I.J. Information Technology and Computer Science, **2019, 4, 18-32**
Published Online April 2019 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2019.04.03

A Novel Dynamic KC$_i$ - Slice Publishing Prototype for Retaining Privacy and Utility of Multiple Sensitive Attributes

N.V.S. Lakshmipathi Raju
Dept. of CSE, GVP College of Engineering (A), Visakhapatnam, Andhra Pradesh, 530048, India
E-mail: suribabu205@gvpce.ac.in

M.N. Seetaramanath
Dept. of IT, GVP College of Engineering (A), Visakhapatnam, Andhra Pradesh, 530048, India
E-mail: seetaramanath@gmail.com

P. Srinivasa Rao
Dept. of CS&SE, A.U. College of Engineering, Visakhapatnam, 530003, India
E-mail: Peri.srinivasarao@yahoo.com

Received: 17 December 2018; Accepted: 13 February 2019; Published: 08 April 2019

***Abstract*—**Data publishing plays a major role to establish a path between current world scenarios and next generation requirements and it is desirable to keep the individuals privacy on the released content without reducing the utility rate. Existing KC and KC$_i$ models concentrate on multiple categorical sensitive attributes. Both these models have their own merits and demerits. This paper proposes a new method named as novel KC$_i$ - slice model, to enhance the existing KC$_i$ approach with better utility levels and required privacy levels. The proposed model uses two rounds to publish the data. Anatomization approach is used to separate the sensitive attributes and quasi attributes. The first round uses a novel approach called as enhanced semantic l-diversity technique to bucketize the tuples and also determine the correlation of the sensitive attributes to build different sensitive tables. The second round generates multiple quasi tables by performing slicing operation on concatenated correlated quasi attributes. It concatenate the attributes of the quasi tables with the ID's of the buckets from the different sensitive tables and perform random permutations on the buckets of quasi tables. Proposed model publishes the data with more privacy and high utility levels when compared to the existing models.

***Index Terms*—**Novel KC$_i$-slice, Data utility, Privacy, Slicing, High sensitive attribute, Low sensitive attribute.

I. INTRODUCTION

Many organizations such as medical institutions, government sectors and business organizations publish their data at regular intervals. The published data becomes available to the public and helps to serve as input for further research. This automatically provides the different directions to the scientists to meet the future requirements. Organizations can also be responsible to secure the confidentiality of the people involved in the published data and should also provide a scope to improve the usefulness of the published data.

Published data should be anonymized by the organizations with appropriate anonymization approaches to meet the objectives of the data publishing. Generally, published data contains attributes such as key attributes, sensitive attributes and quasi identifiers [32]. Key attributes can reveal the individuals identity and hence can't be considered for data publishing. Sometimes, it is also possible to obtain by grouping the values of quasi identifiers may expose the privacy of the individuals included in the data set when they are pointed with external repositories [1]. Sensitive attributes have confidential information when compared with other attributes.

Anonymization of the published data can be done only on the basis of publishing type such as one time publishing and dynamic publishing. Different data publishing techniques are K-anonymity, t-closeness, l-diversity, KC – slice, m-invariance and so on [29]. These techniques use either generalization or suppression or anatomization approaches to prevent the published data from various attacks [23]. It should be necessary to reduce the privacy disclosure rate to an acceptable level and also maximize the benefit of the published data [31]. Utility of the published data can be high, only when it contains more number of sensitive attributes. Some of the PPDP models are limited to the data sets which contain single sensitive attribute only [22]. In general data sets contain multiple sensitive attributes. Hence the PPDP models for multiple sensitive attributes are to be considered at most important.

Publishing the models for multiple sensitive attributes is to be considered with extra care and attention. These models should preserve the confidentiality of the people included in the released content from different types of attacks including back ground knowledge attack, membership attack and various other attacks and also these models should exhibit more useful information. Existing PPDP models on multiple sensitive attributes do not satisfy the goals with respect to privacy and utility. The proposed novel KC_i– slice model also considers multiple categorical sensitive attributes while publishing the data.

Privacy and utility of a sensitive attribute are inter dependent [13]. Dataset is published with high privacy means that it is not possible to expect high utility rate from it. From the published data, it should be necessary to acquire the required utility levels and privacy levels with respect to sensitiveness of the attributes. The proposed novel KC_i – slice model moves in this path by acquiring the necessary utility levels and privacy levels on the sensitive attributes and chooses all the sensitive values of a sensitive attribute having equal priority. This model also considers attribute sensitiveness. Based on the sensitiveness of the sensitive attributes, this model considers the sensitive attributes as either high sensitive attributes or low sensitive attributes to acquire more privacy levels with good utility rates. High sensitive attribute is one which contains more number of high sensitive values with contrast to low sensitive attributes, which can be determined by a threshold level [7].

The remaining part of the paper is organized as follows. Section II explains related work, Section III describes the methodology of the work, Section IV reveals the results analysis, Section V explains the attacks against privacy protection and Section VI focuses on the conclusion with future extensions.

II. Related Work

In today's world, different models are available to publish the data of multiple sensitive attributes. Each model have their own strengths and weaknesses.

Liu et al [9] introduced an approach called as rating for publishing sensitive data and to maintain the correlations among the published data. This approach considers the sensitivity degree of the attributes while publishing the data and also introduces the algorithms to publish the tables. It implements each attribute in an independent manner, which can affect the relations among sensitive attributes.

Liu et al. [7] described a model for numerous sensitive attributes which discriminate the high sensitive attributes from low sensitive attributes using greedy approach and new k-anonymity technique. This model applies the greedy approach to sort the tuples of high sensitive attributes and break the associations among high sensitive values. It can secure the published content from different attacks and requires more execution time.

Yuki et al. [24] proposed an approach for anonymization and analysis of horizontally and vertically divided user profile databases with multiple sensitive attributes. This approach inserts the duplicate values into the original data to protect the data from various types of attacks while publishing. It also uses a separate method to reduce the influence of the noise generated from duplicate insertion and also shows the anonymization method when the participants were only two organizations. It was focused on the analysis of sensitive attributes, but it should necessary to concentrate on the analysis of both quasi attributes and sensitive attributes.

Ashoka et al. [26] described a model on enhanced utility in preserving privacy for multiple heterogeneous sensitive attributes using correlation and personal sensitivity flags. This model concentrated on both numerical and categorical sensitive attributes by also considering the personal sensitivity flags of the record owners. It can also find out the highly correlated sensitive attributes by using correlation. Anonymization operation can be done only on the tuples which require the privacy and publish the remaining tuples as it is to improve the utility rate of the published data. But this work is suitable only for one time publication and does not concentrated on multiple releases.

Maheshwarkar et al. [14] implemented an approach to secure the confidentiality of the published data for numerous sensitive attributes by adding an extra record to maintain diversity and anonymity of the equivalence classes based on the k-factor of the data base. This model can reduce the distortion without performing any tuple suppression. But the k-anonymity may be attacked by using backdrop awareness.

Han et al. [4] demonstrated an approach for multiple sensitive attributes micro data termed as SLOMS. This method uses slicing to split the micro data into different tables and l-diversity algorithm is used to group the tuples. SLOMS approach anonymizes the quasi attributes by using K-anonymity algorithm and MSB-KACA technique is used to anonymize the micro data with numerous sensitive attributes. This model generates large suppression and it takes extra execution time.

Shyamala susan et al. [20] implemented a new model for numerous sensitive attributes using anatomization and slicing. This approach uses vertical partitioning to break the relations between the different sensitive attributes and generate the various sensitive tables by an advanced clustering algorithm. It uses MFA algorithm to bucketize the tuples. This model reduces the suppression rate by using anatomization algorithm and limits the high dimensionality by using slicing algorithm. It safeguards the data against different attacks. This model generates the quasi identifier table and sensitive tables by using slicing technique in an individual manner.

Abou_el_ela abdo hussain et al. [27] introduced a model for multiple published tables privacy preserving data mining. This model introduces techniques to secure the individual privacy when the organization releases more than one table. This paper publishes the two tables by using the concepts of (α,k)-anonymity with lossy join and anatomy technique. It publishes the separate tables for sensitive values and quasi values. But this work considers a single sensitive attribute in each published table. Aldeen et al. [30] also used (a,k) anonymity technique for their experimental investigation to ensure privacy among published tables.

Usha et al. [21] implemented a model by considering the sensitiveness of the attributes using k-anonymity and non-homogeneous anonymization for numerous sensitive attributes. Generalization can be performed on the basis of high sensitive attribute. The anonymization of the quasi attributes can be decided only on the basis of cluster sensitiveness. But this model was not implemented on original data sets. Also, it can be cracked by backdrop awareness.

Yifan ye et al. [28] described an anonymization approach by including both anatomy and permutation for securing the privacy in micro data with multiple sensitive attributes. This paper anonymizes the data by using the algorithms called as naive multi sensitive bucketization permutation and closest distance multi sensitive bucketization permutation. This approach is suitable only for one time release and it was not concentrated on multiple releases.

Qinghai Liu et al. [12] demonstrated an approach by using the techniques of clustering and multi-sensitive bucketization for multiple numerical sensitive attributes named as MNSACM. This method considers only two numerical sensitive attributes and does not explain the methodology for more than two numerical sensitive attributes. This method can be applicable to numerical sensitive attributes and which is not suitable for categorical attributes. It is limited to onetime publication only.

Hassan et al [5] described a model to preserve the confidentiality on multiple independent data publishing. This model concentrates on the new direction and introduces a composition attack. It can also provide a solution called as cell generalization to improve the utility rate and individual's confidentiality of the released content. This method combines the different published data sets by maintaining the l-diversity requirements and it was implemented on single sensitive attribute only.

Anjum et al. [1] introduced a model for numerous sensitive attributes using (p,k) - Angelization technique to protect the data from different types of attacks. This model publishes a sensitive table for all the sensitive attributes and a quasi table for all the quasi attributes. It generalizes the quasi values of a quasi table to protect the data from demographic attacks and also increases the privacy levels and utility levels of the published data. It is confined to the case that only a single tuple of a person shou ld be presented in the data.

Ram Prasad Reddy et al [19] demonstrated a new concept on personalized privacy model for numerous sensitive attributes. This approach uses deterministic anonymization on quasi attributes, generalization on categorical sensitive attributes and fuzzy approach on numerical sensitive attributes. Slicing technique is used to partition the dataset into different slices. This method considers only the static data and not assume that each categorical group should have only one occurrence of a value and duplicate value can be represented by higher level value.

Onashoga et al. [18] implemented a model using the concepts of LKC privacy technique and slicing named as KC - slice. LKC model generalizes traditional privacy models [16]. KC - slice model finally releases a quasi table and sensitive table. It does not prioritize the sensitiveness of the attributes and concentrates on only one sensitive value from each sensitive attribute. Hence, there may be a chance to detect the suppressed sensitive values from the sensitive buckets. This method chooses constant privacy level on all the sensitive attributes and does not use any technique to bucketize the tuples and it may lead to similarity attacks.

Lakshmipathi Raju N.V.S. et al. [6] implemented a model for multiple sensitive attributes to overcome the pitfalls of the KC - slice model named as KC_i - slice to preserve the individuals confidentiality on the released content by considering the attributes sensitiveness. This approach chooses different thresholds to the various sensitive attributes based on their sensitiveness and generates a quasi table and various sensitive tables. This model reduces the similarity attacks by using a semantic l-diversity technique for bucketization of the tuples and produce the acceptable results with respect to privacy and utility. But, the semantic l-diversity technique does not concentrate on the sensitivity levels of the sensitive values and this may affect the individual's privacy on the published content.

III. Methodology of the Models

A. *Previous Models (KC–Slice & KC_i – Slice)*

These models choose multiple sensitive attributes while publishing the data. KC – slice model gives priority to only one sensitive value of each sensitive attribute and does not concentrate on the remaining sensitive values of a sensitive attribute even if it contains other high sensitive values. This model does not consider attribute sensitiveness and gives equal priority to all the sensitive attributes. No proper technique is used for the bucketization of the tuples and also produces a separate table to publish sensitive attributes. This approach considers a same level of threshold to each sensitive attribute. This method suppresses a single value for each sensitive attribute which facilitates a scope to recognize a suppressed sensitive value from the sensitive bucket [18].

KC_i – slice model considers the attribute sensitiveness and also gives equal priority to all the values of a sensitive attribute. This approach uses semantic l-diversity algorithm for grouping the tuples and considers different threshold levels to various sensitive attributes based on their sensitiveness. Semantic l-diversity algorithm considers the semantic levels of the values while forming the groups, but it does not consider the sensitivity level of the sensitive values. Consider a bucket having three disease values like HIV, liver cancer and heart attack. These three values can belong to three distinct semantic levels, which satisfies the rule of the semantic l-diversity algorithm [25]. But these three values come under high sensitive diseases. Assume a victim resides in the group by considering the background knowledge. Even though one need not find the disease of the victim, it is possible to recognize that person is suffering from high sensitive disease. It may lead to various sensitivity attacks and high suppression ratio.

The major drawback of the semantic l-diversity technique is its un-ability to restrict the sensitivity attacks. Finally, this approach produces the results in the form of a single quasi table and multiple sensitive tables [6]. Due to this there may be a possibility to associate the sensitive attributes with quasi attributes.

B. Proposed Novel KC_i – Slice Model

The proposed model uses a novel methodology to publish the data when compared to the previous models. Initially, anatomization approach is used to separate the sensitive attributes and quasi attributes. Until now all the existing models on multiple sensitive attributes concentrated on only similarity attacks and do not concentrate on sensitivity attacks. To overcome this problem, proposed model designs a new technique named as enhanced semantic l-diversity technique to group the tuples in order to effectively eliminate both the similarity attacks and sensitivity attacks. The methodology of the enhanced semantic l-diversity technique is different from semantic l-diversity technique which is used in previous KC_i – slice model. Some of the existing models on multiple sensitive attributes publish data in the form of multiple sensitive tables along with a single quasi table [6]. But the proposed model publishes the data in the form of multiple sensitive tables and multiple quasi tables by using slicing approach in order to completely minimize the all types of linkage attacks. That's why this model is named as novel KC_i – slice model.

Let us discuss the proposed novel KC_i – slice model in an elaborated manner. The proposed novel KC_i – slice model can overcome the pitfalls of the existing models and also produces expected outcomes in terms of utility and privacy. Table 1 shows a slice of the original data. This approach initially splits the entire data set into quasi identifier table and sensitive table by using anatomization approach. Table 2 specifies the quasi attribute table having attribute like hours, lifetime, sex, income, zip and country. Table 3 is the sensitive table having the sensitive attributes like education, disease, relationship and occupation.

The proposed model uses two slots to finish the publication process. As per the first algorithm, initially it can find out the bucket size by all the records on the data set with all distinct sensitive values of each sensitive attribute and also considers each sensitive attribute have a separate threshold level as per the attribute sensitiveness. The proposed approach can also discriminate the sensitive attributes as either high or low based on the number of high sensitive values of each sensitive attribute [7]. It chooses disease attribute and occupation attribute as high sensitive when contrast to the remaining two sensitive attributes.

This paper also introduces a new enhanced semantic l-diversity technique to formulate the tuple groups. The methodology of this technique is included in algorithm 1. This approach considers both the semantic levels and sensitivity levels of the high threshold sensitive attribute values. In this approach, initially it categorizes the high threshold sensitive attribute values into either high or low sensitive values based on their sensitiveness for each semantic level. It also considers the ratio of total high sensitive values and total low sensitive values of all the semantic levels of high threshold sensitive attribute while formulating the groups. This method assigns the tuples in to each bucket from both high sensitive category and low sensitive category of all the semantic levels thereby forming the effective groups and also considers the overall ratio of high sensitive values and low sensitive values in order to evenly distribute both high sensitive values and low sensitive values into each group. When a tuple is placed in to a bucket, then it reduces the tuple count from that semantic category and also reduces the tuple count from the overall ratio. By maintaining all these things, it is possible to see both the high sensitive values and low sensitive values into each bucket there by reducing the sensitivity attacks. By considering the semantic levels and sensitivity levels of the sensitive values it can effectively limit the sensitivity attacks and similarity attacks thereby reducing the suppression ratio and increases utility rate of the released content [11].

Let us demonstrate the example specified in the paper, assume that the threshold levels of the attributes are occupation and disease is 2 and education and relationship is 3 accordingly. In the proposed model bucketization is done based on the disease sensitive attribute. The sensitive attribute education have 4 values of the doctorate, obviously one value is to be reduced from doctorate is depicted in Table 4, because the maximum threshold of the education is considered as only 3. Table 3 specifies that disease attribute have 3 values of tumor and 3 values of stomach cancer. This model can automatically suppress one value from both tumor and stomach cancer specified in Table 5. There is no need to suppress any value of the relationship attribute, because the counts of all the values are in the specified range of relationship. As per Table 4, one value of the exec-managerial in occupation sensitive attribute is to be suppressed. This model also generates different sensitive tables based on their correlation among sensitive attributes to protect the data against various attacks. Highly correlated sensitive attributes of education-occupation and disease-relationship with their corresponding counts are depicted in Table 4 and Table 5.

As per the second algorithm, each sensitive bucket should have ID. Table 4 and 5 depicts the sensitive bucket IDs. This model also concatenates the correlated quasi attributes based on the correlation among quasi attributes. Table 6 presents the concatenated quasi attributes. Novel KC_i – slice model also performs the vertical partition on Table 6 to generate the multiple quasi tables and secure the data from different attacks. The vertical partitioning can be done on the basis of the total sensitive attributes residing on the original data. This method does not anonymizes the quasi values to produce better utility rates.

The proposed model also concatenates the IDs of sensitive buckets with the correlated quasi attributes. As per the example, it concatenates the IDs of Table 4 with one slice of Table 6 resulting in Table 7 and also concatenate the IDs of Table 5 with remaining slice of the Table 6 resulted in Table 8. Finally, this model performs random alterations on the buckets of Table 7 and Table 8 to produce the final tables Table 9 and Table 10. This model finally publishes sensitive tables Table 4 and Table 5 and quasi tables Table 9 and Table 10.

Table 1. Slice of the original data set

Hours	Lifetime	Sex	Income	Pin code	Native	Education	Disease	Relationship	Occupation
24	40	F	34500	55322	India	Doctorate	Tumor	Husband	Other-service
15	28	M	172000	22190	Mexico	Masters	Stomach cancer	Not-in-family	Dam-clerical
55	27	F	54000	25174	England	Doctorate	Malaria	Wife	Craft-repair
37	22	F	64000	55322	India	Bachelors	Motion	Not-in-family	Transport-moving
19	39	M	81500	22190	Mexico	11th	Stomach cancer	Own-child	Exec-managerial
35	33	F	82000	22190	Mexico	Doctorate	Tumor	Wife	Exec-managerial
40	24	F	740000	25174	England	Bachelors	Heart pain	Wife	Craft-repair
27	37	F	133500	55322	India	Doctorate	Stomach cancer	Not-in-family	Craft-repair
26	46	M	61000	22190	Mexico	Hs-grad	Tumor	Husband	Exec-managerial
34	22	M	32500	55322	India	Masters	Malaria	Own-child	Exec-managerial

Table 2. Quasi Table

Hours	Lifetime	Sex	Income	Pin code	Native
24	40	F	34500	55322	India
15	28	M	172000	22190	Mexico
55	27	F	54000	25174	England
37	22	F	64000	55322	India
19	39	M	81500	22190	Mexico
35	33	F	82000	22190	Mexico
40	24	F	740000	25174	England
27	37	F	133500	55322	India
26	46	M	61000	22190	Mexico
34	22	M	32500	55322	India

Table 3. Sensitive Table

Education	Disease	Relationship	Occupation
Doctorate	Tumor	Husband	Other-service
Masters	Stomach cancer	Not-in-family	Dam-clerical
Doctorate	Malaria	Wife	Craft-repair
Bachelors	Motion	Not-in-family	Transport-moving
11th	Stomach cancer	Own-child	Exec-managerial
Doctorate	Tumor	Wife	Exec-managerial
Bachelors	Heart pain	Wife	Craft-repair
Doctorate	Stomach cancer	Not-in-family	Craft-repair
Hs-grad	Tumor	Husband	Exec-managerial
Masters	Malaria	Own-child	Exec-managerial

Table 4. Grouping of SA's at explicit threshold level for education and occupation

ID	Sensitive attributes
01	#####(1), Doctorate(3),Bachelors(2), 11th(1), Masters(2),Hs-grad(1)
02	#####(1), Exec-managerial(3),Other-service(1), Transport-moving(1) ,Dam-clerical(1), Craft-repair(3)

Table 5. Grouping of SA’s at explicit threshold level for disease and relationship

ID	Sensitive attributes
11	#####(2),Tumor(2), Malaria(2), Heart Pain(1), stomach Cancer(2),Motion(1)
12	Wife(3), Own-Child(2),Not-in-family(3), Husband(2)

Table 6. Concatenated quasi attributes based on their correlation

Hours/Income	Lifetime	Sex	Pin code/Native
24,34500	40	F	55322,India
15,172000	28	M	22190, Mexico
55,54000	27	F	25174, England
37,64000	22	F	55322, India
19,81500	39	M	22190, Mexico
35,82000	33	F	22190, Mexico
40,740000	24	F	25174, England
27,133500	37	F	55322, India
26,61000	46	M	22190. Mexico
34,32500	22	M	55322, India

Table 7. Concatenation of correlated Quasi attributes with IDs of SAs after vertical partitioning of QID

Hours/Income	Lifetime
24,34500,01	40,02
15,172000,01	28,02
55,54000,01	27,02
37,64000,01	22,02
19,81500,01	39,02
35,82000,01	33,02
40,740000,01	24,02
27,133500,01	37,02
26,61000,01	46,02
34,32500,01	22,02

Table 8. Concatenation of correlated Quasi attributes with IDs of SAs after vertical partitioning of QID

Sex	Pin code/Native
F,11	55322,India,12
M,11	22190, Mexico,12
F,11	25174, England,12
F,11	55322, India,12
M,11	22190, Mexico,12
F,11	22190, Mexico,12
F,11	25174, England,12
F,11	55322, India,12
M,11	22190. Mexico,12
M,11	55322, India,12

Table 9. Permutated Quasi attributes and IDs of SAs

Hours/Income	Lifetime
15,172000,01	28,02
24,34500,01	40,02
35,82000,01	33,02
27,133500,01	37,02
34,32500,01	22,02
19,81500,01	39,02
37,64000,01	22,02
40,740000,01	24,02
26,61000,01	46,02
55,54000,01	27,02

Table 10. Permutated Quasi attributes and IDs of SAs

Sex	Pin code/Native
F,11	55322,India,12
F,11	55322, India,12
M,11	22190, Mexico,12
F,11	25174, England,12
F,11	22190, Mexico,12
M,11	22190. Mexico,12
M,11	22190, Mexico,12
F,11	25174, England,12
F,11	55322, India,12
M,11	55322, India,12

Algorithm 1. Bucket generation with enhanced semantic l-diversity algorithm and privacy checking

Input: T - data table, s - number of sensitive attributes g- number of semantic levels

Output : $T_i[T_{B1}, T_{B2}, ..., T_{Bn}]$ // Ti - different tables to be published , b_i - bucket i

```
Begin
Find out number of tuples(n) from table T
Find out the distinct values($d_i$) from each
 Sensitive attribute
r=0;
for (j=0;j<s;j++)
        r = r + $d_j$;
  end for
 k= n/r; // Consider k be the bucket size
for (j=0;j<s;j++)
      $As_{jc} = c_j$;
 end for
 Let  p be the highest threshold value
 Let g be the number of semantic levels of the
 sensitive attribute which is having threshold p
 Assign 1 to high sensitive values and 0 to low
 sensitive values of  the sensitive attribute which
 is having threshold p
 //The following steps specify the enhanced
 semantic l-diversity approach
 x=0; y=0;
 for(j=0;j<n;j++)
            for(m=0; m<g;m++)
                 if($s_i[j]\epsilon g_m$)
                      if($s_i[j]$==1)
                           $g_{mh} = s_i[j]$;
                           $x_{mh}$ ++; x++;
                    else
                           $g_{ml} = s_i[j]$;
                           $y_{ml}$ ++; y++;
                  break;
                 end if
            end for
  end for
while (T != 0) do
     Repeat the loop  to assign k number of
     sensitive values  into each sensitive bucket
      from different semantic groups.
         Assign the values in to the buckets from
         $g_{jh}$ and  $g_{il}$ based on the ratio of x : y.
         from each semantic group  assign one
         high sensitive value  and one low
         sensitive value in to the bucket.
              $T_{Bj} = g_{jh[i]}$; $x_{jh}$ --;
            $T_{Bj} = g_{il[i]}$; $y_{jl}$ --;
          Place k records in to a bucket $B_j$ with a
          combination of both high and low
          sensitive values from various semantic
          categories of the sensitive attribute
```

```
            having threshold p
        end loop
end while
Apply suppression on sensitive values of an
    attribute that violates the specified level
    for each bucket $T_{Br}(As_1,As_2,.....As_s)$ of
    $T(B_1,B_2,............B_z)$
for(j=1; j<=s; j++)
    $q_j$ = n(Distinct($As_j[T_{Bj}]$)) ;
            for(p=1 ; p<=$q_i$ ; p++)
        $c_{jp}$ = (n| $As_{jp}$ | / k)* 100;
                if($c_{jp}$>$As_{ip}$)
                    $c_{jp}$=$c_{jp}$-$As_{jp}$ ;
                    supp ($T_{Bj}[As_{jp}]$){$c_{jp}$};
                    update ($T_{Bj}$);
                end if
                else
                    update ($T_{Bj}$);
                end else
            end for
end for
creation of multiple sensitive tables generation
    based on their correlation
for(x=0; x<s; x++)
    for(y=x+1; y<s; y++)
        $t_{xy}$ ←correlate($As_x$ , $As_y$);
    end for
    [e,f] = Index(max[$t_{xy}$]);
    s=s – ($As_e$ , $As_f$);
    $T_i$ = T($As_e$ , $As_f$);
end for
End
```

Algorithm 2. Bucket slicer and table generator

```
Input   : $T_i[T_{B1,}T_{B2},...,T_{Bz}]$, $T_q$
Output: $Bq_{a1}$ , $Bq_{a2}$,..
            $Bs_{a1}$, $Bs_{a2}$,..
Start
Repeat loop for each $T_i(T_{B1,}T_{B2},...,T_{Bz})$
Generate ID for each $Cs_i$ of $T_i[T_{Br}]$
    $Bs_j$ < - cluster ($ID_i$, $Cs_i$);
    $Bs_{ai}$ <− $u_{i=1}^{n}$ $Bs_i$;
Repeat the loop for all quasi buckets of $T_q$
for(m=0; m<q; m++)
    for(n=m+1; n<q; n++)
        $t_{mn}$ ←correlate($Aq_m$ , $Aq_n$);
    end for
    [e,f] = Index(max[$t_{mn}$]);
    if(max[$t_{mn}$] > h )
        concatenate($Aq_e$ , $Aq_f$)
        q=q – ($Aq_e$ , $Aq_f$)
    end if
end for
Perform vertical partition on the quasi attribute
    table based on the number of sensitive attributes
$B_j$ = $Cq_{ij}$ || $ID_i$
$Bq_{ai}$ = $U_{j=1}^{n}$ $JB_j$
$Bq_{ai}$ <− Randpermute ({$Bq_i$}, $B_{1,}B_{2....}B_n$)
publish ($Bq_{a1,}$ $Bq_{a2,}$..........);
publish ($Bs_{a1}$, $Bs_{a2}$……...);
End
```

C. *Proposed Algorithm Complexity*

Initially, this approach partitions the data set into the quasi identifier table and sensitive table, here the complexity is O(n) where n is the number of rows in the data set. It formulate the buckets with size k also needs O(n). To discriminate the high sensitive values from the low sensitive values and to determine the semantic levels of sensitive values needs O(n). To bucketize the tuples with enhanced semantic l-diversity algorithm needs O(n) and suppress the excess values into the specified thresholds and to check the sensitive levels of each bucket can automatically depends on the distinct number of tuples on each block i.e $O(n^2)$. To split the sensitive and quasi tables into the multiple sensitive tables and quasi tables also needs O(n). To perform random permutations on the tables also needs O(n).

IV. Results and Discussion

The following section describes various performance evaluation parameters and their performance on novel KC_i – slice model [15].

A. *Loss Metric (LM)*

This parameter can determine the loss of all the values of each sensitive attribute only when they are either suppressed or generalized[2].

$$Loss\ Metrics = \frac{N-1}{|C|-1} \qquad (1)$$

As per formula (1) Let N indicate the total suppressed values in an attribute domain. |C| specifies the attribute domain size. Table 11 shows the loss metrics of different sensitive attributes at various levels. As per Fig. 1, occupation and disease attributes produce less loss even at high privacy levels. The relationship and education attributes produce a considerable loss up to privacy level 25. Fig. 2 specifies the overall loss metrics for all sensitive attributes. According to the Fig. 1 and Fig. 2, proposed model produces good results with respect to the existing two models.

Table 11. Loss metrics for sensitive attributes

Threshold	5	10	15	20	25	30	35	40
Education	0.372	0.193	0.077	0.032	0.0006	0	0	0
Disease	0.024	0	0	0	0	0	0	0
Relationship	0.32	0.196	0.12	0.065	0.037	0.01	0	0
Occupation	0.274	0.03	0.001	0	0	0	0	0

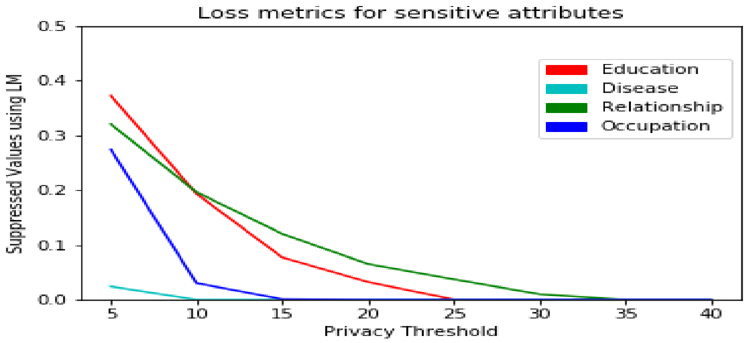

Fig.1. Loss metrics for various SAs

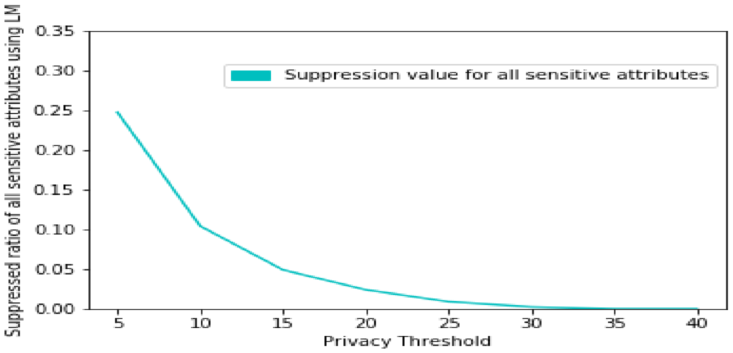

Fig.2. Overall LM for all Sas

B. *Minimal Distortion*

Distortion can be measured by weighted hierarchical distance (WHD), which is used to measure the anonymization at various levels of abstractions. Higher levels produce more amount of distortion with respect to lower levels [10].

$$WHD(x,y)=\frac{\sum_{i=y+1}^{x} W_{i,i-1}}{\sum_{i=2}^{t} W_{i,i-1}} \qquad (2)$$

As per formula (2) hierarchy height of a specific domain as t, this ranges from most general level to most specific level. It ranges from 1,2,3,...,t-1,t. Weight between two consecutive domains can be denoted by 1. The value of value of i ranges from 2 to h.

In uniform weight, the value of

$$W_{i,i-1}=1 \qquad (3)$$

For height weight,

$$W_{i,i-1}=\frac{1}{(i-1)^{\beta}}, \mathit{where}\ \beta \geq 1 \qquad (4)$$

Formula (3) specifies the weight between i and i-1 domains in the case of uniform weight and formula (4) specifies weight between the two consecutive domains in the case of height weight respectively.

$$Distortion(t,t')=\sum_{j=1}^{m} WHD(level(a_i),level(a_i')) \qquad (5)$$

Formula (5) is used to measure the distortion between an ordinary tuple and an anonymized tuple. Formula (3) and Formula (4) are used to uniform weight and height weight respectively.

$$Distortion(D,D')=\sum_{j=1}^{|D|} Distortion(t_j,t_j') \qquad (6)$$

Formula (6) gives the distortion among tables. t_j specifies normal table and t_j' specifies anonymized table.

Fig. 3 shows that the distortion rates produced by this model are less at initial threshold levels and 0 at all other threshold levels in case of high sensitive attributes. Distortion rates are also decreasing while increasing the threshold in the case of low sensitive attributes.

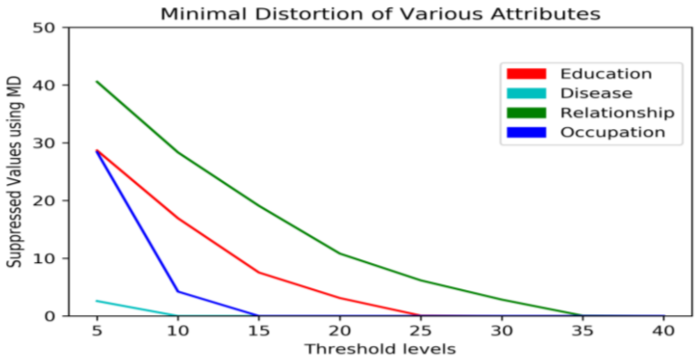

Fig.3. Minimal distortion of attributes at various levels.

C. *Suppressed and Unsuppressed Values*

Fig. 4 and Fig. 5 specify the relation between suppressed values with total values and suppressed values with unsuppressed values respectively. Fig. 4 indicates that proposed novel KC_i – slice model obtains 28.39% suppression rate when compared to 40% in KC_i – slice model at privacy level 5. Education sensitive attribute acquires 28.68% suppression rate in the proposed model when compared to 38% suppression rate in KC_i – slice model. Fig. 5 shows the relation between suppressed and unsuppressed values is also reducing by increasing the threshold for all sensitive attributes.

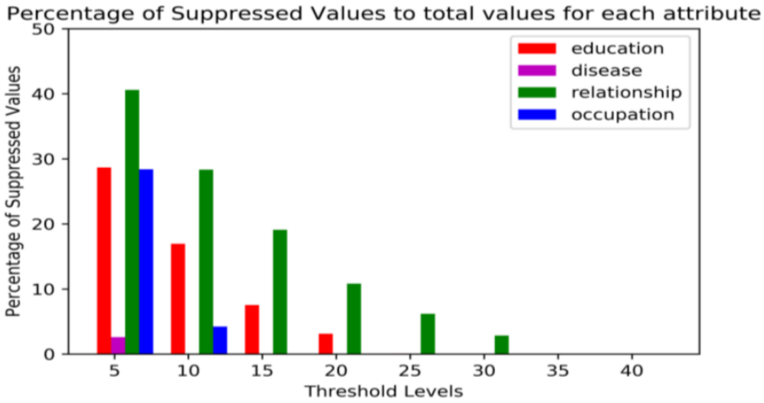

Fig.4. Percentage ratio of suppressed and total values

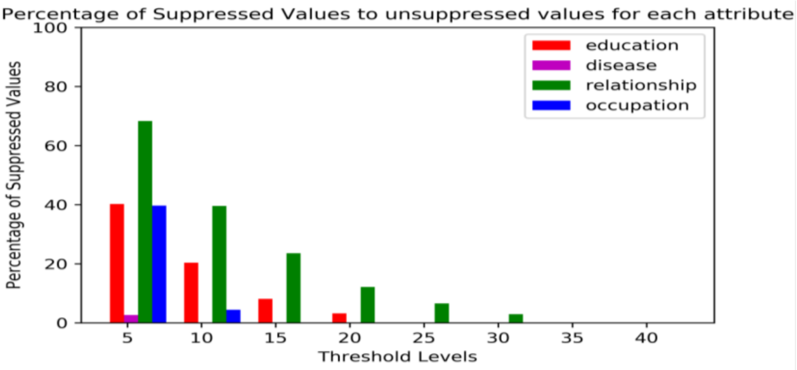

Fig.5. Percentage ratio of suppressed with unsuppressed value

D. *Utility Gain*

Published data utility gain can be measured by the trade-off metrics, which can find out the privacy loss and information gain for each generalization or suppression activity [3]. It can be represented by

$$IGPL(c) = \frac{IG(c)}{PL(c)+1} \tag{7}$$

As per formula (7) IG(c) specifies the information gain, PL(c) specifies the privacy loss of either generalization or suppression operation. This metric is useful to determine important relations among the published data, which can be dependent on the type of the anonymization operation used on the published data [8]. The unsuppressed and suppressed values of education sensitive attribute are depicted in Table 12 and Table 13. As per Fig. 6 and Fig. 7, the suppression rate can be minimized by increasing the privacy level. No need to consider high thresholds to the education attribute, because proposed model considered as a low sensitive attribute. This specifies that it is possible to obtain the maximum utility rate by applying the required threshold level.

The utility gain of heart pain and glaucoma of disease attribute are depicted in Table 14 and Table 15. According to the Fig. 8 and Fig. 9 specifies that suppression rate is almost negligible at even high thresholds and 0 for remaining all thresholds. This concludes that disease sensitive attribute acquires the highest utility rate at initial thresholds only.

The suppressed values of husband and own-child of relationship attribute are depicted in Table 16 and Table 17. From Fig. 10 and Fig. 11 describe that the number of unsuppressed values are increased by increasing the threshold.

Table 18 and Table 19 highlight the other-service and exec-managerial values of occupation sensitive attribute. As per Fig. 12 and Fig. 13 the suppression rate is less at high thresholds and 0 at other thresholds. As per the results, novel KC_i – slice model yields good results when compared to the previous KC - slice and KC_i – slice models.

Table 12. Hs-grad - Suppressed and Unsuppressed values Total : 10,501

Thresholds	5	10	15	20	25	30	35	40
Suppressed	5425	4216	2452	1015	25	0	0	0
Un-suppressed	5076	6285	8049	9486	10476	10501	10501	10501

Table 13. Bachelors - Suppressed and Unsuppressed values Total : 5355

Thresholds	5	10	15	20	25	30	35	40
Suppressed	1310	65	0	0	0	0	0	0
Un-suppressed	4045	5290	5355	5355	5355	5355	5355	5355

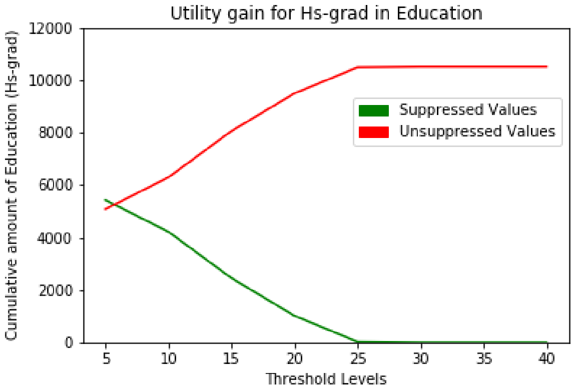

Fig.6. Utility gain of "hs-grad"

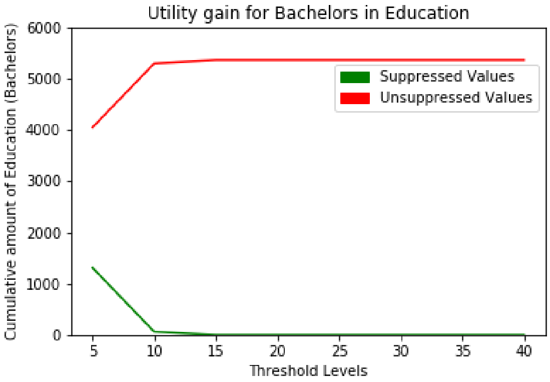

Fig.7. Utility gain of "bachelors"

Table 14. Heart pain - Suppressed and Unsuppressed values Total: 1843

Thresholds	5	10	15	20	25	30	35	40
Suppressed	62	0	0	0	0	0	0	0
Un-suppressed	1781	1843	1843	1843	1843	1843	1843	1843

Table 15. Glaucoma - Suppressed and Unsuppressed values Total: 1853

Thresholds	5	10	15	20	25	30	35	40
Suppressed	45	0	0	0	0	0	0	0
Un-suppressed	1808	1853	1853	1853	1853	1853	1853	1853

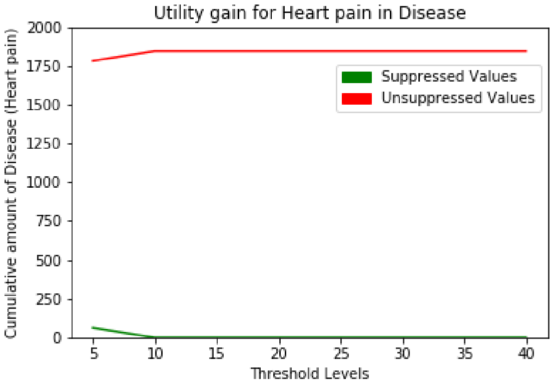

Fig.8. Utility gain of "heart pain"

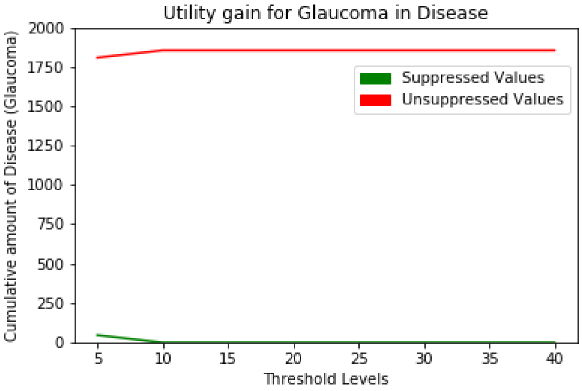

Fig.9. Utility gain of "glaucoma"

Table 16. Husband - Suppressed and Unsuppressed values Total: 13193

Thresholds	5	10	15	20	25	30	35	40
Suppressed	8452	7115	5112	3521	2012	925	21	0
Un-suppressed	4741	6078	8081	9672	11181	12268	13172	13193

Table 17. Own-child - Suppressed and Unsuppressed values Total: 5068

Thresholds	5	10	15	20	25	30	35	40
suppressed	1102	25	0	0	0	0	0	0
un-suppressed	3966	5043	5068	5068	5068	5068	5068	5068

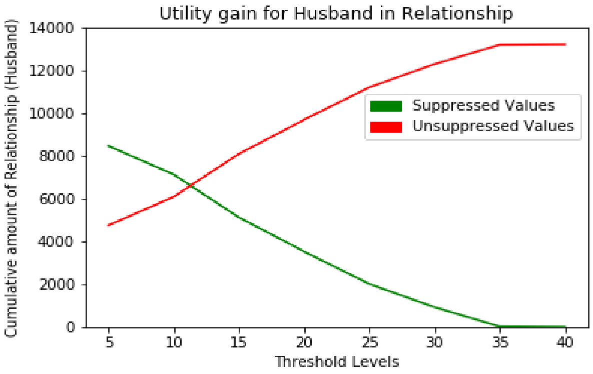

Fig. 10. Utility gain of "husband"

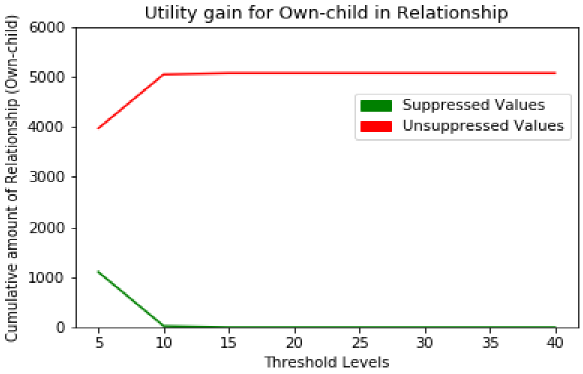

Fig. 11. Utility gain of "own-child"

Table 18. Other-service - Suppressed and Unsuppressed values Total: 5138

Thresholds	5	10	15	20	25	30	35	40
Suppressed	2235	975	80	0	0	0	0	0
Un-suppressed	2903	4163	5058	5138	5138	5138	5138	5138

Table 19. Exec -managerial - Suppressed and Unsuppressed values Total: 4066

Thresholds	5	10	15	20	25	30	35	40
Suppressed	1511	95	0	0	0	0	0	0
Un-suppressed	2555	3971	4066	4066	4066	4066	4066	4066

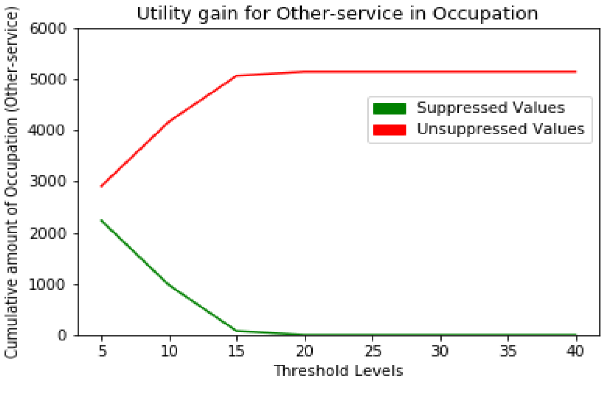

Fig.12. Utility gain of "other-service"

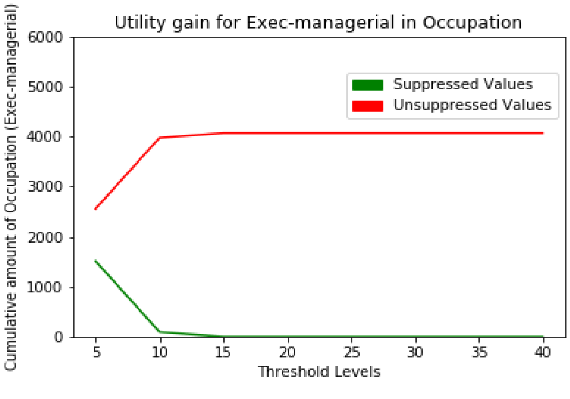

Fig.13. Utility gain of "exec-managerial"

E. Comparison among Novel KC_i-Slice Model with Other Models

The following section describes the effectiveness of the novel KC_i model when compared with the KC-slice and KC_i models. Fig. 14, Fig. 15 and Fig. 16 show the evaluation of different models for various sensitive values. Table 20 shows the comparison of all the three models for not-in-family of relationship at different thresholds. From Fig. 14, when contrast to other two models the novel KC_i model generates less number of suppressed sensitive values at all thresholds. Table 21 shows that utility level of the novel KC_i -slice model is high on all threshold levels against remaining two models.

Table 22 shows the suppressed values of some-college in education attribute. As per Fig. 15, novel KC_i -slice model have more number of unsuppressed values, contrast to other two models. The proposed model also produces more amounts of utility levels than the remaining two models for some-college value is depicted in Table 23. Novel KC_i -slice model acquires more than 16% excess utility rate against KC_i -slice model and 38% extra utility rate against KC - slice model at threshold 5. The proposed model obtains excess utility levels at all thresholds when compared to the other models.

The suppressed values of exec-managerial at different threshold levels are shown in Table 24. As per Fig. 16, the rate of suppression is less at different thresholds in novel KC_i -slice model with respect to the existing two models. Table 25 shows that novel KC_i -slice model obtains more than 15% excess utility rate against KCi -slice model and 22% excess utility rate against KC - slice model at privacy level 5. Novel KC_i -slice approach acquires excess utility rates for exec-managerial value of occupation at different threshold levels.

Table 20. Number of values suppressed for "not-in-family"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	3641	2090	1105	0	0	0	0	0
KC_i-slice	4741	3121	1501	163	0	0	0	0
KC-slice	6361	4741	3121	1501	179	0	0	0

Table 21. Utility levels for "not-in-family"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	56.15%	74.83%	86.69%	100%	100%	100%	100%	100%
KC_i-slice	42.9%	62.42%	81.92%	98%	100%	100%	100%	100%
KC-slice	23.4%	42.2%	61.46%	80.73%	97.84%	100%	100%	100%

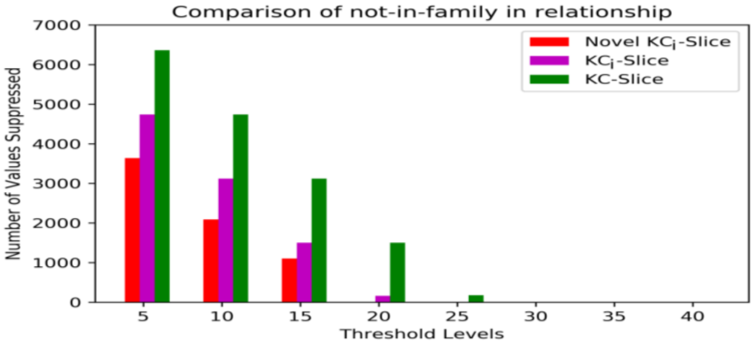

Fig.14. Comparison among various models for "not-in-family"

Table 22. Number of values suppressed for "some-college"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	2605	1232	0	0	0	0	0	0
KC_i-slice	3727	2107	568	539	5	0	0	0
KC-slice	5347	3727	2201	568	45	0	0	0

Table 23. Utility levels for "some-college"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	64.27%	83.10%	100%	100%	100%	100%	100%	100%
KC_i-slice	48.88%	71.1%	92.2%	92.6%	99.93%	100%	100%	100%
KC-slice	26.66%	48.88%	69.81%	92.2%	99.89%	100%	100%	100%

Table 24. Number of values suppressed for "exec-managerial"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	1511	95	0	0	0	0	0	0
KC_i-slice	2122	530	1	0	0	0	0	0
KC-slice	2400	600	15	0	0	0	0	0

Table 25. Utility levels for "exec-managerial"

Thresholds	5	10	15	20	25	30	35	40
Novel KC_i-slice	62.83%	97.66%	100%	100%	100%	100%	100%	100%
KC_i-slice	47.81%	86.96%	99.97%	100%	100%	100%	100%	100%
KC-slice	40.97%	85.24%	99.63%	100%	100%	100%	100%	100%

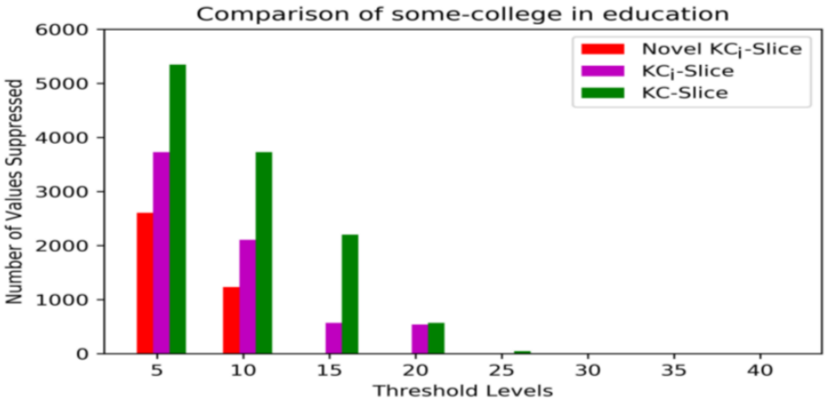

Fig.15. Comparison among various models for "some-college"

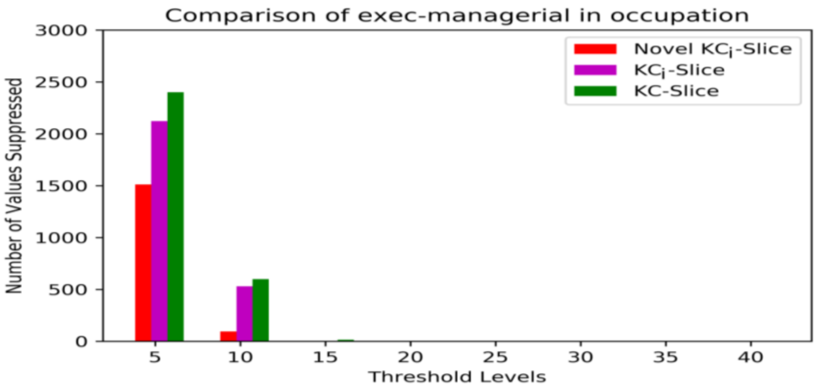

Fig.16. Comparison among various models for "exec-managerial

F. *Working Environment And Data Set*

Novel KC_i - slice model was executed on windows 8 environment of 1.6 -GHz Pentium processor with 16 GB RAM. Python and MYSQL were used to implement this model. UCI machine learning repository published a data set called as adult data set and it was used for the experimental investigation [17].

V. Protection Against Various Attacks

The novel KC_i - slice model produces the results in the form of multiple tables to protect the data against all types of linkage attacks. This model introduces a new enhanced semantic l-diversity approach to safeguard the data against the similarity attacks and sensitivity attacks. It publishes multiple sensitive tables and multiple quasi tables to protect the data against table linkage attacks. This model uses different thresholds on different attributes to safeguard the data against skewness attack and homogeneity attack. Due to the size of the bucket, it protects the data against correspondence attack, minimalist attack and backdrop knowledge attack.

VI. Conclusions and Future Extensions

Novel KC_i - slice model chooses different threshold levels to each sensitive attribute according to the attribute sensitiveness and also all the sensitive values of an attribute are equally prioritized. To acquire the required utility levels and privacy levels, this paper also introduces a new concept called as enhanced semantic l-diversity algorithm for horizontal partitioning the tuples which considers both the sensitivity levels and semantic levels of the values of an attribute.

The proposed model categorizes the sensitive attribute as either high or low sensitive attributes based on the number of high sensitive values resides in a sensitive attribute. Novel KC_i - slice model assigns variable thresholds to the different sensitive attributes with respect to their sensitiveness. It generates multiple sensitive tables for sensitive attributes and multiple quasi tables for quasi attributes based on their correlation to prevent from various types of attacks. Novel KC_i - slice model acquires the excellent results with respect to privacy and utility when contrast to the other two models. This model can reduce all types of similarity attacks.

This work can also be extended in the other ways by considering categorical sensitive attributes along with numerical sensitive attributes. There is also necessary to design an algorithm to discriminate the high sensitive values and low sensitive values. This model also to be extended to the distributed data sets using horizontal partitioning and vertical partitioning approaches.

References

[1] Anjum A, Ahmed Naveed, Malik S.U, Zubair S, Shahzad B, "An efficient approach for publishing microdata for multiple sensitive attributes", *The Journal of Supercomputing*, pp. 1- 29, 2018.

[2] Fung B. C, Wang K, Fu A. W. C, Philip S. Y, "Introduction to Privacy - Preserving Data Publishing: Concepts and techniques", *Chapman and Hall/CRC Press*, 2010.

[3] Fung B. C. M, Wang K., Chen R, Yu P. S, "Privacy-preserving data publishing: A survey of recent developments", *ACM Computing Surveys*, 42(4), article 14, 2010.

[4] Han J, Luo F, Lu J, Peng H, "SLOMS: A Privacy Preserving Data Publishing Method for Multiple Sensitive Attributes Micro data", *Journal of Software* 8(12), pp. 3096-3104, 2013.

[5] Hasan A. S. M, Jiang Q, Chen H, Wang S, "A New Approach to Privacy-Preserving Multiple Independent Data Publishing", *Applied Sciences*, 8, 783, pp. 1-22, 2018.

[6] Lakshmipathi Raju N.V.S, Seetaramanath M.N, Srinivasa Rao P,(in press) "An enhanced dynamic KC-Slice model for privacy preserving data publishing with multiple sensitive attributes by inducing sensitivity", *Journal of King Saud University-Computer and Information Sciences*, 2018.

[7] Liu F, Jia Y, Han W, "A new k – anonymity Algorithm towards multiple-sensitive attributes", *In 2012 IEEE 12th International Conference on Computer and Information Technology (CIT) IEEE* pp 768-772, 2012.

[8] Liu J, "Enhancing utility in privacy preserving data publishing" *Doctoral dissertation, Applied Science: School of Computing Science*, 2010.

[9] Liu J, Luo J, Huang J. Z, "Rating: privacy preservation for multiple attributes with different sensitivity requirements", *In Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference IEEE*, pp. 666-673, 2011.

[10] Li J, Wong R. C.W, Fu A.W.C, Pei J, "Achieving k-anonymity by clustering in attribute hierarchical structures", *In International Conference on Data Warehousing and Knowledge Discovery. Springer, Berlin, Heidelberg*. pp. 405-416, 2006.

[11] Li N, Li T, Suresh V, "t-closeness: Privacy beyond k-anonymity and l-diversity", *In Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on. IEEE*, pp. 106-115, 2007.

[12] Liu Q, Shen H, Sang Y, "Privacy-preserving data publishing for multiple numerical sensitive attributes", *Tsinghua Science and Technology*, 20(3), pp. 246-254, 2015.

[13] Li T, Li N, "On the tradeoff between privacy and utility in data publishing", In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM*, pp 517 - 526, 2009.

[14] Maheshwarkar N, Pathak K, Choudhari N. S, "K-anonymity model for multiple sensitive attributes", *International Journal of Computer Applications (IJCA)*, 2012.

[15] Mendes R, Vilela J. P, "Privacy-preserving data mining: methods, metrics, and applications", *IEEE Access* 5, pp. 10562-10582, 2017.

[16] Mohammed N, Fung B, Hung P. C, Lee, C. K, "Anonymizing healthcare data: a case study on the blood transfusion service", *In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining ACM*. pp. 1285-1294, 2009.

[17] Newman D, Hettich S, Blake C, Merz C, (2006) UCI repository of machine learning databases, *university of california, irvine, Dept of Information and Computer Sciences*,1998.

[18] Onashoga S.A, Bamiro S.A, Akinwale A.T, Oguntuase J.A, "KC-Slice: A dynamic privacy-preserving data publishing technique for multisensitive attributes", *Information security journal: A Global Perspect* 26(3), pp. 121-135, 2017.

[19] Ram Prasad Reddy S, VSVN Raju K, Valli Kumari V, "A

novel approach for personalized privacy preserving data publishing with multiple sensitive attributes", *International journal of engineering & technology* 7(2.20) pp. 197-206, 2018.

[20] Susan V. S, Christopher, T, "Anatomization with slicing: a new privacy preservation approach for multiple sensitive attributes", *Springer Plus*, 5(1), 964, 2016.

[21] Usha P, Shriram R, Sathishkumar S, "Multiple sensitive attributes based privacy preserving data mining using k-anonymity", *Int. J. Sci. Eng.* Res, 5(4), 2014.

[22] Xiao X, Tao Y, "m-invariance: towards privacy preserving republication of dynamic datasets", *In Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data ACM*, pp. 689-700, 2007.

[23] Xu Y, Ma T, Tang M, Tian W, "A survey of privacy preserving data publishing using generalization and suppression. *Applied Mathematics & Information Sciences*, 8(3), 1103, 2014.

[24] Yuki Ina, Yuichi Sei, Yasuyuki Tahara, Akihiko Ohsuga, "Anonymization and analysis of horizontally and vertically divided user profile databases with multiple sensitive attributes", *International conference on service operations and logistics and informatics(SOLI)* pp.262-267, 2018.

[25] Emad Elabd, Hatem Abdulkader, Ahmed Mubark,"L-Diversity-Based Semantic Anonymization for Data Publishing", *International Journal of Information Technology and Computer Science(IJITCS)*, vol. 7, no 10, pp. 1-7, 2015. DOI: 10.5815/ijitcs.2015.10.01.

[26] Ashoka K, Poornima B, "Enhanced Utility in Preserving Privacy for Multiple Heterogeneous Sensitive Attributes using Correlation an Personal Sensitivity flags", *International conference on advances in computing, communications and informatics (ICACCI)*, pp. 970-976, 2017.

[27] Abou_el_ela Abdo Hussain, Nagy Ramadan Darwish, Hesham A. Henfy, "Multiple-Published Tables Privacy Preserving Data Mining: A Survey for Multiple-Published Tables Techniques", *International Journal of Advanced Computer Science and Applications*, vol.6, No.6, pp.80-85, 2015.

[28] Yifan Ye, Lixxia Wang, Jianmin Han, Sheng Qiu, Fangwei Luo, "An anonymization method combining anatomy and permutation for protecting privacy in micro data with multiple sensitive attributes", *Proceedings of the 2017 International Conference on Machine Learning and Cybernetics*, July 09-12, Ningbo, China, pp. 404-411, 2017.

[29] Kavitha D, "A Survey on Privacy Preserving Data Mining Techniques", *International Journal of Computer & Mathematical Sciences(IJCMS)*, vol.7, issue 2, pp. 160-169, 2018.

[30] Aldeen Yousra S, Salleh Mazleena, "A new heuristic Anonymization Technique for Privacy Preserved Datasets Publication on Cloud Computing", *Journal of Physics : Conf. Series*, pp. 1-15, 2018.

[31] Ram Mohan Rao P, Murali Krishna S, Siva Kumar A.P, "Privacy Preservation techniques in big data analytics: a survey", *Journal of Big Data*, pp. 1-12, 2018.

[32] Sandeep Varma Nadimpalli, Valli Kumari Vatsavayi, "Verification in Privacy Preserving Data Publishing", *Computing and Informatics* vol. 35, pp. 1160-1188, 2016.

Authors' Profiles

N.V.S. Lakshmipathi Raju received M.Tech degree in Computer Science and Engineering from JNT University Kakinada in 2007. He is pursuing Ph.D in JNTUK, Kakinada. Presently he is working as Associate Professor in Department of CSE at Gayatri Vidya Parishad College of Engineering (A), Visakhapatnam, Andhra Pradesh, India. His research interests include Data Mining, Big data analytics, Information security and privacy preservation in data publishing. He published research papers in international Journals.

M. N. Seetaramanath is a Professor in Department of IT at Gayatri Vidya Parishad College of Engineering (A), Visakhapatnam, Andhra Pradesh, India. He received his PhD from Andhra University in 1984. He has guided several research scholars. His research interests are in resource and mobility management for wireless adhoc networks, Data Mining, wireless sensor networks, and Network Security.

Peri Srinivasa Rao is presently working as Professor in the Department of Computer Science and Systems Engineering, and also working as a Principal to AU College of Engineering, Andhra University, Visakhapatnam. He got his Ph.D degree from Indian Institute of Technology, Kharagpur in Computer Science in 1987. He published several research papers and delivered invited lectures at various conferences, seminars and workshops. He guided a number of students for their Ph.D and M.Tech degrees in Computer Science and Engineering and Information Technology. His current research interests are Image Processing, Communication networks, Data mining and Computer Morphology.

How to cite this paper: N.V.S. Lakshmipathi Raju, M.N. Seetaramanath, P. Srinivasa Rao, "A Novel Dynamic KCi - Slice Publishing Prototype for Retaining Privacy and Utility of Multiple Sensitive Attributes", International Journal of Information Technology and Computer Science(IJITCS), Vol.11, No.4, pp.18-32, 2019. DOI: 10.5815/ijitcs.2019.04.03