Multilevel Anti-Discrimination and Privacy Preservation Correlativity

Naveena M.S
M. Tech Student, Dept. of CSE
Marian Engineering College
Trivandrum
navchin@gmail.com

Ms. Merlin Shoerio
Asst. Professor, Dept. of CSE
Marian Engineering College
Trivandrum
merlinshoerio@gmail.com

Abstract: In the fast growing technology, most organizations will need to reveal their crucial data which includes the sensitive information that discloses one's identity during their business analytics and to provide services. To limit the access to such sensitive data, various privacy preservation techniques are applied based on the level of priority assumed. The multilevel privacy preserved discrimination free data transmission deals with the correlation of discrimination prevention and privacy preservation. By applying appropriate privacy preservation techniques, it can be shown that the discrimination prevention can be easily accomplished along with secure transmission of data to different levels of users. On the basis of sociological aspect, discrimination is the unfair treatment of an individual or group based on their membership on a particular category. So, the decision attribute that leads to discrimination needs to be hided or transformed. The unified discrimination prevention method is available which avoids both direct and indirect discrimination simultaneously or both at the same time. Although discriminatory biases are eliminated, it results in huge data loss which drops down the data transmission efficiency. The data quality is much preserved since encryption technique is included. The proposed system is dynamic in nature and can be implemented in any organization. The experimental evaluation aids us to conclude that the proposed work is efficient for data transmission without discrimination and with maximum privacy preservation.

Keywords: *Data mining, Microdata, Discrimination Prevention, Privacy Preservation, Data Anonymization*

____*****____

I. INTRODUCTION

In fast growing technology, most organizations will share aggregated data for their business analytics and to provide services. For example, retailers will often share their customer's purchasing information to the product merchants for their effective advertising but they do not need to reveal customer's sensitive data which leads to discrimination. So, such valuable details should be preserved securely without any harm to the data. On the basis of sociological aspect [9], discrimination is the prejudicial treatment of an individual based on their membership in a particular group or category. Discrimination can be on the basis of gender, race, religion, etc. There are certain antidiscrimination laws available, but all of them are reactive, not proactive.

Discrimination can be either direct or indirect [8]. Direct discrimination occurs due to the rules that are built up by considering the sensitive discriminatory attributes such as age, sex, religion, etc. Indirect discrimination (redlining rules) occurs due to the rules that are not explicitly mentioning discriminatory attributes, but intentionally or unintentionally could generate discriminatory decisions. Indirect discrimination could happen due to the availability of certain background knowledge, for example, if one knows the zip code, he can easily identify whether a person is from black community or not. The background knowledge can be accessed from the publicly available data such as census data or might be from original data set because of the existence of non-discriminatory attributes that are highly correlated with the sensitive attributes.

Privacy preserving of microdata is the release of aggregate information about data without leaking the individual information of the participants. Microdata consists of records each of which contains personal information about individual entity such as name, household, organization, etc. which may be the attribute for discrimination. Several microdata anonymization techniques [12] are developed so far. The most popular ones are generalization for k-anonymity, bucketization for l-diversity, slicing.

Consider the case when a sequence of data is extracted from the personal data set of a population of individual persons for some decision making process [2], such as granting or denying loan. Here, from the sequence of data, first the sensitive information about the individual is revealed. Then, the decision rules based on such sequence may lead to unfair discrimination. So, the discrimination risks and the privacy should be tackled together to avoid information loss and to provide security while data transmission.

II. RELATED WORKS

Despite the wide technological development, the issue of anti-discrimination in data mining did not get as much attention until 2008. There are some proposals developed for discrimination discovery and measurement and others for discrimination prevention.

The discriminatory decision discovery was first proposed in[] is based on the mining of classification rules and conducting reasoning on the basis of discrimination measurement. The main issue in discrimination discovery is the accessing of hidden historical data without privacy. This is because the personal data in decision records are highly dimensional and there is complexity in indirect discrimination discovery.

On the basis of the way of eliminating discrimination, the prevention methods are classified into three groups–preprocessing, inprocessing and postprocessing [10]. In preprocessing, the source data is transformed in such a way that the discriminatory biases are eliminated. In inprocessing, the data mining algorithms are changed so that the resulting models does not contain any unfair decisions. In the case of postprocessing, the resulting data mining models are changed instead of cleaning the original data set or changing the data mining algorithms. The preprocessing approach is widely used for discrimination prevention since it is more flexible. The common preprocessing methods are suppression, massaging the dataset, reweighing and sampling.

For massaging of data [5], first a ranker for predicting the class attribute without taking into account the discrimination is learnt. The ranker is used to rank the data objects according to the probability of being in the desired class. The class labels of the most likely victims and the profiteers are changed. Victims are the training instances of the discriminated community with a negative label but with a high positive class probability. Profiteers are the training instances of the favoured community with a positive label but a low positive class probability. The modified data is then used for learning a classifier with no discrimination for future decision making. The drawback of this approach is that it is intrusive in nature.

Preferential sampling [6] changes the distribution of differential data objects for a particular dataset to be discrimination free. The data objects close to the decision boundaries are more prone to discrimination. So, the distribution of the borderline objects is changed to make it discrimination free. The data objects are divided into four groups – the Discriminated community with Positive class labels (DP), the Privileged community with Positive class labels (PP), the Discriminated community with Negative class labels (DN) and the Privileged community with Negative class labels (PN). By using the ranking function, the class probability of each data tuple is calculated. Then, expected size for each group is calculated to make the dataset bias free. Finally, apply sampling with replacement either to increase (duplicate) the size of DP and PN and decrease (remove) the size of DN and PP. Drawback of this approach is low utility rate and minimum discrimination removal.

In decision tree learning approach [7], the solution is based on the integration of discrimination awareness into the induction model process of a decision tree. The techniques used are dependency aware tree construction and leaf relabeling. In dependency aware tree construction, the splitting criterion for a tree node is based on its accuracy contribution and level of discrimination. In leaf relabeling, the label of selected leaves are changed in such a way that the discrimination is lowered with a minimal loss in accuracy. This method gives high accuracy and low discrimination score, but the decision tree construction is complex.

In [3], the discrimination prevention can be done under two phases – discrimination measurement and data transformation. Based on the identified α discriminatory rules and the redlining rules, the frequent classification rules are grouped in Potentially Discriminatory (PD) and Potentially Non Discriminatory (PND) rules. Then, the discrimination is measured. In [3], the crime intrusion is detected and only direct discrimination prevention is addressed. [3] is extended to [4] and it discovers indirect discrimination in databases. Here, only preliminary experimental proofs are given. A unified approach for discrimination prevention of direct as well as indirect discrimination, with finalized algorithms and all possible data transformation methods based on rule protection and/or rule generalization could be applied simultaneously or both at the same time was proposed in [1]. There is no efficient privacy preservation techniques explained which results in huge information loss.

A. Basic Terms

- Data set: a collection of data records and their attributes.
- Item set: a collection of one or more items.
- Classification rule: an expression X -> C, where C is a class item (a yes/no decision), and X is an item set containing no class item.
- Support of an item set, sup(X): the fraction of records that contain the item set X. We say that a rule X -> C is completely supported by a record if both X and C appear in the record. Support is a measure of statistical significance.
- Confidence of a classification rule, conf(X -> C): measures how often the class item C appears in records that contain X. Hence, if sup(X) > 0 then conf(X -> C) =sup(X, C) / sup(X). Confidence is a measure of the strength of the rule. Support and confidence range over [0, 1].
- Frequent classification rule: a classification rule with support and confidence greater than respective specified lower bounds. Let FR be the frequent classification rule under consideration. Frequent classification rules can be classified into two classes: Potentially Discriminatory (PD) rule and Potentially Non-Discriminatory (PND) rule.
  - ○ A classification rule X -> C is potentially discriminatory (PD) when X = A, B with A subset of DIs is a nonempty discriminatory item set and B a non-discriminatory item set.
  - ○ A classification rule X -> C is potentially non-discriminatory (PND) when X = D, B is a non-discriminatory item set.
- Extended Lift (elift): Let A,B -> C be a classification rule such that conf(B -> C) > 0. The

extended lift of the rule is elift(A,B -> C) = conf(A,B -> C) / conf(B -> C)

- elb: Let r : D,B -> C be a PND classification rule, and let γ = conf(r : D,B -> C) and δ = conf(B -> C) > 0. Let A be a discriminatory item set, and let $\beta 1$, $\beta 2$ such that conf(rb1 : A,B ->D) $\geq \beta 1$, conf(rb2 : D,B -> A) $\geq \beta 2 > 0$.
  Compute $f(x) = \beta 1/\beta 2 \; (\beta 2 + x - 1)$
  elb(x, y) = f(x)/y if f(x) > 0, otherwise 0.

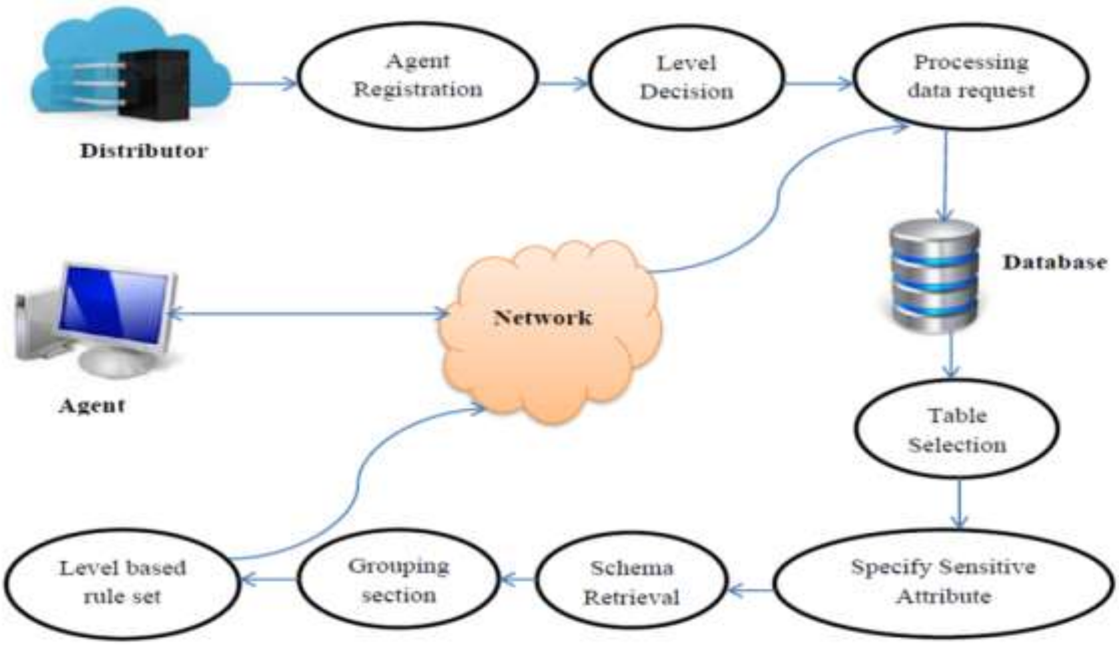

Fig.1. Multilevel Privacy Preserved Anti-Discrimination Data Transmission Architecture

B. Algorithm

Direct Discrimination Measure

```
Assign a threshold, t
Step 1: Find the confidences of the classification rules, C1 and C2
Step 2: Calculate elift (elift = C1/C2)
Step 3: if (elift < t)
            then "the rule is alpha protective"
Step 4: else "the rule is alpha discriminatory"
        Transform the discriminatory items
Step 5: Re-compute confidence
```

Indirect Discrimination Measure

```
Assign a threshold, t
Step 1: Find the confidences of the classification rules, C1 and C2
Step 2: Find the confidences of the background rules, B1 and B2
Step 3: F = (B1/B2)*( B2 + C1 - 1)
Step 4: Calculate elb (elb = F/C2)
Step 5: if (elb < t)
then "the rule is not discriminatory"
Step 6: else "the rule is discriminatory"
Transform the discriminatory items
```

C. Data Transformation

The original data set is transformed in such a way to remove direct and/or indirect discriminatory biases with minimum data loss. The approach used in [1] results in huge information loss. So, we develop another approach which can transform the original data set to remove discrimination without any information loss using privacy preservation approach.

III. MULTILEVEL PRIVACY PRESERVED ANTI-DISCRIMINATION DATA TRANSMISSION

The correlation of discrimination prevention and privacy preservation is dealt in the proposed architecture. Here, the distributor will register the agent and he will decide the level of priority to be given to the agent. For each level, different privacy preservation rule sets are designed which determine the degree of data to be hided or preserved. The distributor will accept the data request from the registered agent only. Then, appropriate database requested by the agent is chosen and required table is selected. The sensitive attributes which lead to discrimination are specified thereby. Then, the schema is retrieved and the resultant structured data is subjected to grouping. Based on the sensitive attributes, the rules are classified into potentially discriminatory and potentially non-discriminatory groups. Then according to the level assigned by the distributor to each agent, the related rule set is applied. The rule sets are the various privacy preservation techniques designed on the basis of the degree of data to be hidden according to the type of agent. The transformed dataset is transmitted through the network and will reach to the agent who requests the dataset.

In the anonymization techniques like generalization, bucketization and slicing [11], the attribute partitioning is the first and foremost step. The attributes are partitioned into:

- Attributes which identify individuals easily. (Social Security Number, Name, Address, etc.)
- Attributes whose values when taken together can identify individuals–quasi identifiers. (Zip code, Birthdate, gender, etc.)
- Sensitive Attributes. (Disease, Salary, etc.)

When releasing microdata, the sensitive details of an individual will be disclosed. There are two types of information disclosure: identity disclosure and attribute disclosure. Identity disclosure occurs when an individual is related to a particular record in the released table. Attribute disclosure occurs when new information about some individuals is revealed. The released data helps in inferring the characteristics of an individual more precisely than it would be possible before the data release. Identity disclosure sometimes leads to attribute disclosure.

The main step of anonymization is removing of explicit identifiers (quasi identifiers). A common anonymization approach is generalization which replaces quasi identifier values with values that are less specific but semantically consistent. So that, more records will have the same set of quasi identifier values. But, even though k-anonymity protects against identity disclosure, it is insufficient to prevent attribute disclosure. A considerable amount of information loss for high dimensional data is the major drawback of generalization.

A new notion of privacy was introduced in [12] called l-diversity, which requires that the distribution of a sensitive attribute in each equivalence class has at least “l” well-represented values. The records are sorted based on the occurrence of sensitive attributes. Then, group the similar records with set of buckets and analyse it. Combine the set of correlated attributes after diversity check is done. Bucketization does not prevent l-diversity is limited in its assumption of adversarial knowledge. It is possible for an adversary to gain information about a sensitive attribute as long as the information about the global distribution of the attribute is known. This assumption generalizes the specific background and homogeneity attacks used to motivate l-diversity.

In slicing [11], after attribute partitioning, the highly correlated attributes are in same column. Then, column generalization is done thereby to protect membership disclosure. Tuple partitioning is done by the use of two data structures-a queue of buckets and a set of sliced buckets. Initially, the queue contains only one bucket which includes all tuples and the sliced buckets are empty. For every execution of the algorithm, a bucket is removed from the queue and splits into two buckets. If the sliced table after split satisfy l-diversity, then the two buckets are placed at the end of the queue. Otherwise, it does not split the bucket further and put it in the sliced bucket. Whenever queue becomes empty, the sliced table is computed.

A. MLPP (MULTILEVEL PRIVACY PRESERVATION) ALGORITHM

Multilevel Privacy Preservation (MLPP) Algorithm details the transformation of the original dataset to privacy preserved discrimination free dataset according to the level of priority of the agent needed. The priority of the agent can be determined by the agent by considering certain factors like agent’s designation in an organization, type of dataset he requested, etc.

MLPP Algorithm

Inputs: DB, $DR_{1\ldots n,}$ $L_{1\ldots n}$

Output: DB’ (transformed data set)

```
for each $DR_i$
$DB_s \leftarrow$ all records that completely supporting SA
 for each $L_i$
   for each $A_i \in DB_s$
       if $A_i = SA_i$ then
         $DB_c \leftarrow$ all records completely supporting SA
         if $L(DR_i = L_i)$ do
          Select first record in $DB_c$
          Modify class item of $DB_c$ with $R_i$
          $DB' \leftarrow$ Recomputed DB
          while (!$DB_c(n)$)
           end do
         end if
```

```
            end if
        end for
    end for
end for
```

The inputs for the MLPP algorithm is the original dataset (DB), the data requests for n agents (DR$_{1\ldots n}$) and their level of privacy required (L$_{1\ldots n}$). DB$_s$ are the set of records that completely supporting the sensitive attributes (SA). DB$_c$ are the corresponding set of records containing the column values which completely supports the sensitive attributes. DB' is the transformed privacy preserved dataset.

IV. PERFORMANCE ANALYSIS

Some of the issues identified in the existing systems are as follows: Firstly, discrimination was detected in the original dataset which is based on only one discriminatory item and on a single measure. Secondly, it cannot guarantee that the transformed dataset is really discrimination free. Thirdly, there is no measure to evaluate how much discrimination has removed and most of the work concentrates only on the direct discrimination. Finally, [1] did not concentrate on the amount of information loss occurred.

In this paper, we tried to overcome these issues by using level based privacy preserved discrimination free data transmission model like differential privacy approaches. It will also provide high privacy preservation rate. The privacy is achieved by using some rule, which is optimized with rule generalization mechanism. These methods provide the component outcome of removing the discrimination with high privacy rate. In this approach, the removal of discrimination from the original dataset by anomalizing the attributes is done. This paper provides high security without any information loss issue. For example, more number of levels can be considered in an organization based on the designation of the agents (employees). According to those levels, the certain data can be made hidden which eliminates the discriminatory biases. The performance analysis can be made on the German Credit Dataset and on some of the dynamic datasets.

A. Data Sets

German Credit Data Set: Data set consists of 1,000 records and 20 attributes of bank account holders. This is a well known real-life data set, containing both numerical and categorical attributes. It has been frequently used in the antidiscrimination literature. The class attribute in the German Credit data set [6] takes values representing good or bad classification of the bank account holders.

Dynamic Data Sets: Three different dynamic data sets are considered for the performance analysis. The first data set is of a Diagnostic Centre which consists of about 14,078 records and 28 attributes of test details. The second data set we consider is that of the Hospital which consists of 3,986 records and 12 attributes of the details of the specimen collection. The third data set is of an Insurance Agency which consists of about 2,735 records and 13 attributes of details of customers.

As stated earlier, the performance analysis can be done on the German Credit Dataset and the three Dynamic Datasets. Firstly, we can analysis the proposed work based on the data loss rate. The unified approach for direct and indirect discrimination removal [1] led to high information loss. After processing of the dataset, we can saw lot of fields in the dataset remains blank without any data. The result of the analysis work of the data loss rate is plotted in a graph (Fig. 2). The German Credit Dataset shows massive information loss rate (86.9%), which means that there will be more blank spaces instead of data after processing. The dynamic datasets under consideration gives considerably less data loss rate, which we cannot able to identify easily.

Secondly, while in the processing stage of the datasets, time is the crucial factor under consideration. The time required for the German Credit dataset and the dynamic datasets for processing of 1000 records is recorded and graph is plotted in Fig.3. The time (in ms) for pre-processing is high for German Credit dataset when compared with proposed dynamic datasets.

Thirdly, the performance analysis can be made on the basis of the time required for discrimination removal and whether the corresponding algorithm can avoid the discrimination by considering the fields that are discriminatory. The output of the analysis of German Credit dataset is shown in the Table 1. Here, there are 1000 rows and 21 attributes and some of the discriminatory data cannot be hided. In the same way, the time and discrimination removal status of the three datasets are given in Table 2. The time required for the discrimination removal is only few milliseconds (ms) and in all the three cases, the discrimination removal status is “Success”. The graphical representation of the analysis that is depicted in Table 1 and Table 2 is shown in Fig. 4.

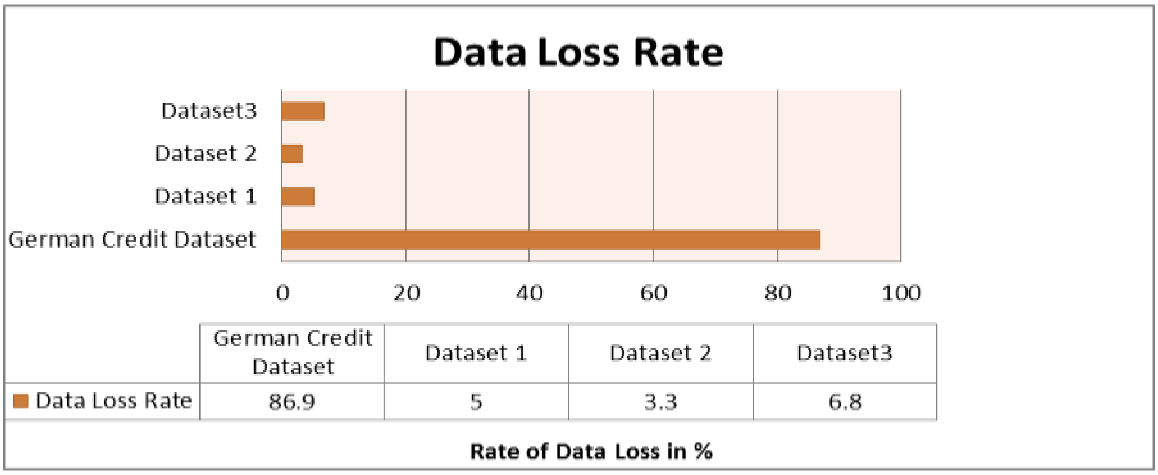

Fig.2. Data Loss Rate Representation

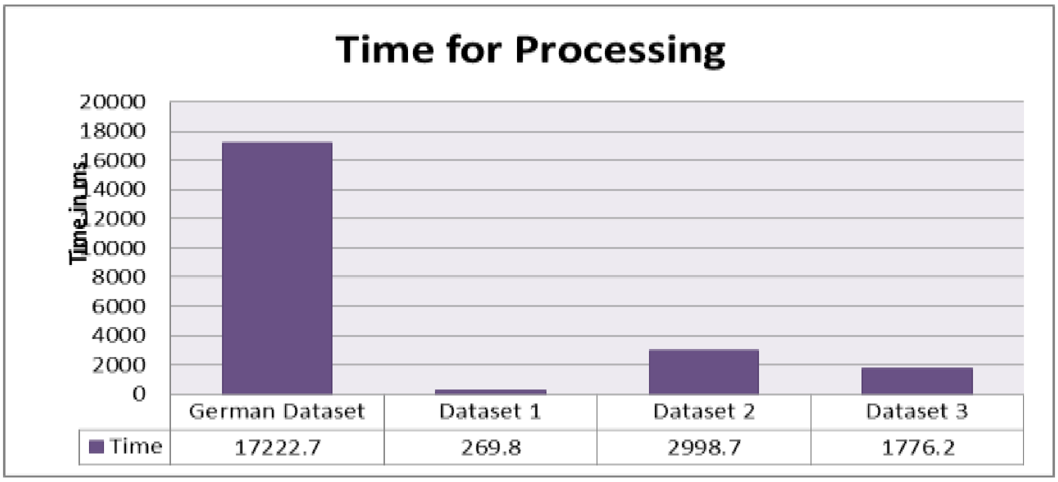

Fig.3. Time for Processing Representation

German Credit Data Set

No of Rows Count: 1000 No of Column Count: 21 Class Name: A21

Table Name	No of columns	No of rows	Discrimination	Discrimination Removal Status	Time taken (ms)
A9	4	0	NA	Fail	-
A10	3	2	100	Success	1000
A12	4	18	100	Success	1244
A14	3	3	100	Success	500
A15	3	2	100	Success	1100
A17	4	1	100	Success	1210
A19	2	0	NA	Fail	-
A20	2	0	NA	Fail	-
A1	4	0	NA	Fail	-
A3	5	0	NA	Fail	-
A4	9	1	100	Success	700
A6	5	0	NA	Fail	-
A7	5	3	100	Success	300

Table 1. Discrimination removal – German Credit Dataset

Dataset Name	Total No of columns	No of Discrimination column affected	No of Rows	Discrimination Removal Status	Time taken(ms)
Dataset 1	28	4	1000	Success	120
Dataset 2	12	5	1000	Success	32
Dataset 3	13	9	1000	Success	110

Table 2. Discrimination removal – Dynamic Datasets

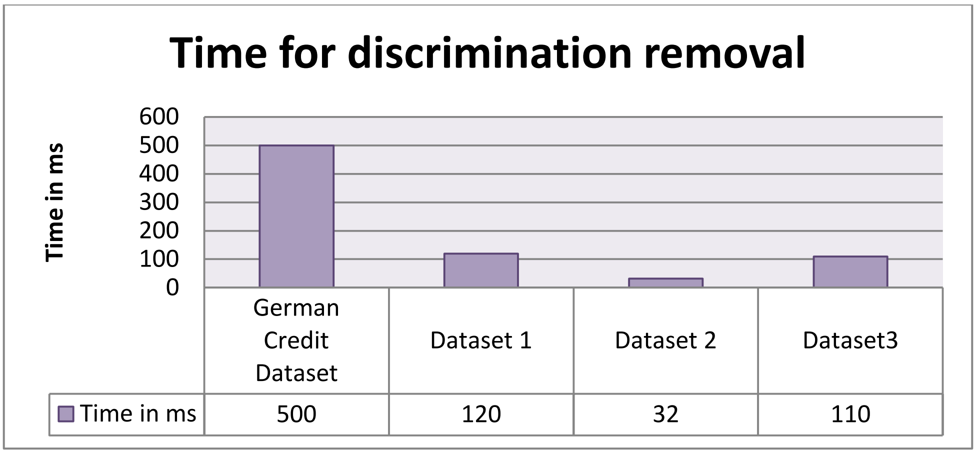

Fig.4. Time for Discrimination Removal Representation

V. FUTURE ENHANCEMENT

While applying privacy preservation to appropriate level of users, slicing with each attribute in exactly one column is considered. To enhance the data quality, overlap slicing which releases more attribute correlations can be used.

VI. CONCLUSION

The description of the multilevel based privacy preserved discrimination free data transmission can be concluded by pointing out the essentiality of the system. The proposed system is dynamic in nature and thus can be installed in any organization and can handle any databases. According to the degree of privacy required to each agent, the different privacy preservation rule sets are applied to it. The main objective that the system focuses is the correlation between the discrimination prevention and privacy preservation. So, we can conclude that by preserving privacy itself, the discrimination can be avoided.

REFERENCES

[6] F. Kamiran and T. Calders, “Classification with no Discrimination by Preferential Sampling,”

[1] Sara Hajian, Josep Domingo-Ferrer, “A Methodology for Direct and Indirect Discrimination Prevention in Data Mining,” IEEE Transactions on Knowledge and Data Engineering, Vol. 25, No. 7, July 2013.

[2] R. Agrawal and R. Srikant,“Fast Algorithms for Mining Association Rules in Large Databases,” Proc. 20th Int’l Conf. Very Large Data Bases, pp. 487-499, 1994.

[3] S. Hajian, J. Domingo-Ferrer, and A.Martı´nezBalleste´,“Discrimination Prevention in Data Mining for Intrusion and Crime Detection,” Proc. IEEE Symp. Computational Intelligence in Cyber Security (CICS ’11), pp. 47-54, 2011.

[4] Sara Hajian, Josep Domingo-Ferrer, and A. Martı´nez- Balleste´, “Rule Protection for Indirect Discrimination Prevention in Data Mining,” Proc. Eighth Int’l Conf. Modeling Decisions for Artificial Intelligence (MDAI ’11), pp. 211-222, 2011.

[5] F. Kamiran and T. Calders, “Classification without Discrimination”, Proc. IEEE Second Int’l Conf. Computer, Control and Comm.(IC4 ’09),2009.

[7] Faisal Kamiran, Toon Calders, and M. Pechenizkiy, “Discrimination Aware Decision Tree Learning,”

Proc. IEEE Int'l Conf. Data Mining (ICDM '10), pp. 869-874, 2010.

[8] D. Pedreschi, S. Ruggieri, and F. Turini, "Discrimination-Aware Data Mining," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (KDD '08), pp. 560-568, 2008.

[9] D. Pedreschi, S. Ruggieri, and F. Turini, "Measuring Discrimination in Socially-Sensitive Decision Records," Proc. Ninth SIAM Data Mining Conf. (SDM '09), pp. 581-592, 2009.

[10] S. Ruggieri, D. Pedreschi, and F. Turini, "Data Mining for Discrimination Discovery," ACM Trans. Knowledge Discovery from Data, vol. 4, no. 2, article 9, 2010.

[11] Tiancheng Li, Ninghui Li, Senior Member, IEEE, Jian Zhang, Member, IEEE, and Ian Molloy, "Slicing: A New Approach for Privacy Preserving Data Publishing", IEEE Transactions On Knowledge And Data Engineering, Vol. 24, March 2012

[12] Ninghui Li, Tiancheng Li, Suresh Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity", Proc. IEEE 23rd In'l Conf. Data Engineering, (ICDE 2007)