ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

Survey Paper of Privacy Preserving Data Analysis in Mental Health Research

Pokharkar Satish. T [1], Dr. Hare Ram Shah [2]

P.H.D Student, Department of Computer Engineering, Oriental University Indore, Madhyapradesh, India[1]

Associate Professor, Department of Computer Engineering, Oriental University Indore, Madhyapradesh, India[2]

ABSTRACT: The anonymization of wellbeing information streams is imperative to secure this information against potential protection breaks. Countless research studies going for offering protection with regards to information streams has been as of late led. Be that as it may, the strategies that have been proposed in these examinations generate noteworthy deferral amid the anonymization procedure, since they focus on applying existing protection models (e.g., k-namelessness and l-decent variety) to bunches of information extricated from information streams in a period of time. In this paper, we present postponement free anonymization, a system for saving the protection of electronic wellbeing information streams. In contrast to existing works, our strategy does not produce a gathering delay, since information streams are anonymized promptly with fake qualities. We further devise late approval for expanding the information utility of the anonymization results and dealing with the fake values .Through tests, we demonstrate the proficiency and viability of the proposed technique for the continuous arrival of information streams

KEYWORDS: Health data stream, Privacy anonymization

I. INTRODUCTION

As of late, information surges of wellbeing records have been broadly used in numerous applications, for example, continuous conclusion frameworks, biomedical information examination, and wellbeing checking administrations [1–4]. Finding frameworks, for instance, analyse patients' maladies by checking the conduct of their continuous wellbeing information, and biomedical information investigators think about the examples of infections by utilizing the

Information streams from bio-signal sensors. By and large, the information streams are endowed to outsiders who break down the information in spot of the information holders, in view of the compelling information use. Then, wellbeing information streams normally contain private characteristics, for example, age, sex, different bio-flag, and conclusions. Security to shield information streams from foundation learning assaults, a few security safeguarding techniques has been as of late proposed [7–14]. These techniques guarantee that the discharged information streams meet the prerequisites of run of the mill security models, for example, k-obscurity [5] and l-decent variety [6]. In spite of the fact that these techniques vary in the manner they change the information streams to offer security insurance, they are altogether founded on the tuples amassing technique. That is, spilling tuples are deferred until they fulfil the given distribution condition, i.e., achieve a specific security level or acceptable information utility. We consider this kind of technique the collection based strategy. Fig. 1 demonstrates a case of offering security assurance in information Streams by utilizing the collection based technique. Accept an info stream of tuples, where a tuple comprises of various qualities that incorporate individual characteristics, i.e., QI (semi identifiers) and SI (touchy data). In the aggregation based strategy, the anonymization framework keeps on gathering and bunch the info tuples until the given protection condition, for example, k-secrecy what's more, l-assorted variety, is fulfilled. In Fig. 1, tuple 1, which has touched base at time T1 and has been gathered in Cluster 1, sits tight for tuple 2? At the point when tuple 2 arrives, it is doled out to Cluster 1, at that point the group moves toward becoming 2-various. At the point when a specific bunch fulfils the security condition, the tuples in the bunch are summed up and discharged. At that point, the QIs of the tuples in the

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

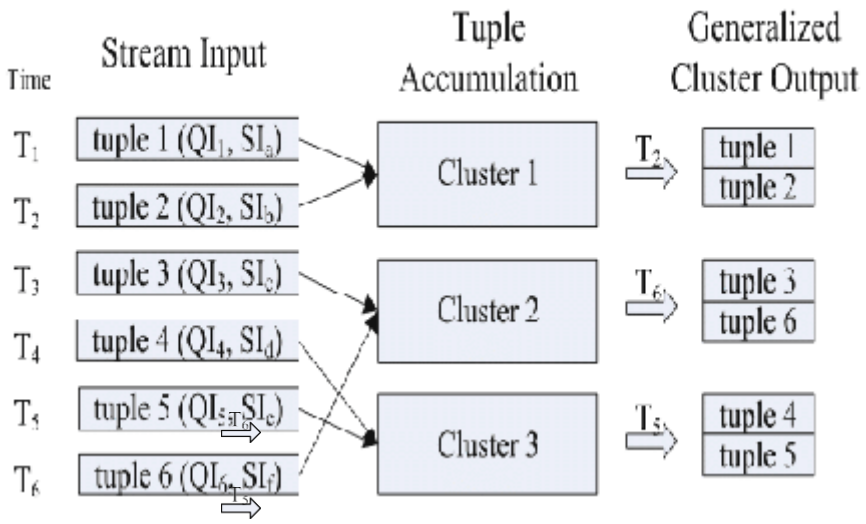

Fig. 1. Privacy protection of a data stream using the accumulation-based method (an example of offering 2-diversity).

bunch become indistinct from one another as indicated by the protection condition. Subsequently, a foe can't distinguish an accurate tuple from the anonymized information by utilizing the QIs (i.e., foundation learning). Therefore, an enemy can't be sure about the objective's SI, despite the fact that he or she can induce the rundown of potential SIs. Notwithstanding utilizing aggregation to fulfil the required security level, most existing techniques go for aggregating more tuples than should be expected, so as to lessen the data misfortune that is brought about by the connected information speculation. For instance, tuple 4 in Fig. 1 is gathered in Cluster 3 rather than Cluster 2, in such a case that it was gathered in Cluster 2 this would prompt expanded data misfortune. There are two unavoidable defer issues in the accumulation based techniques: amassing and calculation delay. The previous is straightforwardly significant to the collection of the info tuples. For model, in Fig. 1, the arrival of tuple 3 is deferred until tuple 6 has arrived. The deferral is because of the time taken to figure the data misfortune. To allot an info tuple to a suitable bunch that creates the base data misfortune, the data misfortune of each group must be determined. Figuring the data loss of all groups takes a lot of time, which forbids the arrival of huge information streams continuously. To take care of these issues, in this paper, we propose a DF (delay free anonymization) structure to safeguard the protection of electronic wellbeing information streams continuously. We consider two sorts of characteristics of an info tuple: the semi identifier (QI) and the delicate data (SI). In the DF structure, QIs are discharged with misleadingly produced l-different SI esteems. For instance, in Fig. 2, tuple 1 is anonymized with ST1, which is the l-various arrangement of the SI. We sort out the set with the first SI from the information tuple and the misleadingly created SI. At that point, QI1 has different SIs (SIa and SId), which fulfils 2-assorted variety. In this manner, assailants can't identify the careful SI of the objective QI. As such, our anonymization procedure safeguards the protection of the wellbeing information streams. We might want to take note of that there is no aggregation delay in DF since it promptly anonymized the info tuples and discharges them. This is conceivable on the grounds that there is no tuple amassing in our anonymization procedure.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

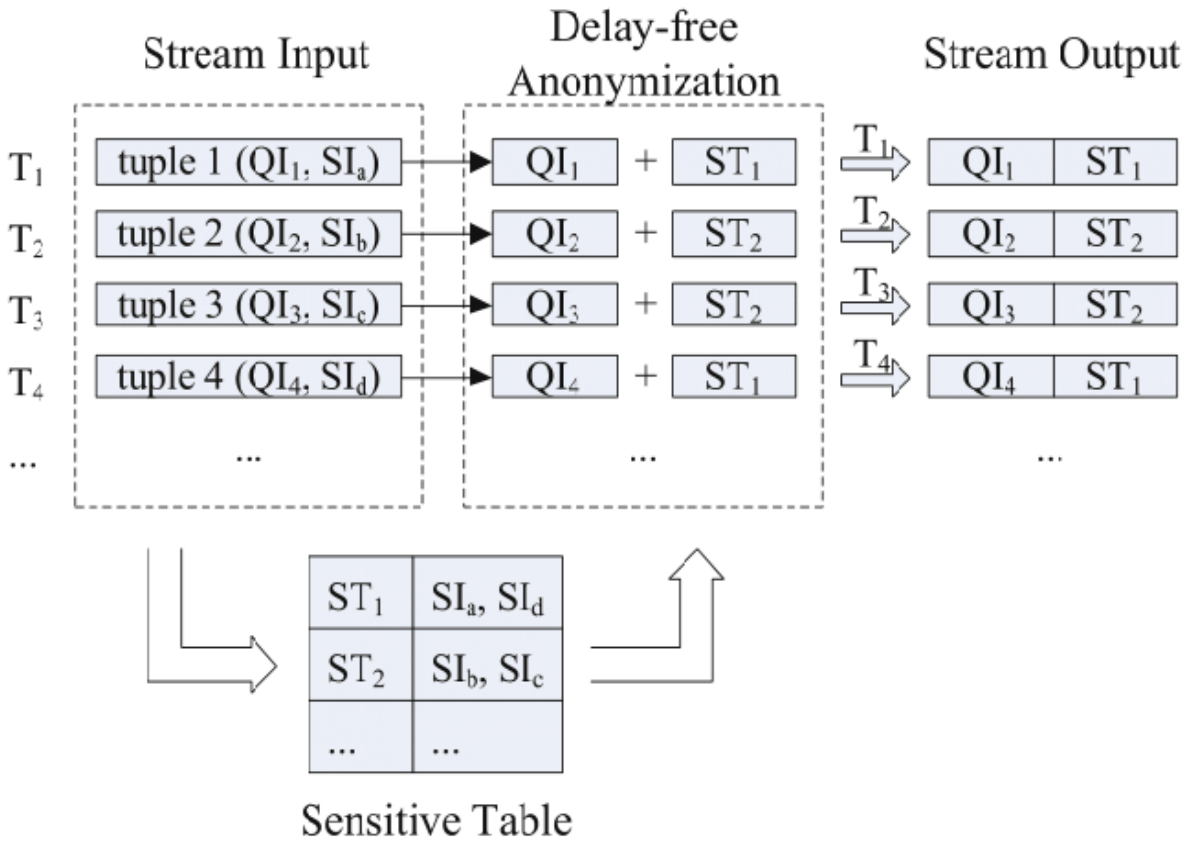

Fig. 2. Privacy protection of a data stream using the delay-free anonymization approach (an example of offering 2-diversity).

Moreover, in correlation with the gathering based strategies, DF produces strikingly little calculation delay, on the grounds that DF executes just straightforward tasks (i.e., alluding to the delicate table and overseeing it) rather than entangled tasks (e.g., computing the data misfortune).We likewise considered the information utility of the anonymization results. Our anonymization strategy, producing l-differing SI esteems, creates fake qualities. For instance, in Fig. 2, SId in ST1 is a fake at T1. Fake qualities work as a security instrument to anticipate the divulgence of the right SI esteems; notwithstanding, they additionally decline information utility. To enhance this issue, we propose late approval, which uses the SI estimations of the recently entered tuples to approve the current fakes. In Fig. 2, SId in ST1 is late approved at T4. After tuple 4, whose SI is SId, is anonymized utilizing ST1; SId in ST1 is viewed as a genuine esteem. To quantify the information utility, counting the utility misfortune from the fakes, we further devise touchy characteristic vulnerability as the quality measurement for DF. The commitments of our postponement free anonymization structure are outlined as pursues:

Immediate release: DF encourages tuple-by-tuple discharge with the certification of l-assorted variety. DF is portrayed by no gathering delay and a low calculation cost, since it right away discharges the information streams and requires just straightforward tasks. The quick discharge is a major preferred position for applications that use continuous information streams.

High-level data utility: To anonymized information streams, DF falsely produces l-different SI esteems as opposed to summing up QIs. Therefore, DF does not create the common data misfortune with regard to QIs. Furthermore, the fakes in the touchy table are likewise limited by late approval.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

II. RELATED WORK

A huge number of research considers in the territory of information protection has been directed in the most recent decade. In what pursues, we first present techniques that have been proposed to offer protection in the setting of static social information. Following that, we talk about methodologies that have been proposed for offering dynamic information protection what's more, information stream protection. The last are legitimately important to
Our methodology.

Privacy for static, relational data: A lot of work has been led in the territory of protection for social information. A few investigations (e.g., [5, 15, and 16]) proposed strategies that utilize k-secrecy to square re-recognizable proof assaults, in which people in the discharged micro data are recognized in light of QIs. To take care of this issue, k-obscurity makes k tuples vague by summing up the QIs. L-assorted variety is a model that explains the absence of assorted variety in k-namelessness [6]. It makes SIs l-different with indistinct QIs. Life systems [17] is a procedure
That likewise utilizes the l-assorted tuples also to the technique proposed in [6]; in any case, it doesn't sum up QIs. Consequently, in spite of the fact that an aggressor knows the accurate target QI is incorporated into the dataset, she
or on the other hand he can't recognize which SI has a place with the objective. The creators of t-closeness [18] proposed a technique that locations certain impediments of l-decent variety. Two potential assaults can be made against l-assorted variety: a skewness assault and similitude assault. So as to avoid these assaults, t-closeness necessitates that the conveyance of the touchy property estimations meets a specific condition. The previously mentioned investigations manage social information and try not to think about any inclusion, erasure or update activities connected on this information. Subsequently, these techniques are not relevant with regards to information streams

Privacy for Dynamic, relational data: Additions, cancellations, and updates much of the time happen in genuine datasets. A few strategies have been proposed to ensure progressively advancing datasets, guaranteeing that the most recent variant of the information can be securely distributed. It is critical to take note of that past privacy protection
Strategies, including static information, are not material in the setting of dynamic datasets on the grounds that they expect just a solitary information discharge. In powerful datasets, another sort of security issue can emerge in light of the potential linkage between various information discharges. Accordingly, a few protection conservation systems [19– 21] secure the protection of dynamic discharges by considering their setting or keeping up additional data.
Byun et al. introduced an anonymization strategy for dynamic information; notwithstanding, this work thought about just inclusion tasks [19]. Xiao et al. in [20] proposed m-invariance, a security strategy which thinks about both inclusion and cancellation tasks. Bu et attended to another protection issue that happens by the lasting
Touchy qualities in the dynamic datasets [21]. The strategies proposed in these examinations manage numerous discharges, which are practically equivalent to information streams. Notwithstanding, they can't
be utilized to anonymized information streams since they manage the anonymization of the whole table and can't be connected on information that ends up accessible progressively.
Data stream Privacy: As of late, there have been a few investigations on protecting the security of information streams. Information streams comprise of arrangements of tuples transmitted progressively, and henceforth, the suspicion concerning the prerequisites of the protection conservation system is unique from that in the past security insurance techniques. To save the protection of information streams, the unit of anonymization must be a tuple, and the anonymization ought to be performed continuously. Supposedly, the principal examine that managed the security safeguarding of information streams is [9], in which the specialization tree for gathering information streams was presented. Hubs in the tree are arranged into hopeful hubs and work hubs, and information streams are summed up by the hub structure. The CASTLE plan creates groups in which the tuples of information streams are gathered [7]. At that point, it discharges the groups at the point when the measure of the gathering achieves the deferral requirement, d. Fundamentally, CASTLE ensures k-namelessness, and furthermore gives a calculation for l-decent variety. The technique introduced in [8] utilizes a likelihood capacity to decide the arrival of information streams; it additionally amasses information from information streams and chooses the discharge time as indicated by this capacity. The capacity advances discharges when the gathered information experiences less data misfortune yet a more extended deferral. For information utility, the strategy proposed in this examination moreover makes an expectation on the

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

information that may show up in the streams by considering the information distribution. The creators of Sabre [14] displayed a calculation for anonym zing micro data to fulfil t-closeness, and furthermore expanded the
Calculation to anonymized information streams. The Sabre structure keeps up sliding windows that capacity as cushions of the info tuples. At the point when tuples in a specific window are lapsed due to recently embedded
tuples, old tuples ought to be discharged. These tuples are anonymized by Sabre and discharged to a yield stream.
The creators of B-CASTLE [10] displayed an improved calculation of CASTLE [7]. The calculation of B-CASTLE thinks about the appropriation of information in information streams when the information is grouped, in
a push to improve information utility. The strategy introduced in [12] too considers the thickness circulation to improve the utility of the information in the anonymized stream. The strategy builds a TDS-tree (top down specialization tree) of QIs, and thinks about the thickness conveyance of tuples in each leaf of the tree. The creators of SANATOMY [11] utilized the bucketization approach (for example Life systems [17]) to anonymized information streams, rather than the speculation approach. SANATOMY jam the information relationship, because of the bucketization approach. The creators of [13] called attention to that the past studies had been founded on tedious calculations due to the k-obscurity model. The calculation introduced in [13] is an effective anonymization calculation, which is bunch based and bolsters numerical information streams

III. THE DELAY FREE ANONYMIZATION FRAMEWORK

In this area, we present the postponement free anonymization (DF) system for securing the protection of the information in wellbeing information streams. To start with, we characterize the information model of information streams. At that point, we depict our proposed strategy in detail.
Data model : We expect ongoing wellbeing information streams containing individual data. Among the different properties in an information stream, we centre on just three fundamental credits so as to improve the issue: tupleID, QI, and SI. As indicated by the characteristics, an information stream can be depicted as ðtupleID; QIs; SIÞ TupleID shows the one of a kind number of a tuple. It is typically evacuated before discharging the information stream, since it tends to be an identifier of the tuple. The QI is a characteristic of a person in the tuple, for example, age, sex, nationality, and zipcode. The SI is the private characteristic, which ought not be legitimately identified with the person,
for example, bio-sign and conclusion. In the accompanying meaning of the information stream, we discard the tupleID, since it is significantly expelled in the anonymization procedure.

Delay free anonymization:

The objective of the delay-free anonymization plot is to organizel-diverse information streams in genuine time, and ensures that the probability of speculating the individual's delicate information should be break even with to or less than 1/l. In arrange to anonymized data streams, we embrace the thoughts behind the Life systems procedure [17], which jam protection by isolating the QI and SI. At that point, we inflate each si of the input tuples into l-diverse values. Accept that a tuples QIt ; sitÞ has arrived. DF produces a QIT (quasi-identifier tuple) and an ST (expanded touchy tuple) and discharges them promptly. The QIT is created from QIt within the tuple t and the ST is an l-diverse setof the counterfeit si that contains sit. The QIT and ST can be joined by the connect key groupID.

So as to ensure protection of information streams, we go for ensuring Privacy preservation:l-decent variety on DF, which enables foes to figure the person's touchy data with a likelihood equivalent to or under 1/l. Since the point of l-assorted variety on DF is marginally unique from the conventional l-assorted variety, in this area, we talk about the protection safeguarding accomplished by DF and characterize l-assorted variety in the unique circumstance of DF. In the first place, so as to ensure the 1/l likelihood, it is basic that the quantity of unmistakable si esteems in a subjective gathering are equivalent to or on the other hand more than l. Think about a subjective gathering j and its touchy tuples STj. The primary condition for l-assorted variety on STj is l 6 STj ð1Þ It ought to be noticed that STj, which means the measure of STj, additionally speaks to the quantity of particular qualities in STj, in light of the fact that there is no copied si in STj (we utilize the tally an incentive to depict rehashed
Delicate qualities). Second, the recurrence of a delicate esteem ought to be considered. Give us a chance to concentrate on the include trait in the ST. In Example 1, the check estimation of every si is 1. Along these lines, the proportion of the Diag.A and the Diag.A is 1:1 for the QI of t1. It is 2-differing, in light of the fact that the attacker who has

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

foundation learning (QI of t1) can't recognize a definite delicate incentive with multiple/2 likelihood. In the mean time, let us expect the tally estimation of Diag.A = 2 and Diag.A = 1. At that point, the proportion winds up 2:1 and the likelihood that the QI of t1 relates to Diag.A moves toward becoming 2/3, which is more noteworthy than 1/2 likelihood. We can infer an end that they include values in the ST ought to be produced with the end goal that they ensure the 1/l likelihood

IV. EFFECTIVE COUNTERFEIT MANAGEMENT

Sensitive group management: Late approval is a procedure that replaces a current fake with a legitimate touchy esteem utilizing an information tuple that will be anonymized. For instance, let us expect an info tuple t1 has been anonymized as Example 1. The anonymized results QIT and ST may be deciphered by applications as The touchy estimation of ð24;MÞ is identified with ST00 There is only one tuple in QIT1, though the aggregate of the include ST1 jjST1jj is 2, which means there is one fake touchy esteem (Diag.A) notwithstanding the genuine esteem (Diag.A) in ST1. At that point, expect an info tuple t2 has arrived and its touchy esteem is a Diag.A. On the off chance that t2 is anonymized as indicated by DF as depicted in Section 3.2, QIT2 and ST2, which incorporates Diag.A as a delicate esteem, are created. Presently, we have to concentrate on the fake estimation of ST1, Diag .B. Utilizing this esteem, rather than anonym zing the information tuple, we can relate the Diag. B to ST1 as The touchy estimation of ð32; Fþ is identified with ST00 This implies t2 has been incorporated into gathering 1 and there are two tuples in QIT1. Since jjST1jj is 2, applications may decipher the outcome as that both Diag.A and Diag .B are legitimate qualities. In this way, the age of fakes can be diminished and existing fakes can be "late approved" with genuine qualities.

Sensitive value generation: the l-different touchy qualities are self-assertively produced at the point when ST is sorted out. The discretionary age of delicate qualities may significantly affect information utility and security assurance.

The produced qualities essentially influence the information utility coming about from the procedure generally approval, on the grounds that late approval is conceivable when the produced touchy esteem coordinates the estimation of the information tuple. For information utility, it is essential to create touchy qualities that are probably going to show up in the information stream. A decent expectation of the information would be the most ideal approach to increment late approvals. To accomplish this, in this work we accept that what's to come is reflected by the past. All the more explicitly, we use the past information for creating the touchy qualities. In the first place, we set up a pool of touchy qualities from the past information streams. A component of the pool comprises of ðsi; countðsiÞþ, which is closely resembling the component of ST. At that point, by arbitrarily picking components from the pool, ldiverse ST is sorted out. It is realized that the l-assorted variety model experiences two assaults (sleekness assault and likeness assault) regarding the protection conservation. Consider a gathering that has two qualities: influenza and hack. Despite the fact that it fulfills 2-assorted variety, the patient's security may not be saved since the qualities are semantically comparable. As another precedent, consider 3-assorted estimations of circulatory strain (150; 151, What's more, 153). The patient's protection may not be saved since the scope of the qualities is excessively thick. The protection of l-differing information is not appropriately saved when the delicate qualities are comparable. All together to square such assaults, Ninghui Li et al. proposed the idea

The algorithm:

1: $QIT=0$; $ST=0$; $RT=0$
2: $gent=O$; $qcnt=O$; $scnt=0$
3: let qit,st ,and rtbeelements of QIT ,ST and RT ,respectively
4: while $t \in$ is not empty do
5: letQtandsitbet[Qna ndt[sz] .respectively
6: letQSmbeasetofallqit[Q]where$qit[grou\ p/D\ ]=ID$
7: letstibeara ndomlychosens$t$where
st[si]=Sit,rti[cou n t]<st; $[\ count\ ]$and$Qc¢QISst[groupID]$
8: ifst; isnot$null$then

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

```
9:          releaseqitqcnt( st[group/D],Qld
10:         insertqitqcntintoQIT
11:         qcnt=qcnt+1
12:         u pda tert;into( rtdgroup/D],rt;§i],rtdcoun t]+1 )
13:     else
14:         releaseqitqcnt( gcnt,Qlt)
15:         insertqitqcntintoQIT
16:         qcnt=qcnt+1
17: letLSTbetheI-diversesetofpairs (si, count) returnedbyl-diverse_5T_Generation ( sic,l)
18:         for     all     lstELSTdo
19:             releasestscnc(gcnt,lst[ si],lst[count])
20:             insertstscntintoST
21:             if   lst[ si]=sicthen
22:                 insertrtscnt( gent,lst[si],1)intoRT
23:             else
24:                 insert rtscnt ( gcnt,lst[si],O)intoRT
25:             end     if
26:             sent=sent+1
27:         end     for
28:         gent=gent+1
29:     end
30: end while
```

Data utility considerations:

Attributes of Adult dataset

Attribute	Type	No. of distinct values
Age	Continuous	74
Education	Continuous	16
Work class	Categorical	8
Marital	Categorical	7
Race	Categorical	5
Sex	Categorical	2
Native country	Categorical	40
Salary occupation	Sensitive	30

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

Attributes of NPS dataset:

Attribute	Type	No. of distinct values
Age	Continuous	107
Length of stay	Continuous	53
Sex	Categorical	2
Location of hospital	Categorical	16
Admission route	Categorical	6
Disease	Sensitive	4848

$$SAU(T) = \frac{\sum_{st \in ST} [Count](T) - \prod_{\in RT} [Count](T)}{\sum_{st \in ST} [Count](T)}$$

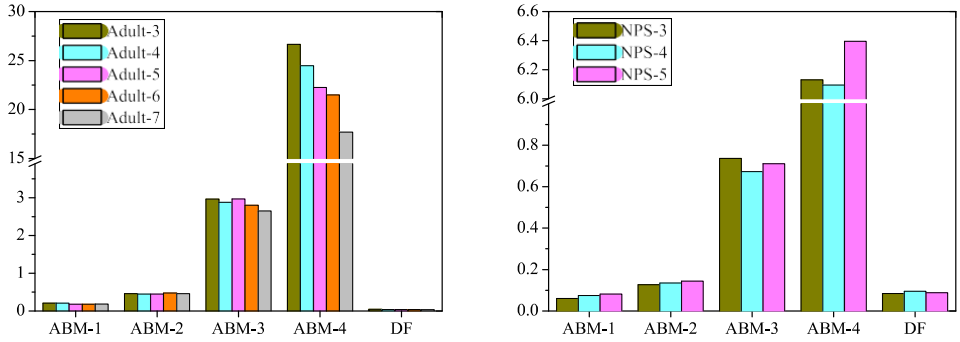

Fig. 3. Information loss.

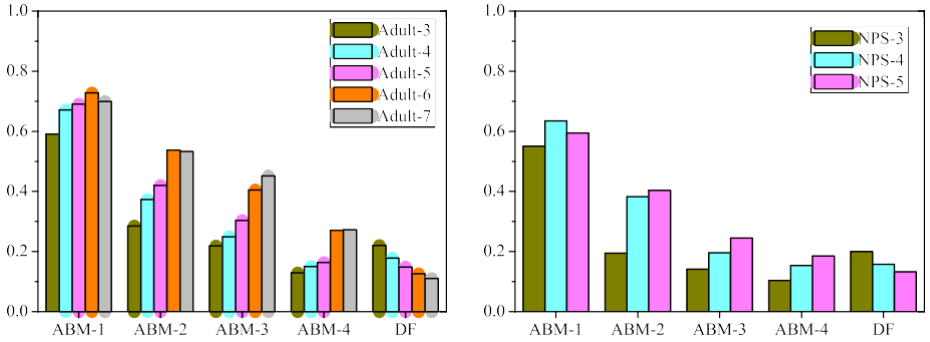

Fig. 4 Average processing time per tuple by No. of records

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

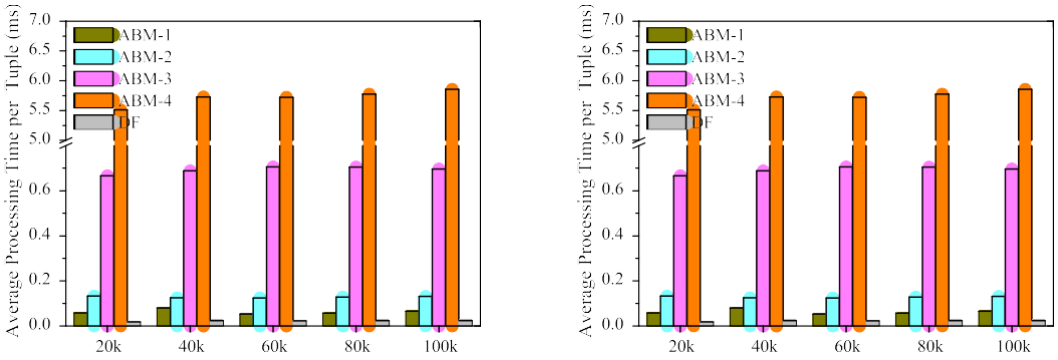

Fig. 4 Average processing time per tuple by No. of records.

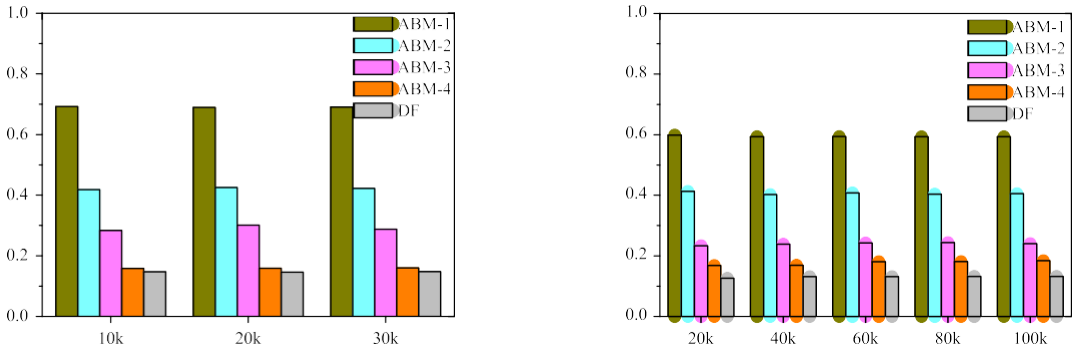

Fig. 5 Information loss by No. of records.

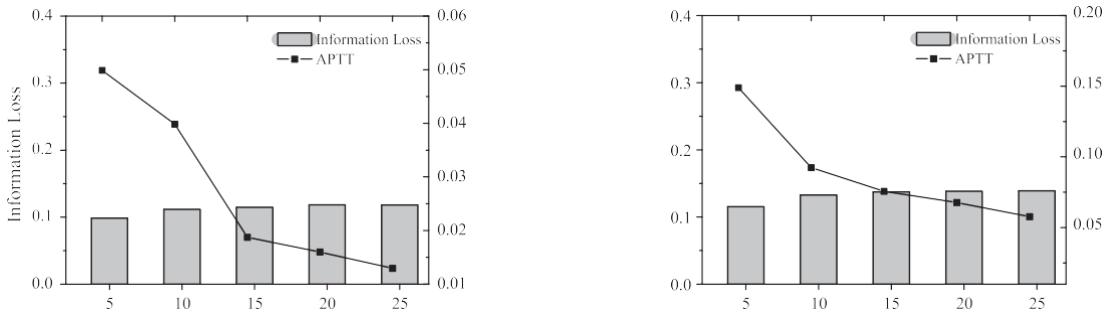

Fig. 6 Effect of L

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

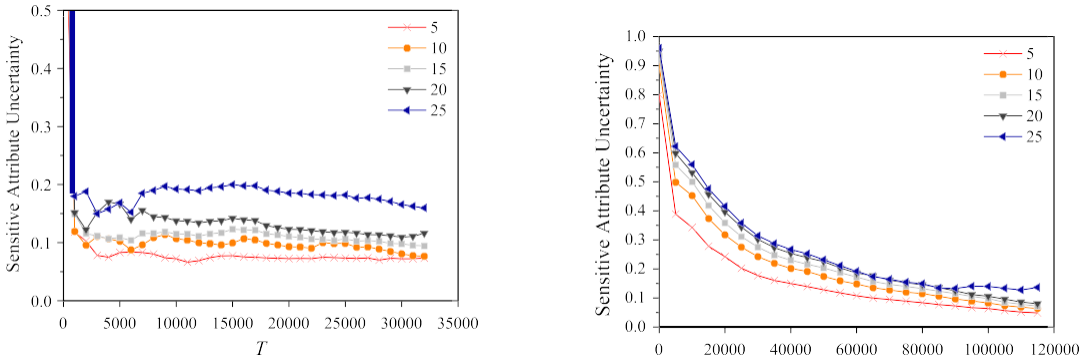

Fig. 7Sensitive attribute uncertainty

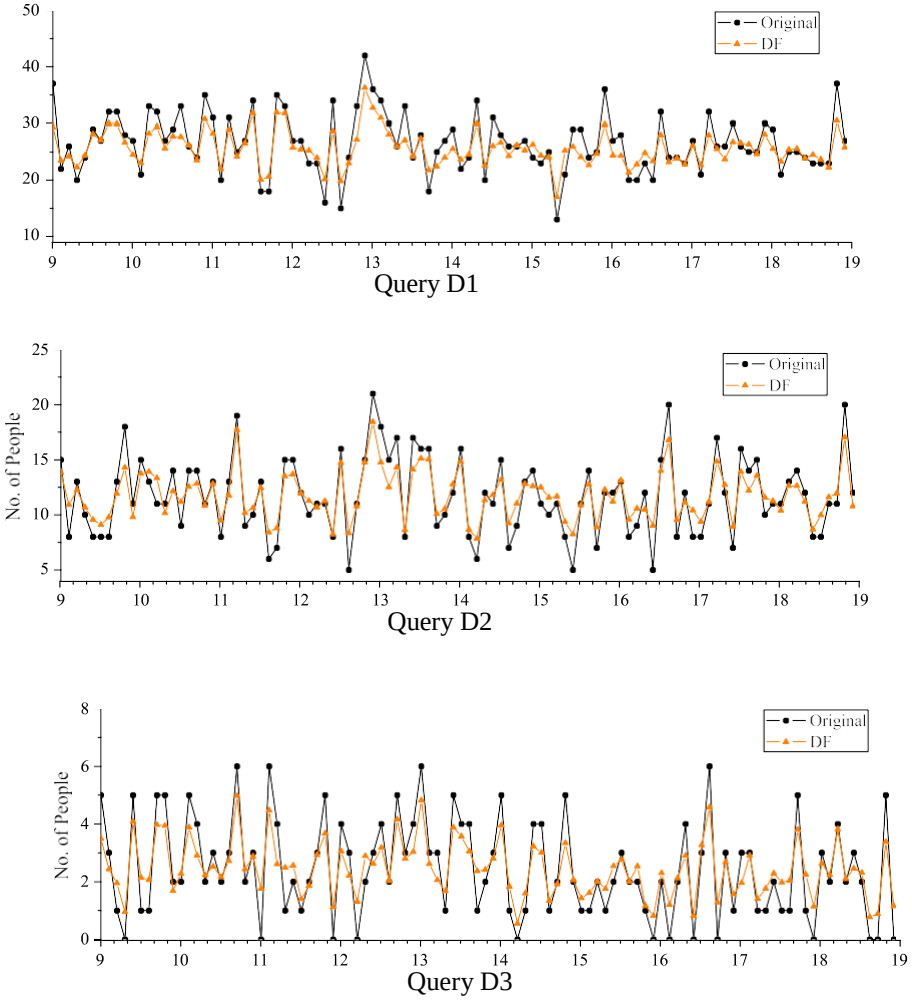

Fig 8 Results of the analysis queries on the demographic data stream

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

V. CONCLUSION

In this paper, we exhibited a postponement free anonymization technique for saving the protection of electronic wellbeing information streams. The proposed strategy does not create a critical deferral during the anonymization procedure and arrival of information streams. Since each QI has l-different touchy qualities, the protection of the information is safeguarded with a likelihood of 1=l. Late approval essentially diminishes the quantity of fake qualities, and thus, the anonymization result has high information utility. Our technique survives the downsides of existing gathering based techniques in real time information stream anonymization. A few future investigations can be considered as an expansion of the concentrate detailed in this paper. In the first place, we will devise a strategy for creating fake and check esteems. In the event that the after effect of measurable investigation or on the other hand past information are considered for producing the qualities, more dependable anonymization results can be figured it out. Second, it would be important to research the planning of the late approval. In this paper, we centre on coordinating info tuples and fakes for the late approval. The adequacy can be improved if the slipped by time of the fakes is considered. Also, we plan to utilize the Slicing [23] rather than the Anatomy [17] method for DF anonymization. We anticipate that the relationship between the QIs and the touchy characteristics will be saved by the Slicing method. At last, we might want to plan and study a conveyed adaptation of our structure. As the volume of electronic wellbeing information streams is quickly developing, information streams progressively should be prepared in a circulated situation.

REFERENCES

[1] Babcock B, Babu S, Datar M, Motwani R, Widom J. Models and issues in data stream systems. In: Proceedings of the twenty-first ACM SIGMOD-SIGACT- SIGART symposium on principles of database systems. ACM; 2002
. p. 1–16.
[2] Abadi DJ, Carney D, Çetintemel U, Cherniack M, Convey C, Lee S, et al. Aurora: a new model and architecture for data stream management. Int J Very Large Data Bases 2003;12(2):120–39.
[3] Gaber MM, Zaslavsky A, Krishnaswamy S. Mining data streams: a review. ACM Sigmod Record 2005;34(2):18–26.
[4] Bifet A, Holmes G, Pfahringer B, Kirkby R, Gavaldà R. New ensemble methods for evolving data streams. In: Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2009. p. 139–48.
[5] Sweeney L. k-anonymity: a model for protecting privacy. Int J Uncertain, Fuzz Knowl-Based Syst 2002;10(05):557–70.
[6] Machanavajjhala A, Kifer D, Gehrke J, Venkitasubramaniam M. l-diversity: privacy beyond k-anonymity. ACM Trans Knowl Discov Data (TKDD) 2007;1(1):3.
[7] Cao J, Carminati B, Ferrari E, Tan K-L. Castle: continuously anonymizing data streams. IEEE Trans Dependable Secure Comput 2011;8(3):337–52.
[8] Zhou B, Han Y, Pei J, Jiang B, Tao Y, Jia Y. Continuous privacy preserving publishing of data streams. In: Proceedings of the 12th international conference on extending database technology: advances in database technology. ACM; 2009. p. 648–59.
[9] Li J, Ooi BC, Wang W. Anonymizing streaming data for privacy protection. In: Proceedings of IEEE 24th international conference on data engineering. IEEE; 2008. p. 1367–9.
[10] Wang P, Lu J, Zhao L, Yang J. B-castle: an efficient publishing algorithm for k-anonymizing data streams. In: Proceedings of 2010 second WRI global congress on intelligent systems (GCIS), vol. 2; 2010. p. 132–6.
[11] Wang P, Zhao L, Lu J, Yang J. Sanatomy: privacy preserving publishing of data streams via anatomy. In: Proceedings of the 2010 third international symposium on information processing (ISIP). IEEE; 2010. p. 54–7.
[12] Zhang J, Yang J, Zhang J, Yuan Y. Kids: K-anonymization data stream based on sliding window. In: Proceedings of the 2010 2nd international conference on future computer and communication (ICFCC), vol. 2; 2010. p. V2-311–6.
[13] Zakerzadeh H, Osborn SL. Faanst: fast anonymizing algorithm for numerical streaming data. In: Data privacy management and autonomous spontaneous security. Springer; 2011. p. 36–50.
[14] Cao J, Karras P, Kalnis P, Tan K-L. Sabre: a sensitive attribute bucketization and redistribution framework for t-closeness. VLDB J 2011;20(1):59–81.
[15] LeFevre K, DeWitt DJ, Ramakrishnan R. Incognito: efficient full-domain k-anonymity. In: Proceedings of the 2005 ACM SIGMOD international conference on management of data. ACM; 2005. p. 49–60.
[16] Samarati P. Protecting respondents identities in microdata release. IEEE Trans Knowl Data Eng 2001;13(6):1010–27.
[17] Xiao X, Tao Y. Anatomy: simple and effective privacy preservation. In: Proceedings of the 32nd international conference on very large data bases, VLDB endowment; 2006. p. 139–50.
[18] Li N, Li T, Venkatasubramanian S. t-closeness: privacy beyond k-anonymity and l-diversity. In: Proceedings of IEEE international conference on data engineering; 2007. p. 106–15.
[19] Byun J-W, Sohn Y, Bertino E, Li N. Secure anonymization for incremental datasets. In: Secure data management. Springer; 2006. p. 48–63.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

[20] Xiao X, Tao Y. M-invariance: towards privacy preserving re-publication of dynamic datasets. In: Proceedings of the 2007 ACM SIGMOD international conference on management of data. ACM; 2007. p. 689–700.
[21] Bu Y, Fu AWC, Wong RCW, Chen L, Li J. Privacy preserving serial data publishing by role composition. Proceedings of the VLDB endowment 2008;1(1):845–56.
[22] Iyengar VS. Transforming data to satisfy privacy constraints. In: Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2002. p. 279–88.
[23] Li T, Li N, Zhang J, Molloy I. Slicing: a new approach for privacy preserving data publishing. IEEE Trans Knowl Data Eng 2012;24(3):561–74