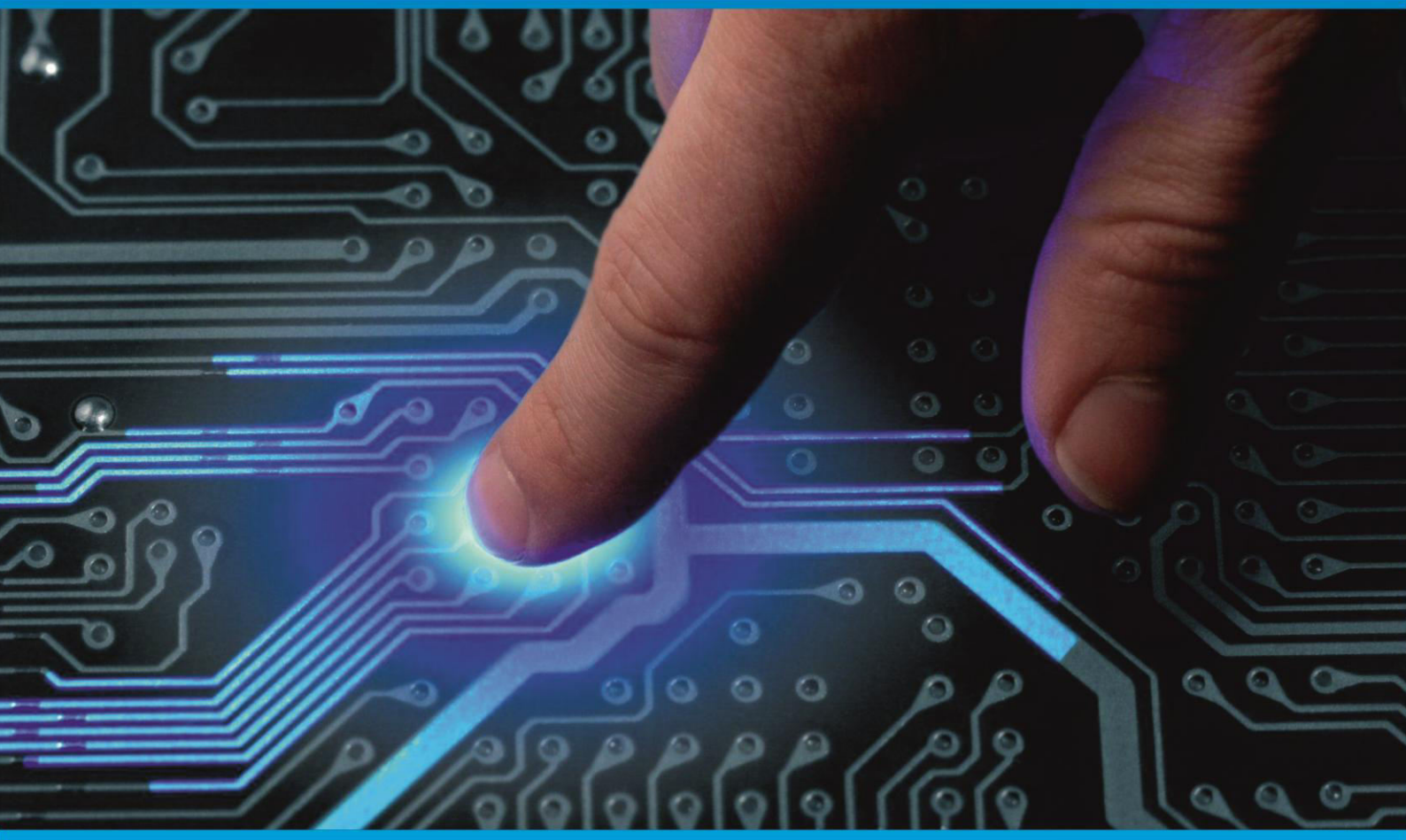
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |

Big Data Security in Cloud to Minimize Privacy Risk

Kounen Fathima[1], Neha Tabassum[1], Umaima Zareen[1], Dr. T Prem Chander[2]

B.E Students, Dept. of I.T., ISL Engineering College, Affiliated to Osmania University, Hyderabad, India [1]

Associate Professor, Dept. of I.T., ISL Engineering College, Affiliated to Osmania University, Hyderabad, India[2]

ABSTRACT: A cloud-based big data sharing system utilizes a storage facility from a cloud service provider to share data with legitimate users. In contrast to traditional solutions, cloud provider stores the shared data in the large data centers outside the trust domain of the data owner, which may trigger the problem of data confidentiality. This paper proposes a secret sharing group key management protocol (SSGK) to protect the communication process and shared data from unauthorized access. Different from the prior works, a group key is used to encrypt the shared data and a secret sharing scheme is used to distribute the group key in SSGK. The extensive security and performance analyses indicate that our protocol highly minimizes the security and privacy risks of sharing data in cloud storage and saves about 12% of storage space.

KEYWORDS: Big Data, Security and Privacy, Cloud Storage, Data Sharing

I. INTRODUCTION

The arising advances about large information like Cloud Processing, Business Intelligence, Data Mining, Modern Information Integration Engineering (IIIE) and Web of-Things have opened a new era for future Enterprise Systems (ES). Distributed computing is another processing model, in which all assets on Internet structure a cloud asset pool and can be assigned to various applications furthermore, benefits progressively. Contrasted and conventional convey framework, a lot of speculation saved furthermore, it brings excellent versatility, adaptability and efficiency for task execution. By using Cloud Computing administrations, the various venture interests in building and keeping up a supercomputing or framework figuring climate for savvy applications can be viably decreased. Regardless of these benefits, security prerequisites drastically rise while putting away close to home identifiable on cloud climate. This raise administrative consistence issues since relocate the touchy information from unify area to disseminate area. To take the benefit empowered by enormous information advances, security and protection issues should be tended to firstly.

Building security instrument for distributed storage isn't an simple errand. Since shared information on the cloud is outside the control space of authentic members, making the common information usable upon the interest of the genuine clients ought to be tackled. Furthermore, expanding number of gatherings, gadgets also, applications engaged with the cloud prompts the unstable development of quantities of passages, which makes it more difficult to take legitimate access control. Ultimately, shared information on the cloud are helpless against lost or inaccurately modified by the cloud supplier or organization aggressors. Ensuring shared information from unapproved erasure, modification and manufacture is a difficult task

Traditionally, there are two separate strategies to advance the security of sharing framework. One is access control; in which just approved client recorded in the entrance control table has the entrance advantage of the common information. The other technique is bunch key administration in which a gathering key is utilized to ensure the common information. In spite of the fact that access control makes the information just be gotten to by real members, it can't shield the assault from cloud suppliers. In the current gathering key sharing frameworks, the gathering key is by and large oversaw by an autonomous outsider. Such strategies expect that the outsider is consistently fair. Not with standing, the supposition that isn't in every case genuine particularly in the climate of distributed storage.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |

II. RELATED WORK

Numerous arrangements have been proposed to settle the security chances of cloud-based capacity. Rao proposed a safe sharing plans of individual wellbeing records in distributed computing dependent on cipherext policy Attribute based (CP-ABE) Encryption. It center on confining unapproved clients on admittance to the confidential information. Liu et al. proposed an entrance control strategy in light of CP-ABE for individual records in distributed computing as well. In and only one completely confided in focal position in the framework is liable for key administration and key age. Huang et al. presented a novel public key encryption with approved balance warrants on the entirety of its cipher text or a specified cipher text. To fortify the getting prerequisite, Wu et al. proposed an efficient and secure personality based encryption conspire with fairness test in distributed computing. Xu et al. proposed a CP-ABE utilizing bilinear blending to furnish clients with looking through capacity on cipher text and fine-grained admittance control. He et al. proposed a plan named ACPC pointed toward giving secure, efficient and fine grained information access control in P2P stockpiling cloud. As of late, Xue et al. proposed another structure, named RAAC, to take out the single-point execution bottleneck of the leaving CP-ABE based admittance control plans for public distributed storage. While these plans use character security by utilizing quality based strategies which neglect to ensure client quality security.

The latest work tending to the security issues in a cloud-based capacity is completed by Pervez et al., who proposed a protection mindful information sharing plan SAPDS. It consolidates the characteristic based encryption alongside intermediary re-encryption and mystery key refreshing ability without depending on any confided in outsider. Yet, the capacity and correspondence overhead of SAPDS is chosen by property encryption conspire.

In SSGK, an efficient arrangement is proposed to settle the secure issues of information sharing on the distributed storage without depending on any trust outsider. Past utilizing symmetric encryption calculation to encode the common information, uneven calculation and secret sharing plan is utilized to forestall the key used to decode the common information from getting by unapproved clients. Secret sharing plans were presented by both Blakley and Shamir freely in 1979 as answer for safe guarding cryptography keys. In a mysterious sharing plan, a mystery is isolated into n shares by a seller and divided between n investors. Any t offers can reproduce this mystery. Chor et al. expanded the idea of the first mystery sharing and introduced a thought of verifiable mystery sharing (VSS). The property of verifiability implies that investors can confirm whether their shares are reliable.

III. PROPOSED SYSTEM

A. *PROTOCOL MODEL*

1) DATA SHARING MODEL*:*

Consider a cloud storage data sharing system with multipleentities and the data sharing model. The protocol model consists of three types of entities: cloudprovider, data owner and group members.The cloud provider: provides a public platform for dataowners to store and share their encrypted data. The cloudprovider doesn't conduct data access control for owners. Theencrypted data can be downloaded freely by any users.

Data owner: defines the access policy and encrypts its data with a symmetric encryption algorithm using a group key. The group members who satisfied the access policy constitute a sharing group. Then secret sharing scheme is used by the owner to distribute the encryption key to the sharing group.

Group members: every group member including the data owner is assigned with a unique and a pair of keys. The group members can freely get any interested encrypted data from the public cloud. However the user can decrypt the data if and only if it gets the data decryption key from the data owner.

2) SECURITY MODEL

In SSGK, we have the following assumptions:

The data owner is totally trusted and will never be corrupted by any adversaries. Cloud provider is semi-trusted, it correctly executes the task assigned to them for profits, but they would try to find out as much secret information as possible based on the data owner's uploaded data. We now describe the security model of SSGK by listing possible attacks. The group key is distributed by running the secret sharing scheme. Parts of the group members can gather their

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |

sub secret shares to reconstruct the group key. Moreover, the communication channel of our protocol is defined as: Every pair of participants have a point-to-point channel to send messages. Additionally, all the participants access to a broadcast channel: when a participant puts a message m on this channel, all the other participants receive m. The group key is distributed on the public channel and the key may be tempered by adversaries.

B. *DEFINITIONS AND NOTATIONS:*

Definition 1 ((t,n)VSS): A verified secret sharing scheme contains four steps:
Sharing Generation Algorithm: An algorithm that, on input a security parameter K and a random polynomial f (x) of degree t 1, output n sub-shares and a verified value v; Distribution: The dealer distributes each sub-share and v to every scheme participant secretly;

Notation	Description for the Notation
O	data owner
D	shared data
K	group key used to encrypt shared data
P_i	Group member P with indentified ID_i
$cipher(x)$	Cipher-text of x
$E_k(P)$	encrypt P with key k using encryption algorithm E
$E^{-1}{}_k(C)$	decrypt C with key k using decryption algorithm E^{-1}
SK_i	Secret key of P_i
PK_i	Public key of P_i

Verify: A verification algorithm that, on input a sub-share and v, output whether the sub-share is tempered during distribution; Secret Reconstructed: For any t sub-shares, the security parameter K can be reconstructed.
Definition 2 (Equity and Availability): Verified secret sharing scheme guaranteeing equity and availability with two conditions: Any participant set in the share group, where the size of the set is less than the total quantity, the participants in the set cannot get any information about K; Only with cooperation of all the legitimate participants, K could be reconstructed.
Definition 3 (Confidentiality): Verified secret sharing scheme guarantees confidentiality if any users outside the sharing group cannot get any information of K even with the knowledge of enough interactive messages.
Definition 4 (Integrity): Once the interactive messages are tempered during VSS, any information about K could be gotten by participants. We said that verified secret sharing scheme guarantees integrity. The notations in Table 1 are used throughout the remainder of this paper.

C. *PROTOCOL DETAILS*

The scene describes as a protocol participant O wish to share data D with the legitimate participants P; i D 1; 2; n. Firstly, O generates a secret key K and uses K to encrypt D, then O stores the encrypted data cipher (D) to the cloud. Secondly, O shares K with the legitimate participants and all participants work together to certify and reconstruct K. Finally, every participant gets K and downloads cipher (D) from the cloud.

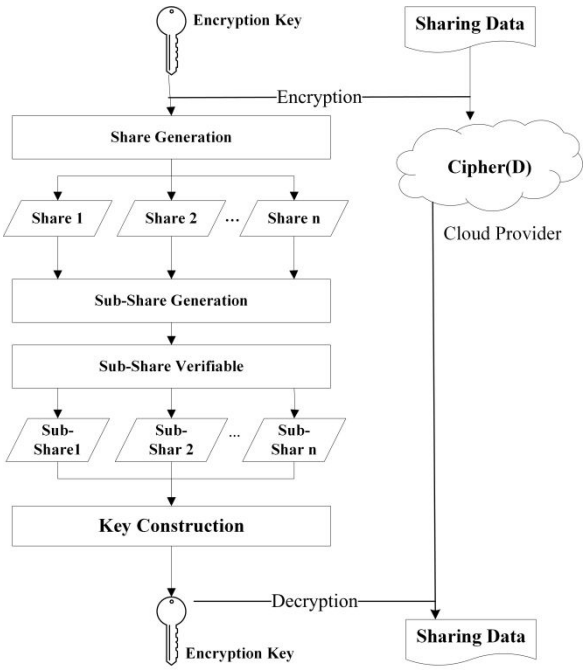
1. KEY DISTRIBUTION AND DATA SHARING

The data owner *O* creates the secret key and encrypts the data using symmetric encryption algorithm AES. Then secret sharing scheme is used by *O*to distribute the secret key. As the public channel is available for communications between every pair of participants, an asymmetric encryption algorithm RSA is used to protect the key sub-shares from known by unauthorized users.

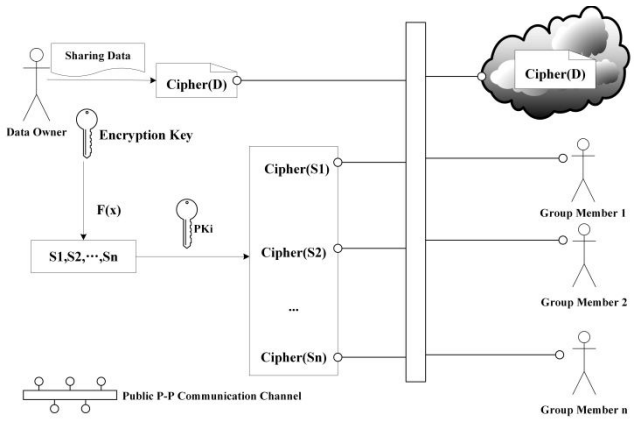
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |



Data sharing model of the proposed SSGK protocol.



Data sharing model of the proposed SSGK protocol

2. KEY RECONSTRUCTION AND VERIFICATION

All the participants may get Cipher (D) and v from the point-to-point public channel. The next goal of the participants is to reconstruct K with collaboration and to verify whether there are any corrupted participants.

IV. RESULTS

We provide the performance assessment of the proposed protocol. The following experiments focus on the storage and computation overhead of SAPDS and the proposed protocol SSGK. These experiments are running on a server with Intel core 2, 2.93GHz Dual Core processor and 2GB RAM.

1) COMPARISON ON SECURITY

This section puts forwards detailed comparison on various security and functionality features of the proposed scheme with some recently developed CP-ABE based schemes. For comparison, we consider related schemes ACPC,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

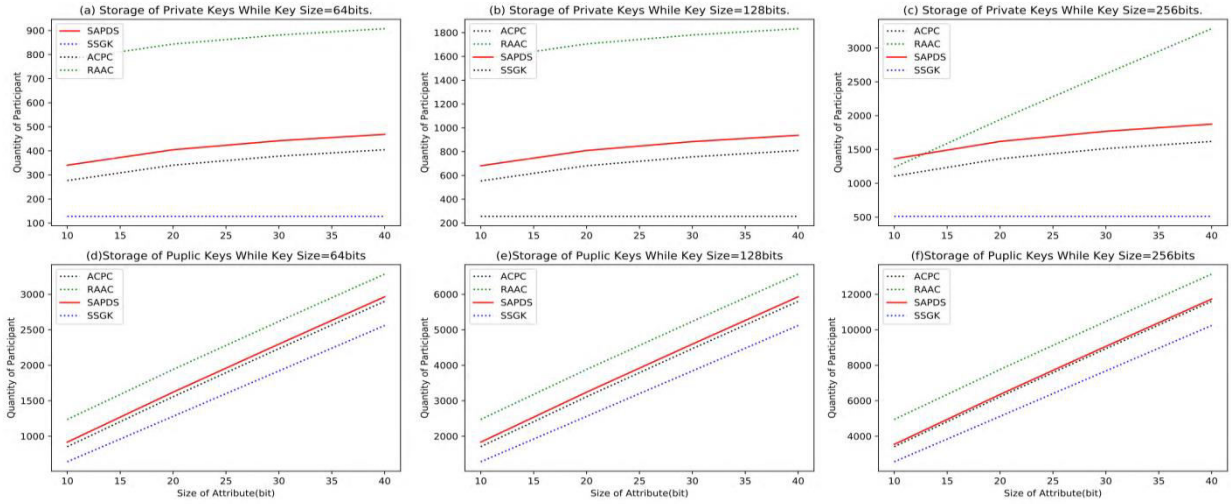
|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |

RAAC, and SAPD. It is noted that our scheme supports many useful properties, such as data equity, confidentiality and integrity protection, collusion resistance and privacy protection.

2) STORAGE OVERHEAD

The storage overhead of ACPC, RAAC, SAPDS and SSGK is tested in order to compare their scalability. The number of private and public keys of these schemes is counted. We assume that the number of the group participants is n and the key size is L bits. Private keys represent the storage consumption on one group participants in protocol. In ACPC, secret key and user attributes are used to compute the encryption key. In RAAC, multiple CAs are used for key generation, four kinds of different keys are kept by users: the symmetric algorithm key to decrypt shared data, user's secret attribute based key, user attributes and CA verified keys(Six CAs are simulated in our experiment). In SAPDS, three kinds of keys are kept to achieve fine-grained access control over the shared data: key used to decrypt shared data, users' secret attribute-based key of the access tree and user attributes. In SSGK, only secret key and sub-share are used to compute the encryption key



V. CONCLUSION AND FUTURE WORK

In this paper, we propose a novel group key management protocol for the data sharing in the cloud storage. In SSGK, we use RSA and verified secret sharing to make the data owner achieve fine-grained control over the outsourced data without relying on any third party. In addition, we give detailed analysis of possible attacks and corresponding defenses, which demonstrates that GKMP is secure under weaker assumptions. Moreover we demonstrate that our protocol exhibits less storage and computing complexity. Security mechanism in our scheme guarantees the privacy of grids data in cloud storage. Encryption secures the transmission on the public channel; verified security scheme make the grids data only accessed by authorized parties. The better performance in terms of storage and computation make our scheme more practical.

The problem of forward and backward security in group key management may require some additions to our protocol. An efficient dynamic mechanism of group members remains as future work.

REFERENCES

[1] P. Zhao, W. Yu, S. Yang, X. Yang, and J. Lin, ``On minimizing energy cost in Internet-scale systems with dynamic data," IEEE Access, vol. 5, pp. 20068_20082, 2017.
[2] D.Wu, G. Zhang, and J. Lu, ``A fuzzy preference tree-based recommender system for personalized business-to-business E-services," IEEE Trans. Fuzzy Syst., vol. 23, no. 1, pp. 29_43, Feb. 2015.
[3] X. Wu, X. Zhu, G.-Q.Wu, and W. Ding, ``Data mining with big data," IEEE Trans. Knowl. Data Eng., vol. 26, no. 1, pp. 97_107, Jan. 2014.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905171 |

[4] X. Shi, L. X. Li, L. Yang, Z. Li, and J. Y. Choi, ``Information _ow in reverse logistics: An industrial information integration study," Inf. Technol. Manage., vol. 13, no. 4, pp. 217_232, Dec. 2012.
[5] N. Bizanis and F. A. Kuipers, ``SDN and virtualization solutions for the Internet of Things: A survey," IEEE Access, vol. 4, pp. 5591_5606, May 2016.
[6] S. Li, L. Xu, X. Wang, and J. Wang, ``Integration of hybrid wireless networks in cloud services oriented enterprise information systems," Enterprise Inf. Syst., vol. 6, no. 2, pp. 165_187, Nov. 2012.
[7] K.-Y. Teng, S. A. Thekdi, and J. H. Lambert, ``Risk and safety program performance evaluation and business process modeling," IEEE Trans. Syst., Man, Cybern. A, Syst. Humans, vol. 42, no. 6, pp. 1504_1513, Nov. 2012.
[8] Z. Fu, X. Sun, S. Ji, and G. Xie, ``Towardsef_cient content-aware search over encrypted outsourced data in cloud," in Proc. 35th Annu. IEEE Int. Conf. Comput. Commun.(INFOCOM), Apr. 2016, pp. 1_9.
[9] J. Han, W. Susio, Y. Mu, and J. Hou, ``Improving privacy and security in decentralized ciphertext-policy attribute-based encryption," IEEE Trans. Inf. Forensics Security, vol. 10, no. 3, pp. 665_678, Mar. 2015.
[10] D. Zou, Y. Xiang, and G. Min, ``Privacy preserving in cloud computing environment," Secur. Commun.Netw., vol. 9, no. 15, pp. 2752_2753 Oct. 2016.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details