# Data Privacy Preserving in Social Media Platforms using Anonymous Revocable Identity-Based Broadcast Encryption

S. Shenbaha
*M.Sc., M.Tech.*
*Assistant Professor, Department of Computer Science with Data Analytics*
*Dr. N. G. P. Arts and Science College, Coimbatore.*

**Abstract-Digital security as a service is a crucial aspect as it deals with user privacy provision and secure content delivery to legitimate users. Most social media platforms utilize end-to-end encryption as a significant security feature. However, multimedia data transmission in group communication is not encrypted. One of the most important objectives for a service provider is to send the desired multimedia data/service to only legitimate subscriber. Broadcast encryption is the most appropriate cryptographic primitive solution for this problem. Therefore, this study devised a construction called *anonymous revocable identity-based broadcast encryption* that preserves the privacy of messages broadcasted and the identity of legitimate users, where even revoked users cannot extract information about the user's identity and sent data. The update key is broadcast periodically to non-revoked users, who can obtain the message using the update and decryption keys. A third-party can also revoke the users. It is proven that the proposed construction is semantically secure against IND-ID-CPA attacks and efficient in terms of computational cost and communication bandwidth.**

**Keywords: instant messaging, ARIBBE, block chain technology, VANET**

## I. INTRODUCTION

Nowadays, the use of social media has become an essential part of today's life. Previously, short messaging service (SMS) was provided by Global System for Mobile Communication (GSM) and Code Division Multiple Access (CDMA) operators. Today, several instant messaging (IM) services offer multimedia data transmission to multiple users simultaneously and facilitate communication with many participants via group chats, which offers a significant advantage over SMS. Based on the IM application and its underlying protocols, groups can be modified either by all the participating users or administrated by some selected users. These applications generate large amounts of data, which are sensitive in nature. For these applications, effective and efficient methods must be deployed to ensure data security and confidentiality of user credentials. Consider a scenario where the data owner wants to transmit multimedia data to only a subset of members from the group of users. For this, many cloud providers support a broadcast feature, which is not end-to-end encrypted. Government, healthcare services, social media networks, associations, businesses, and individuals have been gathering personal information for analysis, decision-making and other reasons for decades. For example, health records are used to monitor disease transmission to uncover the secret connection between diseases and their prevention and control. In addition to social media platforms, smart cities, the Internet of Things (IoT), cloud environments, Pay TV, and, healthcare systems have the potential for the application for proposed scheme.

Consider a scenario of identity-based broadcast encryption in smart city that uses information and communication technologies (ICT) and IoT technologies to improve the life quality of all residents. For example, to identify traffic congestion data of users can be collected by some sensors, and electricity consumption can be monitored by smart meters enabling residents to receive better services than before.

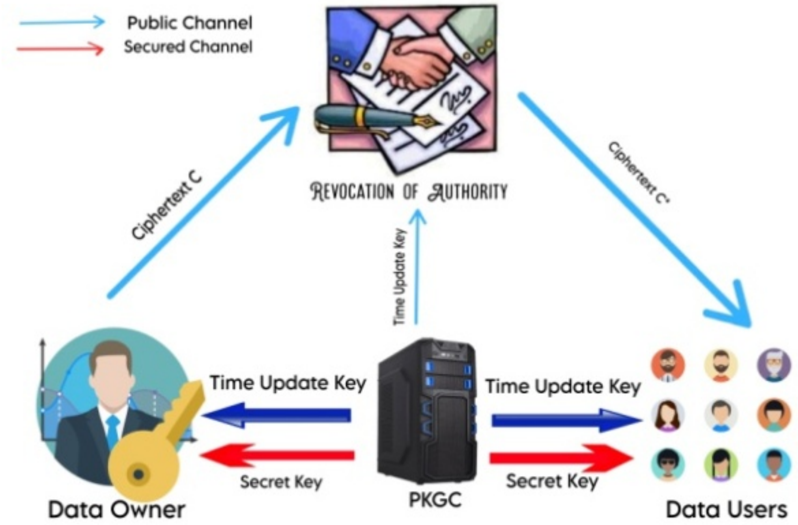


Fig 1. Architecture of Broadcast Cryptosystem

Nevertheless, most new devices, such as smart meters, sensors, traffic and surveillance cameras, traffic lights, and, cellphones are wireless, which are very easy to be attacked if communications are not properly encrypted. Another example of healthcare scenario using an anonymous identity-based broadcast encryption, is the health record sharing system for hospitals, which is supplied by a cloud service. Without losing generality, it can be assumed that the system consists of a cloud server, a data owner, and a group of doctors labelled “S”. The data owner first encrypts a data for a selected group and then stores the encrypted file in the cloud for sharing. When a doctor leaves the hospital, the server must revoke him from accessing all files. Revoked set is denoted as “R”.

If the revoked doctors are not in group S, they cannot decrypt the cipher text after the server conducts revocation. Most importantly, it requires the cloud server to be able to revoke users from a cipher text without knowing the encrypted file and identities of receivers. For this, the server must have the ability to decrypt the cipher text. When some identities should be revoked, the server first decrypts the cipher text and removes them from the original authorized user set. It then re-encrypts the data using the new authorized user set. However, in this trivial solution, the cloud server is able to learn the content and the identity of authorized users who can access the file. A cloud server must revoke the user without knowing the identity of user at a particular time period using time key. Simultaneously, the cloud server must not learn any information pertaining to data owner’s health record. Thus, this paper proposes an anonymous revocable identity-based broadcast cryptosystem (ARIBBE) that overcomes these challenges through the following:

- The proposed scheme preserves the privacy of legitimate subscribers.
- The primary objective of the proposal method is the retrieval of the function without learning the revocation identity set with cipher text evolution and protection of data privacy despite revoked user collisions.
- This paper presents a new cryptosystem with cipher text evolution over cloud, which is appropriate for fine-grained data access control, protects legitimate receiver’s identity, and allows a third party to revoke select receivers.
- The construction is semantically secure under BDHP computational assumption and random oracle model.
- The computation cost in the revocation phase is linear to the number of revoked users and revocation is performed using time key.

## II. LITERATURE REVIEW

The concept of broadcast encryption (BE) was first proposed by Fiat and Naor, however, Bone et al. constructed a broadcast encryption scheme with smaller decryption, encryption key size, shorter encrypted message size, and better computational cost. BE is a cryptographic primitive which provides a solution to the problem of communicating encrypted messages to only legitimate set of users, “S”, over an insecure channel. Only legitimate users from Scan decrypt the encrypted message. In contrast, revoked users of S would not learn anything about the message. Users who obtain access to the cipher text are called legitimate subscribers (member of the set S) and unsubscribed users (non-member of the set S) are called revoked users. The broadcast encryption algorithm works on the demarcation of revoked and legitimate users and this partition can vary for each broadcasted message. In such a scenario, only the broadcaster or the sender acts as the source of a message. The sender shares a common session key with all the subscribed users.

To broadcast an encrypted message to all users in S, message is encrypted using the session key and then, to decrypt it, the legitimate users need the session key, their own secret key, and S to identify all the receivers. In the asymmetric key setting, the broadcast encryption uses public key framework. All users of the set S has a pair of keys, one for encryption function and another for decryption function. The broadcaster and other possible other entities can act as a source of a message, while, only legitimate subscribers or receivers can decrypt and learn the actual message. It also resolves the problem of refreshing the secret keys after any update in the set of legitimate subscribers for symmetric key setting. BE in the public key setting is well studied and can be classified as Fig 1: Identity-based broadcast encryption, attribute-based broadcast encryption, anonymous broadcast encryption, hierarchical broadcast encryption, dynamic broadcast encryption, and distributed broadcast encryption. It has several applications such as in secure email system, digital rights management system, pay TV, database security system, online social network system etc.

Identity-based broadcast encryption (IBBE) was first introduced by C. Delegable, which is an extension of identity-based encryption scheme in public key setting. Instead of public keys each legitimate subscriber is identified by their Identity, such as email-id, passport number, or driving license number (arbitrary strings, alphanumeric values, and numerals) etc., are used as encryption keys. It is a practical cryptographic primitive that allows an exponential number of recipients to exchange messages securely. This implies that the public parameter is not correlated by any means to the decryption key of recipients.

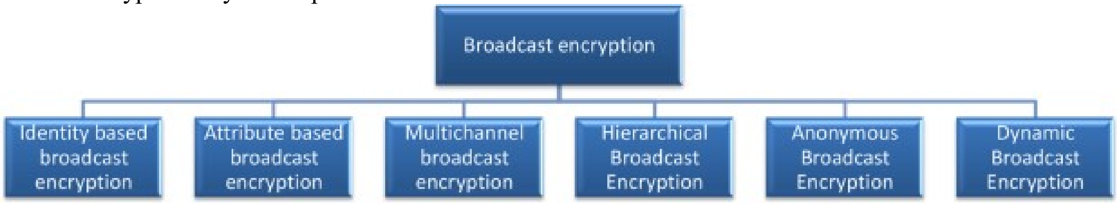

Broadcast encryption
Identity based broadcast encryption
Attribute based broadcast encryption
Multichannel broadcast encryption
Hierarchical Broadcast Encryption
Anonymous Broadcast Encryption
Dynamic Broadcast Encryption

Fig 2. Online e-healthcare system - data related to patients is collected and uploaded to the centralized storage server

Due to high prevalence of IoT technology applications and block chain technology (BCT), security issues such as identity authenticity and data privacy are becoming increasingly important concern. This scheme achieves a proper tradeoff between security and performance, compared with other schemes of IoT but the only drawback is, it doesn't provide an anonymous authentication. Block chain along with attribute based searchable encryption offers decentralized and computationally efficient construction. For e-health proposed secure and energy-efficient IoT model. It enables secure transmission and retrieval of biomedical images over IoT networks.

Various online social networks (e.g. What Sapp, Twitter, Instagram, and Facebook) make the distribution of user's real time data between multiple users over the same and different networks very easy. Many researchers have extensively analyzed the impact of social media on information sharing. However, these schemes does not ensure the anonymity of the receivers. To address this problem has come up with the anonymous broadcast encryption in public key setting and the issue of anonymity has been studied extensively in schemes.

Many applications e.g., vehicular ad-hoc network(VANET) use computationally efficient privacy preserving anonymous authentication scheme based on the use of anonymous certificates and signatures which is an important component of IoT. This scheme is efficient in terms of certificate and signature verification cost and providing anonymity. VANET entities become anonymous to each other until they are revoked from the VANET system. In the e-healthcare domain, data privacy and security of electronic health records are the most prominent challenges with cryptographic primitives playing a vital role in providing privacy and secure access. None of these schemes are able to achieve anonymity and revocation simultaneously with respect to the time key. This paper attempt to solve this problem.

### *2.1 SYSTEM MODEL*

Consider an online e-healthcare system, such as that presented in Fig 2, where the data related to patients is collected and uploaded to the centralized storage server, for example, a cloud server. The patient's data must be secure and the privacy (identity) of the patient must be preserved. If this is not taken seriously, the patients may suffer the consequences of having their medical records leaked online. Recently millions of user's data has been compromised. In online e-healthcare system, a patient acts as data owner and may they choose to share their medical data with various medical professionals or related personnel, such as

1. To various doctors for taking opinion for his medical case.
2. government offices for providing information as they are working in that office many a times, it is necessary to share medical records of the employee with employer.
3. And insurance agencies for medical-claim disbursement purposes.

In other words, if a data owner chooses to share their medical records to various receivers where no receiver is able to learn other receivers' identity. If some receivers are revoked then they cannot learn any message by combining their keys. This revocation list is based on time update keys. Most of the time, online healthcare data resides in shared environments, thus, ensuring that the data is shared and accessed in a secure manner on the cloud and access is a non-trivial task.

One way to share data among group of legitimate subscribers is identity-based broadcast encryption. The privacy of the recipients is an important issue to be addressed in broadcast encryption schemes. Encrypting the message again for the newly formed subscriber set after the revoked user is a trivial but impractical solution. The notion of Recipient revocable identity-based broadcast encryption (R-IBBE) provides efficient solution to this problem.

#### *2.1.1 DESIGN GOALS*

- No probabilistic polynomial-time (PPT) adversary is able to recover plaintext from cipher text after revocation. This scheme is aimed at being secure against chosen plaintext attacks.
- The scheme is collusion resistant. More particularly, if the maximum receivers is set as one, the resultant scheme is an anonymous revocable IBE scheme with timestamp revocation.
- The receiver set is anonymous i.e., the identities of the receivers are hidden from outside world.
- It should be difficult for the cloud server to retrieve the identities of receivers from C*.
- The computational cost of decryption is independent of total number of receivers.
- The public parameters in the proposed scheme is linear in the maximum size of the privileged identity set.

### *2.2 DESIGN ISSUES OF THE IBBE SCHEMES*

There are various essential design issues for the construction of IBBE cryptographic primitives. These are briefly discussed below.

#### *2.2.1 SECURITY MODEL*

In the basic security model of IBBE schemes, an adversary is allowed to obtain the secret keys $S_k$ for a specific set, S, of identities of k-subscribers. The security model also allows adversary to obtain secret keys corresponding to an identity ID $\subseteq$ S. Based on the targeted identity set by an adversary, two types of security notions emerge, namely: selective security and adaptive security. Selective security is a weaker security notion that allows an adversary to specify a target recipient set before learning about the public parameters, however, here the adversary is restricted in raising encryption queries.

Whereas, in the adaptive security, a stronger security notion, allows an adversary to specify target recipient set adaptively. Semantic security and chosen cipher text security notions are defined as stronger security notions and the schemes based on these security notions are highly desirable.

#### **2***.2.2 COMPUTATIONAL COMPLEXITY ASSUMPTION*

Security of cryptographic primitives are based on computationally hard problems. There are several standard computation problems based on which most of the public-key cryptosystems are constructed. Bilinear pairings or mappings are employed in most of the public key broadcast encryption schemes. Moreover, lattice-based or code-based schemes can also serve as a candidate that can resist attacks using quantum computers.

#### *2.2.3 PAIRING CHOICE*

There exist many pairing based efficient construction of IBBE schemes. Pairing based construction requires a function are additive groups and it is a multiplicative group for some large prime q.

All types of pairing have been well studied in literature [34, 35] with the consensus being that Type-III is more suitable to use because it provides designs based on this are most efficient, secure, and have more compact parameter sizes. In addition to bi-linear pairing, lattice is also one of the most powerful tools for constructing post quantum cryptographic primitives.

#### *2.2.4 ADDITIONAL PROPERTIES*

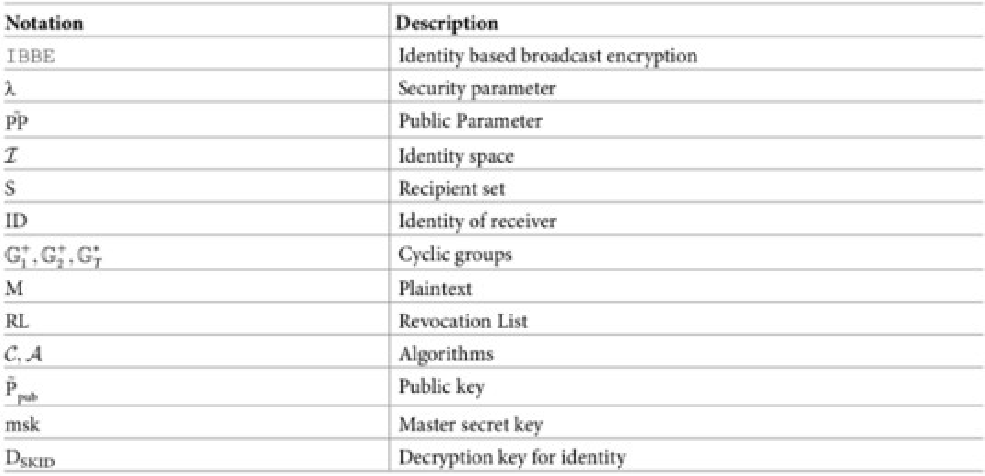
Several efficient algorithms of IBBE have been constructed based on various complexity assumptions, security models, compact size of header, and keys. Apart from all these parameters, some additional properties are also required.

1. **Anonymity property**: This states that the adversary cannot obtain set of receivers from the ciphertext. Leakage of the recipient set reveals the subscriber's information which is a major security concern and may cause a personal attack to receivers such as trolling, bullying, etc. Previous schemes have focused on the confidentiality of the message while the recipient set was openly known to adversaries. Preserving privacy is a serious concern in designing cryptographic algorithms, and therefore, various broadcast encryption schemes consider this issue such as Anonymous-IBBE [21], Private broadcast encryption [4], Attribute-based broadcast encryption (hidden policy) [22, 23], outsider anonymous BE [36], lattice-based BE [24].

2. **Revocation**: This is the intrinsic property of basic broadcast encryption techniques. Consider a scenario where the broadcast encryption setup algorithm and key generation algorithm generate all the parameters as per the construction. The system has reported some malicious receiver or somehow the legitimate receiver's decryption key is leaked. Therefore, in these circumstances, a receiver needs to be revoked. Schemes with revocation property are preferred in the construction of systems.

## III. PRELIMINARIES

This section first describes a relevant IBBE cryptosystem followed by a revocable identity-based broadcast encryption scheme. Secondly, the basic concepts on bilinear pairing and decisional BDHE assumption are also presented. **The notations and acronym used in the paper are described in Table 1.**



| Notation | Description |
|---|---|
| IBBE | Identity based broadcast encryption |
| $\lambda$ | Security parameter |
| $\tilde{P}P$ | Public Parameter |
| $\mathcal{I}$ | Identity space |
| S | Recipient set |
| ID | Identity of receiver |
| $G_1^+, G_2^+, G_T^+$ | Cyclic groups |
| M | Plaintext |
| RL | Revocation List |
| $\mathcal{C}, \mathcal{A}$ | Algorithms |
| $\tilde{P}_{pub}$ | Public key |
| msk | Master secret key |
| $D_{SKID}$ | Decryption key for identity |

### **3**.*1 IDENTITY-BASED BROADCAST ENCRYPTION SCHEME*

The notion of IBBE construction as presented in [5] is described here. An identity-based broadcast encryption method consists of ensemble of three probabilistic algorithms (IBBE.setup, IBBE.extract, IBBE.Enc) and one deterministic algorithm (IBBE.Dec).

### 3.2 RE*VOCABLE IDENTITY-BASED BROADCAST SCHEME*

An extension of identity-based broadcast cryptosystem that allows subscribed receivers to be revoked is called revocable identity-based broadcast encryption R-IBBE [37]. This extension facilitates a legitimate subscriber with an ID to be revoked if their credentials are expired or leaked. In R-IBBE, each user of a recipient set obtains a decryption key from the private key generation center/authority (PKGC) that is related with the user's ID. Once the system is configured, the key generation authority periodically updates revoked recipient set RL with respect to time T and then broadcasts the update key for the remaining non-revoked recipients. The generation of the update key depends on RL and T.

If a legitimate user does not revoke at time T when the update key has been issued then the user can generate their own decryption key corresponding to ID. With the use of decryption key corresponding to their ID and time T, the legitimate recipient is able to decrypt a ciphertext for receiver $ID_P$ and time $T_c$ only if $ID = ID_P$ and $T = T_c$ holds.

## IV. PROPOSED SCHEME

The proposed construction considers the idea of revocation of subscribed receivers used in the IBE scheme of and the ID-based broadcast encryption scheme.

The anonymous revocable identity-based broadcast encryption (ARRIBE) scheme is associated with message space $\mathcal{M}$, identity space $\mathcal{I}$, and time space $\mathcal{T}$. The protocol is an ensemble of probabilistic algorithms, namely, ARIBBE.setup, ARIBBE.Genkey, ARIBBE.Timekey, ARIBBE.encrypt, ARIBBE.decrypt, and ARRIBE.revoke.

Rev is the revocability and m denotes the total number of subscribed users in the system. The scheme does not use pairing for constructing the scheme. We have obtained this for comparing the anonymity and revocation property. Some schemes are anonymous but not revocable, while this scheme ensures both. Further, only proposed provides timestamp key based revocation. Our scheme is highly efficient in terms of computational cost and communication bandwidth.

## V. CONCLUSION

In this paper, a new technique called ARIBBE cryptosystem based on a public key framework using Type-I bi-linear map was proposed. The privacy of user's content identity is one of the primary concerns in data sharing. Hence, this paper first proposed a privacy-preserving (anonymous) revocable ID-based broadcast cryptosystem with timestamp option that facilitates broadcasters to transmit encrypted data to legitimate group participants so that revoked users will not learn anything if they all collide with each other. The proposed construction also provides an access control method in online social networks that offers one-to-one and one-to-many encrypted communication. The scheme also offers a data access control method that permits a third party to revoke any recipient identity without learning the data contents and legitimate user identities. The result indicate that the proposed scheme is extremely efficient in terms of computational cost and communication bandwidth as well as secure under CPA attack, with the ciphertext size being independent of the number of receiver identities. The proposed cryptosystem could be deployed in OSN services for distributing information to provide data access control. Security proofs show that proposed security requirements are met. However, construction of a scheme with the same parameters but without pairing is left as an open problem.

## REFERENCES

[1] Tewari A, Gupta BB. Secure timestamp-based mutual authentication protocol for iot devices using rfid tags. International Journal on Semantic Web and Information Systems (IJSWIS). 2020;16(3):20–34.

[2] Fan Q, Chen J, Deborah LJ, Luo M. A secure and efficient authentication and data sharing scheme for Internet of Things based on blockchain. Journal of Systems Architecture. 2021; 117:102112.

[3] Kumari K, Singh JP, Dwivedi YK, Rana NP. Towards Cyberbullying-free social media in smart cities: a unified multi-modal approach. Soft Computing. 2020;24(15):11059–11070.

[4] Rosler, Paul, Christian Mainka JS. On the End-to-End Security of Group Chats in Instant Messaging Protocols. Proceedings of 3rd IEEE European Symposium on Security and Privacy (EuroS&P 2018). 2018.

[5] Fiat A, Naor M. Broadcast encryption. In: Annual International Cryptology Conference. Springer; 1993. p. 480–491.

[6] Barth A, Boneh D, Waters B. Privacy in Encrypted Content Distribution Using Private Broadcast Encryption. In: Di Crescenzo G, Rubin A, editors. Financial Cryptography and Data Security. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006. p. 52–64.

[7] Delerablée C. Identity-based Broadcast Encryption with Constant Size Ciphertexts and Private Keys. In: Proceedings of the Advances in Crypotology 13th International Conference on Theory and Application of Cryptology and Information Security. ASIACRYPT'07. Berlin, Heidelberg: Springer-Verlag; 2007. p. 200–215. Available from: http://dl.acm.org/citation.cfm?id=1781454.1781471.

[8] Gupta BB, Li KC, Leung VC, Psannis KE, Yamaguchi S, et al. Blockchain-assisted secure fine-grained searchable encryption for a cloud-based healthcare cyber-physical system. IEEE/CAA Journal of AutomaticaSinica. 2021;8(12):1877–1890.

[9] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-TSeries Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372,Dec.2012.

[10] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PIDUsing State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[11] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with FuzzyLogic Controller Using State Space Techniques'- Taylor &amp; Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis

[12] and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools &amp; Technology. 2022; 12(2): 34–41p.

[13] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.

[14] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756

[15] Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical &amp; Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[16] Kaur M, Singh D, Kumar V, Gupta BB, Abd El-Latif AA. Secure and energy efficient-based E-health care framework for green internet of things. IEEE Transactions on Green Communications and Networking. 2021;5(3):1223–1231.

[17] Acharya K, Dutta R. Enhanced Outsider-anonymous Broadcast Encryption with Subset Difference Revocation. IACR Cryptology ePrint Archive. 2017;2017:265.

[18] Lai J, Mu Y, Guo F, Susilo W, Chen R. Fully privacy-preserving and revocable ID-based broadcast encryption for data access control in smart city. Personal and Ubiquitous Computing. 2017;21(5):855–868.

[19] Boldyreva A, Goyal V, Kumart V. Identity-based encryption with efficient revocation. Proceedings of the ACM Conference on Computer and Communications Security. 2008; p. 417–426.

[20] Li X, Yanli R. Efficient Anonymous Identity-Based Broadcast Encryption Without Random Oracles. Int J Digit Crime For. 2014;6(2):40–51.

[21] Ge A, Wei P. Identity-based broadcast encryption with efficient revocation. In: IACR International Workshop on Public Key Cryptography. Springer; 2019. p. 405–435.

[22] Gentry C, Waters B. Adaptive security in broadcast encryption systems (with short ciphertexts). In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer; 2009. p. 171–188.

[23] Gentry C, Halevi S. Hierarchical identity based encryption with polynomially many levels. In: Theory of Cryptography Conference. Springer; 2009. p. 437–456.

[24] Sakai R, Furukawa J. Identity-Based Broadcast Encryption. IACR CryptolePrint Arch. 2007;2007:217.

[25] Hur J, Park C, Hwang SO. Privacy-preserving identity-based broadcast encryption. Information Fusion. 2012;13(4):296–303.

[26] Lai J, Mu Y, Guo F, Susilo W, Chen R. Anonymous identity-based broadcast encryption with revocation for file sharing. In: Australasian Conference on Information Security and Privacy. Springer; 2016. p. 223–239.

[27] Susilo W, Chen R, Guo F, Yang G, Mu Y, Chow YW. Recipient revocable identity-based broadcast encryption: How to revoke some recipients in IBBE without knowledge of the plaintext. In: Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security; 2016. p. 201–210.

[28] Maiti S, Misra S. P2B: Privacy preserving identity-based broadcast proxy reencryption. IEEE Transactions on Vehicular Technology. 2020;69(5):5610–5617.