sensitive information, such as people's medical data, financial details, or private photographs, is kept remotelyon the cloud. Concerns about the safety of sensitive infor- mation including corporate finances, government records, and user databases are heightened when they are outsourced by businesses [2]. Furthermore, the cloud storage provider may access this private information without proper authorization. Thus, before outsourcing, data is encrypted to ensure privacy. However, this leads to several major issues.
Industrial IT is undergoing a paradigm transition toward the delivery of computer resources through the Internet on a pay- as-you-go or subscription basis decade comes to a close. Users get a plethora of benefits from this model, including ability to supply compute resources, access a wide range of networks, pool their resources, and quickly adapt to changing demands while paying only for the resources they use. Data owners may need to encrypt sensitive data like emails, medical records, photo albums, tax documents, financial transactions, etc. before outsourcing to commercial public cloud to protect data privacy and combat unsolicited accesses; however, this renders the traditional data utilisation service based on plain text keyword search obsolete. The simple choice of down- loading all data and decrypting locally is clearly not viable in cloud-scale systems because of massive amount of bandwidth expenditure. unworkable. In addition, putting data onto the cloud serves no function other than removing the local storage management if they are not readily searchable and usable.

II. LITERATURE REVIEW

We take into account the scenario where a user U wishes to encrypt his data before uploading them to a distant file server S. When the time comes, User U needs to quickly recover a subset of the encrypted files that include (or are indexed by) certain keywords without compromising the security of the remotely stored data or revealing the keywords themselves. For instance, a user may want to encrypt their old e-mails, store them on a server operated by Yahoo or another major provider, and then access certain messages from their mobile device while abroad. In this research, we provide approaches to resolving this issue within the context of strict security guidelines. Not requiring a public-key cryptosystem allows our techniques to function quickly and efficiently. Indeed, our solution works regardless of how the distant data are encrypted. Additionally, they are incremental in the sense that additional files may be sent while still being searchable for future searches.

With searchable symmetric encryption (SSE), one may entrust someone with private storage of his data without giving up the capacity to do selective searches over it. Many researchershave been focusing on this issue, and many different security definitions and constructs have been offered. In this study, we first examine previous definitions of security before presenting our own, more robust definitions. After defining two newtypes of safe structures, we present them. Interestingly, our designs are more efficient than any earlier constructions whilestill providing higher security assurances. In addition, previousSSE research has only explored a scenario in which the data owner submits search queries. As a logical next step, we thinkof a scenario in which a group of people who aren't thesite's owner may send in search requests. Using a theoretical definition, we demonstrate a practical implementation of SSE in this multi-user environment. Using a public key encryption scheme, we investigate thechallenge of searching encrypted data. Take the case of Riya sending an encrypted email to Priya using Priya's public key. In order to properly route emails, an email gateway may check for the presence of the term "urgent" in the subject line. However, Priya would rather not provide the gateway the key to her encrypted communications. To ensure that the gateway can determine whether or not the word "urgent" is a keyword in the email without reading it, we develop and build a systemthat allows Priya to supply the gateway with a key. Public KeyEncryption with Keyword Search is the name given to this system. Take Priya's e-mail server as another example; it mayhold communications that have been publicly encrypted for her.By using our system, Priya may provide the mail server with a key that will allow it to recognize all mails containing a given phrase without gaining any other information about the messages. Here, we provide a working definition of public-keyencryption using keyword-based search, along with various implementation details.

III. METHODOLOGY

The following are some of the main design considerations behind the proposed solution.

1) Our suggested multi-keyword search strategy is designedso that users may dynamically update their documentcollection in the cloud.

2) Using a specialized tree-based index and an efficient and effective search algorithm, the multi-keyword-based encryption system developed for the cloud environment seeks to improve search efficiency.

3) The suggested approach is aimed to protect users' privacywhile relying on a semi-trusted cloud service provider.

We investigate the issue of multi keyword ranked search over encrypted cloud data for the first time, and we create a set of stringent privacy criteria for a safe cloud data usage system. In,we offer two MRSE techniques that use the notion of "coordi-nate matching" to satisfy varying degrees of privacy concern across two distinct threat models. Experimental results on a real-world dataset confirm that the suggested methods actually impose little overhead on computation and communication, and a thorough investigation is provided into the privacy and efficiency guarantees of these schemes.

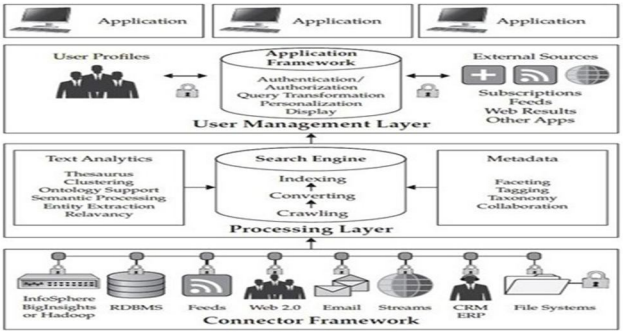


Fig.1. architectural diagram

A. Module Description

1) *Multi- Keyword Searching:* Single-keyword systems were proven to be inefficient since users had to study the whole page in order to search for a specific term, but multi-keyword search became very popular due to its applicability and the freedom it gave users. However, these models relied on the case of a single data owner and a single data user, in which only that one owner has access to the data and may outsource its management. While cloud computing has many clear benefits, it also raises some valid concerns. There are privacy problems associated with the outsourcing and storage of sensitive personal information on the cloud, such as medical data, financial accounts, and private photographs. Concerns about the safety of sensitive information including corporate finances, government records, and user databases are amplified when they are outsourced by businesses.
2) *Attribute Based Encryption:* Due to this, we use an attribute based encryption method to safeguard the files. Information is encrypted so that only people who meet the access policy criteria may read it. Data and real information storedin the cloud may be kept private via attribute based document encryption and decryption. In this work, we outline the issue of system-wide privacypreservation in cloud computing and provide a solution for multi-keyword ranked search over encrypted cloud data. While there are numerous different multi keyword semantics, we've settled on the effective approach of "coordinate matching," or finding as many matches as feasible between the search query and the data documents.
3) *Multiple Users:* Data encryption for maximum secrecy. Recent proposals for searchable encryption methods have mostly been limited to cases where one person owns and uses the encryption key. In this research, we provide a novel methodfor protecting the personal information of users while sharing data in the cloud. Attribute-based encryption and a tree-based index are at the heart of our approach. Through security and performance analysis, we demonstrate that our method is both safe and effective. We improved search performance with a tree-based index and ensured the confidentiality of outsourced data with attribute-based encryption. Security study demonstrates that our methodis private.
4) *Privacy Preserving:* In this research, we provide a novel method for protecting the personal information of users while sharing data in the cloud. Attribute-based encryption and a tree-based index are at the heart of our approach. Through security and performance analysis, we demonstrate that our method is both safe and effective. Design of a Search pattern of data user in MRSE denotesany information that can be deduced by server if it gains the knowledge that two arbitrary searches are made for the same keywords or not, as defined in related work on single keywordsearchable encryption. It would be trivial for the server to figure out the user's search behaviour simply by comparing the trapdoors it receives from different users if the trapdoors are created in a predictable fashion.
5) *Algorithm:* Using highly secure encryption techniques comes with a high computational cost, which might de- crease performance and raise cloud computing costs. Here, we present a safe and effective encryption technique for both cloud storage and queries. In order to facilitate a multi-keyword search on encryptedmaterial without decryption, we want to develop a searchable encryption system. We have combined the Term Frequency – Inverse Document Frequency (TF-IDF) model with the vector space model for index creation and query generation to provide a multi-keyword ranked search. The search performance has been improved because to the tree-based index we have built.
6) *Advantages:*

a) User benefits from this paradigm include resource shar- ing, quick flexibility with metered services, and provide computing capabilities, to name a few.
b) To safeguard data privacy and resist uninvited accesses, data owners may need to encrypt their data before out- sourcing it to commercial public cloud services. However, this makes the standard data usage service based on plain text keyword search outdated.
c) Our multi-user, encrypted data search system uses a combination of keywords to protect user anonymity.
d) For both security and speed, our scheme and design employ a tree-based index and Greedy Breadth First Search(GBFS) to ensure quick retrieval of results.

IV. RESULTS

The designing and implementing a "Data Privacy Using Multi-Keyword search and coordinate matching overencrypted documents" will be accessible to everyone who wants to secure their documents by creating their login credentials. The data owner have the authority to login into their interface and upload documents by indicating their respective files with unique keywords.

The data receiver generates a hatch for search request using their keyword and send to the cloud service provider. The cloud authority generates the unique public key every time when the data receiver tries to open the document and also a private key upon the receivers request and the key will be sent to the data receivers registered mail by the cloud service provider so that document will be accessed only by the authorized users.

Fig. 2. Data owner and receivers login page

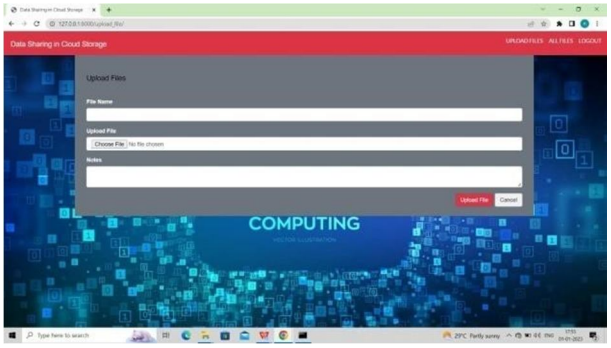


Fig. 3 Data owners dashboard

V. FUTURE SCOPE

Our secure methodology has been proven by a securityanalysis. Performance analysis is all about evaluating the efficacy of our plan. We want to enhance this effort by including real-time data updates in the near future.

Experiments on a real-world data set show that the offered approaches incur little cost in terms of both processing and communication, and a careful study explores the privacy and efficiency guarantees of the proposed systems. Additional research into privacy guarantees under a more severe threat model, integrity checks of rank orders in search results, and support for alternative multi-keyword semantics (such as weighted queries) over encrypted data are all things we wantto do in the future.

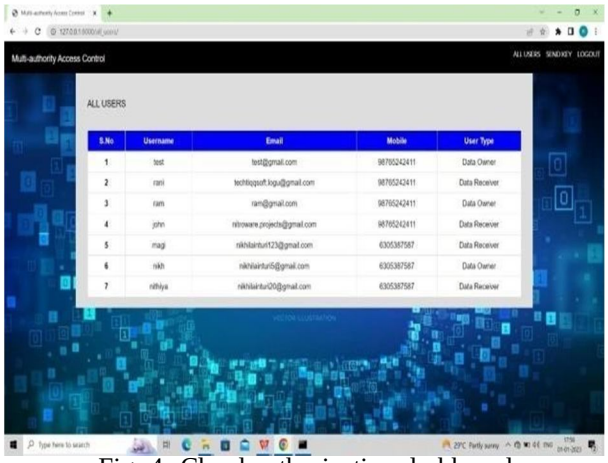


Fig. 4. Cloud authorization dashboard

VI. CONCLUSION

We provide a secure tree-based multi-keyword ranked search method over encrypted data that also allows for dy- namic action on the document collection. There are rising pri- vacy and security concerns related to the increasing amount of sensitive data being kept on the cloud at the current time. Priorto being stored in the cloud, data is encrypted. As a corollary, an efficient search engine is necessary since the volumes of data stored are often rather large. Therefore, we pay special attention to two crucial aspects of cloud computing: safetyand ease of use. We offer a secure and efficient encryption approach for cloud-based data and queries, as well as a multi- keyword search strategy for searching through encrypted clouddata.

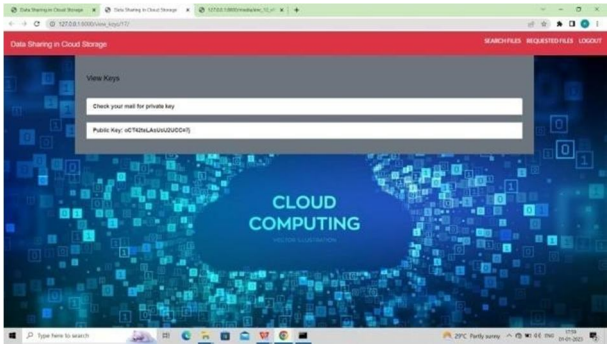


Fig. 5. data receivers dashboard (public key display)

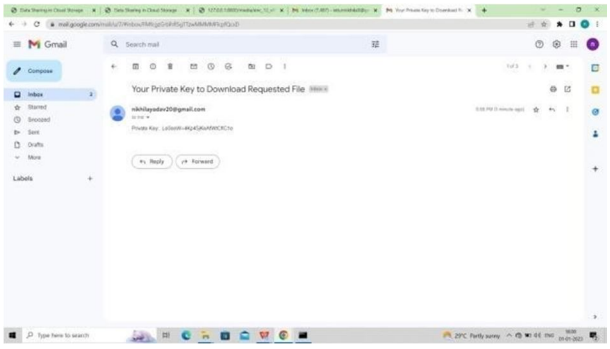
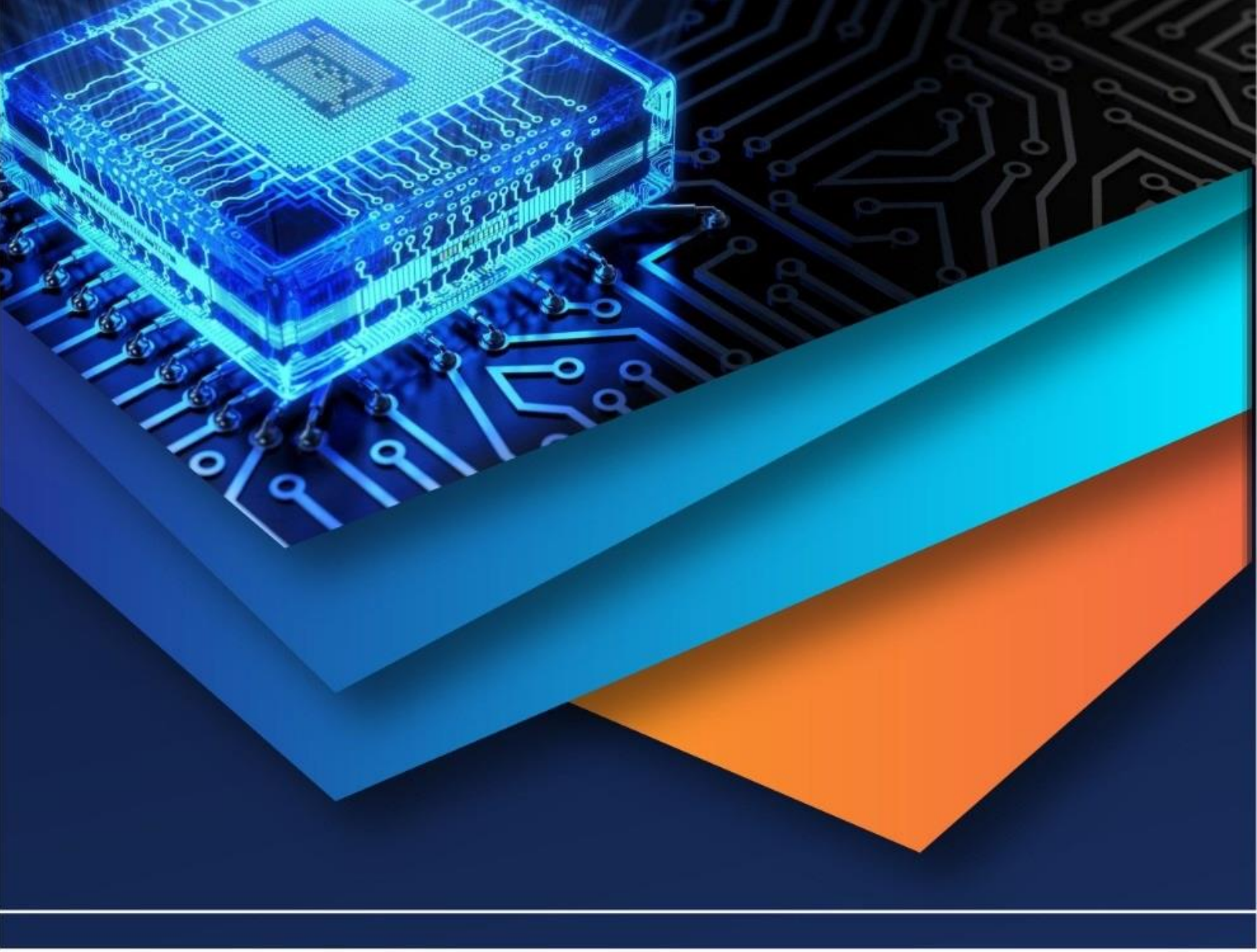


Fig. 6. Private key display in authorized receivers mail

REFERENCES

[1] K. Ren, C. Wang, and Q. Wang, “Security challenges for the public cloud,” IEEE Internet Compute., vol. 16, no. 1, pp. 6973, Jan-Feb. 2012.
[2] S. Kamara and K. Lauter, Cryptographic cloud storage, Proc. Financ. Cryptography Data Secur., 2010, pp. 136149.
[3] C. Gentry, A fully homomorphic encryption scheme, Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 2009.
[4] O. Goldreich and R. Ostrovsky, Software protection and simulation on oblivious rams, J. ACM, vol. 43, no. 3, pp. 431473, 1996.
[5] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, Public key encryption with keyword search, in Proc. Adv. Cryptol.-Eurocrypt, 2004,pp. 506522.
[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, Public key encryption that allows pir queries, in Proc. Adv. Cryptol., 2007, pp. 5067.
[7] D. X. Song, D. Wagner, and A. Perrig, Practical techniques for searches on encrypted data, in Proc. IEEE Symp. Secur. Privacy, 2000, pp. 4455.
[8] E.-J. Goh, Secure indexes, IACR Cryptol. ePrint Archive, vol. 2003, p. 216, 2003.
[9] Y.-C. Chang and M. Mitzenmacher, Privacy preserving keyword searcheson remote encrypted data, in Proc. 3rd Int. Conf. Appl. Cryptography Netw. Secur., 2005, pp. 442455.
[10] C. Wang, K. Ren, S. Yu, and K. M. R. Urs, Achieving usable and privacy-assured similarity search over outsourced cloud data, in Proc. IEEE INFOCOM, 2012, pp. 451459.
[11] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, “A Breakin the Clouds: Towards a Cloud Definition,” ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50-55, 2009.
[12] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, “LT Codes-Based Secure and Reliable Cloud Storage Service,” Proc. IEEE INFOCOM, pp. 693- 701, 2012.
[13] Mehdi Sookhak, Helen Tang, Ying He and F. Richard Yu,“Security and Privacy of Smart Cities: A Survey, Research Issues and Challenges,” 2019.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY