A Survey on metaheuristic nature inspired computations used for Mining of Association Rule, Frequent Itemset and High Utility Itemset.

K Logeswaran¹, R K S Andal², S T Ezhilmathi³, A Harshath Khan⁴, P Suresh⁵ and K R Prasanna Kumar⁶

¹,⁶Assistant Professor, Department of Information Technology, Kongu Engineering College
²,³,⁴UG Students, Department of Information Technology, Kongu Engineering College
⁵Associate Professor, Department of Information Technology, Kongu Engineering College

*Email:klogesbtech@gmail.com

Abstract. Metaheuristics are dilemma-independent methods that are generalized in a variety of problems. In the real world, various problems are solved using generalized dilemma-independent methods called Metaheuristics Computation. Metaheuristic Nature Inspired Computing (MNIC) is a generalized approach to solve NP-hard problems by taking inspirations from the behavior of mother biological nature and their characteristics. Mining of Association rule, Frequent Itemset and High Utility Itemset are strongly interrelated and developing in the field of Data Mining. Metaheuristic nature inspired computation was widely used for the mining association rules of frequent itemsets and high utility itemsets to address the high computation time and optimal solutions. While various articles have been written, there is no systematic review of contemporary metaheuristic nature inspired approaches used in Association Rule Mining (ARM), Frequent Itemset Mining (FIM) and High Utility Itemset Mining (HUIM). This paper explores recent literature on various metaheuristics nature inspired approaches used for ARM, FIM and HUIM.
1. Introduction
Many exact approaches are applied to ARM, FIM and HUIM, which are heuristic in nature. These exact approaches are often problem dependent that exploits problem-dependent information to provide solution for the problem. However, the solutions provided by any heuristic approaches cannot be guessed for goodness and consumes more computation time and memory. On the other side, in recent years many problem independent metaheuristic nature inspired approaches were applied for ARM, FIM and HUIM, that provides approximate solutions to the given problem and these approximate solutions can be proved for how close it is to the optimal solution. Hence these approaches are also called as nature inspired optimization computations.

Many metaheuristics approaches are in the form of stochastic optimization where the results produced depends on the set of random variables. When looking for a wide variety of possible solutions, metaheuristics may also find successful solutions that has less computational work than the optimization algorithms, iterative approaches, or basic heuristics. As such, approaches to optimization problems are useful. A main advantage of metaheuristics nature inspired optimization approaches is that stringent terminating constraints can be imposed to limit the time of computation so an almost optimal solution could be reached.

Metaheuristic natured optimization approaches are divided in two categories as evolutionary computation and swarm intelligent computation. Swarm Intelligence systems usually consist of a community of basic agents or bodies communicating geographically with each other and their surroundings. Inspiration also comes from nature, particularly from biological systems. Agents obey very basic rules, and while there is no hierarchical control system that determines how individual agents can behave, local and, to some degree, unpredictable interactions between those agents contribute to the development of "intelligent" global activity, not understood to the singular agent. The Evolutionary Computation is inspired by natural evolution. In evolutionary computation, a container is defined that imitates what the solutions look like and what are the ingredients. The systems then generate randomly solutions that are viable but not necessarily good because they were obviously put together randomly. These solutions are not going to be very good, so we rank order these solutions using some measure. Then remove the worst ones and retain the ones that are a little better. In addition, in the next generation we create some more candidate solutions. Then mix and match and beg and borrow from those solutions that were less bad in the last generations. This process continues for some time until a certain condition is reached and often the time system is able to come up with near optimal solutions. Due to its Due diligence, it has been used in may research domains like cloud computing [1], data mining [2], IoT[3] and many
more. Figure 1 shows the statistical values of the number of papers published related to ARM, FIM, and HUIM using metaheuristic nature-inspired optimization approaches.

![No of Paper published from 2003 - 2020](image)

**Figure 1.** Papers published based on MNIC - ARM

Figure 2 illustrates the various MNIC approaches used for ARM, FIM, and HUIM. The rest of the paper will be organized as follows: Section 2 presents basic information on ARM and addresses the various metaheuristic nature-inspired methods used for ARM. Section 3 provides the basic details on FIM and explains the common metaheuristic essence of FIM-inspired approaches. Section 4 addresses the fundamental details of HUIM and the different metaheuristic methods used by HUIM. Section 5 ends the document, ultimately.
Figure 2. Pictorial representation of various MNIC used of ARM, FIM and HUIM
2. Association Rule Mining

Association Rule Mining has association rules that works with basic if / then that enable to uncover associations between apparently separate relational databases or other repositories of data. Association Rule Mining, as the name proposes, association rules are basic If/Then proclamations that help find connections between apparently autonomous social information bases or other information stores. Association Rule will be in the form of if A happens, the B is likely to happen. So, is likely conditional dependence relation that your are mining if A happens, that makes B more likely to happen. Such a patterns can be found in sequences looking at time series data, like financial data or looking at fault analysis, where one thing causes fault to occur or in the transactional data column context, which is where it was originally proposed.

Many Artificial Intelligence (AI) calculations work on measurable information bases and accordingly will in general be quantitative. In any case, the association rule for mining is adequate for non-numeric, categorical information and needs just somewhat more than plain checking. Association Rule is one of the significant ideas of Artificial Intelligence being utilized in market container examination. Figure.3 shows the statistical analysis based on the papers published based on MNIC – ARM.

![Figure3. Papers published on MNIC - ARM](image)

2.1. Elemental study of ARM

Market container examination is the investigation of client exchange information bases to decide conditions between the different things they buy at various occasions. Association rule learning is
a rule based AI strategy for finding fascinating relations between factors with regards to huge information bases. It describes regular if-then relationships are referred to as association laws, which consist of an antecedent (if) and a consequential (then).

For instance: "In the event that laptop and mouse, at that point keyboard" ("If Laptop and Mouse are bought, at that point Keyboard would likewise be purchased by the client").

Antecedent: Laptop and Mouse  
Consequent: Keyboard

There are three regular measurements to gauge association:

- **Support** means that how oftentimes the things show up in the information. Numerically, Support is the absolute portion of the number of transactions where the thing happens. Mathematical formula for Support can be derived as in Equation. (1)

\[ \text{support}(\{U\} \rightarrow \{V\}) = \frac{\text{Transactions containing } U \land V}{\text{Total number of Transactions}} = \frac{|U \cup V|}{m} \quad (1) \]

- **Confidence** implies the number of times the if-then arguments are proven to be true. Confidence is the conditional probability of the existence of an outcome given the antecedent. Mathematical formula for Confidence can be derived as in Equation. (2)

\[ \text{Confidence}(\{U\} \rightarrow \{V\}) = \frac{\text{Transactions containing } U \land V}{\text{Transactions containing } U} = \frac{|U \cup V|}{|U|} \quad (2) \]

- **Lift** can be utilized to contrast Confidence and anticipated Confidence. This says how likely thing V is bought when thing U is bought, while controlling for how famous thing V is. Mathematical formula for Lift can be derived as in Equation. (3)

\[ \text{Lift}(\{U\} \rightarrow \{V\}) = \frac{\left( \frac{\text{Transactions containing } U \land V}{\text{Fraction of Transactions containing } U} \right)}{\text{Transactions containing } V} = \frac{|U \cup V|}{|U||V|} \quad (3) \]

If the support value exceeds the value of min_supportthreshold, the itemset is frequent denoted by $\alpha_{\text{min}}$. When its confidence value is greater than min_confidencethreshold, the rule is strong and denoted by $\beta_{\text{min}}$.

The Positive ARs consist of items that are upheld by transactions. Interestingly, Negative ARs consider items from transactions are missing. These kind of rules are utilized in market bin investigation to recognize things supplement one another. The nonappearance of an itemset $U$ is
appeared as \( \neg U \). For example, rule \( U \rightarrow \neg V \) implies that the transactions that incorporate itemset \( U \) don't uphold itemset \( \neg V \).

Rules mined using ARM can be broadly classified into three types as Boolean / Binary ARM, Quantitative / Numeric ARM, Fuzzy ARM and Rare ARM. In Binary ARM, attributes of dataset contains Boolean values. Any item involved in binary ARM can be considered as Boolean variable whose value is either 1 or 0 which corresponds to occurrence or absence of item in an transaction respectively. In most of the cases, categorical or numerical and Boolean ARs based data items are involved. In a numeric AR, there is no restriction to be Binary / Boolean yet it can be measured quantitatively (e.g., age, kilometers and temperature) or emphatic (e.g., sex and brand). In this manner, numeric ARs are expressive and useful when compared to Binary / Boolean ARs. An inaccurate method to address trait with extremely enormous or consistent is by discretizing the continuous attributes to various stretches. After discretization, the attributes are preserved as clear cut traits. The fundamental issue with discretization is loss of data and helpless outcomes. Discretization-based quantitative ARM has an issue between the sharp limit of spans. Fuzziness is introduced to address this challenge. Fuzzy set hypothesis changes quantitative qualities into etymological terms, in this way producing semantic or fuzzy information. Fuzzy ARM id utilizes the fuzzy set and mining QARs through Membership Functions (MFs) ideas. A good progress is created between a part and a non-individual from a set by Fuzzy ARM. This makes the outcome more in detail. Rare ARM are based on Patterns that appear occasionally. Events that occur regularly can be present in several realms Less fascinating than unusual occurrences, as regular patterns reflect known patterns. It is assumed that unusual patterns can reflect uncommon or previously unknown connections that are useful to domain experts.

2.2. Metaheuristic Optimization approaches for ARM

Fuzzy logic uses Association Rule Mining and Whale Optimization algorithms to spot fuzzy recurrent items and rules. Same individual threshold values are given to items and grouped based on recurrent items and rules that are generated. There arises a problem in the large numbers of databases as scanned from huge volumes of datasets. Data analysis is not required for all the transactions and items. So to solve this problem as mentioned in [4], (DRITI) Dimensionality reduction in transactions and (FRUWOA) Fuzzy rules using Whale Optimization algorithms are applied. In the methodology, recurrent itemsets and fuzzy association rules are opened up by fuzzy logic. Hence number of scans to find the irrelevant items are reduced by low variance method and hash table technique. Better results are made when compared to existing techniques and gives better results. A hierarchical heterogeneous Ant Colony optimization is based on the action rule mining approach and it produce time consumedsolutions. It is a risky problem to select the values of threshold, support and confidence. These are mentioned in [5] and to overcome this,
Hierarchical Heterogeneous Ant Colony Optimization (HHACOARM) algorithm is used. Ant agents are used in the methodology to identify the flexible attributes whose values are changed during mining. Values of support and confidence of the action rule are not considered in HHACOARM which is the advantage of this algorithm. HHACOARM algorithm has computational complexity that is less when compared to the other algorithms. Multi-objective association rule mining with Binary Bat algorithm is an optimal algorithm whose convergence is greater than Binary Particle Swarm Optimization (BPSO). The rules generated by traditional Multi-Objective Algorithms are very large and complex to analyze and explore. So to overcome this as mentioned in [6], the Multi-Objective Binary Bat Algorithm (MBBA) is applied. The algorithm is free from of min_support and min_confidence and also the value of lift and comprehensiveness are optimization objectives. Hence this new algorithm has better performance than BPSO, BBA and Apriori in terms of quality of rules and time efficiency. This MBBS algorithm has good convergence and set of non-dominated solutions by running at one time. Association rule mining is used in genetic programming and it gives feedback to the instructors for multiple-choice quiz data. But it does not give suggestions about how to improve student performance. This is the problem mentioned in [7] and to overcome this, grammar-guided genetic programming (G3P) approach based on Evolutionary Algorithm (EA) is applied. The information found is applied to a list of specific quizzes and courses. In addition, this method evaluates the changes introduced in terms of the improvement of the student. So comparison is done between the various groups of students who solve the quizzes before and after the updates of this application. Bat-based algorithm (BA) (ARM Bat) aims to increase the fitness function in the defined dataset. There are some association mining rules that could be considered as optimization problem. Evolutionary algorithms with rules that increase the fitness function are selected. So to overcome this disadvantage as mentioned in [8], Association rule mining based on Bat Algorithm was proposed. The aim of the methodology is to increase the fitness function and generate the best rules in the defined dataset. Thus, the algorithm proved its efficiency in terms of time, memory usage and quality of generated rules. It also maximizes the fitness function as mentioned.

3. Frequent Itemset mining

Nowadays decision making process has large impact on growth of any business. The decision making process can be done from huge volume of data collected from customer and market. Frequent Itemset Mining (FIM), is one of data mining technique used for making decision for any type of business. FIM is the process of extracting itemset from any transaction dataset, which occurs for many numbers of times. There are many approaches developed in recent years to
improve the performance of FIM and to apply the FIM in different domain apart from market basket analysis. Figure 4 shows the statistical analysis of number of paper published based on MNIC – FIM. Figure 4 shows the statistical analysis of number of paper published based on MNIC – FIM.

![Figure 4](image.png)

**Figure 4.** Number of paper published based on MNIC – FIM

### 3.1. Elemental study of FIM

In a transaction database, any items, elements or event that occurs regularly is mentioned as patterns $P$. Precisely, in a database $\varphi$, pattern $P$ interpret the useful insights from the data and it is a subset of items $K = \{k_1, \ldots, k_n\} \in \varphi$. For subset of items defined by $P$, number of occurrence of itemset is denoted as pattern $P$ frequency. Formally, in database $\varphi$, consisting of $K = \{k_1, \ldots, k_n\}$ items and $T = \{t_1, \ldots, t_m\} \in \varphi$ transactions, and each transaction includes set of items which are subset of $K$, frequency of $P$s described as $|\{\forall t \in T : P \subseteq t, \ t \subseteq K \in \varphi\}|$.

### 3.2. Metaheuristic Optimization approaches for FIM

Frequent itemsets are regularly generated by applying association rule mining algorithms on huge datasets. Frequent itemsets is mined using Genetic algorithm. Sometimes algorithms like Apriori, Partition, Incremental, computation time Border take so long for computing frequent item sets. To solve this issue in [9], Genetic algorithm is proposed. The method of Population is implemented here. Collection of individuals, which signify the potential solutions through mapping, called a coding. Fitness function is used for ranking of the population and for each step selected population is applied with genetic algorithms to calculate the fitness. As the method is very simple and efficient, different large data sets are tested and the results are correct and accurate. Selection patterns of iEEG and fMRI epilepsy data is identified using Genetic Programming and Frequent
Itemset Mining. There arises confusion whether stochastic properties with indirect feature selection as it is uncertain. This mentioned drawback was overcome by Genetic Programming and Frequent Itemset Mining in [10]. The method uses GP algorithm for selecting features to classify interracial activity. Separation of fMRI epilepsy patient data and iEEG is performed by subsequent frequent itemset. Hence FIM is used in computation of qualitative confidence intervals through GP algorithm. GP-based feature selection is observed based on subject consistency and across-subject variability for iEEG and fMRI signals. The issue faced while detecting interracial biomarkers for each iEEG and fMRI signal modalities are patient-specific feature extraction, feature selection, and machine learning modules. To solve this problem, Apriori heuristic and genetic algorithms are combined using GA-Apriori. Due to number of frequent itemsets discovered, GA based approach is inefficient in terms of solution’s quality. Apriori heuristic principle is considered to define crossover and mutation operators. So this drawback as mentioned in [11] is overcome by GA-Apriori algorithm. Thus while comparing Apriori and FP growth, GA-Apriori is competitive in the aspect of the quality of solution, i.e., the number of frequency itemsets discovered. This problem is solved by applying the combination of Apriori heuristic and bio-inspired algorithms. There arose difficulties as said in [12] that quality of the solution is less as obtained by bio-inspired approaches than optimal quality and a failure with FIM bio-inspired approaches is faced due to ignorance of intrinsic properties of the FIM solution space. So as a solution the recursive property of FIM exploited by the approaches and k size itemsets produce only individuals. In addition, the PSO-Apriori produce quality of solutions. Polygene-based evolutionary algorithms are used to determine frequent pattern mining. A gene that acts as an independent component during evolution is the primitive evolution unit. It is considered as a drawback in [13] and it is overcome by polygene-based evolutionary algorithms (PGEAs). There are three phases involved in PGEA framework which are polygene discovery, polygene planting and polygenic-compatible evolution. Thus, this algorithm significantly reduces space and retains the quality individuals. It also have good probability to find good solutions. Customized Ant Colony Algorithm helps in mining High Utility infrequent Itemsets. High utility infrequent itemset mining uncovers the notable drawback in [14] is that rarely highly profitable itemsets. So to overcome this, the algorithm high utility infrequent itemset mining (ACHUIIM) on the basis of Ant colony is proposed. Itemsets where rate of occurrence which falls below the threshold value is identified using this methodology. The number of ants present in this system determines the execution time. The execution time of the real time datasets are reduced in this algorithm. A new framework is proposed for metaheuristic-based frequent itemset mining. The problems mentioned in [15] are only few frequent itemsets found and recursive property for guiding metaheuristics approaches. For mining, Multiple algorithms such as (META-GD) Guided Meta Heuristic Framework, (GA-GD) Guided Genetic algorithm, (PSO-
GD) Guided Particle swarm Optimization algorithm and (BSO-GD) Guided Bee Swarm Optimization algorithm are implemented. By guiding FIM metaheuristic based approaches and when searching the itemset space the recursive property was applied. Thus state-of-the-art metaheuristic based approaches performed by PSO-GD has better runtime and solution quality. PSO-GD has superior results compared to GA-GD and BSO-GD. A multi-objective evolutionary approach is implemented for mining frequent and high utility itemsets. There are cases when prior parameters must be set up by the users for mining algorithms. Also the difficulty mentioned in [16] for the users is to set optimal values as the real applications has no prior knowledge. Hence this is overcome by Multi-Objective Evolutionary Algorithm. The methodology includes the transformation of Frequent and High utility itemset mining as 2-objective problem. Advantage of the proposed method is that the prior parameters are not needed. Multiple recommendations for decision makers can be provided by the algorithm and is it also superior over traditional algorithms. User is needed for TKU-Miner to set prior parameters and also the solutions given by HUIM are extremely infrequent. These are the downfalls given in [17] that are solved by (CP-MOEA) - Closed Itemset property based multi objective evolutionary approach. Here improvement of quality of mining is done by two designed strategies USC and USA. Hence these two designed approaches guide the evolution which helps the CPMOEA in acceleration of the convergence and diversity of population. For mining frequent itemsets from streaming transaction data uses Genetic algorithms. In the field of mining there arises many questions as given in [17]. There is no evidence whether frequent itemsets can be mined if streaming data is without permanent master databases. Also there are no techniques that reduce time complexity and solve the challenge that appear in the result of streaming environment. As a solution for the above problems, Genetic algorithm proposed a methodology which consists of. Closed Itemset Property based Multi-objective Evolutionary Approach for mining datasets contains Mining Frequent and High Utility Itemsets. Exploration of relationship between concept drift, sliding window size and genetic algorithm constraints. Genetic algorithm framework is used for mining streaming data by determining concept drift. Overall, the performance of transactions per drift was good with the ratio of the window size. In addition, detection of drift concept, when larger windows were used was performed well.

4. High Utility Itemset Mining
High Utility Itemset Mining (HUIM) is as extension of Frequent Itemset Mining (FIM). In FIM, it is assumed that itemset with high frequency of occurrence will have more interesting value ie., high profit than other itemset. This assumption is not valid for all cases, as itemset with lower frequency may more profitable in some situation. To address this pitfall, HUIM was introduced [14]. In recent
years, many researchers have introduced several approaches to effectively design HUIM. Also due to increasing benefits of MNIC, many HUIM approaches based on MNIC have been evolved in recent years. Fig.5 shows the statistical analysis of the papers published based on MNIC – HUIM.

![Figure 5. Papers published based on MNIC - HUIM](image)

### 4.1. Elemental study of HUIM

In a database $\varphi$, a transaction $T_s$ utility of an item $i_k$ is denoted as in eq. 4

$$u(i_k T_s) = q(i_k, T_s) \times p(i_k) \quad 4$$

where $q(i_k, T_s)$ is the purchase quantity of $(i_k)$ in the transaction $T_s$ and $p(i_k)$ represent the profit value of $(i_k)$.

The utility of an itemset $X$ in a transaction $T_q$ is $u(X, T_q)$ and defined as:

$$u(X, T_q) = \sum_{i_k \in X \land X \subseteq T_q} u(i_k T_s) \quad 5$$

The utility of an itemset $X$ in a database $D$ is denoted as $u(X)$, and defined as:

$$u(X) = \sum_{X \subseteq T_q \land T_q \in D} u(X, T_q) \quad 6$$

### 4.2. Metaheuristic Optimization approaches for HUIM

Group of items is identified by High-Utility Itemset Mining, which gives the high-profit database. It is a extended version of Frequent itemset mining. Particle swarm optimization using high utility mining, which has optimization with high efficiency. Parameters required for genetic algorithm is very high in
particle swarm optimization. Storing of itemsets in database with single objective of high-utility is of huge size which discovers HUIs for the usage of exponential search which is a major drawback. To overcome this disadvantage they introduced an algorithm for binary particle swarm optimization on the basis of named High utility itemset mining with sigmoidal function (HUIM-BPSOsig). Discrete particle swarm optimization based algorithm is the methodology used in [18] and TWU model is also used in this algorithm. As a result, Improved genetic algorithm is an upgraded version of GA, which gives better effectiveness for evolutionary computation based algorithm, and it is an alternative approach. The privacy preserving utility mining method of GA based approach is proposed for hiding sensitive high utility itemsets and to discover suitable transactions to be added into the database. There are few difficulties as given in [19] as it is difficult to modify the quantities of the items in the database. The major issue is that privacy threats may occur. So to solve this problem GA-based privacy preserving utility mining method is applied. To sanitize the database for hiding sensitive information, GA based algorithm is used. Execution time is also reduced by prelarge concepts. Transaction insertion mechanism is used to hide the sensitive high utility itemsets. By three factors hiding failure, missing and artificial high utility itemsets, three user-specific weights are assigned. To mine high-utility itemsets applies an evolutionary algorithm. High profit itemsets cannot be identified when item sets don't appear frequently. Occurrence of frequent itemsets discover only in FIM or ARM. These difficulties as mentioned in [20] are solved by evolutionary algorithm. The methodology uses PSO-based algorithm based on mining sigmoid updating strategy. Improved Genetic Algorithm is used for High-Utility Itemset Mining as it handles large databases. Some HUIM algorithms handle the some problems with huge search space due to the size of the database size or diverse itemsets. EC-based algorithms that find complete (HUIs) in transaction databases take more time. To overcome these disadvantages, HUIM algorithm based on an improved genetic algorithm (HUIM-IGA) is applied [21]. Various strategies including population diversity maintenance, neighbourhood exploration and transaction-weighted utility based individual repair strategy are performed for better performance. The local search ability and search space are improved. The HUIM-IGA performs best in terms of convergence speed, runtime and skill to discover. High-utility itemset mining can be done without minimum utility threshold by BPSO-based method. Minimum utility threshold for High Utility Itemsets are computed in advance by some algorithms. Sometimes, extreme end threshold values may consequence as not getting meaningful itemsets. These major issues are solved by Binary Particle Swarm Optimization approach in [22]. Initially high utility itemsets are mined in this approach and these itemsets are generated according to BPSO algorithm. A minimum utility threshold is forced on the list and the method is compared to TKU and TKO. The process of obtaining high utility itemsets is most favorable as there is no prior setup for minimum utility. Hence, high utility itemsets can be obtained without a minimum utility threshold as it is based on BPSO. High utility itemsets are also
mined by ACO-based approach. When the size of databases and number of distinct items are large, exponential search space must be handled for discovering HUIs. Occurred frequencies of itemsets are only revealed in this process. These drawbacks mentioned in [23] are overcome by Novel algorithm based evolutionary computation technique, ant colony optimization (ACO). The algorithm proposes two pruning rules and the experiments are performed in a number of databases in real life. Specific routing graph and TWU model are implemented by HUIM-ACS approach to find the HUIs. Spare fitness is avoided and global optimal solutions are obtained by ant-based algorithm. Thus the efficient method scans the dataset and increases the performance for revealing utility value in dataset. Based on the Artificial Bee Colony Algorithm, High Utility Itemsets can be discovered. The exponential problem exists the number of items and large database which affects the search space. The main problem mentioned in [24] is premature convergence which is solved by HUI Mining based on the Artificial Bee Colony algorithm (HUIM-ABC). The methodology has the initial step as modeling the problem to mine HUIs. Information representation and search space pruning are done by bitmap. New nectar sources are produced by the size of information from discovered HUIs. Only few iteration cycles are involved in employed bees, onlooker bees and scout bees. Thus within the range of discovered results, new nectar sources are generated. Boolean operators-based modified grey wolf optimization algorithm is applied in High utility itemset mining. It is tough to find out a set of items from a transactional database with high level of utility. Also a difficulty arises the number of items contained in the database as it takes more time when the transaction database increases. So to increase the efficiency with transactional database, grey wolf optimization algorithm which works on the basis of biological behavior of grey wolf is applied in [25]. Solution space is represented by binary representation and GWO has operators that are redefined with unconventional boolean operators such as Adder, Difference, Multiplexer, Circular Shift and De Morgan’s AND. To hold the originality, the flow of GWO was preserved. Thus proposed model has less overall time complexity for decision making as it depends on the exponential component. A Diverse Optimal Value Framework is adopted for mining High Utility Itemsets using Bio-Inspired Algorithms. High values cannot be discovered through traditional FIM algorithms. Even though new algorithms are developed for mining HUIs, there is no guarantee to find all HUIs in a database. So to improve these functionalities in [26], new methodology using the Bio-HUI framework using genetic algorithms is adapted. The framework consists of strategies to discover the HUIs probability. To accelerate the process of HUI discovery, different strategies including bitmap database representation are performed. So this proposed method not only performs well in terms of efficiency but also more efficient in convergence speed. An efficient PSO based algorithm was used for mining High Utility itemsets. Heuristic HUPEumu GRAM algorithm was proposed to mine HUIs based datasets. These algorithms also require only few parameters compared to GA-based approaches. These functionalities fall under the drawbacks of this approach.
to overcome this as given in [27], PSO-based algorithm, namely HUIMBPSO algorithm is applied. A new tree structure with OR/NOR representation is developed as it reduces more number of scans in database. By adding sigmoid updating strategy and TWU model, HUIs are found. Thus, non-trivial tasks are performed along with handling of continuous problems and performing random operations. An efficient maximal pattern (MP)-tree structure was developed and it has better results. The adoption of maximal pattern (MP)-tree structure reduces the computations of multiple database scans. Genetic algorithm along with ranked mutation is used to find high utility itemsets. Suitable minimum utility thresholds should be specified by data analysts even though they might not have knowledge about the database. Vast search space is found due to large number of database. The downfalls mentioned in [28] were modifies by a novel evolutionary approach. High utility pattern extraction using genetic algorithm with ranked mutation has minimum utility threshold is based on GA approach. Thus it is an effective approach as GA plays an important role in mining HUIs from the database. It has the capability to explore large spaces well. Privacy-preserving utility mining is performed for mining high utility-itemset mining. The privacy threats may occur when private or secure information is published in the public place and even information can be misused. Some of the real-life applications are not considered in ARM. So to overcome these difficulties as mentioned in [29], PSO-based algorithm and GA-based approach are used for HUIM and PPUM respectively. These approaches can make sure the exchange between mining performance and preserving for SHUIs. Experiments on various real-life and synthetic datasets proved that this methodology gives better results. Thus the approach has faster magnitude than naive GA based technique. An Optimization based Modified Maximum Sensitive Item-Sets Conflict First Algorithm (MSICF) is applied for hiding Sensitive Item-Sets. Nowadays hiding information from the adversaries is a major issue. The existed approaches such as HHUIF and MSICF are obscure sensitive itemsets as it is not possible to mine the adversaries from the customized database. These difficulties as given in [30] are overcome by Modified maximum Sensitive Item-Sets Conflict First Algorithm (MMSICF). Best items with same high utility value must be selected and they are hidden by modifying the frequency values of items. Hybridization of ABC along with GA is proposed for the selection of threshold values. All sensitive itemset values are reduced lower than threshold value by continuous hiding process. Thus the sensitive itemsets from the adversaries are hidden by the improved MMSICF algorithm even if the items has utility value similar or not. Database transactions is modified by the method containing sensitive itemsets which reduce the utility value lesser than the given threshold by preventing therestoration of the original database. Optimization of Evolutionary Algorithm Using Machine Learning Techniques is applied on Pattern Mining in Transactional Database. There are many challenges as given in [31] that includes expanding increasing search space for utility mining algorithms and existing mining algorithms guess that the database has appropriate threshold and fixing it is not an effortless task by the user. By Reinforcement Learning based Genetic
Algorithm (RLGA), these issues are solved. A new method that has quite a few supporting functions that are added to calculate each target function and optimization of few supporting functions are performed efficiently. Target function is mutually related to the part of supporting functions. Thus the search speed is improved in genetic algorithm and selection of optimal minimum utility threshold value is automated by GA. Metaheuristics approach is applied for Frequent and High-Utility Itemset Mining. Some of the techniques developed in FIM cannot be used in HUIM as the search space depends on anti-monotonicity of the support. These problems as said in [32] are solved by particle swarm optimization, genetic algorithms, ant colony optimization and bee swarm optimization. These approaches consists of various methods including grammar guided genetic programming, artificial neural network and algorithms for finding frequent patterns and mimetic algorithms. Also the Swarm intelligence based approaches include binary PSO, SET-PSO, bat metaheuristic algorithm and some important methods. Thus various metaheuristic approaches are used for exploring search space of itemsets by stochastic search process. Mining High Utility Itemset for Online Ad Placement is done by Particle Swarm Optimization Algorithm. Sometimes due to some privacy reasons, the user’s information may not be offered. There is also another difficulty in mobile apps as compared to web pages where we can go down to find the next information. The mentioned difficulties in [33] are solved by Particle Swarm Optimization algorithm. The comparison between the local to the global best conversion rate of the ad and to the HUI is done by extended PSO. Hence Particle Swarm Optimization a resourceful Bio inspired algorithm determines the high utility ads for the advertisement on placement website with the help of conversion rate.

Conclusion
This paper presented a detailed study of various metaheuristic nature inspired optimization approaches used to solve ARM, FIM and HUIM. This paper provided a list of nature inspired algorithms along with a short overview of them, a comparative analysis of the approaches in terms of various variables and critical problems have also been addressed. According to the survey, each approaches used for ARM, FIM and HUIM performs better at particular environmental setup. Also, it is observed the may modified version of nature inspired approaches have been evolved in recent years.

References
[1] K. R. Prasanna Kumar and K. Kousalya, “Amelioration of task scheduling in cloud computing using crow search algorithm,” Neural Comput. Appl., vol. 32, no. 10, pp. 5901–5907, May 2020, doi: 10.1007/s00521-019-04067-2.
[2] K. Logeswaran, P. Suresh, S. Savitha, K. R. Prasanna Kumar, A. P. Ponselvakumar,
and A. R. Kannan, “Data driven diagnosis of cervical cancer using association rule mining with trivial rule expulsion approach,” *Int. J. Emerg. Technol.*, vol. 11, no. 2, pp. 110–115, 2020.

[3] S. B. Gopal *et al.*, “Minimum Energy Consumption Objective Function For RPL In Internet Of Things,” *Int. J. Sci. Technol. Res.*, vol. 9, p. 1, 2020, Accessed: Oct. 05, 2020. [Online]. Available: www.ijstr.org.

[4] S. Sharmila and S. Vijayarani, “Association rule mining using fuzzy logic and whale optimization algorithm,” *Soft Comput.*, Aug. 2020, doi: 10.1007/s00500-020-05229-4.

[5] N. K. Sreeja and A. Sankar, “A hierarchical heterogeneous ant colony optimization based approach for efficient action rule mining,” *Swarm Evol. Comput.*, vol. 29, pp. 1–12, Aug. 2016, doi: 10.1016/j.swevo.2016.02.004.

[6] A. Song, X. Ding, J. Chen, M. Li, W. Cao, and K. Pu, “Multi-objective association rule mining with binary bat algorithm,” *Intell. Data Anal.*, vol. 20, no. 1, pp. 105–128, Jan. 2016, doi: 10.3233/IDA-150796.

[7] C. Romero, A. Zafra, J. M. Luna, and S. Ventura, “Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data,” *Expert Syst.*, vol. 30, no. 2, pp. 162–172, May 2013, doi: 10.1111/j.1468-0394.2012.00627.x.

[8] K. E. Heraguemi, N. Kamel, and H. Drias, “Association rule mining based on bat algorithm,” *J. Comput. Theor. Nanosci.*, vol. 12, no. 7, pp. 1195–1200, 2015, doi: 10.1166/jctn.2015.3873.

[9] S. Ghosh, S. Biswas, D. Sarkar, and P. Sarkar, “Mining Frequent Itemsets Using Genetic Algorithm,” *Int. J. Artif. Intell. Appl.*, vol. 1, no. 4, pp. 133–143, 2010, doi: 10.5121/ijaia.2010.1411.

[10] O. Smart and L. Burrell, “Genetic programming and frequent itemset mining to identify feature selection patterns of iEEG and fMRI epilepsy data,” *Eng. Appl. Artif. Intell.*, vol. 39, pp. 198–214, Mar. 2015, doi: 10.1016/j.engappai.2014.12.008.

[11] Y. Djennouri and M. Comuzzi, “GA-Apriori: Combining apriori heuristic and genetic algorithms for solving the frequent itemsets mining problem,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10526 LNAI, pp. 138–148, doi: 10.1007/978-3-319-67274-8_13.
[12] Y. Djenouri and M. Comuzzi, “Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem,” *Inf. Sci. (Ny)*., vol. 420, pp. 1–15, Dec. 2017, doi: 10.1016/j.ins.2017.08.043.

[13] S. Wang and Y. Yin, “Polygene-based evolutionary algorithms with frequent pattern mining,” *Front. Comput. Sci.*, vol. 12, no. 5, pp. 950–965, 2018, doi: 10.1007/s11704-016-6104-3.

[14] M. S. Arunkumar, P. Suresh, and C. Gunavathi, “High Utility Infrequent Itemset Mining Using a Customized Ant Colony Algorithm,” *Int. J. Parallel Program.*, 2018, doi: 10.1007/s10766-018-0621-7.

[15] Y. Djenouri, D. Djenouri, A. Belhadi, P. Fournier-Viger, and J. C.-W. Lin, “A new framework for metaheuristic-based frequent itemset mining,” *Appl. Intell.*, vol. 48, no. 12, pp. 4775–4791, Dec. 2018, doi: 10.1007/s10489-018-1245-8.

[16] L. Zhang, G. Fu, F. Cheng, J. Qiu, and Y. Su, “A multi-objective evolutionary approach for mining frequent and high utility itemsets,” *Appl. Soft Comput.*, vol. 62, pp. 974–986, Jan. 2018, doi: 10.1016/j.asoc.2017.09.033.

[17] H. Cao, S. Yang, Q. Wang, Q. Wang, and L. Zhang, “A Closed Itemset Property based Multi-objective Evolutionary Approach for Mining Frequent and High Utility Itemsets,” *2019 IEEE Congr. Evol. Comput. CEC 2019 - Proc.*, pp. 3356–3363, 2019, doi: 10.1109/CEC.2019.8789985.

[18] J. C.-W. Lin *et al.*, “Mining high-utility itemsets based on particle swarm optimization,” *Eng. Appl. Artif. Intell.*, vol. 55, pp. 320–330, Oct. 2016, doi: 10.1016/j.engappai.2016.07.006.

[19] C.-W. Lin, T.-P. Hong, J.-W. Wong, G.-C. Lan, and W.-Y. Lin, “A GA-Based Approach to Hide Sensitive High Utility Itemsets,” *Sci. World J.*, vol. 2014, no. 1, pp. 1–12, 2014, doi: 10.1155/2014/804629.

[20] J. C. W. Lin, L. Yang, P. Fournier-Viger, J. Frnda, L. Sevcik, and M. Voznak, “An evolutionary algorithm to mine high-utility itemsets,” *Adv. Electr. Electron. Eng.*, vol. 13, no. 4, pp. 392–398, 2015, doi: 10.15598/aeee.v13i4.1474.

[21] Q. Zhang, W. Fang, J. Sun, and Q. Wang, “Improved Genetic Algorithm for High-Utility Itemset Mining,” *IEEE Access*, vol. 7, pp. 176799–176813, 2019, doi: 10.1109/ACCESS.2019.2958150.

[22] R. Gunawan, E. Winarko, and R. Pulungan, “A BPSO-based method for high-utility
itemset mining without minimum utility threshold,” *Knowledge-Based Syst.*, vol. 190, no. xxxx, p. 105164, 2020, doi: 10.1016/j.knosys.2019.105164.

[23] J. M.-T. Wu, J. Zhan, and J. C.-W. Lin, “An ACO-based approach to mine high-utility itemsets,” *Knowledge-Based Syst.*, vol. 116, pp. 102–113, Jan. 2017, doi: 10.1016/j.knosys.2016.10.027.

[24] X. Zhu, Y. Liu, J. Li, T. W. B, and Z. Q. B, *Discovering High Utility Itemsets Based on the Artificial Bee Colony Algorithm*, vol. 10939. Springer International Publishing, 2018.

[25] N. Pazhaniraja, S. Sountharrajan, and B. Sathis Kumar, “High utility itemset mining: a Boolean operators-based modified grey wolf optimization algorithm,” *Soft Comput.*, vol. 0123456789, 2020, doi: 10.1007/s00500-020-05123-z.

[26] W. Song and C. Huang, “Mining High Utility Itemsets Using Bio-Inspired Algorithms: A Diverse Optimal Value Framework,” *IEEE Access*, vol. 6, pp. 19568–19582, 2018, doi: 10.1109/ACCESS.2018.2819162.

[27] J. C.-W. Lin, L. Yang, P. Fournier-Viger, T.-P. Hong, and M. Voznak, “A binary PSO approach to mine high-utility itemsets,” *Soft Comput.*, vol. 21, no. 17, pp. 5103–5121, Sep. 2017, doi: 10.1007/s00500-016-2106-1.

[28] S. Kannimuthu and K. Premalatha, “Discovery of High Utility Itemsets Using Genetic Algorithm with Ranked Mutation,” *Appl. Artif. Intell.*, vol. 28, no. 4, pp. 337–359, Apr. 2014, doi: 10.1080/08839514.2014.891839.

[29] J. C.-W. Lin et al., “High utility-itemset mining and privacy-preserving utility mining,” *Perspect. Sci.*, vol. 7, pp. 74–80, 2016, doi: 10.1016/j.pisc.2015.11.013.

[30] D. JayaKumari, N. V. E. S.Murthy, and S. Srinivasa Suresh, “An Optimization based Modified Maximum Sensitive Item-Sets Conflict First Algorithm (MSICF) for Hiding Sensitive Item-Sets,” *Int. J. Comput. Appl.*, vol. 72, no. 4, pp. 1–4, Jun. 2013, doi: 10.5120/12479-8881.

[31] K. Logeswaran, P. Suresh, S. Savitha, and P. K. K. R, “Optimization of Evolutionary Algorithm Using Machine Learning Techniques for Pattern Mining in Transactional Database,” in *Handbook of Research on Applications and Implementations of Machine Learning Techniques*, US, 2019.

[32] Y. Djenouri, P. Fournier-Viger, A. Belhadi, and J. Chun-Wei Lin, “Metaheuristics for Frequent and High-Utility Itemset Mining,” Springer, Cham, 2019, pp. 261–278.
[33] M. Keerthi, J. Anitha, S. Agrawal, and T. Varghese, “Mining High Utility Itemset for Online Ad Placement Using Particle Swarm Optimization Algorithm,” in *Advances in Intelligent Systems and Computing*, vol. 1108 AISC, Springer, 2020, pp. 555–562.