

# Recommendation System for High Utility Itemsets over Incremental Dataset

J K Kavitha<sup>1</sup>, U Kanimozhi<sup>2</sup>, D Manjula<sup>3</sup>Research Scholar, Department of Computer Science and Engineering, CEG, Anna University, Chennai, India<sup>1,2</sup>Professor, Department of Computer Science and Engineering, CEG, Anna University, Chennai, India<sup>3</sup>

**ABSTRACT:** Mining high utility itemsets has gained much significance in the recent years. When the data arrives sporadically, incremental and interactive utility mining approaches can be adopted to handle users' dynamic environmental needs and avoid redundancies, using previous data structures and mining results. The dependence on recommendation systems has exponentially risen since the advent of search engines. This paper proposes a model for building a recommendation system that suggests high utility itemsets over dynamic datasets for a location prediction strategy to predict users' trajectories using the Fast Update Utility Pattern Tree (FUUP) approach. Through comprehensive evaluations by experiments this scheme has shown to deliver excellent performance.

**KEYWORDS:** Data mining, Utility Mining, Incremental Mining, Items recommendation, Semantic prediction

## I. INTRODUCTION

Association rule mining [1,2,3,11,13,14,21,22,23,24] is commonly seen in transactional databases in the field of Data Mining. The above mentioned approaches are based on frequency values of items, which cannot meet the demands of different factors like profit or price or sales based on user preferences in real world applications. Utility mining [5,29,30] was proposed to solve the above mentioned problem by considering the factors like cost, profit or other factors of users' interest. Thus the issue of high utility itemsets mining is raised and many studies [4,9,17,19,20,25,27] have addressed this problem. Liu, Liao & Choudhary [19,20] proposed the two phase utility mining algorithm for efficiently extracting all high utility itemsets based on the downward closure property. Although two phase algorithm reduces search space, it still generates too many candidates and requires multiple database scans. To overcome this problem, Li, Yeh & Chang [17] proposed an isolated items discarding strategy (IIDS) to reduce the number of candidates. To efficiently generate HTWI's and to avoid multiple scans, Ahmed, Tanbeer, Jeong & Lee [4] proposed a tree based algorithm named IHUP. More recently, Tseng, Wu, Shie & Yu [27] proposed UP-Growth for mining high utility itemsets with a set of effective strategies for pruning candidate itemsets.

However, in real world applications, the problem of discovering the frequent itemsets becomes more time consuming if the dataset is incremental in nature. This may introduce new frequent itemsets and some existing itemsets would become invalid. Several approaches [6,7,8,10,11,15,18,28] are proposed to address this issue. Thus designing an efficient algorithm that can maintain association rules as a database grows is thus critically important. One noticeable incremental mining algorithm was the Fast-Updated Algorithm (called FUP) which was proposed by Cheung, Han, Ng & Wong [6] for avoiding shortcomings mentioned above. It primary calculates frequent itemsets from new transactions and compares them with the previous found frequent itemsets from the original database. Different procedures are then done according to the comparison results. For some cases, FUP can avoid or reduce the number of re-scanning the original database, thus saving computation time in incremental mining. Based on FUP a new incremental mining FUUP (Fast Update Utility Pattern) Tree algorithm [31] for efficiently mining high utility itemsets is proposed to handle the above mention situation. It is based on the concept of UP-Growth (Utility Pattern Growth) for mining high utility itemsets with a set of effective strategies for pruning candidate itemsets and Fast Update (FUP) approach, which first partitions itemsets into four parts according to whether they are high-transaction weighted utilization items in the original and newly inserted transactions.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

On the other hand, the recommendation system for mobile users has attracted a lot of attentions in recent years. Recommendation systems offer people expedient access to the products they might be fascinated in. Several items recommendation methods, considering users' current and next locations, is called location-based recommendation and has been discussed in many existing work (Bao, Zheng, & Mokbel, 2012; Eagle, Pentland, & Lazer, 2009; Lu, Lee, & Tseng, 2012; Lu, Tseng, & Yu, 2011; Taiwan Tourism Bureau, 0000). The location-based recommendation methods usually use the frequent moving behaviours of users to predict the next move of a user and recommend the items which are related to that location. To make accurate location prediction, the location-based recommendation systems always not only record users' GPS trajectories but also mine the frequent moving behaviours from the users' GPS trajectories. A novel approach for recommending items for mobile users based on both the geographic and semantic features of users' trajectories. The core idea of the recommendation system is based on a novel cluster-based location prediction strategy, namely TrajUtiRec [32] to improve items recommendation model. The cluster-based location prediction strategy evaluates the next location of a mobile user based on the frequent behaviours of similar users in the same cluster determined by analysing users' common behaviours in semantic trajectories.

According to above-mentioned reasons, in this paper, we propose a recommender which can not only predict users' next movable location but also recommend items which are sold by stores located in users' next predicted locations. The traditional item recommendation methods do not fully consider both the users' next locations and high utility itemsets. Thus a new recommendation system is proposed to predict user next movable location based on TrajUtiRec [32] and to suggests high utility itemsets over an incremental datasets based on FUUP [31] tree algorithm. Experimental results of our proposed recommendation system for suggesting high utility itemset over incremental datasets is shown to deliver excellent performance.

Paper is organized as follows. Section II describes review of related works. The system overview and flow diagram represents the step of the recommendation system is given in Section III. Section IV presents experimental results showing results. Finally, Section V presents conclusion.

## II. REVIEW OF RELATED WORKS

In this section, some related researches are briefly reviewed. They are the mining high utility itemsets and the concept of location-based prediction based on GPS trajectories.

### A. Mining High Utility Itemsets

In the Past, several mining algorithms were proposed for efficiently discovering high utility itemsets. Yao, Hamilton & Butz [29] proposed an algorithm for efficiently mining high utility itemsets based on mathematical properties of utility constraints. Two pruning strategies based on utility upper bounds and expected utility upper bounds respectively were adopted to reduce the search space. These pruning strategies were then incorporated into the mining approach UMining and its heuristic successor, UMining\_H [30]. Liu, Liao & Choudhary[19,20] designed a two phase algorithm for efficiently discovering all high utility itemsets. It consisted of two phases to generate and test high utility was first used as effective upper bound of each candidate itemset in the transaction such that the "transaction-weighted downward closure property" could be kept in the search space to decrease the number of candidate itemsets. An additional database scan was then performed to find out the real utility values of the remaining candidates and identifies the high utility itemsets. Although two phase algorithm reduces search space, it still generates too many candidates and requires multiple database scans. To overcome this problem, Li, Yeh & Chang [17] proposed an isolated items discarding strategy (IIDS) to reduce the number of candidates. To efficiently generate HTWI's and to avoid multiple scans, Ahmed, Tanbeer, Jeong & Lee [4] proposed a tree based algorithm named IHUP. Although IHUP achieves better performance than IIDS & Two-phase, it still produces too many HTWI's in phase I. Such a large number of HTWUI's will degrade the mining performance in phase I substantially in terms of execution time and memory consumption. As advancement, Tseng, Wu, Shie & Yu [27] proposed UP-Growth for mining high utility itemsets with a set of effective strategies for pruning candidate itemsets. Correspondingly, a compact tree structure, called *UP-Tree (Utility Pattern Tree)*, was designed to maintain the important information of the transaction database related to the utility patterns. An incremental mining algorithm FUUP tree [31] for efficiently mining high utility itemsets is proposed to handle dynamic datasets. It is based on the concept of UP-Growth (Utility Pattern Growth) for mining high utility itemsets with a set of

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

effective strategies for pruning candidate itemsets and Fast Update (FUP) approach, which first partitions itemsets into four parts according to whether they are high-transaction weighted utilization items in the original and newly inserted transactions. An FUUP tree must be built in advance from the initially original database before new transactions come. Its initial construction is similar to that of an UP tree according to the strategy of DGU and DGN. The database is first scanned to find the items with their TWU larger than a minimum utility threshold, which called promising items. Other items are called unpromising. Next, the promising items are sorted in descending order and reorganized transaction utility is evaluated. At last, the reorganized transaction is scanned again to construct the tree according to the sorted order of promising items. The construction process is executed tuple by tuple, from the first transaction to the last one. After all transactions are processed, the final UP tree for the original database is completely constructed.

When new transactions are added, the incremental maintenance algorithm will process them to construct the FUUP-tree. The new transactions are first scanned to find the promising and unpromising items according to the TWU of newly inserted transactions. Then, it partitions items into four parts according to whether they are large or small in the original database and in the new transactions. Each part is then processed in its own way. The Header-Table and the FUUP tree are correspondingly updated whenever necessary.

In the process for updating the FUUP tree, item deletion is done before item insertion. When an originally large item becomes small, it is directly removed from the FUUP tree and its parent and child nodes are then linked together. On the contrary, when an originally small item becomes large, it is added to the end of the Header-Table and then inserted into the leaf nodes of FUUP tree. It is reasonable to insert the item at the end of the Header-Table since when an originally small item becomes large due to the addition of new transactions; its updated support is usually only a little larger than the minimum support. The FUUP-tree can thus be least updated in this way, and the performance of the proposed incremental algorithm can be greatly improved. The entire FUUP-tree can be re-constructed in a batch way when a sufficiently large number of transactions are inserted.

## B. Location based prediction

Many data mining studies have discussed the problems of predicting the next location where a mobile user moves to. Personal-based prediction (Jeung, Liu, Shen, & Zhou, 2008; Yavas, Katsaros, Ulusoy, & Manolopoulos, 2005; Ye, Zheng, Chen, Feng, & Xie, 2009) and general-based prediction (Monreale et al., 2009; Morzy, 2006, 2007; Zheng, Zhang, Xie, & Ma, 2009a,b) are two approaches often adopted in this problem domain. The personal-based prediction approach considers movement behavior of each individual as independent and thus uses only the movements of an individual user to predict his/her next location. On the contrary, the general-based prediction makes a prediction based on the common movement behavior of general mobile users. In Jeung et al. (2008), propose an innovative approach which forecasts future locations of a user by combining predefined motion functions, i.e., linear or non-linear models that capture object movements as sophisticated mathematical formulas, with the movement patterns of the user, extracted by a modified version of the Apriori algorithm. In Yavas et al. (2005), mine the movement patterns of an individual user to form association rules and use these rules to make location prediction. Additionally, they consider the support and confidence in selecting the association rules for making predictions.

In Ye, Zheng, Chen, Feng, and Xie (2009), propose a novel pattern, called Individual Life Pattern, which is mined from individual trajectory data, and they use such pattern to describe and model the mobile users' periodic behaviors. In Morzy (2006), uses a modified version of Apriori algorithm to generate association rules. Morzy (2007), uses a modified version of PrefixSpan algorithm to discover frequent patterns of users' movements for generating the prediction rules. The matching functions employed in these previous works are based on the notions of support and confidence. Although all of Morzy's approaches have considered temporal information and location hierarchy, they do not take into account the semantic tags of locations. In Monreale et al. (2009), proposes a method aiming to predict with a certain level of accuracy the next location of a moving object. The movement patterns extracted for prediction covers three different movement behaviors, including order of locations, travel time, and frequency of user visits. In Zheng et al. (2009a), uses a HITS based model to mine users' interesting location and detect users' travel sequence to make locations prediction, and in Zheng et al. (2009b), they consider the location correlation for generating the users' interesting locations and travel sequence. Note that the above-mentioned prediction methods are based on geographic

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

information only. On the contrary, our proposal predicts the next location of a user based on both geographic and semantic information in trajectories.

In recent years, a number of studies on semantic trajectory data mining have appeared in the literature (Alvares et al., 2007; Bogorny et al., 2009). In Alvares et al. (2007), propose to explore the geographic semantic information to mine semantic trajectory patterns from mobile users' movement histories. First, they discover the stops of each trajectory and map these stops to semantic landmarks to transform geographic trajectories into semantic trajectories. By applying a sequential pattern mining algorithm on semantic trajectories, they obtain frequent patterns, namely, semantic trajectory patterns, to represent the frequent semantic behaviors of mobile users. In Bogorny et al. (2009), use a hierarchy of geographic semantic information to discover more interesting patterns. Notice that the notion of stops in the above-mentioned works only considers the aspect of 'stay' in stops but not the 'positions' of these stops in geographic space. As a result, many unknown stops are generated. Thus, a recent recommendation system TrajUtiRec [32], by taking into account the geometric distribution of these stops is grouped together such that the Trajectory is transformed as the sequence instead. Besides, a feature vector is proposed by Zheng to describe the semantics of each location. Based on the feature vector, the semantic similarity between two mobile users could be calculated. In addition to the GPS trajectory, Ying, Lu, Lee, Weng, and Tseng (2010) also exploit the cell trajectory to derive the semantic similarity between two mobile users. The cell trajectory consists of a sequence of spatio-temporal points in the form of cell station ID, arrive time, and leave time. They propose a novel similarity measurement, namely, Maximal Semantic Trajectory Pattern Similarity (MSTP-Similarity) to evaluate the user similarity. As such, the similarity of two mobile users, even if they live in different cities, may be evaluated based on their similar semantic trajectory patterns.

### III. SYSTEM OVERVIEW

We propose a system using novel cluster based location prediction and High utility items recommendation framework, based on both the geographic and semantic features in trajectories. The proposed approach works for locations where the users may have never visited, e.g., a location in other cities. The overall framework consists of (1) an offline training module, and (2) an online location prediction and high utility items recommendation module, Fig.1 shows the framework and its flow of data processing. The main motive is to explore the activities of mobile users, captured in semantic trajectories, to improve accuracy of location prediction and High utility items recommendation. As shown, the training module includes two parts. The first part is for identifying user next location based on trajectories. It involves three steps. The first step, called data preprocessing, transforms each user's trajectories as stay location sequences. The second step, called semantic mining, extracts users' semantic behaviors (as 'semantic trajectory patterns' which will be detailed later). It also obtains user clusters based on the semantic behavior similarity of users. The third step, called geographic mining, extracts the geographic behaviors of users in each cluster (as 'stay location patterns' which will be detailed later). The second part is for identifying high utility itemsets over an incremental datasets.

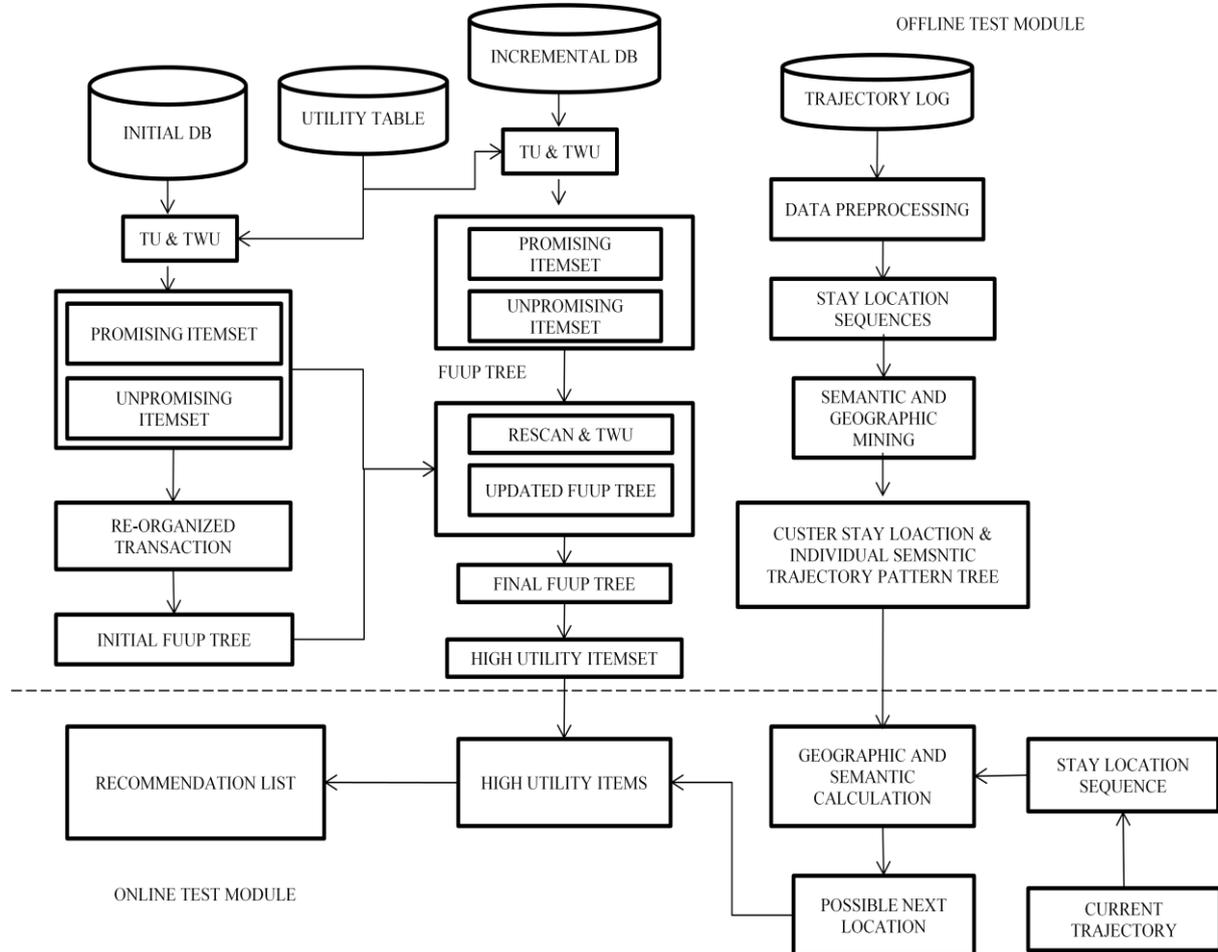


Fig. 1 Framework and the flow of data processing

In the online module, a scoring function is used to evaluate the probability for a location to be the next location. Here, we consider not only geographic information but also semantic information. First, we calculate the geographic score and derive several candidate paths. Then, the semantic score of each candidate path is evaluated. Finally, we compute a weighted average of geographic score and semantic score for each candidate path to select the most probable path for predicting the next location in a user's move. The high utility items are suggested for the predicted next location of the user.

### A. Offline Test Module

#### Location Prediction

In this section, we propose an approach to extract the users' frequent movement behaviors which includes the semantic behavior information for individual users and the geographic behavior information for clusters of similar users. We mine a kind of frequent patterns, called semantic trajectory patterns (Alvares et al., 2007; Ying et al., 2010), from trajectories of individual users and adopt a prefix tree, called semantic trajectory pattern tree, to compactly represent a collection of semantic trajectory patterns. Based on individual semantic information (i.e., the semantic trajectory patterns and their support values), we cluster mobile users. For each cluster, the sequential pattern mining is used to extract cluster geographic information, called stay location patterns. Similarly, we also adopt a prefix tree to compactly

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

represent a collection of stay location patterns. As mentioned earlier, this mining module consists of (1) Data Preprocessing step, (2) Semantic Mining step, and (3) Geographic Mining step.

The data preprocessing step transforms each user's GPS trajectories into stay location sequences. We argue that most activities of a mobile user are usually performed at where the user stays. Our framework is able to deal with both the GPS trajectories and cell trajectories (Ye et al., 2009). For GPS trajectory, we follow Zheng et al.'s work (Zheng, Zhang, & Xie, 2010) to discover stay points from users' GPS trajectories. Then, a density-based clustering algorithm is performed on these stay points to obtain stay locations. For cell trajectories, we follow Ying et al.'s approach (Ying et al., 2010), which treats a cell as a geographic location. The stay time in a cell is derived by calculating the difference between the time a user arrives in and leave from the cell. Finally, the stay locations (i.e., the cells with stay time equal or greater than the time threshold and the number of visitors equal or greater than the crowd threshold) are obtained and each trajectory is transformed into a stay locations sequence.

Semantic mining is used to extract semantic trajectory patterns from a user's stay location sequences and build semantic trajectory pattern tree based on the discovered patterns. There are two main steps. First, we mine semantic trajectory pattern from each user's stay location sequence set. Then, we perform a hierarchical clustering method to cluster users, where the user's similarity is based on MSTP-Similarity (Ying et al., 2010). Although semantic mining discovers users' semantic trajectory patterns, they cannot be used directly for location prediction since locations are not deductable from the semantic labels. To overcome this problem, we mine the geographic information from users' stay location sequences. While we aim to take into account the common frequent behaviors of mobile users, considering the frequent behavior of all general users may cause imbalanced data problem. Hence, we consider the clusters resulted from the semantic mining to aggregate the stay location sequences of mobile users. We then perform a sequential pattern mining algorithm Prefix-Span (Pei et al., 2001) on each cluster's semantic stay location sequences to mine the frequent stay location sequence, called stay location pattern. Similarly, the longer patterns we discover the more subsequences are generated due to the downward closure property (Pei et al., 2001). It leads to a loss of efficiency because all the subsequences of a long pattern are to be checked in the next location prediction. Therefore, we also adopt a prefix tree, called stay location pattern tree (SLP-Tree), to compactly represent a collection of stay location patterns. We also perform the STP-Tree Building algorithm, on each stay location pattern set of each cluster to build an SLP-Tree. Similarly, the paths with only one node are not included in the pattern tree.

## High Utility Itemsets Mining

To find the high utility itemsets, we adopt the FUUP tree mining algorithm to find high utility itemsets from transaction database [31]. An FUUP tree must be built in advance from the initially original database before new transactions come. Its initial construction is similar to that of an UP tree according to the strategy of DGU and DGN. The database is first scanned to find the items with their TWU larger than a minimum utility threshold, which called promising items. Other items are called unpromising. Next, the promising items are sorted in descending order and reorganized transaction utility is evaluated. At last, the reorganized transaction is scanned again to construct the tree according to the sorted order of promising items. The construction process is executed tuple by tuple, from the first transaction to the last one. After all transactions are processed, the final UP tree for the original database is completely constructed.

When new transactions are added, the proposed incremental maintenance algorithm will process them to construct the FUUP-tree. The new transactions are first scanned to find the promising and unpromising items according to the TWU of newly inserted transactions. Then, it partitions items into four parts according to whether they are large or small in the original database and in the new transactions. Each part is then processed in its own way. The Header-Table and the FUUP tree are correspondingly updated whenever necessary.

In the process for updating the FUUP tree, item deletion is done before item insertion. When an originally large item becomes small, it is directly removed from the FUUP tree and its parent and child nodes are then linked together. On the contrary, when an originally small item becomes large, it is added to the end of the Header-Table and then inserted into the leaf nodes of FUUP tree. It is reasonable to insert the item at the end of the Header-Table since when an originally small item becomes large due to the addition of new transactions; its updated support is usually only a little larger than the minimum support. The FUUP-tree can thus be least updated in this way, and the performance of the

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

proposed incremental algorithm can be greatly improved. The entire FUUP-tree can be re-constructed in a batch way when a sufficiently large number of transactions are inserted. Based on the FUUP tree, the desired association rules can then be found by the UP-Growth mining approach (Tseng, Wu, Shie & Yu, 2010).

## B. Online Prediction and Recommendation Module

Given a mobile user, the on-line location prediction and items recommendation module predicts her next stay location based on the stay location pattern tree of her cluster and her own semantic trajectory pattern tree. Given these two pattern trees, the geographic information (i.e., the stay location patterns) of the cluster which the mobile user belongs to and the semantic information (i.e., the semantic trajectory patterns) of the mobile user herself can be incorporated in the location prediction and items recommendation. Thus, given the trajectory of a user's recent moves, we compute the best matching scores of candidate paths in these two pattern trees by following [32]. After location scoring, we can easily predict user's possible next locations. By the off-line module, we have discovered the high utility itemset which is sold by retailers in user's possible next locations. Therefore, we can make the recommending list in which the items are ranked according to their utility values

## IV. EXPERIMENTAL EVALUATION

In this section, we conduct a series of experiments to evaluate the performance for the proposed location prediction and items recommendation technique. All the experiments are implemented in Java JDK 1.6 on an Intel Core Quad CPU Q6600 2.40 GHz machine with 1 GB of memory running Microsoft Windows XP. The data preparation task on the MIT reality mining dataset is done first and then introduces the evaluation methodology. Finally, we present our experimental results.

The simulation framework by following [32] is done. All of the parameters can be classified into four categories, i.e., travel network building, classical trajectory generation, travel trajectory simulation and mobile transaction modeling. The simulation framework can be divided into three phases. The first phase is to build the travel network. We use a mesh network to represent the travel network. After building the travel network, the next task is to generate the classical trajectory. The purpose of this phase is to generate classical trajectories as candidate trajectories. The second phase is to simulate the travel trajectory. The last phase is to model the mobile transaction. The Precision, Recall, and F-measure are the main measurements for the experimental evaluation.

Items recommendation model is more effective. We have changes in the minimum utility to render hit rate results, and the horizontal axis for the minimum utility, the vertical axis is the hit rate, to prove that we look at different effective threshold. As shown in the Fig. 2, the proposed recommendation system using EIRM and FUUP Mining shows better performance than the already existing recommendation system [32] using two phase mining algorithm. We can observe that hit ratio of both are shown in straight line after minimum utility 42. The reason is that our recommender is to recommend high utility itemsets to users. It leads the recommending list would not be changed while minimum utility is set high enough because only the highest utility itemset is filtered from database. Thus, we can conclude the critical value of minimum support in this dataset is 42.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

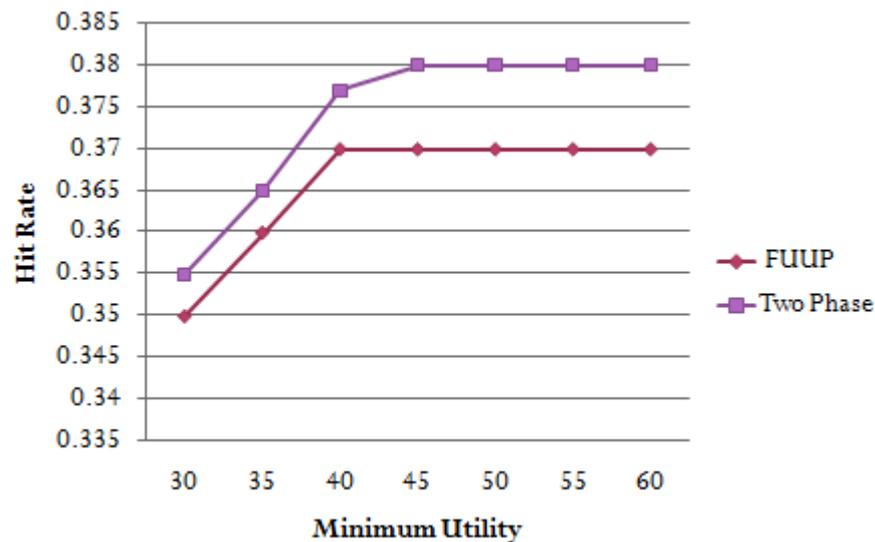


Fig. 2 Performance of the Proposed System

## V. CONCLUSION AND FUTURE WORK

In this paper, we propose a recommendation system, which not only predict users' next locations but also recommend high utility items for the users over dynamic datasets. It is based on the concept of recommendation system using cluster based location prediction strategy to predict users' next movable location and mining high utility itemset to find high utility itemset over dynamic datasets using Fast Update Utility Pattern Tree (FUUP) approach. The core idea of the cluster-based location prediction technique is to group users according to their similarity of semantic trajectory. By adopting the high utility itemset mining, we can find high utility products of each location according to recent user preferences and current trends. Through a comprehensive evaluation by experiments, our proposed recommendation system for suggesting high utility itemset is shown to deliver excellent performance. Although our recommendation system has excellent performance, some research issues still have not been addressed in this paper. The location prediction is based on the user trajectory whereas not on the basis of user preferences. We leave these issues as future works and plan to design more advanced recommendation strategies to address these issues in location-based services.

## REFERENCES

- [1] Agrawal, R., Imielinski, T., and Swami, A., "Mining association rules between sets of items in large database," The ACM SIGMOD Conference, pp. 207-216, 1993.
- [2] Agrawal, R., and Srikant, R., "Fast algorithm for mining association rules," The International Conference on Very Large Data Bases, pp. 487-499, 1994.
- [3] Agrawal, R., Srikant, R., and Vu, Q., "Mining association rules with item constraints," The Third International Conference on Knowledge Discovery in Databases and Data Mining, pp. 67-73, 1997
- [4] Ahmed, C.F., Tanbeer, S.K., Jeong, B.S., and Lee, Y.K., "Efficient tree structures for high utility pattern mining in incremental databases," IEEE Transactions on Knowledge and Data Engineering, Vol. 21, Issue 12, pp. 1708-1721, 2009.
- [5] Chan, R., Yang, Q., and Shen, Y.D., "Mining high utility itemsets", In The 3rd IEEE international conference on data mining, pp. 19-26, 2003.
- [6] Cheung, D. W., Han, J., Ng, V. T., and Wong, C. Y., "Maintenance of Discovered Association Rules in Large Databases: An Incremental Updating Technique", 12<sup>th</sup> International Conference on Data Engineering, New Orleans, Louisiana, 1996.
- [7] Cheung, D. W., Lee, D. W., and Kao, S. D., "A General Incremental Technique for Maintaining Discovered Association Rules", In Proceedings of the Fifth International Conference on Database System for Advanced Applications, Melbourne, Australia, 1997.
- [8] Das, A., and Bhattacharyya, D.K., "Rule Mining for Dynamic Databases", AJIS, 13, No.1, pp 19-39, 2005.
- [9] Erwin, A., Gopalan, R.P., and Achuthan, N.R., "Efficient mining of high utility itemsets from large datasets," in Proc. of PAKDD 2008, LNAI 5012, pp. 554-561.
- [10] Ezeife, C. I., and Su, Y., "Mining Incremental Association Rules with Generalized FP Tree", Proceedings of 15th Canadian Conference on Artificial Intelligence, AI2002, Calgary, Canada, May 2002.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2015

- [11] Feldman, R., Aumann, Y., and Lipshtat, O., "Borders: An Efficient Algorithm for Association Generation in Dynamic Databases", Journal of Intelligent Information System, Pages 61–73,1999.
- [12] Fukuda, T., Morimoto, Y., Morishita, S. and Tokuyama, T., "Mining optimized association rules for numeric attributes," The ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, pp. 182-191, 1996.
- [13] Han, J., and Fu, Y., "Discovery of multiple-level association rules from large database," The Twentv-first International Conference on Ve-v Large Data Bases, pp. 420-431, 1995.
- [14] Han, J., Pei, J., Yin, Y., and Mao, R., "Mining frequent patterns without candidate generation: a frequent-pattern tree approach". Data Mining and Knowledge Discovery, 8, 53–87,2004.
- [15] Hong, T-P., Lin, C-W., and Wu, Y-L., "Incrementally Fast Updated Frequent Pattern Trees", Expert Systems with Applications: An Int'l Journal, 34, no. 4, pp 2424-2435, 2008.
- [16] IBM Quest Data Mining Project, " Quest Synthetic Data Generation Code", <http://www.almaden-ibm.com/cs/quest/syndata.html>,1996.
- [17] Li, Y.-C., Yeh, J.-S., and Chang, C.-C., "Isolated items discarding strategy for discovering high utility itemsets," Data & Knowledge Engineering, Vol. 64, Issue 1, pp. 198-217, 2008.
- [18] Lin, C. M., Hsieh, Y. L., Yin, K. C., Hung, M. C., and Yang, D. L., "ADMiner: An Incremental Data Mining Approach Using a Compressed FP-tree", Journal Of Software, VOL. 8, NO. 8, AUGUST 2013.
- [19] Liu, Y., Liao, W., and Choudhary, A., "A fast high utility itemsets mining algorithm," in Proc. of the Utility-Based Data Mining Workshop, 2005.
- [20] Liu, Y., Liao, W.K., and Choudhary, A., "A Two Phase Algorithm for Fast Discovery of High Utility of Itemsets," Proc. Ninth Pacific- Asia Conf. Knowledge Discovery and Data Mining (PAKDD '05), pp. 689-695, 2005.
- [21] Mannila, H., Toivonen, H., and Verkamo, A.I., "Efficient algorithm for discovering association rules,"The AAAI Workshop on Knowledge Discovery in Databases, pp. 181-192, 1994.
- [22] Park, J.S, Chen, M.S., and Yu, P.S, "Using a hash-based method with transaction trimming for mining association rules," IEEE Transactions on Knowledge and Data Engineering, Vol. 9, No. 5, pp. 812-825, 1997.
- [23] Srikant, R., and Agrawal, R., "Mining generalized association rules,"The Twenty-first International Conference on Very Large Data Bases, pp. 407-419, 1995.
- [24] Srikant, R., and Agrawal, R., "Mining quantitative association rules in large relational tables," ACM SIGMOD International Conf on Management of Data, pp. 1-12, Montreal, Canada, 1996.
- [25] Shie, B.E., Tseng, V.S., and Yu, P. S., "Online mining of temporal maximal utility itemsets from data streams," in Proc. of the 25th Annual ACM Symposium on Applied Computing, Switzerland, Mar., 2010.
- [26] Tseng, V. S., Chu, C. J. ,and Liang, T., "Efficient Mining of Temporal High Utility Itemsets from Data streams," in Proc. of ACM KDD Workshop on Utility-Based Data Mining Workshop (UBDM'06), USA, Aug., 2006.
- [27] Tseng, V. S., Wu, C.W., Shie, B.E., and Yu, P. S., "UP-Growth: An Efficient Algorithm for High Utility Itemsets Mining," in Proc. of the 16th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD 2010), pp. 253-262, 2010.
- [28] Yafi, E., Al-Hegami A., Alam A., and Biswas, R., "YAMI:Incremental mining of interesting Association Patterns",International Arab Journal of Information technology, Vol-9, No-6, November 2012.
- [29] Yao, H., Hamilton, H J., and Butz, C. J., "A foundational approach to mining itemset utilities from databases". In The 4th SIAM international conference on data mining (pp. 211–225) 2004.
- [30] Yao, H., & Hamilton, H. J. "Mining itemset utilities from transaction databases". Data and Knowledge Engineering, 59, 603–626, 2006.
- [31] Kavitha, J.K., Manjula, D., and Kasthuri Bha, J.K., 'Fast Update Utility Pattern Tree (FUUP-Tree)', International Conference on Mathematical Computer Engineering-2013 (ICMCE-2013), pp. 180-187.
- [32] Jia-Ching Ying ., Huan-Sheng Chen ., Kawuu W. Lin ., Eric Hsueh-Chan Lu ., Vincent S. Tseng ., Huan-Wen Tsai ., Kuang Hung Cheng ., Shun-Chieh Lin., "Semantic trajectory-based high utility item recommendation system", Expert Systems with Applications, 41 (2014), pp. 4762–4776

## BIOGRAPHY



**Kavitha J.K.**, Assistant Professor. She received the M.Tech. Degree from the Department of Computer Science and Engineering at Anna University, Chennai, in 2009. She is now a Ph.D. candidate in the Department of Computer Science and Engineering at Anna University, Chennai. Her research interests include Data mining, Image Processing, Cloud Computing.



**Manjula D.**, Associate Professor. She received her B.E Electronics & Communications Engineering degree in 1983 from Thiagarajar College of Engineering, Madurai and M.E Computer Science & Engineering in 1987 and Ph.D (Computer Science & Engineering) in 2004 from Anna University Chennai. Her research interests include Data mining, Image Processing, Cloud Computing, Network security.