# SECURED HEALTH PROTECTIVE SERVICES IN SOCIAL MEDIA NETWORK USING K-ANONYMITY

**N.A.S.Vinoth, M.Yohapriyaa, K.Janani, V.Vijaypriya**

***Abstract:Healthcare media provides a capable model in order to fascinate consumer to discuss their health information and access health protection facilities from connected caretakers. Since the traditional health care service is time consuming process, we are switching to online health care practice, which can reduce the gap between care takers and patients. Due to the frankness of the social network the trust between the patients and care takers is a challenging issue and there is a chance of revealing personal information of the patients. Here, we intended to propose a reliable health protection service to facilitate user in social media networks, we deploy Bloom Filter technique for suitable personalized caretakers, in order to assure trustworthy rankings and critiques for caretakers, we include Sybil attack detection technique to identify users' fake rankings and critiques using various false name. It incorporates generalization and suppression techniques to protect individual's private data. For this purpose, k-anonymity Technique is implemented to anonymize the data.***

***KEYWORDS:Sybil Attack, Bloom Filter, K-Anonymity, Linkage Attack, Local Recoding.***

## I.INTRODUCTION:

Health care systems must provide trusted information between patient and caretaker. People post about everything online, including their health. Hash tags can reveal when diseases are popping up in new locations. The newly joined patients can easily able to identify or filter the caretakers based on the rating which was given by the other patients. When the rating is high then the caretaker will have high reputation. Due to the false ratings and reviews, Existing system faces challenges in guarantee the trusted service for the patients. The result may be counterfeit by sibyl attacks [1]. It may happen in subsequent ways. 1) Critiques may be false and negative when the user gives several critiques towards a caretaker 2) Caretaker mislead the user to give several optimistic critiques. The status of caretakers becomes unfair, confusing the users. The status of the caretaker will not assure the facilities which are suitable for a particular person in sick.

Here, we incorporate a modified and reliable health protection system to ensure reliable health protection with privacy. Our system comprises the following; we design a framework for healthcare service. The user can provide health related data to others having related problems, and get suitable suggestions from caretakers we suggest a privacy model for finding relationship among users with the help of bloom filter. The user's health related problems are mapped into vectors, to find out the relationship among users.

To prevent, our system incorporates a Sybil discovery method for defending against fake rankings and critiques from fake users by validating the signatures of various false name with same critiques and rankings, this identifies the unwanted activities, and inform the authority for identifying the Sybil attackers. By this, it assures genuine caretaker's data for the patients.

Recent trends in technology lead to an enormous raise in the amount of data generated. This big data is used for analysis by Businesses, Healthcare Organizations, Government, etc. the healthcare organization also contains user-specific information, releasing this data publicly becomes serious threats to user's privacy. Linking attack is one of the main drawbacks in existing system; Identifiers like Name, location is removed. Even though some of the attributes includes quasi identifiers which results in obtaining the information about the patient easily.

The remaining part of the paper is structured as follows. Segment II evaluates the Sybil attack and providing reliability in health protection, Data Anonymity. Segment III presents prefaces. Then, proposed methodology in Segment IV, followed by a conclusion in Segment V.

## II.RELATED WORKS:

### Trust estimation of Social media Health protection:

Online health protection attracts several researches works [3]– [4]. Parham et al. [2] propose a reliable system to calculate rankings based upon the grouping of various similar data and the confidential information. Xu et al. [5] Study conviction result by evolutionary game theory, and suggested a method to enhance shared systems. Jiang et al. [6] suggest a typical model and which assures accomplishment and system lifespan.

### Providing confidentiality in online Health protection

User's health information has wide separate confidential information, providing confidentiality is essential in online health protection system. Li et al. [7] introduces a new diverse specified scheme, aims at inter-related health protection system. This method incorporates the system table configuration to extensively decrease the scope of contender set, then exploits user information to recognize the right mapping users with a high assurance.

### Data Anonymity

Anonymity in data got significant awareness, because of the necessity of

**Revised Manuscript Received on July 05, 2019.**

**N.A.S.Vinoth[1], M.Yohapriyaa[2], K.Janani[3] , V.Vijaypriya[4]**
[1]AssistantProfessor, Department of Software Engineering, SRMIST, Kattankulathur.
[2,3,4]Teaching Associate, Department of Software Engineering, SRMIST, Kattankulathur.

several organizations to publish data, by excluding the identity of specific data. Even though the recognizing attributes (e.g., name) are eliminated, it is easy for an attacker to correlate archives per persons by means of other attributes (e.g., date of birth, Pin code, sex), called quasi identifiers (QI). A table is k-anonymized if various data were impossible to differentiate from at least k−1 other records. The methods used to provide privacy are Suppression and Generalization [5]. Suppression eliminates some QI attributes or whole records from the micro data. Generalization substitutes their actual QI with universal values.

## III.PRELIMINARIES

### Linearity based Grouping

In linearity coupling viz. Weil coupling and Tate coupling of numerical curvature is a definite plot E : g1 × g1 → gt , where g1 is recurrent additive set, produced by g, it has an order - prime P, and gt is a recurrent additive group, it has an order Q. Discrete logarithm problems (DLP) in both g1 and gt are rigid. linearity pairings have the following properties: Linearity: for any X, Y ∈ g1 , and C, D ∈zP , it has e(XA , YB ) = E(X, Y)AB, e(G,G) = 1, 1 is the unit constraint in gt. Computability: for all U, V ∈ g1, has an effective algorithm to calculate e (U, V).

**K-anonymity**:

This methodology prevents Linking Attacks. With respect to k-anonymity rule, a single data in the available data are impossible to differentiate from k-1 other data in that particular data set. Thus, an attacker who knows the values of quasi-identifier attributes of an individual are unable to differentiate the data from the k-1 other data's [9]. Generalization and suppression techniques are used in order to hide the identity of a particular person.

## IV.PROPOSED METHODOLOGY

### A. Architecture Model

Our architecture model consists of five entities Trustworthy server, Caretaker, Patients, Service database and Attacker. Trustworthy server provides general registration for the patients and the caretakers based on their identity, and validate the system database. Caretaker's employments are full-time as well as part-time until they are eligible as a legitimate health protection facility sources by the authority. Caretaker gives health protection guidance and obtains critiques and rankings to show their reputations and to maintain their standards. User can find other user with related body problems, and contact expert health protection facilities from proper and reliable caretakers. User can use several nick names to converse with other patients, as well as critiques the care takers after receiving related health protection guidance. The database may be an unreliable server, it incorporates the user to find related user and provide appropriate caretakers having high rankings. In the meanwhile, the system database checks whether the various critiques for a specific doctor are given by the annoying users to discover Sybil attacks, and sends back the details to the trustworthy server in order to expose the original identity of the annoying user. The attacker can hack the sensitive information of the patients. The service database provides access control to the valid users in the system to secure the private information from unauthorized users.

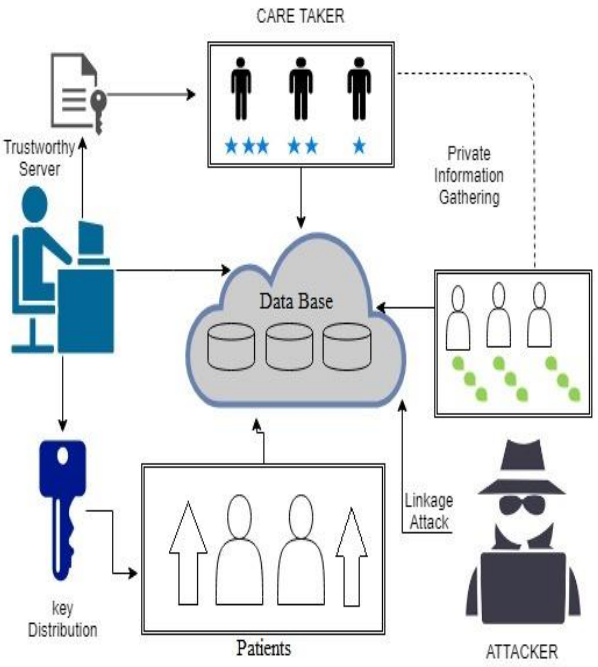


**Fig a. Architecture Model**

### B. Security Model

This model prevents Linking Attacks. Linking attack is one of the main drawbacks in existing system; Identifiers like User's Name are removed. Even though the recognizing attributes (e.g., name) are eliminated, it is easy for an attacker to correlate archives per persons by means of other attributes (e.g., date of birth, Pin code, sex), called quasi identifiers (QI). So, it is impossible for an attacker to distinguish the data [9]. The methods used to provide privacy are Suppression and Generalization [5]

| Name | Age | Sex | Pin code | Disease |
|---|---|---|---|---|
| Arun | 25 | Male | 626100 | Small pox |
| Seetha | 26 | Female | 624101 | Jaundice |
| vijay | 24 | Male | 520101 | Chicken Pox |
| Siva | 25 | Male | 520100 | Dengue |

Fig b: Original Dataset

| Age | Sex | Pin code | Disease |
|---|---|---|---|

| | | | |
|---|---|---|---|
| 25 | Male | 626100 | Small pox |
| 26 | Female | 624101 | Jaundice |
| 24 | Male | 520101 | Chicken Pox |
| 25 | Male | 520100 | Dengue |

Fig c: Dataset After Eliminating Identifiers

| Name | Age | Sex | Pin code |
|---|---|---|---|
| Arun | 25 | Male | 626100 |
| Seetha | 26 | Female | 624101 |
| Vijay | 24 | Male | 520101 |
| Siva | 25 | Male | 520100 |

Fig d: Publicly Available Dataset

On relating fig c by fig d, the attacker came to identify that Arun is suffered by Small pox. Even though the identifiers are eliminated, the patient can easily identify through publicly accessible data. Relating the unconfined table with publicly accessible table is called as Linking Attack.

**C. K-ANONYMITY ALGORITHM**

This algorithm is used to avoid Linking Attacks. In Existing system Private information's which are anonymized using global recoding technique generalizes or suppresses all the attributes equally for all the entries. A value of an attribute generalizes to another value for all of its occurrences. For e.g., a PIN Code value 626101 will be generalized to 626*** for all of its occurrences. The disadvantage of using this technique is that it results in more information loss.

According to k-anonymity rule, a data in the issued data set is indistinguishable [10]. For e.g., fig e represents that technique, i.e., Datasets which are anonymized using local recoding technique suppresses attributes on per cell basis. In other words, local recoding map individual data values to generalized values. The advantage of Local Recoding is that it results in less information loss. Local Recoding has better utility than Global Recoding [10].

| Age | Sex | Pin code | Disease |
|---|---|---|---|
| [20-35] | Male | 62**** | Small pox |
| [20-35] | Female | 62**** | Jaundice |
| [20-35] | Male | 520*** | Chicken Pox |
| [20-35] | Male | 520*** | Dengue |

Fig e: Anonymous Table

**D. ANONYMIZATION TECHNIQUE**

The main motive of this technique is to discover the change in the generalization hierarchy which protects the private information of the patients.

**e. Local recoding**

Local recoding is implemented based on generalizing existing values; The Substitution should be partial; only some occurrences of the tuples are replaced with another tuple. This can possibly decrease the alterations in information, by substituting strictly a neighborhood data. The discovery of an appropriate neighborhood is challenging to accomplish the task and is rigid to discover information with associated values.

Local recoding relates individual patient data values to generalized values. The advantage is that it results in less information loss. The following tables illustrate the concept of local recoding.

| Age | Sex | Pin code | Disease |
|---|---|---|---|
| 25 | Male | 626100 | Small pox |
| 26 | Female | 624101 | Jaundice |
| 24 | Male | 520101 | Chicken Pox |
| 25 | Male | 520100 | Dengue |

Fig f: Table for Demonstrating Local Recoding

| Age | Sex | Pin code | Disease |
|---|---|---|---|
| [20-35] | M | [626000-629100] | Small pox |
| [20-35] | * | [624000-626100] | Jaundice |
| [20-35] | M | [520000-526100] | Chicken Pox |
| [20-35] | * | [520000-626100] | Dengue |

Fig g: Result of Local Recoding on fig f

Fig f shows a dataset whereas fig g shows the result of local recoding on the dataset shown in fig f that it remains useful for research and analysis. A good generalization should focus on preserving data for future use while achieving k-anonymity [10].

**V.CONCLUSION & FUTURE WORK**

This paper incorporates a reliable health Protection system to ensure reliability in social media health protection, which improves the reliability between the user having health related problems and caretakers with higher rankings and critiques with confidentiality. Our system protects the information about the patient and avoids Sybil attacks. Performance assessment shows that our method achieves well-known progress towards care taker and Sybil

attack conflict to preserve privacy. we proposed a finest algorithm, we implemented the system as two partitioning-based approaches, first targeting to protect from Sybil attacks by identifying fake critiques and rankings, second is to avoid linkage attack by implementing local recoding anonymization algorithm. Even though, this method couldn't accomplish best in practice, due to inability in creating best groups is a problem In Future we are aiming to use a refined Gray-ordering method for partitioning, it is also very expensive in nature to anonymize data and become a bottleneck of the entire process.

## REFERENCES

1. X. Liang, X. Lin, and X. S. Shen, "Enabling trustworthy service evaluation in service-oriented mobile social networks," IEEE Trans. Parallel Distributed System., vol. 25, no. 2, pp. 310–320, Feb. 2014.
2.P. Moradi and S. Ahmadian, "A reliability-based recommendation method to improve trust-aware recommender systems," Expert Syst. with Appl., vol. 42, no. 21, pp. 7386–7398, 2015.
3. S. Shen, S. Wang, L. Jiayue, and J. Zhao, "Review of social trust mechanism and doctor-patient trust evaluation," Chin. Med. Ethics, vol. 30, no. 9, pp. 1098–1102, 2017.
4. N. B. Truong, H. Lee, B. Askwith, and G. M. Lee, "Toward a trust evaluation mechanism in the social Internet of Things," Sensors, vol. 17, no. 6, 2017, Art. no. e1346.
5. Q. Xu, Z. Su, S. Yu, and Y. Wang, "Trust based incentive scheme to allocate big data tasks with mobile social cloud," IEEE Trans. Big Data, to be published, doi:10.1109/ TBDATA. 2017.2764925.
6. J. Jiang, G. Han, L. Shu, S. Chan, and K. Wang, "A trust model based on cloud theory in underwater acoustic sensor networks," IEEE Trans. Ind. In format., vol. 13, no. 1, pp. 342–350, Feb. 2017.
7. H. Li et al., "Privacy leakage via de-anonymization and aggregation in heterogeneous social networks," IEEE Trans. Dependable Secure Comput., to be published.
8. B. Wang, W. Song, W. Lou, and Y. T. Hou, "Privacy-preserving pattern matching over encrypted genetic data in cloud computing," in Proc. IEE, INFOCOM, 2017, pp. 1–9.
9. Zakerzadeh, Hessam, Charu C. Aggarwal, and Ken Barker. "Privacy-preserving big data publishing. "Proceedings of the 27th International Conference on Scientific and Statistical Database Management. ACM, 2015.
10. Russom, Yohannes. Privacy preserving for Big Data Analysis. MS thesis. University of Stavanger, Norway, 2013.

## AUTHORS PROFILE

**Mr.N.A.S.Vinoth is** currently working as an assistant professor at software engineering department in SRMIST. He has 4 years of teaching experience. While working in Kalasalingam Institute of Technology he worked in research projects for Networking Research Lab. His research interest include network security, Big Data, machine learning. Vinoth presented nearly 4 research articles in National and international conferences. He has organized IET sponsored National work shop on VANET and its security, International Conference on Artificial intelligence and evolutionary computations in engineering systems. He is an IELTS Scorer and also an member of IEI.

**Mrs. M. Yohapriyaa is** currently working as a Teaching Associate at software engineering department in SRMIST. She has 1 year of teaching experience. Her research interest include network security, Big Data, machine learning. Yohapriyaa presented nearly 2 research articles in National and international conferences. She has organized International Conference on Artificial intelligence and evolutionary computations in engineering systems.

**Mrs. K. Janani is** currently working as a Teaching Associate at software engineering department in SRMIST. She has 1 year of teaching experience and 5 years of industrial experience in the area of software testing. Her research interest include network security Cloud Computing, machine learning. Janani presented nearly 2 research articles in National and international conferences. She has organized International Conference on Artificial intelligence and evolutionary computations in engineering systems.

**Mrs. VijayPriya is** currently working as a Teaching Associate at software engineering department in SRMIST. She has 1 year of teaching experience and 5 years of industrial experience in the area of software testing. Her research interest include network security Cloud Computing, machine learning. VijayPriya presented nearly 2 research articles in National and international conferences. She has organized International Conference on Artificial intelligence and evolutionary computations in engineering systems.