ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

Investigations on Evolution of Approaches Developed for Data Privacy

R.S.Venkatesh[1], P.K.Reejeesh[2], Prof.S.Balamurugan[3], S.Charanyaa[4]

Department of IT, Kalaignar Karunanidhi Institute of Technology, Coimbatore, TamilNadu, India[1,2,3]

Senior Software Engineer Mainframe Technologies Former, Larsen & Tubro (L&T) Infotech, Chennai, TamilNadu, India[4]

ABSTRACT: This paper reviews methods developed for anonymizing data from 1984 to 1988 . Publishing microdata such as census or patient data for extensive research and other purposes is an important problem area being focused by government agencies and other social associations. The traditional approach identified through literature survey reveals that the approach of eliminating uniquely identifying fields such as social security number from microdata, still results in disclosure of sensitive data, k-anonymization optimization algorithm ,seems to be promising and powerful in certain cases ,still carrying the restrictions that optimized k-anonymity are NP-hard, thereby leading to severe computational challenges. k-anonimity faces the problem of homogeneity attack and background knowledge attack . The notion of l-diversity proposed in the literature to address this issue also poses a number of constraints , as it proved to be inefficient to prevent attribute disclosure (skewness attack and similarity attack), l-diversity is difficult to achieve and may not provide sufficient privacy protection against sensitive attribute across equivalence class can substantially improve the privacy as against information disclosure limitation techniques such as sampling cell suppression rounding and data swapping and pertubertation. This paper aims to discuss efficient anonymization approach that requires partitioning of microdata equivalence classes and by minimizing closeness by kernel smoothing and determining ether move distances by controlling the distribution pattern of sensitive attribute in a microdata and also maintaining diversity.

KEYWORDS: Data Anonymization, Microdata, k-anonymity, Identity Disclosure, Attribute Disclosure, Diversity

I. INTRODUCTION

Privacy-An important factor need to be considered while we publishing the microdatas. Usually government agencies and other organization used to publish the microdatas. On releasing the microdatas, the sensitive information of the individuals are being disclosed. This constitutes a major problem in the government and organizational sector for releasing the microdata. In order to sector or to prevent the sensitive information, we are going to implement certain algorithms and methods. Normally there two types of information disclosures they are: Identity disclosure and Attribute disclosure. Identity disclosure occurs when an individual's linked to a particular record in the released Attribute disclosure occurs when new information about some individuals are revealed.(i.e)the released data make it possible to infer the characteristics of an accurately than it would be possible before the data released. The Knowledge of identity disclosure would often allow us to know about attributes disclosure .Once the identity disclosure comes into exists ,the individuals sensitive information is reidentified. Due to the effects of false attributes ,an observer of a release table may incorrectly perceive that an individuals. sensitive attribute takes a particular value. This can harm the individuals even if the perception is incorrect. When the table is released, it present disclosure risk to the individual who are all in table.

The remainder of the paper is organized as follows. Section 2 deals about basic definition and primitives of data anonymization. Protection and resource control in operating systems are discussed in Section 3. Section 4 portrays Computer Security System For a Time Shared Computer Accessed Over Telephone Lines. Section 5 briefs about Computer Security Systems. Section 6 briefs about network security without observability. Section 7 Concludes the paper and outline the direction for Future Work.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

II. BASIC DEFINITION AND PRIMITIVES

Data refers to organized personal information in the form of rows and columns. Row refers to individual tuple or record and column refers to the field. Tuple that forms a part of a single table are not necessarily unique. Column of a table is referred to as attribute that refers to the field of information, thereby an attribute can be concluded as domain. It is necessary that attribute that forms a part of the table should be unique. According to L.Sweeney et.al., (2001) [26] each row in a table is an ordered n-tuple of values $<d_1,d_2,....d_n>$ such that each value d_j forms a part of the domain of j^{th} column for j=1,2,...n where 'n' denoted the number of columns.

Attributes

Consider a relation $R(a_1, a_2,... a_n)$ with finite set of tuples. Then the finite set of attributes of R are $\{a_1, a_2,... a_n\}$, provided a table $R(a_1, a_2,... a_n)$, $\{a_1, a_2,... a_i\} \subseteq \{a_1, a_2,... a_n\}$ and a tuple $l \in R$, $l[a_i,... a_n]$ corresponds to ordered set of values $v_i,...v_j$ of $a_i... a_j$ in l. R $[a_i... a_j]$ corresponds to projection of attribute values $a_1, a_2,...,a_n$ in R, thereby maintaining tuple duplicates.

According to Ningui Li, Tiancheng Li et.al., [17] (2010), attributes among itself can be divided into 3 categories namely

1. Explicit identifiers- Attributes that clearly identifies individuals. For eg, Social Security Number for a US citizen.
2. Quasi identifiers- Attributes whose values when taken together can potentially identify an individual. Eg., postal code, age, sex of a person. Combination of these can lead to disclosure of personal information.
3. Sensitive identifiers- That are attributes needed to be supplied for researchers keeping the identifiers anonymous. For eg, 'disease' attribute in a hospital database, 'salary' attribute in an employee database.

TABLE 1:
MICRODATA DATABASE CONTAINING SENSITIVE INFORMATION

Race	Birth	Gender	Zipcode	Disease (Sensitive Information)
Black	1965	M	0213	Shortbreath
Black	1965	M	0213	Shortbreath
Black	1965	M	0214	Hypertension
White	1964	F	0213	Obesity
White	1965	F	0214	Chestpain
White	1967	M	0213	Shortbreath
White	1964	M	0214	Chestpain

Quasi- Identifiers

As proposed by L.Sweeney et.al., (2001) [26], A single attribute or a set of attributes that, in combination with some outside world information that can identify a single individual tuple in a relation is termed as quasi-identifier. Given a set of entities E, and a table $B(a_1,\ldots,a_n)$, $f_a: E \rightarrow B$ and $f_b: B \rightarrow E'$, where $E \rightarrow E'$. A quasi-identifier of B, written as U_E, is a set of attributes $\{a_i,\ldots,a_j\} \rightarrow \{a_1,\ldots,a_n\}$ where: $\exists s_i \in U$ such that $f_a(f_b(s_i)[U_E]) = s_i$.

k-Anonymity

Let RT(A1,A2,….An) be a table and QI_{RT} be the Quasi identifier. RT is said to be k-anonymous [26] if and only if each sequence of values in $RT[QI_{RT}]$ appears atleast k-times in $RT[QI_{RT}]$. In short, the Quasi identifier must appear atleast 'k' times in RT, where k=1,2,3,... where 'k' is termed to be the anonymity of the table.

l-diversity

Since k-anonymity failed to secure the attribute disclosure, and is susceptible to homogeneity attack and background knowledge attack A.Machanavajjhala et.al, (2006) [38] introduced a new privacy notation called 'l-diversity'[20]. An equivalence class is said to possess l-diversity if there are atleast 'l' well represented values for the sensitive attribute. A table is said to have l-diversity if every equivalence class of the table has l-diversity. Here the technique is the sensitive attribute in each equivalence class is distributed with l-well represented values. Generally there are four types of l-diversity.

1) Distinct l-diversity: This ensures that there are atleast l-distinct values for the sensitive attribute in each equivalence class. The biggest disadvantage here is that distinct l-diversity fails to prevent probabilistic inference attacks.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

2) Probabilistic l-diversity: An anonymised table is said to be probabilistic l-diversity if the frequency of the sensitive value in each group is atmost 1/l.

3) Entropy l-diversity: It is defined by, Entropy (E)= $-\sum_{s\in S} P(E,s)\log p(E,s)$, where 's' is the sensitive attribute.

4) Recursive(c,l) diversity: This technique proceeds by making, the value appearing most frequently, not appear too frequently and less frequently appearing value not to appear too rarely.

One problem with l-diversity is that it is limited in its assumption of adversarial knowledge. l-diversity fails to prevent attribute disclosure and is susceptible to two types of attacks.

t-closeness

Privacy is measured by the information gain of an observer. Before seeing the released table the observer may think that something might happen to the sensitive attribute value of a single person. After seeing the released table the observer may have the details about the sensitive attributes. t-closeness [17] should have the distance between the class and the whole table is no more than a threshold t, Ningui Li et.al., (2010)[17].

III. PROTECTION AND RESOURCE CONTROL IN DISTURBUTED OPERATING SYSTEMS

Sape J. Mullender, Andrew S.Tanenbaum (1984) Local networks have cable snaking along with sockets which is used by the users to plug their PC, intelligent terminals, file servers, etc. These devices are called as "producers" since they provide service to "consumers". It is difficult to build a secure Operating System where the system administrator is unable to prevent malicious users from plugging. This paper describes how to build a secure Operating System without restrictions. We need end-to end encryption protection method to avoid passive or active wire tappers.

The main problem in anonymous network is authentication. In this paper, authentication doesn't produce any attack. It describes about the knowledge of some critical information to control access to services.

The basic model of the distributed system is called as "service". The main part of the service is that one of servers is used to access each object in a service. From a port on the client process, a message (request) is sent to each server. In return, the server sends a reply from a port on the server process to the client. Each port consists of 2 names – 1.put-port :-used by clients,2.get-port:used by servers.

"Signature" is often used to determine the user who signed in. Signature has some properties: only the "owner" of the document has right to sign in the document. If we user ports as signature, there are two types: 1.Public signature- checks whether the signature is genuine or not. 2.Private signature – required to sign in.

In this proposed protection scheme, a "capability" mode of bit-string is used to perform operations on objects. It consists of 4 parts: 1.Server field of capability – used by put-port. 2.Object field used a index,3.Rights field-Specifies the type of operation,4.Random field-sparse capability. In a protected Operating System, kernel can be used to maintain capabilities. But we can't trust the kernel's in user machines. So, we are forced to use a protection mechanism where port names are selected randomly and port names are selected randomly and port name space is made sparse.

The directory server is a tree-like structure that provides a private directory to each user to maintain capabilities. In order to use the directory service, a client may encrypt the ports. A login server is used to authenticate a user.

Open System Interconnection [OSI] was proposed by ISO with 7 layers of protocol. LAN doesn't contain network layer which is often used for routing and congestion control. But our idea is to make these capabilities as address to create another form of network layer. This newly formed network layer is used to receive messages in the destination with the help of put-port and get-port. Even we can also use "public key encryption" to make this new network layer to be more secure under certain considerations.

Lower protocol layers are used by port layer to implement this message passing technique. And the system calls or subroutines are used to implement port layer. In a user Interface, there are 2 calls for client and server which is same as that is used for sending and receiving messages.

Eg: put(var put port, scrport , signature : PORT ; var buffer: MESSAGE);
get(var getport, srcport, signature : PORT; var buffer: MESSAGE);

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

Object, rights and random fields are combined together to build a library layer to make message passing easier.

In network protection, we establish a hardware interface between user machines and cable. In order to provide certain control, this network interface is included in a logic board of user's machine. The interface should be tamperproof. Interface should be capable of monitoring all kinds of traffic that occurs in a cable. This network protection scheme works on one-way ciphers where a function F is associated with get-port and put-port by the formula.

Put-port=F(get-port)

When we create a service, "make port" selects a random number from the address space to use it as get-port. The client uses a put-port to send message to the server. In order to receive message, the server has to use the function F by including get-port. The put-port which is computed is stored in a table. Each time when the message is passed, the interface checks the table for destination port. It is found, then the interface copies the message to user machine. Since only the put-port is known publicly, the intruder will find difficult to determine the server's get port number which is kept as secret. Also, we are using port space to select get-ports randomly. So, there is no way for an intruder to obtain the port used by the server. "Signed messages" can also be sent using one way function, with the help of private and public signatures. Both the signatures field and the srcport field are encrypted on transmitted message. Eg: consider server A and B. Assume X is sending a message from B's put-port A. In turn, A sends a reply to B without knowing that the sender is X. Now B, after receiving message from A, it will assume the message as request and will send a reply to A. Again A will and will assume the message as request and so on.

In case, if the network user's physical machine numbers for routing then a "locate message" is sent to all interfaces. After receiving the replies, the sender interface will select the exact physical address to be used. In a star-shaped network, the one-way function mechanism is used in such a way that the functions are performed by switch instead of interfaces. Only if the secret ports are revealed, the intruder can steal the message which is completely impossible.

In a network protection problem, the Operating System important solution is public-key cryptography in which a kernel or interface is not needed. Both the messages and ports are encrypted to transfer information in a secure way between the server and the client. A public key is used for encryption. This key pair is created using "make port" with port-identifier, encryption/decryption key. All the messages are encrypted and decrypted using these keys and sent to the server. The entire message need not be encrypted; it is enough to encrypt the header alone. We need to use checksum when we encrypt the entire message.

The servers and put-port key is used to encrypt the requests and the clients put-port key is used to encrypt the replies.

Object-oriented protection system based on capability is easily built on this protection mechanism. When a server creates a new object, the server returns a message along with object capability. It makes use of object field, server field, rights field and random field.

There may be causes existing in which the capabilities are able to perform all the operations with partial rights and no rights field.

In object-oriented protection system, it is easy to detect the vulnerable activities because the capabilities are modified with their address spaces. The clients can also have redundant copy of its capabilities are returned back to the client.

The above discussed are the mechanisms used to enhance the protection in distributed and network Operating System. Compared with other protection schemes, here we don't need any special components for authentication. A controlled environment will be efficient only if a trustworthy Operating System is used. But here, these protection schemes can support any untrustworthy Operating System. One-way function is less expensive compared to public-key cryptography. One-way function is computed by assigning each interface a processor. Also we require a table to maintain put-port numbers.

Whenever a message is sent on a cable interface will store the message I a buffer. The destination port of the message is determined and it is compared with the port table. If it is similar, the message is passed or it is ignored. When compared to public-key cryptography, one-way function can easily generate keys with more efficiency. The keys

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

in public key cryptography are layer than in one-way function. It's efficiency is also less because its needs signed messages to be retrieved.

Implementing distributed and network Operating System will enhance the protection and improves security. No central authority is needed. OSI model is more simple because it derives separate channels for communication between the lower layers of protocol. Wire-tapping is avoided by using end-to-end encryption method. The one-way mechanism and public-key cryptography can be implemented using any software thus providing more secure way of communication.

IV. COMPUTER SECURITY SYSTEM FOR A TIME SHARED COMPUTER ACCESSED OVER TELEPHONE LINES

Paul J.Levine (1985) The time-shared computer uses an automatic tracing equipment of a telephone switching systems to prevent unauthorized access. The telephone system whenever a user uses the telephone system to access the central computer. At the same time, a password that addresses a memory is provided to the time-shared, central computer by the user. A comparison is made between the telephone no of authorized users and telephone no of calling user. If the result shows that the calling user is from proper telephone station, then the access is permitted or it will be denied.

By using unchangeable telephone no's the authorized user identity cannot be altered. And also avoiding access to the system program will prevent tampering.

Remote users accessing a time-shared computer system will avoid un authorization and invention of security computer systems, software passwords and codes were used. The biggest disadvantage is that it leads to large coding and decoding combinations. They also tried inventing unbreakable codes. Later used unchangeable passwords to improve security.

Each and every former invention of security systems has some disadvantages like direct access by a remote user, trying out all the combinations of passwords to find exact passwords , etc.

Security computer systems is the recent invention to prevent un authorization . The time-shared system acts as a central computer which is accessed using telephone switching systems with automatic tracing equipment. The information of a remote caller is obtained and is compared with file information using a comparator associated with security computer. If the comparison results are similar, then the security computer produces and transfers a acknowledgement signal to the time-shared system. Otherwise a negative acknowledgement is produced. Unless the access of a remote caller is from a proper station he cannot access the time-shared system. The security computer remains inaccessible to remote caller. Even though if a remote caller access a central computer, he cannot change the telephont no and data produced.

It is not necessary to create complex cryptographic techniques to improve the security of a system if we omit offsite security elements in a system.

Modification done by a remote user doesn't affect the security since the security computer is not dependent on real-time computers. Along with automatic tracing equipment, the telephone switching system also has conventional switching equipment. It is combined with remote users keyboard through which the information is sent to the telephone switching system. The phone lines is combined with an online computer with the help of buffer present in real-time computer systems. So the telephone switching system establishes connection in a normal way. At the same time , the tracer will find the caller's phone no and address and passes it to the telephone switching system. Additionally , it also finds time and date of call.

Information generated from a tracer is compared with file information present in real-time computers. Real-time computers also use passwords to scan the complete record of a file for caller's information. As per the result of comparison, an acknowledgement signal or negative signal is produced.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

If a shift register is used to store the callers phones no and address, then the comparison is made between the phone no and address within the shift register itself, According to the result, ACK or NACK is produced by the comparator. Memory access is controlled by a controller. The controller generates a password to address the file information. It helps in storing the required information in the comparator. Again ACK or NACK signal is produced.

The file information consists of lists of phone no's and address from which a computer can be accessed from any location. Any location will compromise either single password or multiple passwords. Every combinations in the lists will be tried to find the exact passwords.

For every incoming call, the information like phone number , address, name, date and time will be gathered and sent to the security computer systems.

An internal security system is built inside a security computer. A user from a single location tries to access a real-time computer systems may times then that location will be ignored from the memory. Further access from the same location will produce a NACK signal.

Many security schemes were derived in order to improve the security mechanism. Every time when a user access the security computer, an unchangeable information is produced automatically.

V. COMPUTER SECURITY SYSTEMS

John G.Campbell,Carl F.Schoeneberger (1986) The computer security systems are widely used in data processing systems for maintaining confidentiality and integrity of information in terms of security, data processing system has many characteristics. That is, it can be executed using passwords and usernames or by improving the computer's internal architecture. Hardware, firmware and software together forms an internal architecture. The system is designed in such a way that any user can access the system resources effectively.

There are 3 layers in the earlier computer systems. The first layer is called as "supervisor" and each layer is controlled by other layers. Firewalls were also intended to provide good environment for these layers to perform all the functions. But still many demerits exist in the system which reduces integrity of data.

In order to achieve security, we can distinguish 3 layers by introducing interfaces between them. "Graphics Package" can be encouraged to reduce the system failure. This type of system Is so effective that the database can be accessed only by the administrator.

The recently designed systems concentrates mainly on improving security by avoiding use of system resources by an authorized user. That is, an attacker can't access others data. The main objective of this system is to provide a good environment of computer system based on capabilities.

In a system, domain's resolves as a virtual machine. Each domain has minimum of 1 node. The keys used in the system are stored inside the node. Each key point to an object in the system using a pointer. A domain can make use of kernel functions to create keys(capability). A "factory" allows 2 domains to their resources. A "hole comparator" defines the ability of a factory whether it is trustworthy or not.

Data Processing s/y is also called as capability s/y. Here, objects are used executing an architecture where an Operating System is constructed. An object may be a key or domain. These objects can be inherited in any one of the following: hardware, firmware, software. The computer memory is divided into two parts:

1.Core memory-high speed memory ranges between 3ns to 250ns. Used for arithmetic calculation by data processing system.

2.Disk memory-Slow memory, consist of one or two magnetic disk memory units.

As mentioned earlier, keys are stored inside a node. These keys allows only authorized user to access the s/f resourses. A domain can't manipulate a key or node. Only a kernel has the ability to manipulate a key or a domain. Thus a s/y intergrity is improved by using kernel functions. It also avoids manipulationg one domain by another domain.

The disk memory is further categorized into:

1. Disk-page space
2. 2.Disk-node space

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

Coded disk address(CDA)-determines the position of an object in the disk memory allocation count(ALC)- it is similar to allocation number of an object. When a frame is removed, the ALU value is increased.
ALC-INC-is a flag representing the core table for implementing ALC.
There are 2 pointers-forward and backward pointers pointing the index space. Both the pointers are added for an item space entry.

A memory key comprises of page keys and segment keys. Page keys provide access to a page. Segment keys provide access to set of pages. Mode keys gives access to perform the operation read or write. Domain keys provide access to manipulate a domain. Device I/O keys allows only authorized I/O devices.

VI. NETWORKS WITHOUT USER OBSERVABILITY

Peter P. Gombrich, Richard J. Beard, Richard A. Griffee (1987) pointed the main risk that we are facing today in networks field is that an attacker can easily gain access and steal user's information. Though the end-to-end encryption method is implemented, the network still remains unprotected. The user of both source and destination should be kept away from attacks. A public network station can be safer for a user rather than private network station.

The intention of this system is to meet the ISDN requirements, to avoid malicious and vulnerable attacks, to increase security and reliability. An attacker can steal one's data in many ways. So the system is designed in a way to convert these attackers into stations and lines. To note is that the stations must be controlled only by a user who access it.

There are two possible ways to maintain security:
1.The message exchanged between the sender and the receiver bond is kept secret.
2.The sender and receiver bond is kept secret.
But the former option is considered as better option. The intruder will find difficult to gain the information exchanged between sender and the receiver.

A message is delivered to all the stations and so that the message remains unknown to the network. Implicit address are used based on their visibility state, that is , an implicit address may be visible or invisible. The "public-key cryptographic" system uses invisible implicit address where the message is encrypted using public key. Every station has to use a private key to decrypt all messages received.

The invisible implicit addresses also makes use of a secret way when compared visible implicit addresses are simple. Here, the user encrypts the message be prefixing it with a randomly chosen name. The implicit address acts as private if only sender knows the address. Implicit addresses acts as public if the addresses is known to all users.
Another mechanism "unlinkability" contains a station called MIX. It gathers all the message from the senders and transfers it to the receiver in a different order by altering the encoding scheme of the messages.

The relationship between a sender and receiver remains unknown for an attacker when we use multiple MIX's.

To protect the confidential data, the sender can also produce a key for each and every message. Then the sender includes all the keys produced with the message. This is passed to the receiver in a secure channel. Usually, in LAN, the stations are connected in the form of rings. This also prevents access of unauthorized users.

The system performance depends on two major categories:
1.Transfer delay
2.Throughput.

A RING network can encourage only limited stations. More than 1000 stations will reduce its efficiency. But a RING network seems to be a better choice than a star or bus network.
In a MIX-network, only a specific station is chosen as MIX. Since we have to mix up each message, the content of the message becomes longer. Encouraging too many MIX'S may result in collision. The station that serves as MIX should have high throughput. Hence this method is more expensive.

We can divide the stations into groups in a networks to increase the performance. This type of network is referred as switched/broadcast networks(SBNS). If the acknowledgement signal is not received, the sender one transmits the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

message. But this end-to-end retransmission may decrease the network performance. To avoid retransmission, the MIX-network can have backups or one MIX can replace the damaged MIX.
The sender and receiver can exchange their message in a secure way. The content of the message remains unknown for an attacker. Integrity of the data is preserved. The network system proves to be stronger against malicious or vulnerable attacks.

TF Lunt (1988) in this paper discusses about the key problem of storage which is related with the process of having secured communication between each pair of user in a big network. The authors have used a method of finite incident structure with important features which is known as key distribution patterns. The extensive formulation of the process of storage scheme enables to use the theory of block designs and hence through this theory a number of examples of key distribution system are extracted. Through this paper the authors have introduced a number of new concepts and thus show how the theory of finite incident structure work with the key management problems.

VII. CONCLUSION AND FUTURE WORK

Various methods developed for anonymizing data from 1984 to 1988 is discussed. Publishing microdata such as census or patient data for extensive research and other purposes is an important problem area being focused by government agencies and other social associations. The traditional approach identified through literature survey reveals that the approach of eliminating uniquely identifying fields such as social security number from microdata, still results in disclosure of sensitive data, k-anonymization optimization algorithm ,seems to be promising and powerful in certain cases ,still carrying the restrictions that optimized k-anonymity are NP-hard, thereby leading to severe computational challenges. k-anonimity faces the problem of homogeneity attack and background knowledge attack . The notion of l-diversity proposed in the literature to address this issue also poses a number of constraints , as it proved to be inefficient to prevent attribute disclosure (skewness attack and similarity attack), l-diversity is difficult to achieve and may not provide sufficient privacy protection against sensitive attribute across equivalence class can substantially improve the privacy as against information disclosure limitation techniques such as sampling cell suppression rounding and data swapping and pertubertation. Evolution of Data Anonymization Techniques and Data Disclosure Prevention Techniques are discussed in detail. The application of Data Anonymization Techniques for several spectrum of data such as trajectory data are depicted. This survey would promote a lot of research directions in the area of database anonymization.

REFERENCES

1. Pieter Van Gorp and Marco Comuzzi “Lifelong Personal Health Data and Application Software via Virtual Machines in the Cloud” IEEE Journal of Biomedical and Healthcare Informatics, Vol. 18, No. 1, Jan 2014
2. Sape J. Mullender, Andrew S.Tanenbaum, "Protection and Resource Control in Distributed Operating Systems", 1984.
3. Paul J.Levine, "Computer security system for a time shared computer accessed over telephone lines US 4531023 A, 1985
4. John G.Campbell,Carl F.Schoeneberger,"Remote hub television and security systems", US 4574305 A, 1986.
5. A Pfitzmann, "Networks without user observability", Computers & Security 6/2 (1987) 158-166, 1987
6. TF Lunt, " Automated audit trail analysis and intrusion detection: A survey" In Proceedings of 11th National Conference on Security, 1988
7. Lichtenstein Eric Stefan 1984 a, Computer control medical care system US4464172.
8. ARalph R.Frerichs, Dr. PH.Robert A. Miller 1985, Introduction of a Microcomputer for Health Research in a Developing Country.
9. Steven P.Brown 1986, Combinational Medical Data, Identification and health Insurance card.
10. Peter P. Gombrich, Richard J. Beard, Richard A. Griffee, Thomas R. Wilson, Ronald E. Zook, Max S. Hendrickson 1989,A Patient care system,US4835372 A.
11. Neil Bodick, Andre L. Marquis1990, Interactive system and method for creating and editing a knowledge base for use as a computerized aid to the cognitive process of diagnosis,US4945476 A.
12. Angela M. Garcia, Dr.,Boca Raton 1991 a, System and Method for scheduling and Reporting Patient related services including prioritizing services,US5974389 A.
13. Clark Melanie Ann, John Finley, Huska; Michael Edward, Kabel; Geoffrey Harold, Graham, Marc Merrill 1991 b,System and Method for scheduling and Reporting Patient Related services.
14. Robert W. Kukla1992,Patient care communication system, US5101476 A
15. Mark C. Sorensen 1993, Computer aided medical diagnostic method and apparatus, US5255187 A.
16. Edward J. Whalen, San Ramon, Olive Ave Piedmont 1994,Computerized file maintenance System for managing medical records including narrative patent documents reports.
17. Desmond D. Cummings 1994b,All care health management system, US5301105 A.
18. Woodrow B. Kesler Rex K Kesslerin 1994 c,Medical data draft for tracking and evaluating medical treatment.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

19. Joseph P. Tallman, Elizabeth M. Snowden, Barry W. Wolcott 1995, Medical network management system and process, US5471382 A.
20. Peter S. Stutman, J. Mark Miller 1996,Medical alert distribution system with selective filtering of medical information
21. Edwin C. Iliff1997,computerized medical diagnostic system including re-enter function and sensitivity factors, US5594638 A.
22. Timothy Joseph Graettinger, Paul Alton DuBose 1998, Computer-based neural network system and method for medical diagnosis and interpretation. US5839438 A.
23. Melanie Ann Clark, John Finley Gold, Michael Edward Huska, Geoffrey Harold Kabel, Marc Merrill Graham1999,Medical record management system and process with improved workflow features, US5974389 A.
24. Richard S. Surwit, Lyle M. Allen, III, Sandra E. Cummings 2000 a, Systems, methods and computer program products for monitoring, diagnosing and treating medical conditions of remotely located patients, US6024699 A.
25. Jeffrey J. Clawson 2000 b, Method and system for giving remote emergency medical counsel to choking patients, US6010451 A.
26. Marc Edward Chicorel 2001, Computer keyboard-generated medical progress notes via a coded diagnosis-based language, US6192345 B1.
27. Charlyn Jordan2002, Health analysis and forecast of abnormal conditions.
28. Jeffrey J. Clawson2003, Method and system for an improved entry process of an emergency medical dispatch system
29. PekkaRuotsalainen 2004, A cross-platform model for secure Electronic Health Record communication.
30. Roger J. Quy2005, Method and apparatus for health and disease management combining patient data monitoring with wireless internet connectivity, US6936007 B2.
31. Avner Amir, Avner Man2006 a, System and method for administration of on-line healthcare, WO2006006176 A2.
32. Paul C.Tang, Joan S. Ash, David W. Bates, J. Marc overhage and Daniel Z.Sands 2006 b, Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption.
33. Christopher Alban, KhiangSeow2007, Clinical documentation system for use by multiple caregivers.
34. Brian A. Rosenfeld, Michael Breslow2008, System and method for accounting and billing patients in a hospital environment.
35. Jacquelyn Suzanne Hunt, Joseph Siemienczuk 2009, Process and system for enhancing medical patient care.
36. Richard J. Schuman2010, Health care computer system, US7831447 B2.
37. Kanagaraj, G.Sumathi, A.C.2011,Proposal of an open-source Cloud computing system for exchanging medical images of a Hospital Information System
38. AvulaTejaswi, NelaManoj Kumar, GudapatiRadhika, SreenivasVelagapudi 2012 a, Efficient Use of Cloud Computing in Medical Science.
39. J. Vidhyalakshmi, J. Prassanna 2012 b, Providing a trustable healthcare cloud using an enhanced accountability framework.
40. Carmelo Pino and Roberto Di Salvo 2013, A Survey of Cloud Computing Architecture and Applications in Health.
41. K.S. Aswathy, G. Venifa Mini 2014 a, Secure Alternate Viable Technique of Securely Sharing the Personal Health Records in Cloud.
42. Abhishek Kumar Gupta, Kulvinder Singh Mann 2014 sharing of Medical Information on Cloud Platform.
43. D. C. Kaelber, A. K. Jha, D. Johnston, B. Middleton, and D. W. Bates, "Viewpoint paper: research agenda for personal health records (PHRs),"J. Amer. Med. Inform. Assoc., vol. 15, no. 6, pp. 729–736, 2008.
44. J. Ahima, "Defining the personal health record," vol. 76, no. 6, pp. 24–25, Jun. 2005.
45. W. Currie and M. Guah. "Conflicting institutional logics: a national programme for it in the organizational field of healthcare:, Journal of Information Technology, 22:235–247,2007.
46. M. Gysels, A. Richardson, and J. I. Higginson "Does the patient-held record improve continuity and related outcomes in cancer care: a systematic review", Health Expectations,10(1):75–91, Mar. 2007.
47. International Organization for Standardization. ISO TR20514:2005 Health Informatics - Electronic Health Record Definition, Scope and Context Standard. International Organization for Standardization (ISO). Geneva, Switzerland,2005.
48. B.Powmeya , Nikita Mary Ablett ,V.Mohanapriya,S.Balamurugan,"An Object Oriented approach to Model the secure Health care Database systems,"In proceedings of International conference on computer , communication & signal processing(IC^3SP)in association with IETE students forum and the society of digital information and wireless communication,SDIWC,2011,pp.2-3
49. Balamurugan Shanmugam, Visalakshi Palaniswami, "Modified Partitioning Algorithm for Privacy Preservation in Microdata Publishing with Full Functional Dependencies", Australian Journal of Basic and Applied Sciences, 7(8): pp.316-323, July 2013
50. Balamurugan Shanmugam, Visalakshi Palaniswami, R.Santhya, R.S.Venkatesh "Strategies for Privacy Preserving Publishing of Functionally Dependent Sensitive Data: A State-of-the-Art-Survey", Australian Journal of Basic and Applied Sciences, 8(15) September 2014.
51. S.Balamurugan, P.Visalakshi, V.M.Prabhakaran, S.Chranyaa, S.Sankaranarayanan, "Strategies for Solving the NP-Hard Workflow Scheduling Problems in Cloud Computing Environments", Australian Journal of Basic and Applied Sciences, 8(15) October 2014.
52. Charanyaa, S., et. al., , A Survey on Attack Prevention and Handling Strategies in Graph Based Data Anonymization. International Journal of Advanced Research in Computer and Communication Engineering, 2(10): 5722-5728, 2013.
53. Charanyaa, S., et. al., Certain Investigations on Approaches forProtecting Graph Privacy in Data Anonymization. International Journal of Advanced Research in Computer and Communication Engineering, 1(8): 5722-5728, 2013.
54. Charanyaa, S., et. al., Proposing a Novel Synergized K-Degree L-Diversity T-Closeness Model for Graph Based Data Anonymization. International Journal of Innovative Research in Computer and Communication Engineering, 2(3): 3554-3561, 2014.
55. Charanyaa, S., et. al., , Strategies for Knowledge Based Attack Detection in Graphical Data Anonymization. International Journal of Advanced Research in Computer and Communication Engineering, 3(2): 5722-5728, 2014.
56. Charanyaa, S., et. al., Term Frequency Based Sequence Generation Algorithm for Graph Based Data Anonymization International Journal of Innovative Research in Computer and Communication Engineering, 2(2): 3033-3040, 2014.
57. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Certain Investigations on Strategies for Protecting Medical Data in Cloud", International Journal of Innovative Research in Computer and Communication Engineering Vol 2, Issue 10, October 2014
58. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Investigations on Remote Virtual Machine to Secure Lifetime PHR in Cloud ", International Journal of Innovative Research in Computer and Communication Engineering Vol 2, Issue 10, October 2014
59. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Privacy Preserving Personal Health Care Data in Cloud" , International Advanced Research Journal in Science, Engineering and Technology Vol 1, Issue 2, October 2014
60. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, "Investigations on Evolution of Strategies to Preserve Privacy of Moving Data Objects" International Journal of Innovative Research in Computer and Communication Engineering, 2(2): 3033-3040, 2014.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

61. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Certain Investigations on Securing Moving Data Objects" International Journal of Innovative Research in Computer and Communication Engineering, 2(2): 3033-3040, 2014.
62. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Survey on Approaches Developed for Preserving Privacy of Data Objects" International Advanced Research Journal in Science, Engineering and Technology Vol 1, Issue 2, October 2014
63. S.Jeevitha, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Privacy Preserving Personal Health Care Data in Cloud" International Advanced Research Journal in Science, Engineering and Technology Vol 1, Issue 2, October 2014.
64. K.Deepika, P.Andrew, R.Santhya, S.Balamurugan, S.Charanyaa, "Investigations on Methods Evolved for Protecting Sensitive Data", International Advanced Research Journal in Science, Engineering and Technology Vol 1, Issue 4, Decermber 2014.
65. K.Deepika, P.Andrew, R.Santhya, S.Balamurugan, S.Charanyaa, "A Survey on Approaches Developed for Data Anonymization", International Advanced Research Journal in Science, Engineering and Technology Vol 1, Issue 4, Decermber 2014.
66. S.Balamurugan, S.Charanyaa, "Principles of Social Network Data Security" LAP Verlag, Germany, ISBN: 978-3-659-61207-7, 2014
67. S.Balamurugan, S.Charanyaa, "Principles of Scheduling in Cloud Computing" Scholars' Press, Germany,, ISBN: 978-3-639-66950-3, 2014
68. S.Balamurugan, S.Charanyaa, "Principles of Database Security" Scholars' Press, Germany, ISBN: 978-3-639-76030-9, 2014

APPENDIX

S.no	YEAR	AUTHORS	TITLE
1	1984	Sape .MULLENDER and Andrew S TANENBAUM	PROTECTION AND RESOURCE CONTROL IN DISTRIBUTED OPERATING SYSTEMS
2	1985	Paul j.Levine	COMPUTER SECURITY SYSTEM FOR TIME SHARED COMPUTER ACCESSED OVER TELEPHONE LINES
3	1986	Norman Hardy	COMPUTER SYSTEM SECURITY
4	1987	Andreas Pfitzmann, Michael Waidner	NETWORKS WITHOUT USER OBSERVABILITY
5	1988	Chris J. Mitchell	KEY STORAGE IN SECURED NETWORK
6	1989	Fred C. Piper	VOICE NETWORK SECURITY SYSTEM
7	1990	Donald Graji Mohnish Pabrai Uday Pahrai	METHODOLOGY FOR NETWORK SECURITY DESIGN
8	1991	L. Todd Heberlein	NETWORK SECURITY MONITOR
9	1992	John R. Corbin	APPARATUS AND METHOD FOR LICENSING SOFTWARE ON A NETWORK OF COMPUTERS
10	1993	Michael P.	COMPUTER NETWORK ABUSE
11	1994	Bruce E. McNair	SYSTEM AND METHOD FOR GRANTING ACCESS TO A RESOURCE
12	1995	Scott D. Hammersley, Arthur D. Smet, Peter M. Wottreng	METHOD AND APPARATUS FOR INTRAPROCESS LOCKING OF A SHARED RESOURCE IN A COMPUTER SYSTEM
13	1995	Daniel B. Clifton	RESOURCE ACCESS SECURITY SYSTEM FOR CONTROLLING ACCESS TO RESOURCES OF DATA PROCESSING SYSTEM
14	1996	Wei-Ming Hu	METHOD AND APPARATUS FOR AUTHENTICATING A CLIENT TO A SERVER COMPUTER SYSTEMS WHICH SUPPORT DIFFERENT SECURITY MECHANISMS

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

15	1997	Mark S. Miller, E. Dean Tribble, Norman Hardy, Christopher T. Hibbert	DIVERSE GOODS ARBITRATION SYSTEM AND METHOD FOR ALLOCATING RESOURCES IN A DISTRIBUTED COMPUTER SYSTEM
16	1998	Ian Foster, Carl Kesselman,Gene Tsudik, Steven Tuecke	A SECURITY ARCHITECTURE FOR COMPUTATIONAL GRIDS
17	1999	Daniel S. Glasser, Ann Elizabeth McCurdy, Robert M. Price	METHOD AND SYSTEM FOR CONTROLLING USER ACCESS TO A RESOURCE IN A NETWORK COMPUTING ENVIRONMENT
18	2000	Rajkumar Buyya, David Abramson, and Jonathan Giddy	AN ARCHITECTURE FOR A RESOURCE MANAGEMENT AND SCHEDULING SYSTEM IN A GLOBAL COMPUTATIONAL GRID
19	2001	Lalana Kagal, Tim Finin and Anupam Joshi	MOVING FROM SECURITY TO DISTRIBUTED TRUST IN UBIQUITOUS COMPUTING ENVIRONMENT
20	2002	Farag Azzedin and Muthucumaru Maheswaran	TOWARDS A TRUST-AWARE RESOURCE MANAGENT IN GRID COMPUTING SYSTEM
21	2003	Von Welch1 Frank Siebenlist2 Ian Foste	SECURITY FOR GRID SERVICES
22	2004	Ivan Krsul, Arijit Ganguly, Jian Zhang	VMPLANTS:PROVIDING AND MANAGING VM EXECUTION ENVIRONMENTS FOR GRID COMPUTING
23	2005	Daniel Olmedilla1, Omer F. Rana2, Brian	SECURITY AND TRUST ISSES IN SEMANTIC GRIDS
24	2006	David S. Linthicum	MOVING TO CLOUD COMPUTING STEP BY STEP
25	2007	Uzi Dvir	SECURITY SERVER IN THE CLOUD
26	2008	Mladen A. Vouk	CLOUD COMPUTING-ISSUES,RESEARCH AND IMPLEMENTATIONS
27	2009	Meiko Jensen,	ON TECHNICAL ISSUES OF CLOUD COMPUTING
28	2010	S. Subashini n, V.Kavitha	SECURITY ISSUES FOR CLOUD COMPUTING
29	2011	Luis M. Vaquero	SECURITY ISSUES IN CLOUD COMPUTING
30	2012 I	Deyan Chen1, Hong Zhao	DATA SECURITY AND PRIVACY PRESERVATION IN CLOUD COMPUTING
31	2012 A	Mohammed A. AlZain ,	CLOUD COMPUTING SECURITY SINGLE-MULTI CLOUDS
32	2013 C	Ming Li,	SCALABLE AND SECURE SHARING OF PERSONAL HEALTH RECORDS
33	2013 B	Miltiadis Kandias,	INSIDER THREAT IN CLOUD COMPUTING
34	2013 A	Niroshinie Fernando	MOBILE CLOUD COMPUTING-SURVEY
35	2014 D	Diogo A. B. Fernandes	SURVEY ISSUES IN CLOUD COMPUTING
36	2014 B	Md Whaiduzzaman	SURVEY ON VEHICULAR CLOUD COMPUTING

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

37	2014 A	A.Madhuri1, T.V.Nagaraju	RELIABLE SECURITY IN CLOUD COMPUTING ENVIRONMENT
38	2015 A	IbrahimAbaker	RISE OF BIG DATA ON CLOUD COMPUTING-REVIEW AND OPEN ISSUES
39	2015	TargioHashem	RISE OF CLOUD COMPUTING ARCHITECTURE IN BIG DATA
40	2015D	Gavin O Donnell,	CLOUD COMPUTING
41	2016	Sundas Iftikhar, Anum Tariq,	OPTIMAL TASK ALLOCATION ALGORITHM FOR COST MINIMIZATION AND LOAD BALANCING OF GSD TERMS
42	2016	Hamed Rezaei, Behdad Karimi, and Seyed Jamalodin	EFFECT OF CLOUD COMPUTING SYSTEM IN TERMS OF SERVICE QUALITY OF KNOWLEDGE MANAGEMENT SYSTEM
43	2017	Thanh Dat Dang	A FRAMEWORK FOR CLOUD BASED SMART HOME
44	2018	Christian Biener, Martin	INSURABILITY OF CYBER RISK