

# Discovering Sequential Rental Patterns by Fleet Tracking

Xinxin Jiang, Xueping Peng, and Guodong Long

Quantum Computation and Intelligent Systems  
University of Technology Sydney, Australia

Xinxin.Jiang@student.uts.edu.au,  
{Xueping.Peng, Guodong.Long}@uts.edu.au

**Abstract.** As one of the most well-known methods on customer analysis, sequential pattern mining generally focuses on customer business transactions to discover their behaviors. However in the real-world rental industry, behaviors are usually linked to other factors in terms of actual equipment circumstance. Fleet tracking factors, such as location and usage, have been widely considered as important features to improve work performance and predict customer preferences. In this paper, we propose an innovative sequential pattern mining method to discover rental patterns by combining business transactions with the fleet tracking factors. A novel sequential pattern mining framework is designed to detect the effective items by utilizing both business transactions and fleet tracking information. Experimental results on real datasets testify the effectiveness of our approach.

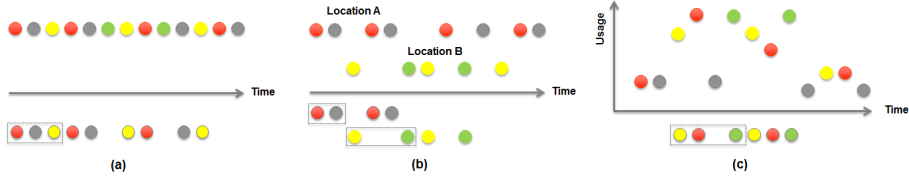
**Keywords:** Sequential Pattern Mining, Fleet Tracking, Item Detection;

## 1 Introduction

In the rental business, the quality of services and products depends on how successful we are in satisfying the customers need for service, and the ability to get the equipment to customer and where they need it [1]. Business transactions and fleet tracking information, such as working location and operating hours, have been widely recognized as important features of improving work performance, predicting consumer preferences, and increasing business competition [2][3].

Sequential pattern mining, as one of the most well-known method on customer behavior analysis, finds out frequent subsequences as patterns in a given sequence database [4][5][6]. For sequential pattern mining in rental industry, combining it with fleet tracking information can efficiently identify high utility items and focus on more relevant circumstance.

As shown in Fig.1. (a) General sequential pattern mining, which compose items into a sequence by different customer. The rental pattern discovered in rectangle (red, grey, yellow) is the highest frequency, which has ignored the actual equipment circumstances in customer sites. For example, the rental fleets of one customer might be worked for two different work sites. It might decrease the



**Fig. 1.** Three sequential pattern mining approaches (a) General; (b) Location-based (c) Usage-based

pattern accuracy to put them into one sequence. (b) Location-based sequential pattern mining, describes this case. The sequence is divided into two different ones by location A and B. Two rental patterns of location A (red, grey) and location B (yellow, green) replace the previous traditional pattern with more accuracy. (c) Usage-based sequential pattern mining, analyzes equipment usage in time frequency, e.g. daily. Detecting high utility items in usage-time figure can help find out more accurate pattern (yellow, red, green) in rental behavior prediction.

To implement high efficient sequential pattern mining by fleet tracking in rental industry, we provide a novel framework to conduct rental pattern mining with proposed modeling algorithms. We propose two sequential rental pattern mining algorithms to identify frequent itemsets by utilizing fleet tracking information as location and usage. Further, our approach is demonstrated by real-world datasets in an rental industry case study. Specifically, we make the following contributions in this paper: 1) Analyzed the problems of discovering sequential rental pattern by fleet tracking; 2) Proposed sequential pattern mining framework to discover efficient rental pattern; 3) Provided fleet tracking data and algorithms to compose location-based and usage-based sequences; 4) Experiments on real-world industry dataset testify the effectiveness of the approach.

The paper is organized as follows. Section 2 reviews the related work. Section 3 proposes a sequential rental pattern mining framework. Section 4 details the approach and algorithm. Experimental results are presented in Section 5. Section 6 concludes the work.

## 2 Related Work

In the analyses of customer purchase behavior, sequential pattern mining with the algorithms such as SPADE [4], Prefixspan [7] and SPAM [8], proposed on the support/frequency sequences. The algorithms of frequent sequences often result in many patterns being mined; most of them may be hardly understood by business, while those related to real business with frequencies lower than the given minimum support are ignored [6][9][10]. Much work has been done in the area of sequential pattern discovery and periodicity detection, multiple minimum supports utilized to enhance the performance of sequential pattern mining [11],

DFSP in biological sequences analysis [13], and utility measure to discover high utility patterns [14].

In the rental industry, in terms of problem definition and utilizing other information such as fleet tracking to discover high quality sequential rental patterns in the real-world cases[2][3], a major common shortcoming among previous work is losing the understanding of what customer actual needs to services and products.

### 3 Technical Preliminaries and Framework

#### 3.1 Basic Concepts and Definitions

**Sequential Pattern Mining** . Given a sequence database and a min-support threshold, the problem of sequential rental pattern mining is to find the complete set of sequential rental patterns in the IoT data. Let  $I = \{i_1, i_2, \dots, i_n\}$  be a set of all items, An itemset is a subset of items. A sequence is an ordered list of itemsets. A sequence  $s$  is denoted by  $\langle s_1 s_2 \dots s_n \rangle$ , where  $s_j$  is an itemset, i.e.,  $s_j \subseteq I$  for  $1 \leq j \leq l$ .  $s_j$  is also called an element of the sequence, and denoted as  $(x_1 x_2 \dots x_m)$ , where  $x_k$  is an item, i.e.,  $x_k \subseteq I$  for  $1 \leq k \leq m$ . The number of instances of items in a sequence is called the length of the sequence. A sequence with length  $l$  is called an  $l$ -sequence.

A sequence database  $S$  is a set of tuples  $\langle sid, s \rangle$ , where  $sid$  is a sequence-id and  $s$  is a sequence. A tuple  $\langle sid, s \rangle$  is said to contain a sequence  $\alpha$ , if  $\alpha$  is a subsequence of  $s$ , i.e.,  $\alpha \sqsubseteq s$ . The support of a sequence  $\alpha$  in a sequence database  $S$  is the number of tuples in the database containing  $\alpha$ , i.e.,  $support_s(\alpha) = |\langle sid, s \rangle | (\langle sid, s \rangle \in S) \wedge (\alpha \sqsubseteq s)|$ . Given a positive integer  $\xi$  as the support threshold, a sequence  $\alpha$  is called a sequential pattern in database  $S$  if the sequence is contained by at least  $\xi$  tuples in the database, i.e.,  $support_s(\alpha) \geq \xi$ . A sequential pattern with length  $l$  is called an  $l$ -pattern.

**Fleet Tracking** . With the integration of machines, sensors, information, software instruction, and communications technologies, fleet tracking creates connectivity between machines and business transactions [3]. In the rental industry, fleet tracking has attracted considerable interest as a way to evaluate how they do business, optimize services and bring value to their customers. By knowing the location or usage of every vehicle in a fleet, a company can manage their vehicles in a more efficient and effective manner. The standard fleet tracking features include: equipment location with date and time , the current total operating hours of the equipment with date and time, the quantity of fuel that was used by the equipment during the 24-hour period that ended at the specific date and time [2][3].

#### 3.2 Sequential Rental Pattern Mining Framework

The proposed framework combines sequential pattern mining techniques with fleet tracking information. A general framework is shown in the left of Fig.2

with four following steps: Data Acquisition, Item Detection, Pattern Mining, and Prediction. With the evaluation and updated model parameters, these steps can be re-executed to peruse higher performance in future prediction. All components described on the right of the graph, give more specific design on how to make sequential pattern mining with fleet tracking data in rental industry.

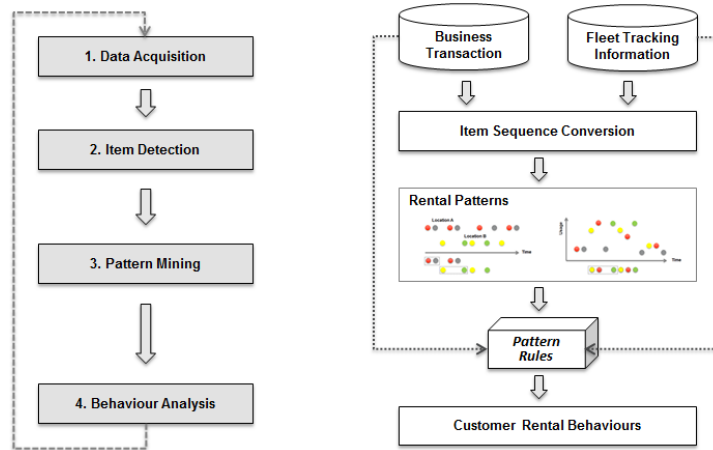


Fig. 2. Sequential rental pattern mining framework

In the framework above, the transaction dataset comes from business transaction and fleet tracking records. Item sequence conversion is designed to operate on business transaction with relevant tracking information. Each transaction is seen as a set of items (an itemset) with equipment tracking data. Given the threshold and parameters, the sequential rental pattern algorithms identify the itemsets which are subsets of the transactions in the database. The discovered rental patterns are stored as prediction rules in knowledgebase, which can be used to predict customer rental behaviors in future, with business transactions and fleet tracking information as inputting values.

## 4 Sequential Rental Pattern Mining

Sequential pattern mining, which discovers frequent subsequences as patterns in a sequence database, has proven to be a useful approach on handling order-based business problems as customer behavior analysis. The rental transactional data dealing with what the customers have bought in past will effectively reflects their behavior patterns around what they might rent in the future. To avoid the low efficient patterns from traditional sequential mining methods, our approach focusing on high performance events, utilizes both transactional data and fleet tracking information to get customer rental behaviors in real circumstance.

#### 4.1 Data Acquisition and Item Detection

The data sources of data acquisition include two parts: Business Transaction and Fleet Tracking Information. It is the process of sampling and converting real-world business transactions and equipment information into series database tables that can be manipulated by following item detection phase. Data acquisition applications are controlled by Extract, Transform and Load (ETL) programs developed using various general purpose programming languages.

**Item Sequence Detection by Location.** Traditional sequential rental item detection focuses on the rental behaviors of a customer, however in practical, the customer rental behaviors depend on special characteristics of each project running in different locations. Detecting location-based items will capture the special features on different worksite and then will effectively improve the prediction accuracy. The detecting results will be stored in a sequence database with customer id and project location.

**Item Sequence Detection with Usage.** Equipment utilization is the core of the equipment rental business. Usage information of the equipment is another important feature to identify high utility rental items. Analyzing customer behaviors on renting high utility equipment becomes hot requirements for many rental companies. Usage can be captured by fleet tracking that we described in Section 3.

#### 4.2 Pattern Mining and Behavior analysis

**Objective of Discovering Rental Patterns.** The aim of the approach is to discover high utility sequential rental patterns. We still need set up the detail objectives to define suitable parameter and fulfill the algorithm. In our case, by discussing with the equipment rental company, several objectives have been described for the project. The objectives are listed as follows: 1) To find out the products a customer prefers to hire sequentially and with high frequency, 2) To predict the product a customer might be interested to hire in the near future.

**Sequential Pattern Mining Algorithm.** Sequential pattern mining algorithm in sequential rental patterns discovering is described in the algorithm table. Given a sequence  $s = \langle s_1 s_2 \dots s_k \rangle$ , we denoted  $k$ -sequences as a sequence with  $k$  items.  $L_k$  is the set of frequent  $k$ -sequences, while  $C_k$  is the set of candidate  $k$ -sequences. Our goal is to generate a candidate set of all frequent  $k$ -sequences, given the set of all frequent  $(k - 1)$ -sequences.

**Algorithm begin**

1. Generate the candidate sequences in  $C_1$
  2. Save the frequent sequences in  $L_1$
- Iteratively find the sequences with  $k$ th pass:

3. Generate the candidate sequences in from the frequent sequences in  $L_{k-1}$ .

**Join Phase:** Join  $L_{k-1}$  with  $L_{k-1}$ .

if ( $s_1$  first item) is the same as ( $s_2$  last item),  $s_1$  join with  $s_2$ .

**Prune Phase:** Delete candidate sequences  $C_k$  that have a contiguous  $(k-1)$  subsequence whose support count is less than the minimum support.

Terminated until

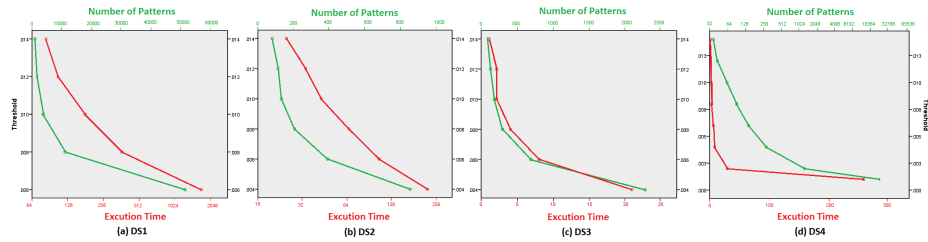
4. No more frequent sequences  $L_k$  are found. No candidate sequences  $C_k$  are generated.

**End**

**Behavior Analysis.** In the current setting learning is performed off-line. The result of learning is a set of Pattern Rules which give information what customer rental behavior might occur with some probability when certain preconditions are satisfied. On the other hand, Business Transaction and Fleet Tracking Information are used as input to the analysis algorithm as well. These rules and data sources are applied to analyze what future rental behaviors are likely to happen.

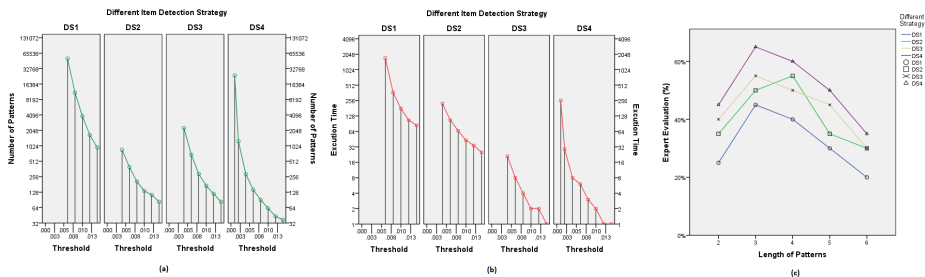
## 5 Testing and Results

We conduct intensive experiments on the real dataset from real-world rental industry, which holds 180,613 customer transactions with related fleet tracking from January 2014 to December 2014. The experiments evaluates the performance of Sequential Rental Pattern Mining in terms of computational cost, memory usage, number of patterns, and length of patterns on different item detection strategy from DS1 to DS4. In the experiments, DS1 focus on the rental transactions by customer only; DS2 uses the rental transactions by customer and fleet tracking with location; DS3 uses the rental transactions by customer and fleet tracking with usage; DS4 analyzes the patterns on the rental transactions by customer and fleet tracking with both location and usage.



**Fig. 3.** Evaluation of Execution Time and Number of Patterns on the Four Item Detection Strategies

The execution times of mining rental sequential patterns on DS1 to DS4 are shown in Figure 3; the figure also since we can obtain many more sequential



**Fig. 4.** Comparison of (a) Number of Patterns and (b) Execution Time on the Four Item Detection Strategies (c) Evaluation of Length of Patterns on the Four Item Detection Strategies

rental patterns. Especially for DS4, with the combination of fleet includes the number of patterns. When the minimum threshold decreases, more execution time is required tracking with both location and usage, the results show that it can extract more sequential rental patterns with less execution cost. The results also show that compared with DS1 without fleet tracking, all the other three strategies with fleet tracking have taken less execution time under same threshold (for instance, 1719 for DS1 and 106 for DS2). Figure 4 (a) and (b) show the effectiveness of the approaches with the fleet tracking. The execution time has dramatically decreased with fleet tracking approaches as DS2~DS4, and the number of patterns under suitable threshold still can keep on a high level (for instance, 1719 with above 25000 patterns for DS1 and 258 with 24196 patterns for DS4). The results also show that utilizing multiple fleet tracking features has the better performance than the simple fleet tracking approaches. Compared with the simple fleet tracking DS2 and DS3, the complex fleet tracking DS4 with both location and usage has discovered much more rental patterns with the reasonable execution time (for instance, 258 with 24196 patterns).

We validated the accuracy of discovering rental patterns on different strategies through the evaluation from domain experts. The experts from rental industry randomly chose groups of the patterns of length (2~6) from different strategy (DS1~DS4) and evaluated the patterns with their domain knowledge. Then they picked up the qualified patterns from each group. The value of expert evaluation means the percentage of qualified patterns in each group. In Figure 4(c), although the value of expert evaluation varied by length of patterns, the strategies with fleet tracking approaches show higher percentage of valuable patterns in every length group. The results show the effectiveness of the fleet tracking strategies.

## 6 Conclusion

Sequential rental pattern mining leads to patterns which discover the customers need for service and with the fleet tracking features we can know about the detail

on when get the equipment to customer and where they need it. In this paper, we have proposed a novel framework on how to discover sequential rental patterns by fleet tracking and testified its performance with different fleet tracking strategies in the rental industry. The results demonstrate the effectiveness of our approach.

**Acknowledgments.** This work is funded by the Australian Research Council Linkage grant (LP120100566).

## References

1. Rieser, C. *Service Success! Lessons from a Leader on How to Turn Around a Service Business*. John Wiley & Sons.(1994).
2. Trimble, T. E., and Bowman, D. S. *Market Guide to Fleet Telematics Services: Creating a Consumer's Guide to Currently Available Aftermarket Solutions*. (2012).
3. Andriesson, J. E., and Roe, R. A. (Eds.). *Telematics and work*. Psychology Press. (2013).
4. Zaki, M. J. SPADE: An efficient algorithm for mining frequent sequences. *Machine learning*, 42(1-2), 31-60. (2001).
5. Pinto, H., Han, J., Pei, J., Wang, K., Chen, Q., and Dayal, U. Multi-dimensional sequential pattern mining. In *Proceedings of the tenth international conference on Information and knowledge management* (pp. 81-88). ACM. (2001).
6. Mooney, C. H., and Roddick, J. F. Sequential pattern mining—approaches and algorithms. *ACM Computing Surveys (CSUR)*, 45(2), 19. (2013).
7. Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., and Hsu, M. C. Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In *2013 IEEE 29th International Conference on Data Engineering (ICDE)* (pp. 0215-0215). IEEE Computer Society. (2001).
8. Ayres, J., Flannick, J., Gehrke, J., and Yiu, T. Sequential pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 429-435). ACM. (2002).
9. Zhao, Q., and Bhowmick, S. S. *Sequential pattern mining: A survey*.ITechnical Report CAIS Nayang Technological University Singapore, 1-26. (2003).
10. Pei, J., Han, J., and Wang, W. Constraint-based sequential pattern mining: the pattern-growth methods. *Journal of Intelligent Information Systems*,28(2), 133-160. (2007).
11. Kumar, K. M., Srinivas, P. V. S., and Rao, C. R. Sequential pattern mining with multiple minimum supports by MS-SPADE. *International Journal of Computer Sciences*, 9(5). (2012).
12. Kumar, K. M., Srinivas, P. V. S., and Rao, C. R. Sequential Pattern Mining With Multiple Minimum Supports in Progressive Databases.*International Journal of Database Management Systems (IJDMS) Vol, 4*. (2012).
13. Liao, V. C. C., and Chen, M. S. DFSP: a Depth-First SPelling algorithm for sequential pattern mining of biological sequences. *Knowledge and information systems*, 38(3), 623-639. (2014).
14. Lan, G. C., Hong, T. P., Tseng, V. S., and Wang, S. L. Applying the maximum utility measure in high utility sequential pattern mining. *Expert Systems with Applications*, 41(11), 5071-5081. (2014).