中正嶺學報 第四十六卷 第一期 民國 106.6
JOURNAL OF C.C.I.T., VOL.46, NO.1, JUN., 2017

A Frequent Itemset Mining Algorithm for Solving Facility Partition Location Covering Problem

Chienwen Wu[1], Heng-Hsing Chu[2*], and You-Shyang Chen[3]

[1]*Department of Information and Finance Management, National Taipei University of Technology*

[2] *College of Management Ph.D. program in Management, National Taipei University of Technology*

[3]*Department of Information Management, Hwa Hsia University of Technology*

ABSTRACT

In this paper, we consider a location covering problem for a partitioned facility planning. Due to the size constraint of a space, a large facility cannot be well accommodated in a location and must be partitioned into several types of sub-facilities to set up in different locations. The different types of sub-facilities have to work collaborated to fulfill orders from all demands. This study proposes a mathematical model of addressing facility partition location covering problem (FPLCP) and uses the maximal frequent itemset algorithm (MAFIA) to solve this FPLCP. Experiments are performed to compare the solving efficiency with well-known CPLEX, and the empirical results show that the MAFIA gains more efficient than CPLEX.

Keywords: location covering, facility partition, frequent itemset mining (FIM), maximal frequent itemset algorithm (MAFIA)

頻繁項目集探勘演算法解決設施分置選址覆蓋問題

吳建文[1]　朱恒興[2*]　陳祐祥[3]

[1]台北科技大學資訊與財金管理系

[2]台北科技大學管理學院管理博士班

[3]華夏科技大學資訊管理系

摘　要

在本研究中，我們探討設施分割設置的選址覆蓋問題；該問題是由實際應用而啟發。在現實生活中，大型設施的設置將因為空間不足等條件限制，而無法完整設置在同一個位址，以致

文稿收件日期 105.11.10;文稿修正後接受日期 105.11.10; *通訊作者
Manuscript received November 10, 2016; revised November 10, 2016; * Corresponding author

必須被分割成為不同功能與性質的子設施設置，因此藉由設置不同功能與性質子設施，透過協同服務的方式最大化覆蓋所有顧客，使顧客能被完整的提供服務和滿足需求，成為最佳化設施設置空間規劃的重要策略。在本研究中，我們提出最大頻繁項目集探勘演算法(MAFIA)的數學模型以解決設施分置選址覆蓋問題；同時，我們運用此MAFIA與著名的CPLEX進行求解效能結果比較。實驗結果顯示，此頻繁項目集探勘演算法比使用CPLEX具有更佳的求解效率。

關鍵詞: 位址覆蓋, 設施分區，頻繁項目集探勘演算法，最大頻繁項目集探勘演算法

Ⅰ. INTRODUCTION

The location covering problems are important planning problems for a wide range of public and private facilities. In the location covering problems, a service of customer is considered covered if it is within some critical distances from a facility. The public location covering problem has many applications, such as fire stations, medical centers, and police stations. Traditional location covering problems assume that a facility can be well accommodated in one location [1-3]. However, in many real situations, a facility may be too large for a single location so that it is not accommodated in one location and has to be partitioned into several sub-facilities in order to conform the realistically environment constraints.

An example uses the medical facilities. A medical facility may need to contain several different sub-facilities, such as ophthalmology facility, otolaryngology facility, obstetrics and gynecology facility, pediatrics facility, and dentistry. Using different types of collaborative sub-facilities work can not only increase space utilization but also improve management performance. Another example is the warehousing facilities. The density of convenience stores in Taiwan is the highest in the world, and the relevant industries are also prospering continually. The convenience stores always emphasize on supplying fresh, timely, high-quality, and diversified products and services to satisfy the customer demands [4].

In consideration of the limited space of locating warehouses and management convenience, different types of sub-warehousing facilities are often established, such as frozen food, prepared food, daily commodity, beverage, books, magazines, and snacks warehousing facilities, to meet the demands of convenience stores and customers at different time points. In addition, in military deployment aspect, the military facility usually considers the mission implementation, and defense needs may need to partition and separately locate several different types of sub-facilities, such as command center, missile base, combat tank base, logistic facility, personal center, and medical station. However, especially in urban area, where the development density and land cost are very high, it is more difficult to find a location to accommodate all these sub-facilities than rural area. Therefore, proper planning of different types of facility locations can reduce the complexity of management and transportation costs and can achieve the goal of enhancing high standard of services.

In solving multi-objective optimization

problems, many kinds of exact algorithms in mathematical models can be used to solve these problems and obtain the optimal solution. The purpose to select an appropriate and efficient algorithm to solve these problems motivates this study. In this study, we understand a facility partition location covering problem (FPLCP) belonged to NP-hard problems [5-6], which can be addressed as the mixed integer linear programming (MILP) model. In the MILP and NP-hard problem, although to select most appropriate and efficient algorithm can always obtain the optimal solutions, the solving efficiency will affect the solution quality. When the problem becomes more complicated, the computing efficiency of the algorithm will become extremely important. This study further introduces an optimal facility location planning problem, which is according to the presetting candidate locations and constraints to generate the optimal feasible solution. Such a searching solution for logistic purpose is similar to the processes of data mining. Therefore, we use a maximal frequent itemset algorithm (MAFIA) [7], which is based on data of frequent itemset mining (FIM) to address the FPLCP, solve this problem, and compare the solving efficiency with a well-known CPLEX [8]. The CPLEX is a high computing performance decision planning tools, which is designed in achieving optimization model and applied in many fields with a satisfied result. To the best of our knowledge, such a study with same problem and same solving model has not been explored and reviewed until now from limited literature reviews.

The following Section 2 is the literature review. The mathematical model of addressing FPLCP used is discussed in Section 3. The description of FIM is presented in Section 4. The proposed approach to solve the FPLCP is present in Section 5. Section 6 analyzes the research results, and Section 7 is the conclusions of this study.

Ⅱ. LITERATURE REVIEW

The first facility location model discussed is a location set covering model (LSCM), which was provided by Toregas et al. [9]. The aim of LSCM is to minimize the number of ambulances required to cover all the demands in an area. Accordingly, the solution of maximal covering location problem (MCLP) was proposed by Church and ReVelle [4]. The MCLP solution was determined by a set of facility locations, which would maximize covering the demand points serviced by the facilities within a critical service criterion and was belonging to a NP-hard problem. The modeling MCLP is one of the most important positioning patterns and complies a non-polynomial function [10]. The MCLP model is usually applied in solving location planning problems regarding retailing facilities, emergency services, ambulances systems, stores of chains, cellular telephony antennae, etc. [11, 12]. Schilling et al. [13] and Farahani et al. [14] contributed reviews on covering models, classifications, analyzing, and applications over recent decades. In the past, many researches discussed the approaches, such as linear programming relaxation [4], greedy heuristics [15, 16], Lagrangean relaxation [17, 18], heuristic concentration (HC) [19], and Tabu search algorithms [20, 21], to solve MCLP. ReVelle et al. [19] proposed a two-step HC algorithm to solve the large-scale MCLP. Başar et al. [21] presented the multi-period maximal covering model, which was applied in emergency medical services system, and surveyed a case study in Istanbul, Turkey. Otto and Boysen [22] proposed dynamic programming-based heuristic algorithm to solve public transport service networks coverage location problems. Segura et al. [23] used the Lingo (2014 version) to solve the milk distribution of MCLP. Farahani et al. [24] presented a hybrid artificial bee colony (HABC) algorithm to solve the hierarchical MCLP. Pereira et al. [25] also presented a neighborhood search heuristic algorithm to solve probabilistic maximal covering location–allocation problem. He et al. [26] introduced a new heuristic mean-shift algorithm to solve large-scale planar MCLP. Murawski and Church [27] introduced a maximal covering network improvement problem, which was formulated as an integer-linear programming model and assumed that the demand points would be covered when the facilities are established in fixed location, and a case study in Ghana is surveyed. Curtin et al. [28] proposed a maximal covering location model to determine the distribution of police patrol areas and applied an example in Dallas.

Ⅲ. FORMULATION

In this instance of modeling FPLCP is represented by $<\boldsymbol{C}, \boldsymbol{F}, \boldsymbol{L}, \{\boldsymbol{d}_{ik}\}, \{\boldsymbol{U}_{ijk}\}, \boldsymbol{S}, \boldsymbol{x}_i, \boldsymbol{y}_j>$. Let $\boldsymbol{F} = \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_k\}$ be a set of partitioned $\boldsymbol{k}$ different service departments (e.g., sub-facilities) that provide an unique service, $\boldsymbol{L} = \{\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_m\}$ be a set of $\boldsymbol{m}$ possible locations for the facilities, and $\boldsymbol{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n\}$ be a set of $\boldsymbol{n}$ clients.

(1) Parameters

$\boldsymbol{w_i}$: The weight (or importance) of client $\boldsymbol{c}_i$ and $\boldsymbol{w}_i \geq 0$.

$\boldsymbol{d_{ij}}$: The distance between client $\boldsymbol{c}_i$ and location $\boldsymbol{l}_j$.

$\boldsymbol{S}$: the critical distance.

(2) Decision variables

$\boldsymbol{u_{ij_k}}$= 1 if the client $\boldsymbol{i}$ is assigned to the each type of facility $\boldsymbol{j}_k$; $\boldsymbol{u_{ij_k}}$ = 0 otherwise.

$\boldsymbol{x_i}$ = 1 if the client $\boldsymbol{c_i}$ is covered ($\boldsymbol{x_i}$ = 0 otherwise).

$\boldsymbol{y_j}$ = 1 if a service department (sub-facilities) is established at location $\boldsymbol{l}_j$; $\boldsymbol{y}_j$ = 0 otherwise.

In addition, we let $\boldsymbol{N}_i = \{\boldsymbol{l}_j \,|\, \boldsymbol{d}_{i,j} \leq \boldsymbol{S}\}$. The mathematical formulation of original MCLP is as follows:

$$\boldsymbol{Max} \sum_{i} \boldsymbol{w_i}\, \boldsymbol{x_i}$$

Subject to

$$\sum_{1 \leq j \leq m} \boldsymbol{y_j} = \boldsymbol{k} \qquad (\mathbf{1})$$

$$\sum_{j \in N_i} \boldsymbol{y_j} \geq \boldsymbol{x_i} \qquad (\mathbf{2})$$

$$\boldsymbol{x_i} \in \{\mathbf{0}, \mathbf{1}\}\ \boldsymbol{for}\ \mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{n} \qquad (\mathbf{3})$$

$$\boldsymbol{y_j} \in \{\mathbf{0}, \mathbf{1}\}\ \boldsymbol{for}\ \mathbf{1} \leq \boldsymbol{j} \leq \boldsymbol{m} \qquad (\mathbf{4})$$

Restriction (1) states that we want to locate $\boldsymbol{k}$ service departments (sub-facilities). Restriction (2) stipulates that when one or more facilities are established at sites in the set $\boldsymbol{N_i}$ (that is, one or more facilities are located within $\boldsymbol{S}$ distance units of demand point $\boldsymbol{i}$). Restrictions (3) and (4) define the 0-1 nature of the decision variables.

The mathematical formulation of the FPLCP is as follows:

$$\boldsymbol{Max\sum_{i} w_i x_i}$$

Subject to

$$\sum_{1\le j\le m} y_j = k \tag{5}$$

$$x_i + y_j \le 1 \quad if\ l_j \notin N_i \quad for\ i, j \tag{6}$$

$$\sum_{j_k} u_{ij_k} \le 1 \ \ for\ all\ i. \tag{7}$$

$$x_i \in \{0, 1\}\ for\ 1 \le i \le n \tag{8}$$

$$y_j \in \{0, 1\}\ for\ 1 \le j \le m \tag{9}$$

Restriction (5) also states that we want to locate k service departments (sub-facilities). Restriction (6) also stipulates that each client is well served if the client is covered by all the service departments (sub-facilities). Restriction (7) ensures that each client can be served by single sourcing facilities. Restrictions (8) and (9) also define the 0-1 nature of the decision variables. The objective of original MCLP model is to maximize the number of demands served or "covered" within the desired service distance. The traditional MCLP model is to covering the maximum possible demand with a known number of pre-setting candidate facilities within the covering distance. The FPLCP model means that the partitioned different types of facilities are all needed for each client and are not only covering the maximum possible demand with a known number of pre-setting candidate facilities locations within the covering distance but also working collaboratively to meet the various demands from clients. The critical distance was based on the reasons of considering transportation costs and timely management response. Thus, we assume that each sub-facility has the same service ability in critical distance. When the critical distance is greater, the service ranges of facilities and clients are broader. This study assumes that the ranges of the critical distance are within 2 and 3 km.

Ⅳ. DESCRIPTION OF FIM

The FIM-based algorithm can be used to solve the FPLCP. The purpose of addressing frequent itemset problem is to find subsets of items that occur together frequently in a large database [29]. The FIM-based algorithm is useful in solving many business decision-making problems, such as cross-selling, store layout, add-on sales, and customer segmentation. The related procedure of FIM is introduced in detail in the following statement.

Firstly, assume that $\boldsymbol{I} = \{\boldsymbol{i_1}, \boldsymbol{i_2}, \ldots, \boldsymbol{i_n}\}$ be a set of items. $\boldsymbol{X} \subseteq \boldsymbol{I}$ is called an itemset. $\boldsymbol{X}$ is called a $\boldsymbol{k}$-itemset if it contains $\boldsymbol{k}$ items and $|\boldsymbol{X}| = \boldsymbol{k}$. A transaction database $\boldsymbol{DB} = \{\boldsymbol{T_1}, \boldsymbol{T_2}, \ldots, \boldsymbol{T_n}\}$ is a set of transactions, where $\boldsymbol{T_j} \subseteq \boldsymbol{I}$. We say that a transaction $\boldsymbol{T_j}$ contains an itemset $\boldsymbol{X}$ if $\boldsymbol{X} \subseteq \boldsymbol{T_j}$. Given a transaction database $\boldsymbol{DB}$, the support of an itemset $\boldsymbol{X}$, denoted as $\boldsymbol{sup}(\boldsymbol{X})$, is the number of transactions in $\boldsymbol{DB}$ supporting $\boldsymbol{X}$. Let the ***min_sup*** be the minimum support specified by the decision-maker, and $\boldsymbol{X}$ is a frequent itemset in $\boldsymbol{DB}$ if $\boldsymbol{sup}(\boldsymbol{X}) \ge$ ***min_sup***. The set of all frequent itemsets can be denoted as FI. If it is not a subset of any other frequent itemset, the set of all maximal frequent itemsets is denoted as MFI. There are many efficient algorithms reported for the frequent itemset problem. Frequent itemset was generated by candidate set and further generated the maximum frequent itemset. Figure 1 shows relationship among the candidate set, frequent itemset, and maximal frequent itemset. Furthermore, MAFIA means that an itemset is maximal frequent if none of its immediate supersets is frequent. The MAFIA provides a compact representation of frequent itemset, mines a superset of the MFI, and utilizes pruning strategies to eliminate non-maximal patterns and remove non-maximal sets. In particular, when the itemset is very long, MAFIA can always obtain the high efficiency in mining data.

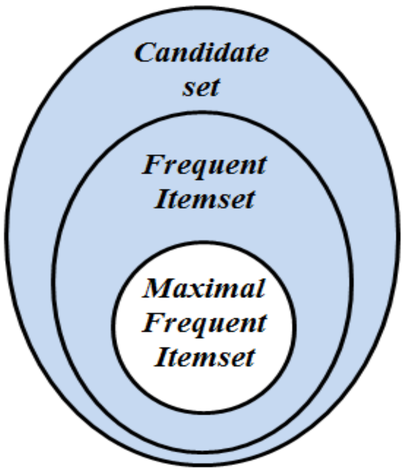

Fig.1. Relationship among candidate set, frequent itemset, and maximal frequent itemset

Due to the importance of these problems, several FIM-based algorithms, such as Apriori [29], FP-growth [30], and Smart Miner [31], have been successfully reported with a good performance. In addition, sophisticated commercial software package for addressing FIM has been developed and is available in the solution market.

V. THE PROPOSED APPROACH

Given a FPLCP instance, the proposed approach is first used to construct an instance of addressing FIM problem $\boldsymbol{a} = <\boldsymbol{C}, \boldsymbol{F}, \boldsymbol{L}, \{\boldsymbol{d}_{ij}\}, \{\boldsymbol{u}_{ijk}\}, \boldsymbol{S}>$ and $\boldsymbol{b} = <\boldsymbol{I}, \boldsymbol{TDB}, \boldsymbol{r}>$. Let $\boldsymbol{I} = \boldsymbol{L}$, $\boldsymbol{TID} = \boldsymbol{C}$ and $\boldsymbol{TDB} = \{<\boldsymbol{c}_i, \boldsymbol{N}_i> \mid \boldsymbol{c}_i \in \boldsymbol{C}\}$. We say that $\boldsymbol{Y}$ is an itemset if $\boldsymbol{Y} \subseteq \boldsymbol{I} = \boldsymbol{L}$. Given a transaction database $\boldsymbol{TDB}$, the support set of an itemset $\boldsymbol{Y}$, denoted as $\boldsymbol{SUP}(\boldsymbol{Y}, \boldsymbol{TDB})$, is defined as the set, containing the transaction identifiers of all the transactions in $\boldsymbol{TDB}$, which support $\boldsymbol{Y}$.

Formally, Let $\boldsymbol{X} = \boldsymbol{SUP}(\boldsymbol{Y}, \boldsymbol{TDB}) = \{\boldsymbol{c}_i \mid \boldsymbol{Y} \subseteq \boldsymbol{N}_i\}$. Suppose we give $\boldsymbol{Y} \in \boldsymbol{L}$ and a sequence $\{\boldsymbol{y}_j\}$ such that $\boldsymbol{y}_j = 1$ if $\boldsymbol{l}_j \in \boldsymbol{Y}$ or $\boldsymbol{y}_i = 0$ otherwise. Let $\{\boldsymbol{x}_i\}$ be the sequence such that $\boldsymbol{x}_i = 1$ if $\boldsymbol{c}_i \in \boldsymbol{X}$ or $\boldsymbol{x}_i = 0$ otherwise.

Theorem 1: $\boldsymbol{Y}$ is a $\boldsymbol{k}$-itemset if and only if Restriction (7) in the FPLCP is satisfied.

Proof:

$|\boldsymbol{Y}| = \boldsymbol{k}$.

$$\Leftrightarrow \sum_{1 \le j \le m} y_j = k$$

$\Leftrightarrow$ Restriction (7) is satisfied.

Theorem 2: $\boldsymbol{X} = \boldsymbol{SUP}(\boldsymbol{Y}, \boldsymbol{TDB})$ if and only if Restriction (8) is satisfied.

Proof:

$\boldsymbol{X} = \boldsymbol{SUP}(\boldsymbol{Y}, \boldsymbol{TDB})$

$\Leftrightarrow$ if $\boldsymbol{c}_i \in \boldsymbol{X}$, $\boldsymbol{Y} \subseteq \boldsymbol{N}_i$ for $1 \le \boldsymbol{i} \le \boldsymbol{n}$.

$\Leftrightarrow$ if $\boldsymbol{c}_i \in \boldsymbol{X}$ and $\boldsymbol{l}_j \in \boldsymbol{Y}$,
$\boldsymbol{l}_j \in \boldsymbol{N}_i$ for $1 \le \boldsymbol{i} \le \boldsymbol{n}$ and $1 \le \boldsymbol{j} \le \boldsymbol{m}$.

$\Leftrightarrow$ if $\boldsymbol{x}_i = 1$ and $\boldsymbol{y}_j = 1$,
$\boldsymbol{l}_j \in \boldsymbol{N}_i$ for $1 \le \boldsymbol{i} \le \boldsymbol{n}$ and $1 \le \boldsymbol{j} \le \boldsymbol{m}$.

$\Leftrightarrow$ $\boldsymbol{x}_i + \boldsymbol{y}_j \le 1$ if $\boldsymbol{l}_j \notin \boldsymbol{N}_i$ for $1 \le \boldsymbol{i} \le \boldsymbol{n}$
and $1 \le \boldsymbol{j} \le \boldsymbol{m}$.

$\Leftrightarrow$ Restriction (8) is satisfied.

Theorem 3: $\boldsymbol{Y}$ is a $\boldsymbol{k}$-itemset if and only if Restriction (9) is satisfied.

Proof:

$|\boldsymbol{Y}| = \boldsymbol{j_k}$.

$$\Leftrightarrow \sum_{j_k} \boldsymbol{u_{ij_k}} \leq \mathbf{1}$$

$\Leftrightarrow$ Restriction (9) is satisfied.

Based on the above three theorems, the proposed approach for solving the FPLCP problem is shown in detail below.

Step 1. Given a FPLCP, we construct the corresponding FIM instance as statements mentioned-above.

Step 2. Use any existing-based algorithm to discover all the frequent itemsets of the FIM problem.

Step 3. Let $\boldsymbol{F}$ be the set of all the discovered frequent itemsets and select $\boldsymbol{S}$ as the itemset in $\boldsymbol{F}$ that has the best support.

Step 4. Present improvement rate compared with CPLEX.

Ⅵ. EXPERIMENTAL RESULTS

In this section, we report the experimental results on the performance of the MAFIA and the proposed approach. The parameters include the numbers of facilities and critical distances. We consider a location space of 100 locations organized as a 10 × 10 grid. Each corner point in the grid is associated with a location. The distance between two locations is measured as the Euclidean distance of the corresponding corner points. In the experiments, the numbers of facilities were assumed to be 5 and 10. The impact coverage distance was assumed to be 2 to 4 when the numbers of facilities were 5 and 10. Populations ($\boldsymbol{w_i}$) on the nodes were randomly generated and were uniformly random in [0, 100]. The experiments were performed on a 1.2 GHz PC with 2 GB of memory running Windows 10. For our proposed approach, we employed a version of the MAFIA for FIM to compare with CPLEX. The CPLEX is a widespread used mathematical tool in solving the linear programming problem, mixed-integer linear programming problem, quadratic programming problem, and many other optimization problems. We also used Visual C++ and the CPLEX callable library to realize our mathematical model for the comparison purpose. Through a series of experimental computation, we obtain evaluations of effectiveness. The experimental results are shown in Table 1 below.

Table 1. Summary of experiment results

$\boldsymbol{n}$	$\boldsymbol{q}$	$\boldsymbol{S}$	CPLEX	MAFIA	Improvement
			Time	Time	rate
100	5	2	85.174	0.011	**7,742.09**%[+]
100	5	2.25	142.819	0.011	12,982.55%
100	5	2.5	176.369	0.012	14,696.42%
100	5	2.75	239.621	0.014	17,114.79%
100	5	3	352.374	0.016	**22,022.38**%
100	10	3	137.382	0.113	**1,214.77**%
100	10	3.25	149.534	0.122	1,224.69%
100	10	3.5	166.720	0.131	1,271.67%
100	10	3.75	226.141	0.175	1,291.23%
100	10	4	278.569	0.213	**1,306.84**%
Average			195.470	0.082	**8,086.74**%

Note: [+]improvement rate (%) = (85.174 – 0.011) / 0.011 = 7,742.09%

Figs. 2-5 show the improvement rates when computing in the location scopes of 100 grids, the numbers of facilities are 5 and 10, and the critical distance is 2 to 3, respectively.

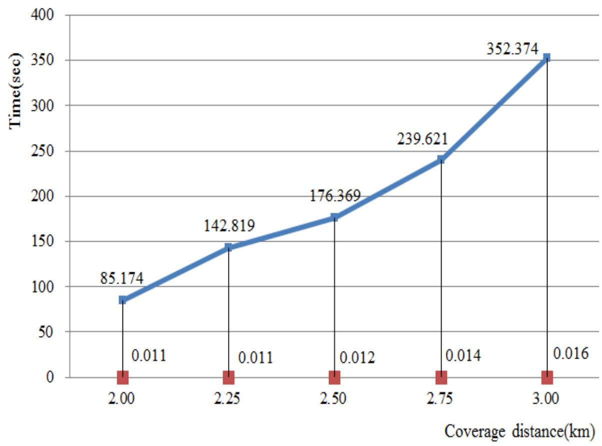

Fig.2. The computing time when numbers of facilities are 5, and coverage distance is 2 to 3

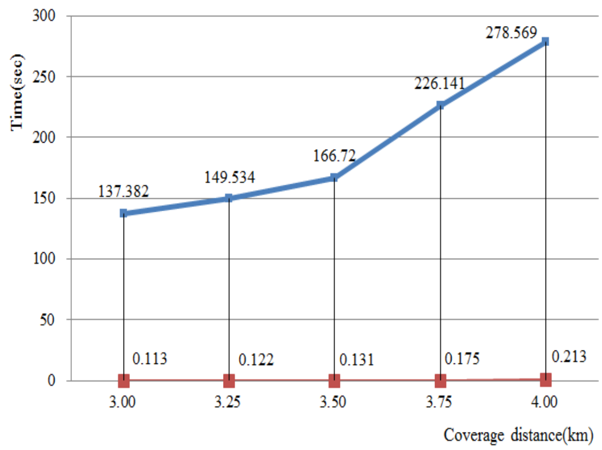

Fig.3. The computing time when numbers of facilities are 10, and coverage distance is 3 to 4

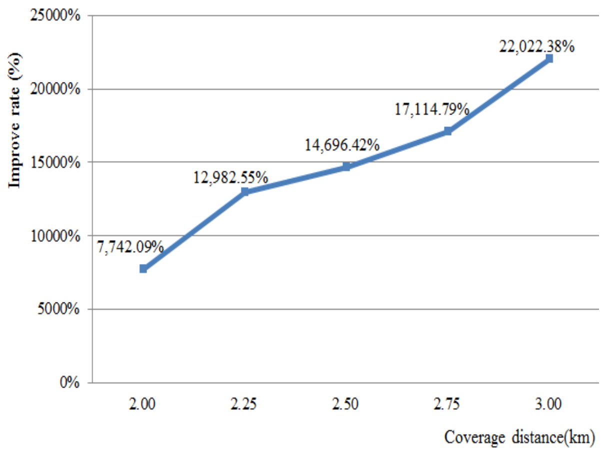

Fig.4. Performance results when the numbers of facilities are 5, and coverage distance is 2 to 3

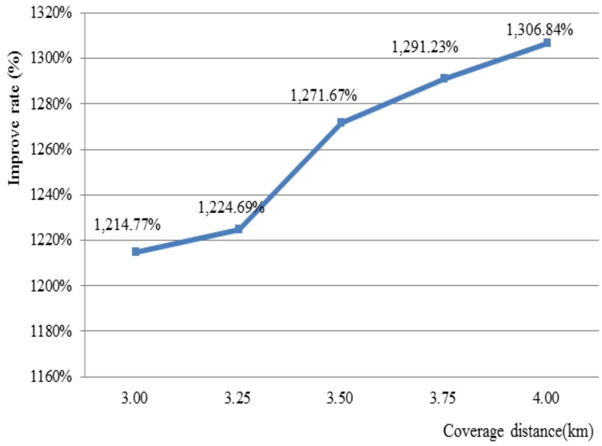

Fig.5. Performance results when the numbers of facilities are 10, and coverage distance is 3 to 4

According to Table 1 and Figs. 2-5, the experimental results show that the MAFIA yields better performance in terms of computational time to obtain optimal results in the 100 candidate locations than CPLEX model. Regarding the experiments, the analytical results are described in detail, as follows:

(1) In the experiment of the setting (***q*** = 5 and ***S*** = 2): The computational time for CPLEX is 85.174 seconds, and the MAFIA is only 0.011 seconds with the improvement rates of 7,742.09%.

(2) In the experiment of the setting (***q*** = 5 and ***S*** = 3): The computational time for CPLEX is 352.374 seconds, and the MAFIA is only 0.016 seconds with the improvement rates of 22,022.38%.

(3) In the experiment of the setting (***q*** = 10 and ***S*** = 3): The computational time for CPLEX is 137.382 seconds, and the MAFIA is only 0.113 seconds with the improvement rates of 1,214.77%.

(4) In the experiment of the setting (***q*** = 10 and ***S*** = 4): The computational time for CPLEX is 278.569 seconds, and the MAFIA is only 0.213 seconds with the improvement rates

of 1,306.84%.

Conclusively, this study learns that in all the experimental operations, the MAFIA has superior computational efficiency in comparison with CPLEX, and the improvement rates have significantly increased. The greatest improvement rate is 22,022.38%. When compared with CPLEX, the overall average optimal improvement rate is obtained from the calculations of the MAFIA 8,086.74%. The average optimal improvement ratio is 14,911.65% when the number of preset candidate locations is 100, the number of facility $\boldsymbol{q}$ is 5, and the critical distance $\boldsymbol{S}$ is 2 to 3. The average optimal improvement ratio is 1,261.84% when the number of preset candidate locations is 100, the number of facility $\boldsymbol{q}$ is 10, and the critical distance $\boldsymbol{S}$ is 3 to 4. To compare the Figures 2 and 3 can apparently find each step of the assumed facilities number when the critical distance becomes greater, and the effects of time improvement are greater. According to the results of comparisons for each experiment, the MAFIA always gains an increased performance when candidate locations and conditions are increased and complex. The performance of MAFIA is more efficient than that of the CPLEX approach.

VII. CONCLUSIONS

In this paper, we provided a mathematical model of MAFIA for solving a FPLCP. The MAFIA is based on FIM data method. Experiments were performed, and the analytical results show that the overall average optimal improvement rate obtained from the calculations through the MAFIA is 8,125.04%. When first time running the MAFIA, the frequent itemset is needed to store in memory, and then the subsequent improvement rates sequentially gain higher than CPLEX. When the number of facility is fewer, and the critical distance is shorter, the computing results show that the improvement rate has more apparent improvement effects. When the number of facility and the critical distance are increased, the experimental results show that winder of overall collaborative services range is offered. The experimental results also show that when the candidate locations and conditions are increased, the MAFIA can gain better performance than CPLEX. A crucial aspect of facility location is prompt supply services and convenient management. With the prompt supply services and convenient management, locating and tracking a given service within a facility becomes possible to facilitate quick and accurate supply services for order fulfillment. The multiple products and services can be collaboratively supplied to fulfill multiple orders. The MAFIA could efficiently solve the facility location optimization problems by expected spatial distribution planning and help the manager make decision exactly. Therefore, this study proposes a suitable model for effectively solving FPLCP that involves the practical problem in the real situation helpful to enterprises in grasping the trend of business, enhancing management performance, and gaining better efficiency to resolve important facility location and supply issues in the real life.

Further work of research on this topic can be addressed in the extension of the computational experiments on the facility location problems that may focus on the development of heuristic methods in order to response with very large scale instances of the problems considered. Moreover, the MAFIA can even be combined with the concept of shared space design to have more diversified and effective utilizations under limited spaces in the future.

VIII. CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests regarding the publication of this article.

REFERENCES

[1] Brandeau, M. L., & Chiu, S. S., "An overview of representative problems in location research," Management science, Vol. 35, No. 6, pp. 645-674, 1989.

[2] ReVelle, C. S., & Eiselt, H. A., "Location analysis: A synthesis and survey," European Journal of Operational Research, Vol. 165, No. 1, pp. 1-19, 2005.

[3] Church, R., & Velle, C. R., "The maximal covering location problem," Papers in regional science, Vol. 32, No. 1, pp. 101-118, 1974.

[4] Lee, W. I., Shih, B. Y., & Chen, C. Y., "A hybrid artificial intelligence sales - forecasting system in the convenience store industry," Human Factors and Ergonomics in Manufacturing & Service Industries, Vol. 22, No. 3, pp. 188-196, 2012.

[5] Kallrath, J., "Planning and scheduling in the process industry," OR spectrum, Vol. 24, No. 3, pp. 219-250, 2002.

[6] Kallrath, J., "Solving planning and design problems in the process industry using mixed integer and global optimization," Annals of Operations Research, Vol. 140, No. 1, pp. 339-373, 2005.

[7] Burdick, D., Calimlim, M., Flannick, J., Gehrke, J., & Yiu, T., "MAFIA: A maximal frequent itemset algorithm," IEEE transactions on knowledge and data engineering, Vol. 17, No. 11, pp. 1490-1504, 2005.

[8] IBM Inc., IBM ILOG CPLEX optimization studio getting started with CPLEX, Retrieved from http://cedric.cnam.fr/~lamberta/MPRO/ECMA/doc/Interface.pdf on March 01, 2016.

[9] Toregas, C., Swain, R., ReVelle, C., & Bergman, L., "The location of emergency service facilities," Operations Research, Vol. 19, No. 6, pp. 1363-1373, 1971.

[10] ReVelle, C., & Hogan, K., "The maximum availability location problem," Transportation Science, Vol. 23, No. 3, pp. 192-200, 1989.

[11] Serra, D., & Marianov, V., "New trends in public facility location modeling," UPF Economics and Business Working Paper, No. 755, 2004.

[12] Erdemir, E. T., Batta, R., Rogerson, P. A., Blatt, A., & Flanigan, M., "Joint ground and air emergency medical services coverage models: A greedy heuristic solution approach," European Journal of Operational Research, Vol. 207, No. 2, pp. 736-749, 2010.

[13] Schilling, D. A., Jayaraman, V., & Barkhi, R., "A review of covering problems in

facility location," Computers & Operations Research, Vol. 1, No. 1, pp. 25-55, 1993.

[14] Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., & Goh, M., "Covering problems in facility location: A review," Computers & Industrial Engineering, Vol. 62, No. 1, pp. 368-407, 2012.

[15] Daskin, M. S., Network and discrete location: models, algorithms, and applications, John Wiley & Sons, Chap. 5, pp. 23-29, 2013.

[16] Erdemir, E. T., Batta, R., Rogerson, P. A., Blatt, A., & Flanigan, M., "Joint ground and air emergency medical services coverage models: A greedy heuristic solution approach," European Journal of Operational Research, Vol. 207, No. 2, pp. 736-749, 2010.

[17] Galvão, R. D., & ReVelle, C., "A Lagrangean heuristic for the maximal covering location problem," European Journal of Operational Research, Vol. 88, No. 1, pp. 114-123, 1996.

[18] Senne, E. L. F., Pereira, M. A., & Lorena, L. A. N., "A decomposition heuristic for the maximal covering location problem," Advances in Operations Research, Article ID 120756, pp. 1-12, 2010.

[19] ReVelle, C., Scholssberg, M., & Williams, J., "Solving the maximal covering location problem with heuristic concentration," Computers & Operations Research, Vol. 35, No. 2, pp. 427-435, 2008.

[20] Rajagopolan, H. K., Saydam, C., & Xiao, J., "A multi-period expected covering location model: formulation, heuristic solution and application," Archives of Transport, Vol. 20, Iss. 1-2, pp. 113-131, 2008.

[21] Başar, A., Çatay, B., & Ünlüyurt, T., "A multi-period double coverage approach for locating the emergency medical service stations in Istanbul," Journal of the Operational Research Society, Vol. 62, No. 4, pp. 627-637, 2011.

[22] Otto, B., & Boysen, N., "A dynamic programming based heuristic for locating stops in public transportation networks," Computers & Industrial Engineering, Vol. 78, pp. 163-174, 2014.

[23] Segura, E., Benítez, R. B. C., Lozano, A., & Flores, I., "Optimization of Milk Distribution for Maximum Demand Coverage in Chihuahua," Mexico. Procedia-Social and Behavioral Sciences, Vol. 160, pp. 519-528, 2014.

[24] Farahani, R. Z., Hassani, A., Mousavi, S. M., & Baygi, M. B., "A hybrid artificial bee colony for disruption in a hierarchical maximal covering location problem," Computers & Industrial Engineering, Vol. 75, pp. 129-141, 2014.

[25] Pereira, M. A., Coelho, L. C., Lorena, L. A., & De Souza, L. C., "A hybrid method for the probabilistic maximal covering location–allocation problem," Computers & Operations Research, Vol. 57, pp. 51-59, 2015.

[26] He, Z., Fan, B., Cheng, T. C. E., Wang, S. Y., & Tan, C. H., "A mean-shift algorithm for large-scale planar maximal covering location problems," European Journal of Operational Research, Vol. 250, No. 1, pp. 65-76, 2016.

[27] Murawski, L., & Church, R. L., "Improving accessibility to rural health services: The maximal covering network improvement problem," Socio-Economic Planning Sciences, Vol. 43, No. 2, pp. 102-110, 2009.

[28] Curtin, K. M., Hayslett-McCall, K., & Qiu, F., "Determining optimal police patrol areas with maximal covering and backup covering location models," Networks and Spatial Economics, Vol. 10, No. 1, pp. 125-145, 2010.

[29] Jin, R., & Agrawal, G., "An efficient implementation of apriori association mining on cluster of SMPs," In Proceedings of the workshop on High Performance Data Mining, held with IPDPS 2001.

[30] Han, J., Pei, J., & Kamber, M., Data mining: concepts and techniques, Elsevier. Chap. 6, pp. 257-264, 2012.

[31] Agarwal, R. C., Aggarwal, C. C., & Prasad, V. V. V., "Depth first generation of long patterns," In Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 108-118, 2000.