International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Privacy Preserving with Association Rule Mining using Evolutionary Algorithm

Sasanko Sekhar Gantayat, Bichitrananda Patra, Niranjan Panda, Manoranjan Parhi

Abstract*: *Privacy-Preserving-Data-Mining (PPDM) is a novel study which goals to protect the secretive evidence also circumvent the revelation of the evidence through the records reproducing progression. This paper focused on the privacy preserving on vertical separated databases. The designed methodology for the subcontracted databases allows multiple data viewers besides vendors proficiently to their records securely without conceding the secrecy of the data. Privacy Preserving Association Rule-Mining (PPARM) is one method, which objects to pelt sensitivity of the association imperative. A new efficient approach lives the benefit since the strange optimizations algorithms for the delicate association rule hiding. It is required to get leak less information of the raw data. The evaluation of the efficient of the proposed method can be conducting on some experiments on different databases. Based on the above optimization algorithm, the modified algorithm is to optimize the association rules on vertically and horizontally separated database and studied their performance.

Keywords:* *Privacy Preserving, Optimization, Vertically Separated Databases, Horizontally Partitioned Databases.

I. INTRODUCTION

Now days, the PPDM takes an important anxiety owing to the fast growing of electronics records in government, corporation and altered organization. Such records may indirectly encompass delicate evidence and can prime to secrecy or refuge extortions if they are altered. Due to rapid changes of the data mining technologies, getting user's sensitive information from the data is very easy.

Association rule hiding is a subarea of PPDM, which trainings the lateral effect of data mining method, which produced since disclosing the penetrating evidence, fit to users or administrations. There are several protracted customary of methods, that composed records or information pattern removed from the record has to be united through others to assist possessor or association actual resolves[8]. The distributing of records or data do at a cost to secrecy owed to dual focal aims: (i) uncertainty records mention to entities, at that time of confession may disrupt the confidentiality of the objects, which is verified in that records, and (ii) the condition the records respect to commercial/ administrations evidence, the exposer of this records. Otherwise, some information excavated as of the records may expose delicate profession confidences that can suggest a vital benefit to commercial participants and therefore may source the record holder to drop corporate ended his earls. The method of beating association rules are a category of record implication mechanism, however the situation fore-most objectives are toward defend the complex rule not the complex records [1]. Popular association rules are beating, the sets of complex association rule is stated through the safety administrators or records holder, to clean the records that will be incapable to excerpt the complex rule and excavation completely the noncomplex rule.

Some technique has been recycled to hide complex ARs through undertaking several deviations in the novel datasets ([2], [3]). In place of the usage of data mining procedure to excerpt, valuable outlines for huge records are growing. Concerns nearby the expose of secluded evidence through this method also raised. PPDM are useful in completely data mining techniques that are assembling, classifications and association rules. Algorithms of this area avoid the depiction of secluded information's, although maintaining the convenience of noncomplex evidence as such as probable by adaptation and alteration of the databank.

Several algorithms are available in its expanse, all of that have dissimilar strength and faults ([4], [5], [20]). The directed activity in PPARM, can be divided into three main categories, Broad based, exact and heuristic approach. One of the met heuristic algorithms is Cuckoo optimization algorithm. Seeing the above stated particulars, the foremost anxiety and detached of this training was to usage modified Cuckoo optimization algorithm. Using this algorithm in privacy preserving ARM the sensitive information completely or partially hidden successfully and by minimum cross effect by, (i) No consuming beating catastrophe, (ii) Reducing trace statute quantity and, (iii) Declining missing rules extra-ordinarily. The collections of numerical evidence by governments, businesses, and entities have created a condition that enables large-scale data mining and statistics enquiry.

Manuscript published on November 30, 2019.
* Correspondence Author
Sasanko Sekhar Gantayat*, Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh, India,
Bichitrananda Patra*, Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India,
Niranjan Panda, Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India,
Manoranjan Parhi, Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India,

Furthermore, by common benefit, or by regulation that requires assured records to be distributed, there is a request for distribution records amongst several revelries. Information participating has an extensive past information technology. Traditional data sharing rises to interactions of data between a data holder and a data recipient [6].

The present privacy practice depends on plans and rules to confine the sorts of publishable information, and concurrences on the utilization and capacity of sensitive information [9]. The confinement of this methodology is that it either mutilates information unreasonably or needs a trust level that is unrealistically great in numerous data-sharing circumstances. What's more, arrangements and rules cannot prevent adversaries who don't observe manages in any case. Agreements and agreements cannot ensure that sensitive data will not be indiscreetly lost and end up in inappropriate hands.

II. PRESENT WORK

A.Collections and Reproducing of Data

An ordinary situation of information collections and re-producing is depicted in Fig-1. Here dataset assembly stage, the dataset frame gathers information from records owner. The data reproducing stage, the data frame discharges the gathered information to a dataset collier or general society, called the dataset receiver, who will at that point lead in-formation mining on the distributed information. Take one instance, a clinic collects records from patient and reproduces the patient record to an outside health center [13]. At here model says, the clinic is the records vendor, patient is records owner, and health-center is the records receiver. The data mining directed at the health-center might be any investigation charge from a straightforward total of the sum of people with diabetic to a cultured cluster study.

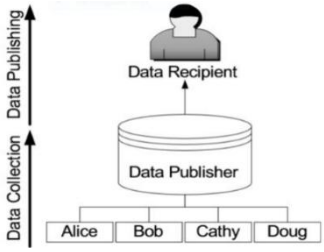


Fig. 1. Data Collections and Reproducing of Data

Here dual model of record holders. This untrusted model, the record vendor is not reliable and it can endeavor to recognize delicate evidence from records owner. Different cryptographies arrangements, unidentified communications, and statistical approaches are projected to gather record namelessly from their owners deprived of uncovering the owner character. Confided in this prototypical, the information holder is dependable and records owner are happy toward give their individual evidence to the records holder; be that as it may, the trust isn't transitive to the data recipient.

B. Privacy-Preserving Data Publishing

The record frames have a counter of the clear identifiers, pseudo identifiers, sensitive attribute and non-sensitive at-tribute. Where explicit identifiers are sets of attribute i.e. Name, SSN and information of record owner. Quasi identifiers are sets of attribute i.e. identification records of owner. Sensitive attribute consists of infection, income, and infirmity status. Non-sensitive attribute contains those attributes are not in preceding groups. Maximum works assume which four set are separate besides every record represent a discrete records owner ([10], [11], [12]).

Anonymization alludes to the PPDP approaches look to conceal character as well as the delicate information of record owner, accepting that sensitive information must be held for information analysis (Fig-2).

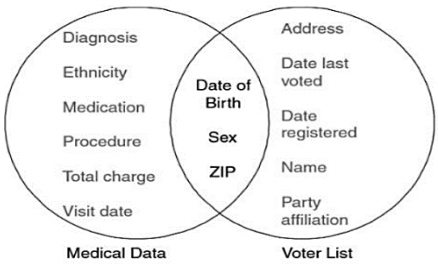


Fig. 2. Linking to re-identify record owner [10]

C. Attack Models and Privacy Models

Admittance to the distributed records ought not to empower the opponent to pick up everything additional almost some objective target with no access to the database, uniform through the nearness of some opponent's related experience got from different source ([14], [15], [20]). Popular a genuine application, complete confidentiality defense is incredible owed to the occurrence of the opponent's related information. Maximum works on privacy-preserving-data-publishing (PPDP) reflects an extra relaxed, extra-applied idea of confidentiality defense by supposing the challenger has restricted related awareness. Under, the word "victim" mentions to the records holder embattled by the opponent. We may approximately categorize secrecy copies to dual category centered on their occurrence ideologies (Fig 3).

Fig. 3. Privacy Model [10]

Privacy Model	Attack Model			
	Record linkage	Attribute linkage	Table linkage	Probabilistic attack
k-Anonymity [201, 217]	✓			
MultiR k-Anonymity [178]	✓			
ℓ-Diversity [162]	✓	✓		
Confidence Bounding [237]		✓		
(α, k)-Anonymity [246]	✓	✓		
(X, Y)-Privacy [236]	✓	✓		
(k, e)-Anonymity [269]		✓		
(ϵ, m)-Anonymity [152]		✓		
Personalized Privacy [250]		✓		
t-Closeness [153]		✓		✓
δ-Presence [176]			✓	
(c, t)-Isolation [46]	✓			✓
ϵ-Differential Privacy [74]			✓	✓
(d, γ)-Privacy [193]			✓	✓
Distributional Privacy [33]			✓	✓

D. Technical Description

Input Dataset

In this experiment the census Income dataset is used ([19], [20]). The data set has the attributes {Age, Work_class, Education_num, Occupation, Marital_status, Race, Capital_loss, Capital_gain, Nat_country, Hr_per_week, Sex, Income}.

Preprocesses of Novel Datasets

Preprocesses is the first process on the original dataset in the proposed algorithm which is greatest significant and powerful phases of the scheduled method. Then the complex substances occur impartial in convinced communications of the datasets, altering and deploying all communications is a tedious and futile assignment which expands the scope of living spaces, yet in addition aims the manufacture pointless and inconsequential explanations an expansion in the TIME devoured on behalf of acquiring progressively viable solutions; subsequently expanding the sanitization time. To maintain a strategic distance from such negative impacts, a preprocess activity have been viewed as that incorporates two stages.

Table I. Input Sam ple dataset of Census Income

Age	Work_Class	Qualification	Educ_ Num	Marital_Status	Occupation	Category	Race	Sex	Cap_Gain	Cap-Loss	HR_perWeek	Nat_Country	Income (K)
39	State-gov	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=50
50	Self-emp-not-inc	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=50
38	Private	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=50
53	Private	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=50

Accordingly, directing pre-process activity arranged novel datasets would prompt abatement of natural surroundings, decrease un-vital change in datasets, inhibition of delivering unrelated arrangements and earlier admittance to optimum solutions.

Novel Algorithm

The fundamental stages of the planned procedure entitled as "Cuckoo Optimization Algorithm for Association Rule Hiding (COA4ARH)" is shown on the flowchart.

COA4ARH Algorithm

1. Pre-processes "Original Dataset"
2. Initiate "Primary Population"
3. Run "Fitness Function"
4. Run "Best Solution Function"
5. Repeat // create novel solution for every solutions
6. K = (Max NNS – Min NNS) × Rand[0,1] + Min NNS //Dedicate K
7. MR = [∝ × Current solution _ sK Total l of all solution _ sK]×(Vhi –Vlow)
8. Run "Fitness Function"
9. Run "Best Solution Function"
10. Fix Rand [0 ,1] to the designated items
11. Bound numbers of solution to Nmax
12. for each solution in population
13. Run "Immigration Function"
14. Run "Fitness Function"
15. Run "Best Solution Function"
16. till the end of state is fulfilled

Terminate the algorithm if the desired conditions achieved

Fitness Function

The fitness value is computed for every standing solution in early populace ([16], [17]). The suitability standard for other solution is deliberate by eq. (1), (4) and (6) that mention in below:

$$minfit_1 = |HF| \quad (1)$$

where |HF| is the quantity of beating failure.

$$|HF| = |Rs|_i = 1\ r_i \text{ state} \quad (2)$$

where |Rs| is the quantity of complex rules

$$r_i \text{ state} = 0 \text{ if } sup(I_i) < MST \text{ or } Conf(r_i) < MCT$$
$$= 1 \text{ otherwise} \quad (3)$$

where Ii is the item set of ri.

The second fitness value is calculated by Eq. (4)

$$minfit_2 = |LR| \quad (4)$$

The procedure of mining the present solution in a populace for every iteration is a time consuming process. Henceforth the functions minfit1, minfit2 and minfit3 have been used in such a way that the mining of the solution is not compulsory.

Best Solution Function

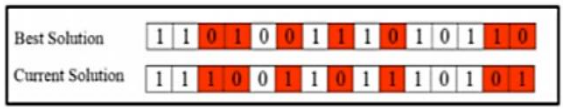
$$K = (Max\ NNS - Min\ NNS) \times Rand\ [0,1] + Min\ NNS \quad (5)$$

where K is the total of innovative solution produced by every paternal solution.

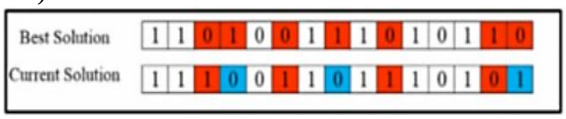
Max NNS is maximum Number of New Solutions. Min NNS is minimum Number of New Solutions. This number is Modifications Radius shown with MR.

$$MR = \propto \times \text{Current solution K Total } l \text{ of all solution K} \times (V_{hi} - V_{low}) \quad (6)$$

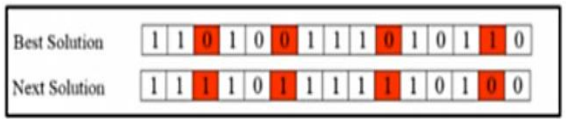
Where MR is the quantity of every paternal solution's items that changes. ∝ shows the quantity of total iterations. V_{hi} and V_{low} are respectively high and low limit of every adjustable in optimizations problems [18].



a) Different Items for best and Current Solutions



b) Arbitrary Selection of N-items from Different Items of Current Solutions



c) Modification of Selected Items and Generation of Next Solutions

Fig. 4. Preparation of Best Solution

Immigration Function

In directive to recover the outstanding results, through the resolve of refining the capability of the projected algorithm to discharge from native goals and outcome better answers associated to what is already expanded; all explanations are altered in a method to be additional comparable and faster to the finest result. This alteration may lead to the compeers of improved solution likened to the premium existing solution. Nevertheless, the best current solution will be considered as the best global solution. Based on the Hamming distance, the quantity of dissimilar objects amongst the current and best solution is identical to 7 that are demonstrated in red as exposed in Fig. 4 (a). The rate of N may currently be resolute that is identical to an arbitrary quantity of dissimilar objects. It means that 3 objects of dissimilar objects of the current solution would be altered. Henceforth after those 7 objects, 3 objects will be selected arbitrarily. The rate of all certain object would be changed. By this method, the following results are created that are alike to the best solution. Now the quantities of unmatched objects have decreased.

III. METHODOLOGY

The analysis of the proposed method is carried out using the software R and MATLAB using Intel Processor. Census Dataset is used to get the Sensitive Association Rules.

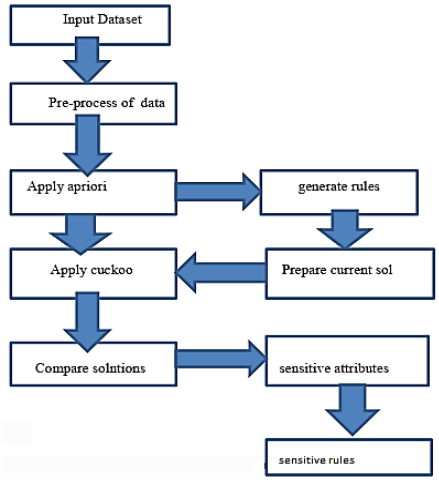
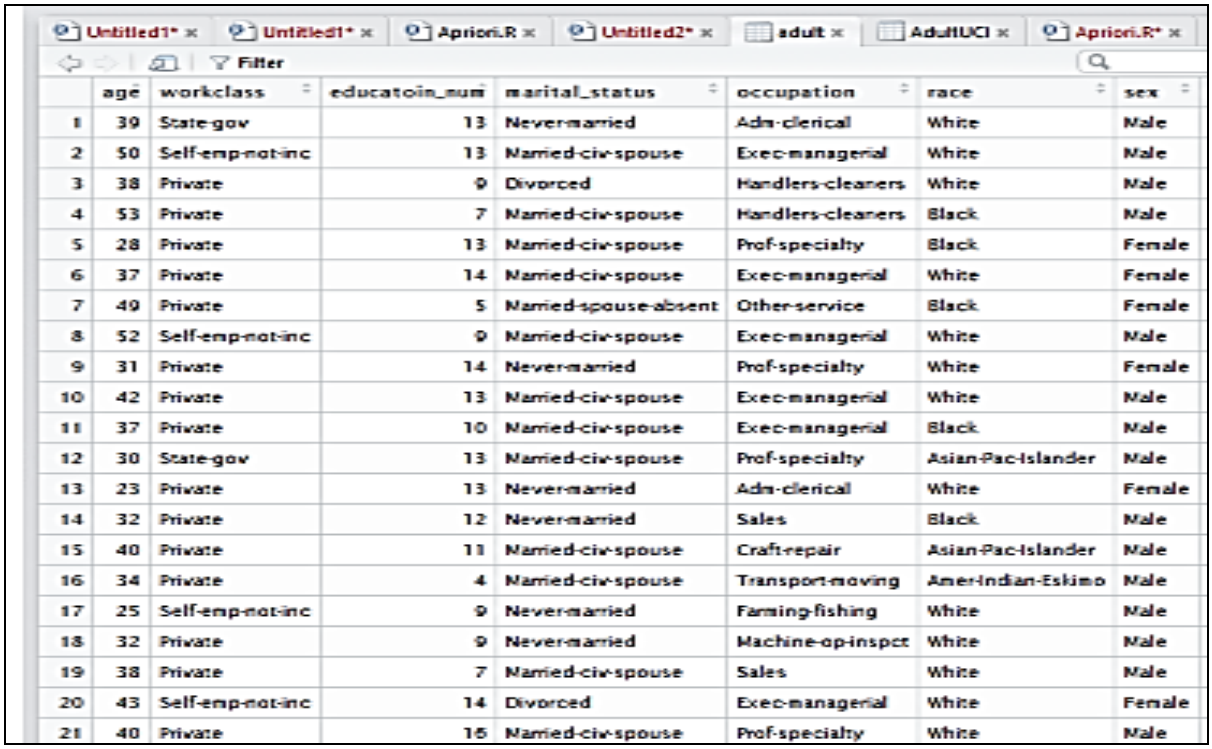


Fig. 5. Flowchart of the Process

Processing Steps

- Taking census data set and pre-processes the original dataset
- Import dataset into R and apply apriori algorithm generate Association rules
- Process of data into Excel sheet and delete unwanted attributes
- Prepare current solution
- Apply cuckoo algorithm and find best solution from graph
- Compare current and best solutions
- Identify the sensitive attributes
- Hide the sensitive rules and generate non-sensitive rules

Import Dataset into R



	age	workclass	educatoin_num	marital_status	occupation	race	sex
1	39	State-gov	13	Never-married	Adm-clerical	White	Male
2	50	Self-emp-not-inc	13	Married-civ-spouse	Exec-managerial	White	Male
3	38	Private	9	Divorced	Handlers-cleaners	White	Male
4	53	Private	7	Married-civ-spouse	Handlers-cleaners	Black	Male
5	28	Private	13	Married-civ-spouse	Prof-specialty	Black	Female
6	37	Private	14	Married-civ-spouse	Exec-managerial	White	Female
7	49	Private	5	Married-spouse-absent	Other-service	Black	Female
8	52	Self-emp-not-inc	9	Married-civ-spouse	Exec-managerial	White	Male
9	31	Private	14	Never-married	Prof-specialty	White	Female
10	42	Private	13	Married-civ-spouse	Exec-managerial	White	Male
11	37	Private	10	Married-civ-spouse	Exec-managerial	Black	Male
12	30	State-gov	13	Married-civ-spouse	Prof-specialty	Asian-Pac-Islander	Male
13	23	Private	13	Never-married	Adm-clerical	White	Female
14	32	Private	12	Never-married	Sales	Black	Male
15	40	Private	11	Married-civ-spouse	Craft-repair	Asian-Pac-Islander	Male
16	34	Private	4	Married-civ-spouse	Transport-moving	Amer-Indian-Eskimo	Male
17	25	Self-emp-not-inc	9	Never-married	Farming-fishing	White	Male
18	32	Private	9	Never-married	Machine-op-inspct	White	Male
19	38	Private	7	Married-civ-spouse	Sales	White	Male
20	43	Self-emp-not-inc	14	Divorced	Exec-managerial	White	Female
21	40	Private	16	Married-civ-spouse	Prof-specialty	White	Male

Fig. 6. Census Dataset imported into R

Output Analysis of Apriori Algorithm

```
Console C:/Users/vin8.1/Desktop/98dca27a7e48694506db6ae413d7570e-84fa919a033c87c997e67c4958c74c7934022369/
ce      lift count
[1]  {}                                  => {capital-gain=None}                0.9173867  0.91738
67 1.0000000 44807
[2]  {}                                  => {capital-loss=None}                0.9532779  0.95327
79 1.0000000 46560
[3]  {sex=Male}                          => {capital-gain=None}                0.6050735  0.90514
55 0.9866565 29553
[4]  {sex=Male}                          => {capital-loss=None}                0.6331027  0.94707
50 0.9934931 30922
[5]  {workclass=Private}                 => {capital-gain=None}                0.6413742  0.92390
73 1.0071078 31326
[6]  {workclass=Private}                 => {capital-loss=None}                0.6639982  0.95649
74 1.0033773 32431
[7]  {race=white}                        => {native-country=United-States}     0.7581127  0.92172
31 1.0270761 38493
[8]  {race=white}                        => {capital-gain=None}                0.7817862  0.91432
40 0.9966616 38184
[9]  {race=white}                        => {capital-loss=None}                0.8136849  0.95163
07 0.9982720 39742
[10] {native-country=United-States}      => {capital-gain=None}                0.8219565  0.91590
62 0.9983862 40146
[11] {native-country=United-States}      => {capital-loss=None}                0.8548380  0.95254
61 0.9992323 41752
[12] {capital-gain=None}                 => {capital-loss=None}                0.8706646  0.94907
05 0.9955863 42525
[13] {capital-loss=None}                 => {capital-gain=None}                0.8706646  0.91333
76 0.9955863 42525
[14] {workclass=Private,
      capital-gain=None}                 => {capital-loss=None}                0.6111748  0.95291
45 0.9996188 29851
[15] {workclass=Private,
      capital-loss=None}                 => {capital-gain=None}                0.6111748  0.92044
65 1.0033354 29851
[16] {race=white,
      native-country=United-States}      => {capital-gain=None}                0.7194628  0.91289
```

Fig. 7. Association Rules Generation

Process of Data into Excel Sheet

AGE	WORKCLASS	EDUC_NUM	MARITAL_STATUS	OCCUPATION	RACE	SEX	CAP_LOSS	CAP_GAIN	HR_PER_WEEK	NAT_COUNTRY	INCOME	SUM
0	0	0	0	0	0	0	1	0	0	0	0	16
0	0	0	0	0	0	0	0	1	0	0	0	8 minimum
0	0	0	0	0	0	0	0	1	0	0	0	131080
0	0	0	0	0	0	0	0	1	0	0	0	131080
0	0	0	0	0	0	0	0	1	0	0	0	4,194,312
0	0	0	0	0	0	0	0	1	0	0	0	4194312
0	0	0	0	0	0	0	0	0	0	1	0	262146
0	0	0	0	0	0	0	1	0	0	0	0	262160
0	0	0	0	0	0	0	0	1	0	0	0	262152
0	0	0	0	0	0	0	0	1	0	0	0	8200
0	0	0	0	0	0	0	0	1	0	0	0	65544

Fig. 8. Prepare Data in Binary String

Delete unwanted Attributes

	WORKCLASS	RACE	SEX	CAP_LOSS	CAP_GAIN	NAT_COUNTRY	CAP_LOSS	CAP_GAIN	NAT_COUNTRY	SUM
CURRENT SOL(1)	0	0	0	0	0	0	1	0	0	4
CURRENT SOL(2)	0	0	0	0	0	0	0	1	0	2 MINIMUM
CURRENT SOL(3)	0	0	1	0	0	0	0	1	0	66
CURRENT SOL(4)	0	0	1	0	0	0	0	1	0	66
CURRENT SOL(5)	1	0	0	0	0	0	0	1	0	258
CURRENT SOL(6)	1	0	0	0	0	0	0	1	0	258
CURRENT SOL(7)	0	1	0	0	0	0	0	0	1	129
CURRENT SOL(8)	0	1	0	0	0	0	1	0	0	132
CURRENT SOL(9)	0	1	0	0	0	0	0	1	0	130
CURRENT SOL(10)	0	0	0	0	0	1	0	1	0	10
CURRENT SOL(11)	0	0	0	0	0	1	0	1	0	10

Fig. 9. Minimize the Sum Value

Output Analysis of Cuckoo Algorithm

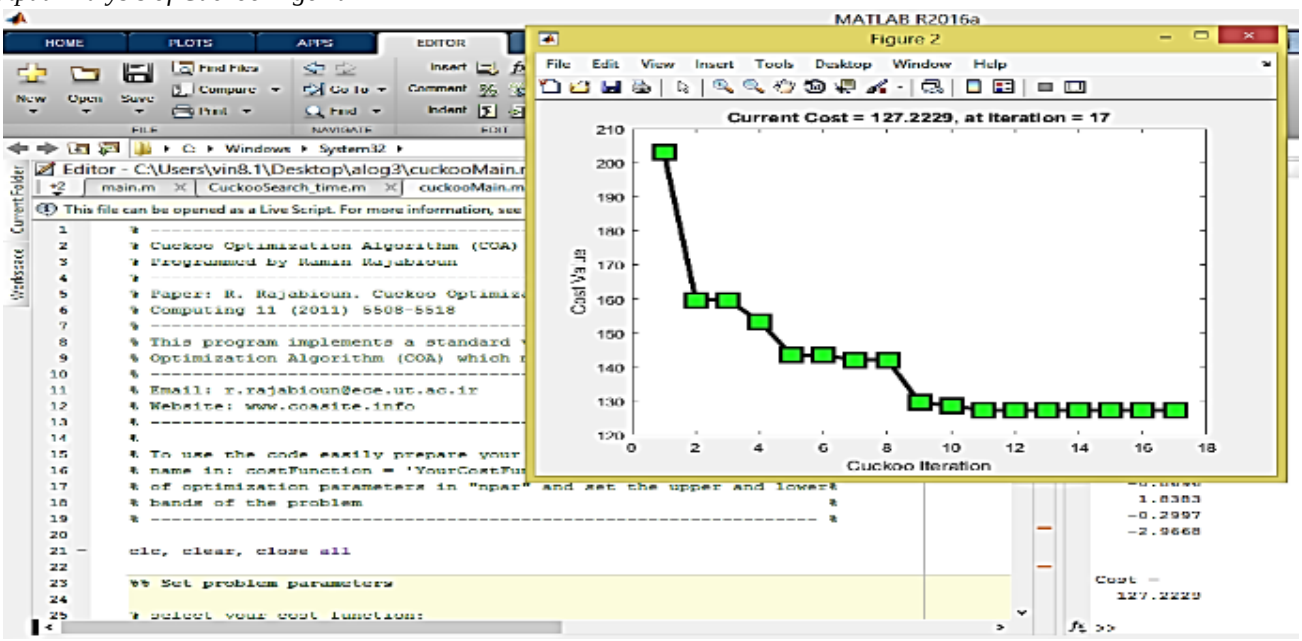
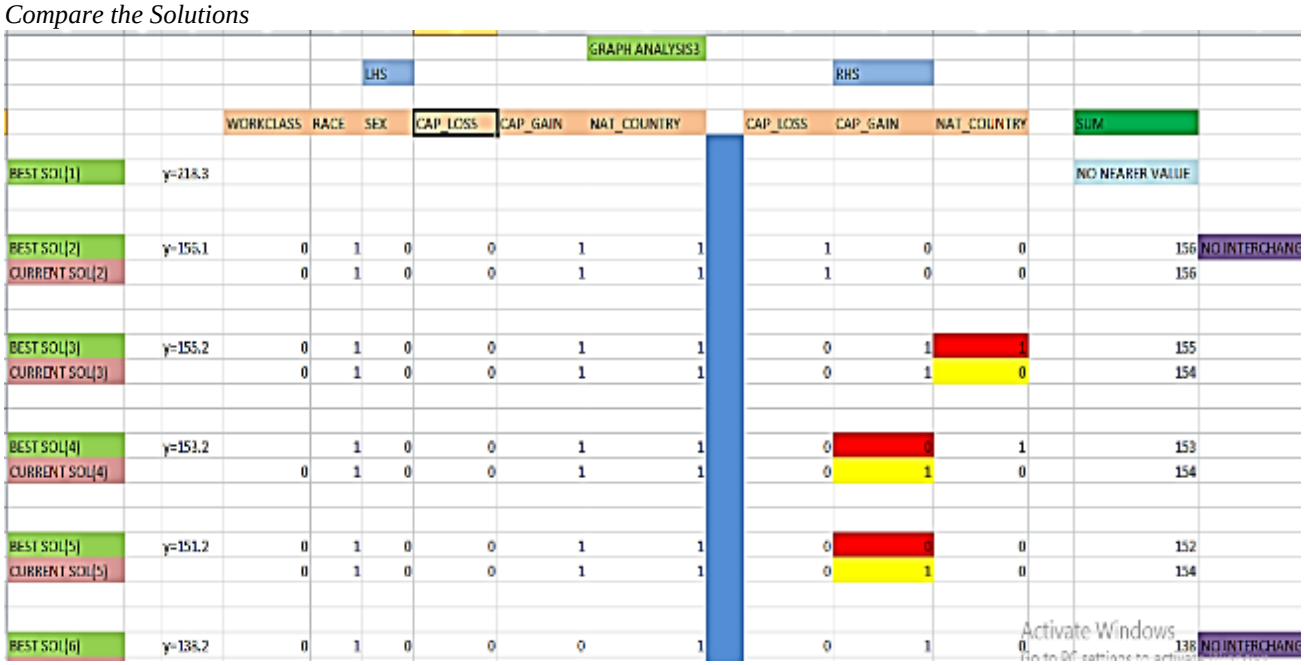


Fig. 10. Finding Best Solution from Cuckoo Algorithm

Compare the Solutions



		WORKCLASS	RACE	SEX	CAP_LOSS	CAP_GAIN	NAT_COUNTRY	CAP_LOSS	CAP_GAIN	NAT_COUNTRY	SUM	
				LHS		GRAPH ANALYSIS3			RHS			
BEST SOL(1)	y=218.3										NO NEARER VALUE	
BEST SOL(2)	y=156.1	0	1	0	0	1	1	1	0	0	156	NO INTERCHANG
CURRENT SOL(2)		0	1	0	0	1	1	1	0	0	156	
BEST SOL(3)	y=155.2	0	1	0	0	1	1	0	1	1	155	
CURRENT SOL(3)		0	1	0	0	1	1	0	1	0	154	
BEST SOL(4)	y=153.2		1	0	0	1	1	0	0	1	153	
CURRENT SOL(4)		0	1	0	0	1	1	0	1	0	154	
BEST SOL(5)	y=151.2	0	1	0	0	1	1	0	0	0	152	
CURRENT SOL(5)		0	1	0	0	1	1	0	1	0	154	
BEST SOL(6)	y=138.2	0	1	0	0	0	1	0	1	0	138	NO INTERCHANG

Fig. 11. Identify Sensitive Attributes

Sensitive Association Rules Hidden

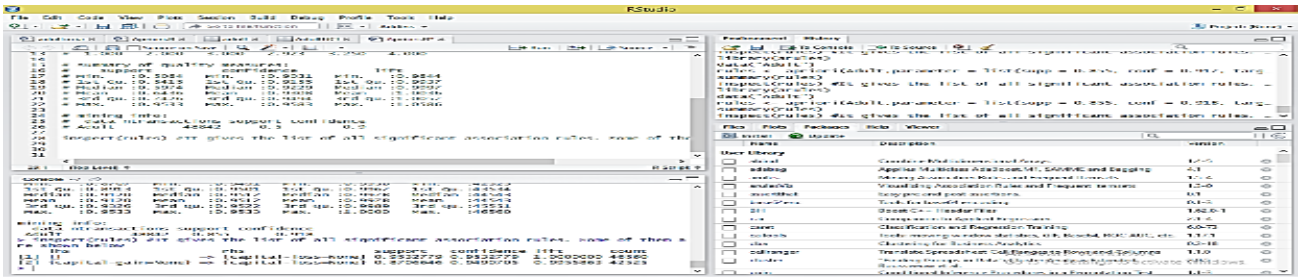


Fig. 12. Non-sensitive Rules

IV. CONCLUSION AND FUTURE WORK

The current work has projected a modified met heuristic technique for beating complex association rule by cuckoo optimization algorithm and considered few attributes to check the algorithm. The projected procedure is skilful of instantaneously beating some profound association rule. The greatest vital and significant feature of the current training is describing pre-processes procedures that include dual stages in the opening of planned process algorithm. This pre-processes procedure origins a incredible fall in the amount of repetitions and immediate access to the optimum solutions. In this paper, there are three suitability procedures have been presented that can invent the result with least lateral effect. Additional an immigration function has been well-defined that enhanced the capability of the pro-jected procedure to emission from limited optimum.

Describing a novel "fitness function" which may shrinkage the extent of Lost Rules and reserve the process's proficien-cy of beating profound rule and evading group of presence rules shall be the resolve of upcoming studies. Also, finished certain computation, the number of subtle item that should be removed for beating subtle rules can be designed; only this quantity of complex item would be erased to shrinkage the numbers of lost rule.

REFERENCES

1. Lichun Li, Rongxing Lu, Kim-Kwang Raymond Choo, Anwitaman Datta, Jun Shao, "Privacy-Preserving-Outsourced Association Rule Mining on Vertically Partitioned Databases", IEEE Trans. Information Forensics and Security 11(8), 2016, pp. 1847-1861.
2. Peipei Sui, Xian Xxian Li,"A Privacy preserving Ap-proach for Multi-Transaction Data Integrated Analysis", Neurocomputing, Vol. 253, 2017, pp.56-64.
3. Gehad Ahmed Sultan , Laila Abd_Ellatif, "Association Rules Hiding for Privacy Preserving Data Mining", Interna-tional Journal of Computer Applications, Vol. 150, No.12, 2016, pp.34-43.
4. Shuo Qiu, Boyang Wang, Ming Li, Jiqiang Liu, and Yanfeng Shiohra,Toward, "Practical Privacy-Preserving Frequent Itemset Mining on Encrypted Cloud Data", IEEE Trans-actions on Cloud Computing, Vol. 1, Issue.1, 2017, pp.99.
5. Madhu V. Ahluwalia, Aryya Gangopadhyay, Zhiyuan Chen, and Yelena Yesha, "Target-Based, Privacy Preserving, and Incremental Association Rule Mining", IEEE Transactions on Services Computing, Vol.10, Issue. 4, 2017, pp.633-647.
6. N. Kowsalya, M. Saraswathi, "Privacy Preserving and Secure Mining of Association Rules in Distributed Database", IJCSMC, Vol. 4, Issue. 1, 2015, pp.94-101.
7. Ashwini, S.Chavan, Rahul, P. Mirajkar,"Privacy Pre-serving Data Mining Using Inverse Frequent Item Set Mining Approach", IRJET, Vol. 4, Issue. 2, 2017.
8. Shubnum Rehaman, Anil Sharma, "Privacy Preserving Data Mining Using Association Rule Based Apriori Algorithm", Advanced Informatics for Computing Research: First International Conference, ICAICR 2017, Jalandhar, India, March 17–18, Revised Selected Papers, 2017,pp. 218-226.
9. Khan, A., Qureshi, M. S., Hussain, A., Improved genetic algorithm approach for sensitive association rules hiding, World Applied Sciences Journal, Vol.31, 2014, pp. 2087–2092.
10. Benjamin C. M. Fung, Ke Wang, Ada Wai-Chee Fu, and Philip S. Yu, Introduction to Privacy-Preserving Data Publishing-Concepts and Techniques, Taylor & Francis Group, Boca Raton, FL, 2011.
11. Yang, X.-S. , & Deb, S., Multiobjective cuckoo search for design optimization, Computers & Operations Research, Vol.40, 2013, pp.1616–1624.
12. Walton, S., Hassan, O., Morgan, K., & Brown, M., Modified cuckoo search: A new gradient free optimisation algo-rithm, Chaos, Solitons & Fractals, Vol.44, 2011, pp.710–718.
13. Gkoulalas-Divanis, A., & Verykios, V. S., Exact knowledge hiding through database extension, IEEE Transac-tions on Knowledge and Data Engineering, Vol. 21, 2009, pp. 699–713.
14. Gkoulalas-Divanis, A., Verykios, V. S., Hiding sensi-tive knowledge without side effects, Knowledge and Infor-mation Systems, Vol. 20, 2009, pp.263–299.
15. Moustakides, G. V., Verykios, V. S., A max-min ap-proach for hiding frequent Itemsets, Data & Knowledge Engineering, Vol.65, pp. 75–89, 2008.
16. Sun, X., Philip, S. Y., Hiding sensitive frequent item-sets by a border-based approach. Journal of Computing Science and Engineering, Vol 1, 2007, pp.74–94.
17. Wang, S.-L., Parikh, B., Jafari, A., Hiding informative association rule sets, Expert Systems with Applications,Vol. 33, 2007, pp. 316–323.
18. Wu, Y.-H., Chiang, C.-M., & Chen, A. L., Hiding sensitive association rules with limited side effects, IEEE Transactions on Knowledge and Data Engineering, Vol.19, 2007, pp. 29–42.
19. Verykios, V. S., Pontikakis, E. D., Theodoridis, Y., Chang, L., Efficient algorithms for distortion and blocking tech-niques in association rule hiding, Distributed and Parallel Databases, Vol. 22 , 2007, pp. 85–104.
20. Kumar, B.S., Karthik, S. & Arunachalam, V.P, Upkeeping secrecy in information extraction using 'k' division graph based postulates, Cluster Computing, Vol. 57, Issue. 22 (Suppl 1), 2019.
21. https://doi.org/10.1007/s10586-018-1705-2
22. https://archive.ics.uci.edu/ml/datasets/census+income, 2019.

AUTHORS PROFILE

Dr. Sasanko Sekhar Gantayat is working as Associate Professor in the Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Prades, India. He received his master degree in Computer Science and Ph. D. in Computer Science from Berhampur University, Odisha, India. His research interests are in Soft Computing, Data Mining, Social Networks, Rough Set, Statistical and Machine Learning. He has published more than 24 research papers in international journals and conferences and having membership for different professional bodies like IEEE, ISTE, CSI, ACM, IACSIT, WSEAS, IRS, IEANG, IOSRD etc.

Dr. Bichitrananda Patra is an Associate Professor at the Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, He received his master degree in Computer Science from the Utkal University, Bhubaneswar, and Ph. D. in Computer Science from Berhampur University, Orissa, India. His research interests are in Soft Computing, Algorithm analysis, statistical neural nets. He has published 20 more research papers in international journals and conferences organized national workshops and conference and also having membership for different professional bodies like ISTE, CSI etc.

Dr. Niranjan Panda is an Assistant Professor in the Computers Science and Engineering Department, Siksha 'O' Anusandhan University, Bhubaneswar, India since 2011. He got his Ph.D. degree from Siksha 'O' Anusandhan University, M.Tech. degree in Computer Science and Engineering from KIIT University, Bhubaneswar, India and B.Tech. degree in Computer Science and Engineering from Utkal University, Bhubaneswar, India. His research interests include Ad hoc networks, Computer Security, Intelligent Systems and Image Processing.

Dr. Manoranjan Parhi is currently working as an Associate Professor in the Department of Computer Science and Engineering, ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. He has completed his doctoral degree in the area of Service Oriented Computing and Cloud Computing. He has been actively involved in the research in the area of Distributed Computing, AI and Machine Learning and has more than 20 papers to his credit.