Received: July 14, 2017 156

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Krill Based Optimal High Utility Item Selector (OHUIS) for Privacy Preserving Hiding Maximum Utility Item Sets

Ketthari Thandapani[1*] **Sugumar Rajendran[2]**

[1]*St.Peter's University, Avadi, Chennai, India*
[2]*Velammal Institute of Technology, Chennai, India*
* Corresponding author's Email: kettharithandapani12@gmail.com

Abstract: Privacy Preserving Data Mining (PPDM) has turned into a well-known research region. Step by step instructions are adjusted between privacy assurance and knowledge discovery in the sharing procedure is a critical issue. Likewise, currently, not more strategies are accessible in the literature to hide the sensitive itemsets in the database. One of the existing privacy-preserving utility mining strategies uses two algorithms, HHUIF and MSICF to disguise the sensitive item sets so that the foes can't mine them from the modified database. This paper concentrates the privacy for the sensitive data actualize in a hybrid approach that is Optimal High Utility Item Selector (OHUIS) method. This Item Selector (IS) procedure upgrades the threshold utilizing inspired Krill Herd Optimization (KHO) method. With a specific end goal to hide the sensitive item sets, the frequency estimation of the items is changed. On the off chance that the utility estimations of the items are same, the OHUIS algorithm chooses the precise items and after that, the frequency estimations of the selected items are altered. The proposed OHUIS reduces the computation many-sided quality and also enhances the hiding performance of the item sets. The algorithm is actualized and the resultant item sets are thought about against the item sets that are acquired from the traditional privacy preserving utility mining algorithms. The test comes about demonstrate that OHUIS accomplishes the lower miss costs compared to the existing procedures for various manufactured datasets.

Keywords: Utility mining, Data mining, Privacy, Sanitized data, Sensitive data, Optimization, Hiding high utility data.

1. Introduction

A standout amongst the most vital approaches in data mining, conventional association rule mining finds all item sets which bolster qualities are more prominent than the given threshold [1]. It has been relied upon to acknowledge privacy preserving data mining with a specific end goal to obtain profitable knowledge from the joined data wellsprings of a few specialist organizations [2]. Particularly, utility mining will probably bring about privacy breaks during the time spent business data investigation since it can discover important data including delicate data escaped substantial scale business data and organizations may noxiously use the data to improve their own benefits [3, 4]. Privacy preserving data mining has increased expanding prevalence in the data mining research group. Accordingly, another arrangement of methodologies was acquainted with considering data mining, while, in the meantime, disallowing spillage of private and delicate data [5]. Most existing methodologies can be ordered into two classifications: systems that ensure the touchy data itself in the mining procedure, and strategies that secure the delicate data mining comes from, i.e., the extricated knowledge, that was delivered by the utilization of data mining [6]. Despite the fact that, these strategies are more suited to mining high utility patterns from a database that is incrementally expanded in a dynamic situation when contrasted with existing static ones, they create hopeful patterns in the mining procedure and require an extra procedure to distinguish genuine examples since they apply the overestimation idea, or they play out an example mining process without

competitor era, yet there is a hindrance in that they require various database filters for unique or extra data[7,8]. Consequently, this examination concentrates on privacy preserving utility mining and introduces one novel algorithm as Optimal High Utility Item Selector (OHUIS), to accomplish the objective of hiding sensitive item sets. We design a novel assessment capacity to consider the symptoms in PPDM, for example, hiding disappointments, missing expenses, and artificial expenses. Whatever is left of this paper is sorted out as takes after. Section 2 are surveyed Recent Literature. At that point, section 3 proposes the OHUIS algorithms to enhance the security insurance and knowledge discovery. Section 4 introduces the trial comes about and assesses the execution of the proposed algorithms. At last, segment 5 introduces the finishes of our work.

2. Literature review

In 2016 J. C. W. Lin *et al*. [9] have proposed the two novel algorithms, in particular, Maximum Sensitive Utility-Maximum item Utility (MSU-MAU) and Maximum Sensitive Utility-Minimum item Utility (MSU-MIU), was separately planned to limit the reactions of the purification procedure for concealing SHUIs.

Privacy-Preserving Utility Verification of the Data Published by Non-interactive Differentially Private Mechanisms by J. Hua *et al.* in 2016 [10] have proposed the protection of safeguarding utility confirmation instrument in view of cryptographic strategy for different Part – a differentially private plan intended for set-esteemed data.

In 2016 D. Brettschneider *et al*. [11] have recommended the Private, a protection safeguarding algorithm for DEM. It uses homomorphism encryption to secretly accumulate collected data and perform vitality administration in view of the maximum, minimum decency rule.

In Y. A.A. Aldeen *et al*. [12] had recommended an all-encompassing diagram on the latest point of view and methodical translation of a run down distributed written works by means of their fastidious association in subcategories. Some literary works in light of privacy preserving data mining are talked about in the literature. The conventional association rule mining neglects to acquire the infrequent itemsets with a higher benefit. Since conventional association, rule mining strategy treats every one of the items in the database similarly by considering just the presence of items inside the transaction. The above issue can be fathomed utilizing the Utility Mining system. The major thoughts of the current protection saving data mining techniques, their benefits, and weaknesses are displayed. Amid mining process, we ought not to recognize either visit or uncommon item sets but rather distinguish itemsets which are more valuable to us. The proposed technique diminishes the computational complexity and also enhances the hiding execution of the item sets. The algorithm is actualized and the resultant item sets are looked at against the item sets that are acquired from the ordinary privacy preserving utility mining algorithms. This attentive examination uncovers the past improvement, introduce inquire about difficulties, future patterns, the crevices, and shortcomings. Advance noteworthy improvements for more powerful security assurance and protection are affirmed to be obligatory.

3. Proposed methodology

Privacy protecting against mining algorithms is a new examination zone that looks at the reactions of data mining systems that acquired from the privacy diffusion of people and associations. In our present work, we have utilized Hiding High Utility Item First (HHUIF) figuring. The HHUIF count covers the sensitive item sets having high utility regard. However, the weakness of this calculation is that, if the items in the sensitive item set having same utility regard, then it will reduce the hiding execution.

For instance, is a sensitive item set, having utility regard level, To cover the item set, here the repeat estimation of the item in the item set having high utility value is changed. To comprehend this drawback, we have proposed HHUIF estimation with Item Selector (IS) with optimization techniques considered that is Optimal High Utility Item Selector (OHUIS) it's appeared in Fig. 1. The Item Selector is used to pick the high utility esteem thing set by using the obtainable calculation, and in this way the repeat estimation of the picked things is balanced. The high utility esteem thing sets are concealed by adjusting the repeat estimations of things contained in the fragile thing sets in light of the base utility threshold regard. This threshold optimization Process animated Krill Herd Optimization (KHO) edge based technique utilized, in light of this ideal IS with threshold used to keep up the security data.

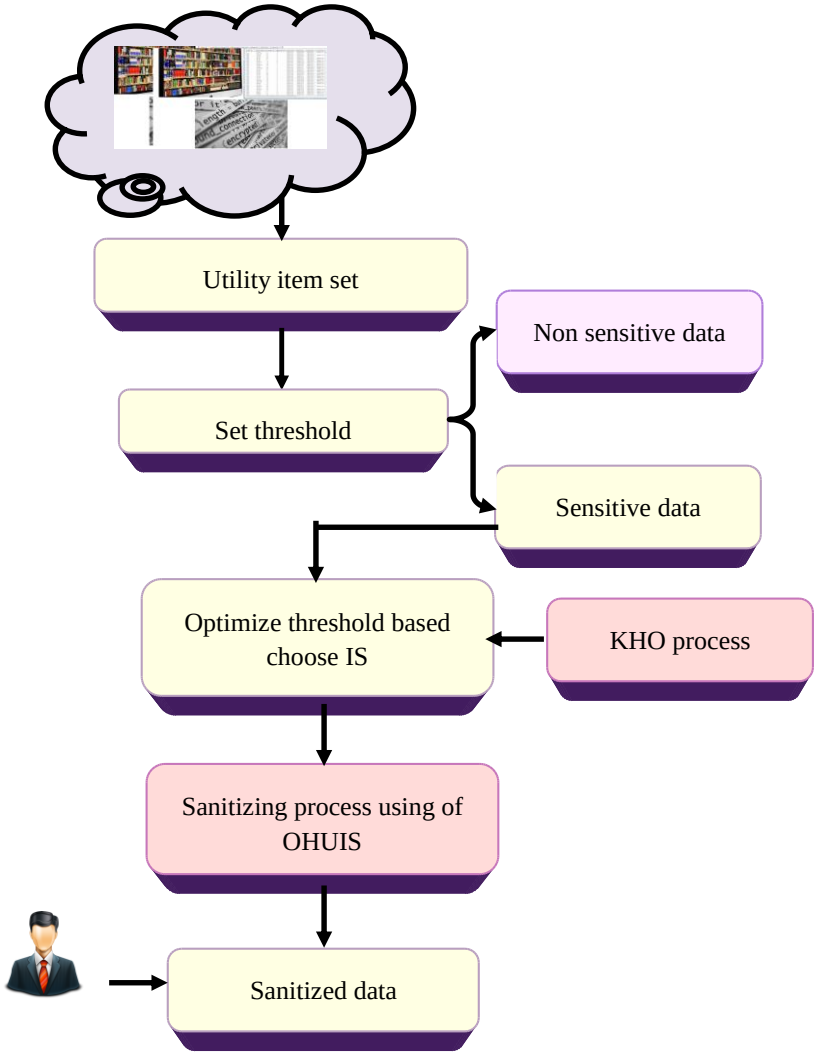


Figure.1 Block diagram for proposed research model

The main contribution of research model is described below

- Conversion of Original database into sanitized database
- Plan the HHUIF algorithm to shroud sensitive item.
- Upgrade the threshold to discover IS of the algorithm used KHO model, so the proposed approach known as OHUIS model.

3.1 Utility mining process

The limitations of incessant or uncommon item set mining enlivened specialists to consider a utility based mining approach, which enables a client to helpfully express his or her viewpoints concerning the value of item sets as utility values and then discover item sets with high utility qualities higher than an edge.

3.2 Hiding High Utility Item First (HHUIF) algorithm

The Utility item set mining, likewise, for the most part, called the utility pattern mining, was first presented in every item in the item sets is related to an extra esteem, called the interior utility which is the amount (i.e. number) of the item. This procedure is utilized to locate the utility values in the item sets.

The utility value of item set I_s in transaction N is defined as,

$$U(I_s, N) = k(I_s) \times m(I_s, N) \quad (1)$$

The utility value of an item set S in transaction T_O is defined as,

$$U(S, N) = \sum_{i_m \epsilon S} U(I_s, N) \quad (2)$$

From that point, such item sets are found out whose utility value surpasses the user-defined threshold value τ, where τ speaks to the minimum utility threshold. The item set S portrays a high utility itemset, if $U(S) \geq \tau$. These high utility itemsets

are amassed in $H=\{s_{1,} s_{2,\ldots,} s_{i,}\}$ and these item sets are known as the sensitive item sets. Encourage, it is fundamental to head the sensitive item sets in view of certain security tricks.

3.3 Item selector in HHUIF algorithm

The Item Selector is utilized to choose the high utility value item set by utilizing the accompanying algorithm, and in this manner, the frequency estimation of the chose items is altered. The created IS will diminish the calculation multifaceted nature and enhances the concealing execution of the item sets.

3.4 Optimal High Utility Item Selector (OHUIS) algorithm

The primary target of the OHUIS algorithm is to choose the best items from the sensitive item sets having same high utility esteem. The high utility esteem item sets are covered up by altering the recurrence estimations of items contained in the sensitive item sets in view of the base utility threshold value$^{\alpha}$. Improve this threshold and discover the item selector Krill Heard Optimization (KHO) model is considered. Steps required in OHUIS appeared in underneath.

Input: Actual database *D*, utility threshold, sensitive item set.
Output: sanitized database *DB*
Step 1: For each sensitive item set $s_g \epsilon D$
Step 2: $diff=U(s_g)-\alpha$ // the utility value needs to be reduced
{
If s_i contains two item sets $s_i \subseteq (g_x, g_y)$
Step 3: Compare $u(g_x, T)$ and $u(g_y, T)$
If $u(g_x, T)$ and $u(g_y, T)$
Yes// go to step 4
Else // go to step 13
}
Step 4: Select g_x, g_y item frequency values $v(g_x, T)$ and $v(g_y, T)$
Step 5: Find the frequency value optimize IS with threshold using KHO model
Step 6: From the optimal IS based sort values $v(g_x, T)$ and $v(g_y, T)$ frequency values in descending order and stored in S_x and S_y.
Step 7: Select top $h^{v(gx, T)}$, $h^{v(gy, T)}$ values from S_x and S_y.
Step 8: Compute frequency values $fh^{v(gx, T)}$, $fh^{v(gy, T)}$ for all each values.
Step 9: Compute L^{gx}, L^{gy}

$$L^{g_x} = \sum_{j=1}^{1} j^{v(g_x,T)} . fh^{v(g_x,T)} \quad (3)$$

$$L^{g_y} = \sum_{j=1}^{1} j^{v(g_y,T)} . fh^{v(g_y,T)} \quad (4)$$

Step 10: If $L^{gy} \geq L^{gyy}$ the change $v(g_x, T)$ otherwise change $v(g_k, T)$

$$o(g_x, T) = \max(g_{x \in s_i, T \in h^{v(gx,T)}})^{(u(g,T))} \quad (5)$$

Step 11: modify $o(g_y, T)$ with

$$o(g_y, T_n) = \begin{cases} 0, \ if \ u(g_y, T) < diff \\ o(i_y, T) - \left[\frac{diff^2}{s(i_y).\alpha}\right], \quad if \ u(g_y, T) > diff \end{cases} \quad (6)$$

Step 12:

$$o((g_y, T), (g_x, T)) = max_{(g \in s_i, s_i \subseteq T)}(u(g, T)) \quad (7)$$

Repeat step 11

Step 13:

$$diff = \begin{cases} diff - u(g_y, T), \ if \ u(g_y, T) < diff \\ 0, \quad if \ u(g_y, T) > diff \end{cases} \quad (8)$$

Step 14: Return the sanitized database DB
Parameter Explanation of above equation
D - Actual Database
α - Threshold
DB - Sanitized database
s_g - Sensitive item set
U - Utility
(g_x, g_y) - Two item set
T – Transaction
$fh^{v(gx, T)}$, $fh^{v(gy, T)}$ - Frequency values

This proposed model having the item set contains two items. We locate these two items utility values and checks which one is high. On the off chance that both the items utility esteem are same we change that items frequency esteem which is depicted in step 7-11 else we go to step 12. Similar utility items frequency values are sorted in ascending order and select top generally esteem.

3.5 Threshold optimization for choose Item Selector (IS)

In the privacy preserving procedure, the Privacy threshold which is the extent of sensitive patterns is as yet found from the sanitized database. Simultaneously, all sensitive patterns can be found. The upside of the threshold component is that the

clients can adjust the privacy and exposure of data. From the ideal threshold based finding the Item Selector (IS) for sensitive databases. This reason just will consider Krill group Optimization (KHO) for privacy preserving data process. The high utility esteem item sets are covered up by changing the frequency estimations of items contained in the sensitive item sets in light of the minimum utility threshold value.

3.5.1 Krill Herd Optimization (KHO) Model

This KHO methodology considers two guideline goals, for instance, expanding krill thickness, and Reaching sustenance, so the swarming behavior of growing thickness and finding food. Fig. 2 exhibits the stream layout of KHO technique.

Fitness for optimization technique

The object is to hide the sensitive high utility itemsets with the negligible reactions. The fitness is meant by methods for the accompanying condition:

$$F = T - [(\sum_{i=1}^{N}(\sum_{j=1}^{m} U_j D(i,j)).\omega_i \varepsilon] \quad (9)$$

Where T represents the best threshold value, U represents the external utility data value, D has the utility data value and ω, ϵ characterized the variables, N is the number of transaction and m represents the number of item sets.

Updating Procedure

The area of a krill individual is influenced by the accompanying three variables

- Movement induced by other krill individuals
- Foraging activity
- Random diffusion

The area of krill is communicated by the accompanying Lagrangian display

$$\frac{dR_i}{dt} = E_i + Q_i + W_i \quad (10)$$

Where I_i represents the motion induced by other krill individuals; K_i is the foraging motion and D_i is the physical diffusion of the ith krill individuals.

Movement induced by other krill individuals

In the improvement, the course of development of a krill individual is settled both by the area swarm thickness (nearby effect), a target swarm thickness (target affect) and a frightful swarm thickness (unpleasant impact). The krill improvement can be portrayed as

$$E_i^{new} = E^{max_{\gamma_i}} + \omega_n E_i^{old} \quad (11)$$

Where,

$$\gamma_i = \gamma_i^{local} + \gamma_i^{target} \quad (12)$$

In above conditions the documentations are clarified just like E_{max} the greatest incited rate, ω_n is the idleness weight of the movement instigated in the extent [0, 1], E_i^{old} is the last movement actuated, γ_i^{old} is the nearby impact gave by the neighbors and γ_i^{target} is the objective heading impact gave by the best krill individual and NN is the quantities of people [13].

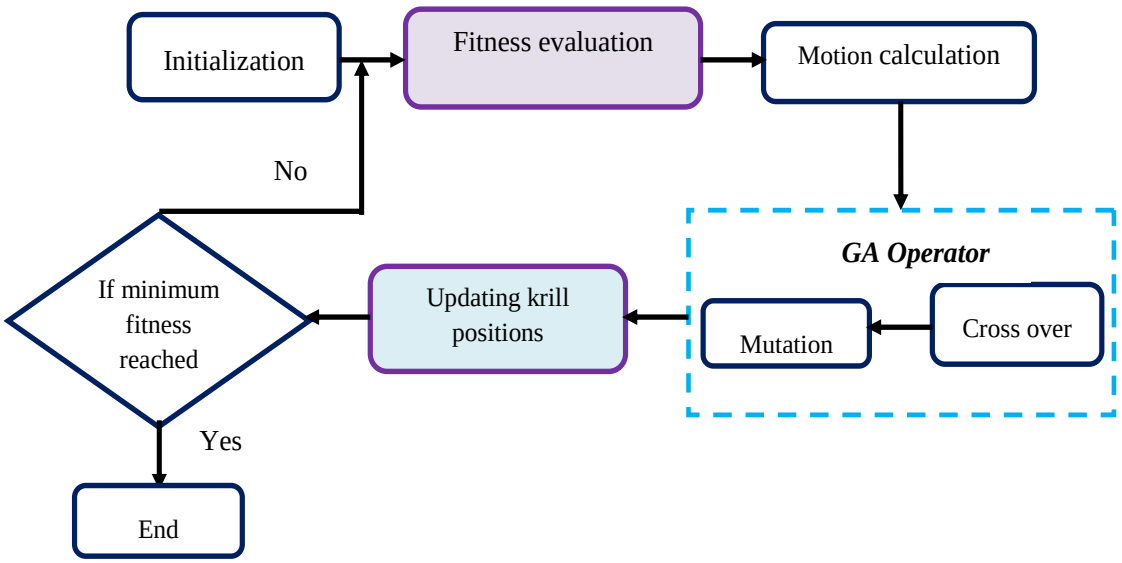


Figure.2 Flow diagram for KHO process

Foraging motion

The scavenging development is figured similarly as two major fruitful parameters. The first is the sustenance region and the second one is the past experience about the food region. This development can be conveyed for the i^{th} krill individual as takes after:

$$Q_i = F_m \delta_i + \omega_m Q_i^{old} \quad (13)$$

Where, $\delta_i = \delta_i^{food} + \delta_i^{best}$

Here F_m is the scavenging velocity, ω_m is the dormancy weight of the searching movement in the reach [0, 1], is the last scrounging movement, δ_i^{food} is the nourishment appealing and δ_i^{best} is the impact of the best wellness of the krill as such. As indicated by the deliberate estimations of the searching rate it is taken 0.02 (ms-1) [13].

Physical diffusion

The physical scattering of the krill individuals is thought to be an unpredictable methodology. This development can be express the extent that a most outrageous dispersion speeds and a sporadic directional vector. It can be characterized as takes after:

$$W_i = W^{max} \lambda \quad (14)$$

Here, W^{max} is the maximum diffusion speed, and λ is the random directional vector and its arrays are random values between -1 and 1.

Crossover

The crossover operator is first utilized as a part of GA as a practical strategy for overall improvement. A vectorized type of the hybrid is furthermore used as a piece of DE which can be considered as a further change to GA. The hybrid rate calculation as takes after.

$$cr = 0.2F_i \quad (15)$$

Mutation

The mutation assumes an essential part in developmental algorithms, for example, ES and DE. The mutation is controlled by mutation probability (*Mr*).

$$Mr = 0.5/F_{ibest} \quad (16)$$

Utilizing this new mutation probability, the mutation probability for the worldwide best is equivalent to zero and its expansion with diminishing the fitness. This KHO methodology based acquire the ideal threshold value to select item selector for hiding sensitive databases.

3.6 Data sanitization process

Data sanitization system with a specific end goal to safeguard privacy against a few components, for example, proximity and divergence attack and furthermore to save the utility of the data for a mining errand in security prepare the insurance algorithm for data sanitization to abstain from uncovering the sensitive data.

4. Result and discussion

This proposed privacy preserving exploration strategy is performed in the working stage of MATLAB form 2015a with an i5 processor and 4GB RAM. Test result examination distinctive engineered database is produced. This proposed show contrasted with other existing strategies.

4.1 Database description

This privacy preserving data mining examination, distinctive four databases used in OHUIS process. The database contains 100 and 200 transactions with various item sets. It's appeared in table 1 as beneath.

4.2 Performance measures

The efficiency of the novel procedure is assessed by methods for computing certain effectiveness measurements.

Hiding Failure (HF)

The Hiding failure measures the rate of the sensitive item sets found from *D*. It is measured by the sensitive itemsets of both the original database and the sanitized database, which is expressed as takes after,

$$HF = \frac{|SI(D')|}{|SI(D)|} \quad (17)$$

Where, *SI*(*D*)and *SI*(*D*) represents sensitive item set from the original database D and sensitive item set from sanitized database D'.

Miss Cost (MC)

The Miss cost measures the distinction proportion of the substantial item sets introduced in the original database and the sanitized database. Its esteem is processed as:

$$MC = \frac{|NS(D) - N(D')|}{|NS(D)|} \quad (18)$$

Table 1. Database transaction and item

Dataset	Transactions	Distinct items
Dataset I	100	10
Dataset II	200	10
Dataset III	300	10
Dataset IV	400	10

Where, *NS(D)* and *NS(D')* denotes the non-sensitive itemsets discovered from the original database *D* and the sanitized database D', respectively.

Dissimilarity (DIS)

The dissimilarity between the original database *D* and the sanitized database D' is ascertained as outfitted underneath.

$$DIS = \frac{1}{\sum_{i=1}^{m} f_D(i)} \left(\sum_{i=1}^{m} [f_D(i) - f_D{}'(i)] \right) \quad (19)$$

Where, $f_{D(i)}$ and fD' represents the frequency of the i^{th} item in the database *D* and the frequency of the i^{th} item in the database D'.

Table 2 demonstrates the threshold range and diverse databases with ideal threshold values. In this four types of datasets are taken and find the accuracy for every threshold value. On the off chance, the threshold fluctuating in the range of 100-500 means the ideal threshold as 488, 435 for hiding sensitive data. The highest accuracy obtained in dataset III it reaches 95.58%. The adequacy of the OHUIS algorithms is measured and the investigations were directed on two manufactured datasets. All analyses were performed on a Dell workstation with a 3.40 GHz Intel center processor and 2 GB of fundamental memory, running Windows 7 Professional.

Table 2. Accuracy and optimum threshold analysis for different database

Databases / Threshold Range	100-500	500-1000	1000-1500	1500-2000	Accuracy (%)
	Optimal Threshold τ_{opt}				
Dataset I (100X10)	488	988	1458	1884	94.52
Dataset II(200X10)	435	999	1450	1789	92.56
Dataset III (300X10)	490	978	1458	1958	95.58
Dataset IV (400X10)	458	945	1471	1945	90.23

Performance metrics for different databases

Fig. 3 demonstrates the HF for proposed strategy with various databases for threshold varying. In the dataset 2, the hiding failure values for various threshold qualities are 0.38636, 0.10121, 0.14543, 0.1579, 0.14286 and 0.16667 for the proposed algorithm and (0.45, 0.55, 0.89). The execution measure concealing failure (HF) estimation of the proposed system is lower and additionally, the miss cost and difference are greatest when contrasted and existing procedure and these qualities are delineated that the proposed OHUIS-KHO algorithm plays out the sanitization procedure.

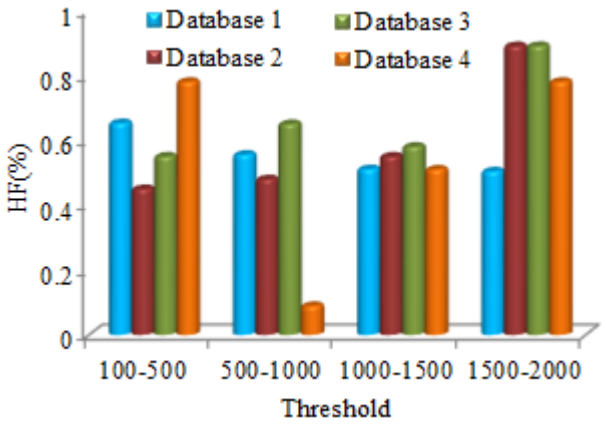


Figure.3 HF for different databases

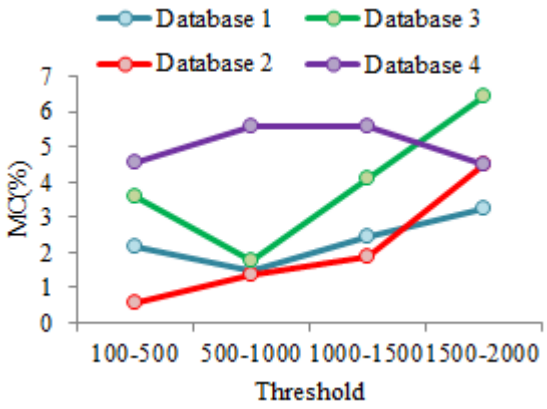


Figure.4 MC for different databases

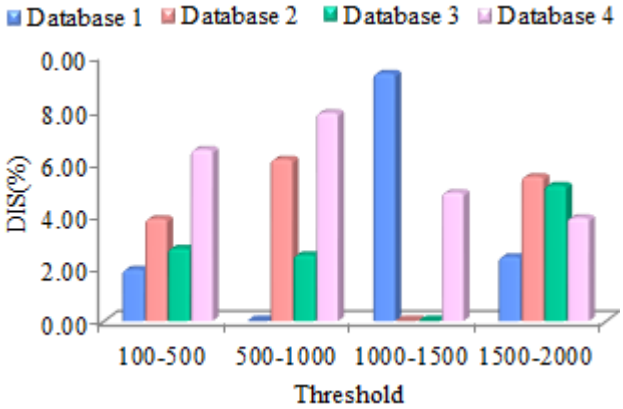


Figure.5 DIS for different databases

From the Fig. 4, the threshold value is low for the lower scope of the threshold value and the proposed strategy is giving the optimal outcome which is lower contrasted and alternate algorithms and furthermore, it is lower for low range threshold values. In the meantime, the other two qualities are Miss Cost (MC) and Dissimilarity (DIS) that is most noteworthy contrasted and the existing algorithm.

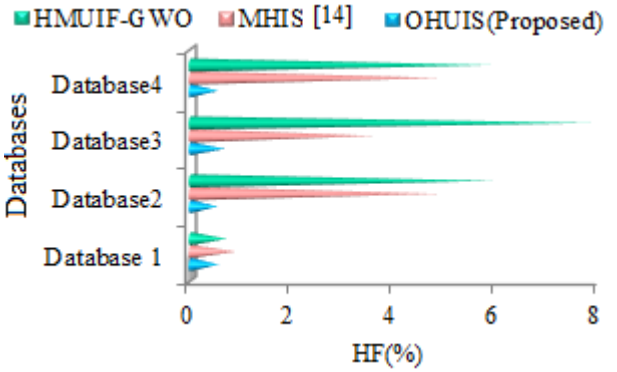
Fig. 5 demonstrates the Dissimilarity for four different databases. For the threshold level 100-500, the DIS level for database 1 is 2.032, database 2 reaches 389, database 3 obtains 3.265, and the database 4 as 6.852. Similarly, the other threshold levels also reached the same type of results. For every threshold value, the dissimilarity changes as per the techniques. Here, the performance measure of DIS investigation results the proposed system achieves lower compared to existing procedures.

Table 3 demonstrates the algorithm is productively executed and the calculation time of the noteworthy item sets is surveyed and stood out from those of the item sets yielded by the preservationist privacy-defending utility mining algorithms. The adequacy of the altered OHUIS algorithms is measured and the analyses were directed on two manufactured datasets. From the threshold esteem for different dataset measurements, the threshold qualities are upgraded when the prevalent strategies are utilized for surveying the execution assessment requirements like the hiding failure (HF), miss cost (MC) and dissimilarity (DIS).

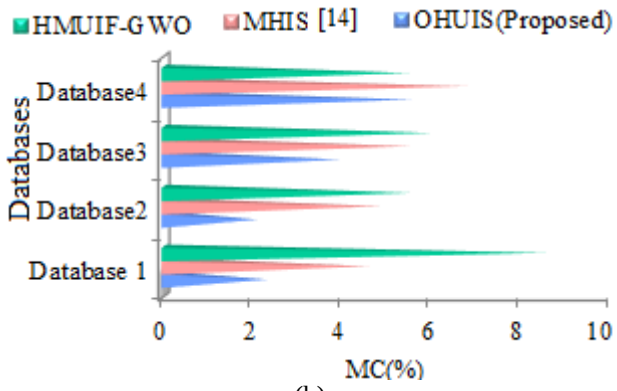
Table 3. Performance measures for proposed method (OHUIS-KHO)

Database /Threshold	**100-500**			**500-1000**			**1000-1500**			**1500-2000**		
	HF	**MC**	**DIS**	**HF**	**MC**	**DIS**	**HF**	**MC**	**DIS**	**HF**	**MC**	**DIS**
Dataset I (100X10)	0.65	2.14	1.95	0.56	1.48	0	0.51	2.42	9.35	0.50	3.27	2.48
Dataset II(200X10)	0.45	0.58	3.86	0.48	1.37	6.11	0.55	1.89	0.1	0.89	4.50	5.45
Dataset III (300X10)	0.55	3.56	2.75	0.65	1.74	2.50	0.58	4.10	0.5	0.89	6.44	5.12
Dataset VI (400X10)	0.78	4.56	6.47	0.09	5.58	7.89	0.51	5.56	4.85	0.78	4.50	3.90

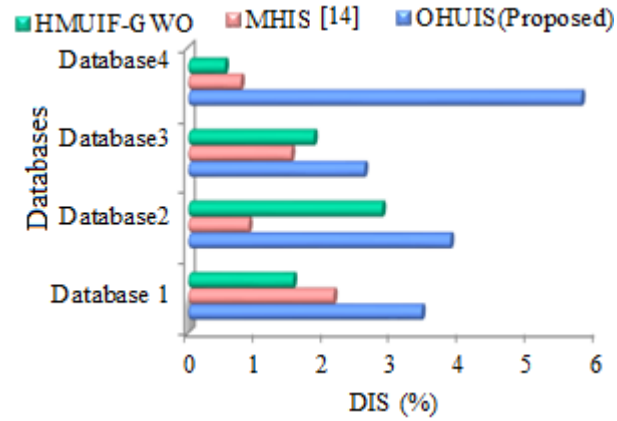
Comparative analysis



(a)



(b)



(c)

Figure.6 Comparative analysis: (a) HF, (b) MC, and (c) DIS

Fig. 6 demonstrates that the diverse performance metrics near investigation of proposed technique and existing strategies for some reference paper like [14]. Fig. 6 (a) demonstrates the HF of proposed model OHUIS and existing HMUIF - GWO and MHIS procedure. The low estimation of HF demonstrates that our proposed strategy hides the sensitive items more productively than the customary OHUIS algorithm. Similarly, 6 (b) and (c) illustrates the comparative analysis for three algorithms and find the optimal value of MC and DIS in percentage. In the existing method, HMUIF-GWO achieves in the range as 0 to 9.63 in MC and 0 to 3.5% in DIS. Then for proposed OHUIS reaches the MC in the range as 0 to 6.5% and DIS as 0 to 6%. For the comparison of all the three techniques, the proposed method achieves an optimal result. It has less computational time and low miss cost, DIS, and HF. It also gives optimal value compared to existing algorithms.

5. Conclusion

Data quality assumes a vital part in the mining procedure. Precise input data brings important mining comes about. Then again, preserving privacy is likewise a fundamental issue. In this manner, explore requirements to offer and research a successful security protecting mining model. This review initially talks about a Privacy Preserving Utility Mining (PPUM) model and exhibits the Optimal High Utility Item Selector (OHUIS) algorithms to decrease the effect on the source database of security safeguarding utility mining. This algorithm changes the database transactions containing sensitive itemsets to limit the utility value below the given threshold while anticipating recreation of the original database from the sanitized one. The effectiveness of the optimal threshold is evaluated by utilizing the hiding failure, miss cost, and dissimilarity factors and it is found to yield the finest threshold value. The proposed method achieves inspired optimization technique I.e. KHO for updating the threshold value. The results proved that the performance of proposed OHUIS algorithm was better than the conventional algorithms. Hence, the KHO algorithm attains the optimal performance measures and the reaches the maximum accuracy as 95.2%. In future new databases are considered for the assessment procedure alongside new approach.

References

[1] G. Narang, A. Shaikh, A. Sonawane, K. Shegar, and M. Andhale, “Preservation Of Privacy In Mining Using Association Rule Technique”, *Journal Of Scientific & Technology Research* Vol.2, No.3, pp.219-222, 2013.

[2] T. Takeuchi, T. Kawamura, and A. Ohsuga, “Hiding of User Presence for Privacy Preserving Data Mining”, In: *Proc. of Advanced Applied Informatics (IIAIAAI), IIAI International Conference on. IEEE*, pp.133-138, 2012.

[3] A. Chouhan and P. Sinha, “A Novel Approach of Data Sanitization using Privacy Preserving Data

Mining", *Journal of Engineering Research &Technology*, Vol.4, No.6, pp.1137-1140, 2015.
[4] X. Yi and D. C. Zhang, "Privacy-preserving naive Bayes classification on distributed data via semi-trusted mixers", *Journal of Information Systems*, Vol. 34, pp.371-380, 2009.
[5] A. Gkoulalas and V. S. Verykios, "An Overview of Privacy Preserving Data Mining", In: *Proc. of Procedia Environmental Sciences*, Vol. 12, pp. 1341-1347, 2012.
[6] C. Clifton, M. Kantarcioglou, X. Lin, and M. Y. Zhu, "Tools for privacy preserving distributed data mining", *Journal of ACM Sigkdd Explorations Newsletter*, Vol.4, No.2, pp.28-34, 2002.
[7] U. Yun, H. Ryang, G. Lee, and H. Fujita, "An efficient algorithm for mining high utility patterns from incremental databases with one database scan", *Journal of nowledge-Based Systems*, pp.1-19, 2017.
[8] C. W. Lin, T. P. Hong, G. C. Lan, J. W. Wong, and W. Y. Lin, "Incrementally mining high utility patterns based on pre-large concept", *Journal of Applied Intelligence*, Vol. 40, No.2, pp. 343-357, 2014.
[9] J. C. W. Lin, T. Y. Wu, P. Fournier-Viger, G. Lin , J. Zhan, and M. Voznak, "Fast algorithms for hiding sensitive high-utility itemsets in privacy-preserving utility mining", *Journal of Engineering Applications of Artificial Intelligence*, Vol.55, pp. 269-284, 2016.
[10] J. Hua, A. Tang, Y. Fang, Z. Shen, and S. Zhong, "Privacy-Preserving Utility Verification of the Data Published by Non-interactive Differentially Private Mechanisms", *Journal of IEEE Transactions on Information Forensics and Security*, Vol. 11, No.10, pp.2298-2311, 2016.
[11] D. Brettschneider, D. Hölker, A. Scheerhorn, and R. Tönjes, "Preserving privacy in Distributed Energy Management", *Journal of Computer Science-Research and Development*, pp.1-13, 2016.
[12] Y. A. Alsahib, Aldeen, M. Salleh, and M. A. Razzaque, "A comprehensive review on privacy preserving data mining", *Journal of Springer Plus*, pp.1-36, 2015.
[13] S. Huang, L. He, X. Si, Y. Zhang, and P. Hao, "An Effective Krill Herd Algorithm for Numerical Optimization", *Journal of Hybrid Information Technology*, Vol.9, No.7, pp.127-138, 2016.
[14] R. Selvaraj and V. M. Kuthadi, "A Modified Hiding High Utility Item First Algorithm (HHUIF) With Item Selector (MHIS) For Hiding Sensitive Item sets", *Journal of Innovative Computing, Information and Control*, Vol.9, No.12, pp.4851-4862, 2013.