Supply chain management using fp-growth algorithm for medicine distribution

A Wahana, D S Maylawati, M Irfan and H Effendy

Departement of Informatics, UIN Sunan Gunung Djati Bandung, Indonesia

Email: wahana.agung@gmail.com

Abstract. Distribution of drugs evenly in accordance with the needs of Public Health Center (Puskesmas) become one of the responsibilities by the Health Office in Indonesia. This study aims to provide recommendations for distribution of drugs from the Department of Health according to the needs of each Puskesmas. Because often the distribution of drugs is not in accordance with the needs of medicines stock of each Puskesmas. This causes the possibility of drug stock void in Puskesmas in need, while there is excess stock of drugs in Puskesmas that do not need. Supply Chain Management (SCM) is applied as a controlling drug stock at Puskesmas in order to avoid drug vacuum or excess drug that eventually unused. In addition, the Frequent Pattern Growth (FP-Growth) algorithm that generates frequent item sets is used to provide drug distribution recommendations by looking at the highest frequency of drug occurrences and frequent drug frequencies. Based on testing black box system conducted in Health Office Purwakarta regency of Indonesia which oversees 20 Puskesmas with 100 data of drug distribution transactions, it can be concluded that system functionality is running well and SCM successfully implemented to arrange distribution process of medicine well. Furthermore FP-Growth algorithm was able to provide recommendations for distribution of drugs with a high success rate. This is evidenced by the test results with various combinations of input parameters, FP-Growth is able to produce the right frequent item sets.

1. Introduction

The Health Office as a government apparatus is in charge of organizing, fostering and supervising the implementation of equitable, quality and affordable health efforts, mobilizing community participation in health development, and providing health facilities and resources accessible to all levels of society. Distribution of drugs evenly to the health center became one of the responsibilities of the Health Office. The process of managing and controlling incoming and outgoing drugs at the Puskesmas needs to be done automatically so that errors in managing the drug data. In addition, the Health Office can control the stock of drugs directly at each Puskesmas so that the distribution of drugs performed according to the needs of each Puskesmas.

Distribution and control of drugs needs to be managed properly, one of them by applying SCM. SCM differs from logistics management where logistics management is only part of SCM [1] [2]. More than simply controlling incoming and outgoing goods, SCM connects all business processes from the provider of goods to services and information to consumers and related stakeholders [1]. Along with its development, SCM has 8 processes within its framework, including Customer Relationship Management, Supplier Relationship Management, Customer Service Management, Demand Management, Order Fulfillment, Manufacturing Flow Management, Product Development and Commercialization, and Return Management [2].
In this research, SCM is used to manage all drug procurement activities among drug providers in this case the Health Office and Puskesmas as recipients of the drug. In addition to SCM, drug distribution control is assisted by track record of drug transactions at Puskesmas so that the distribution of drugs is evenly in accordance with the needs of each Puskesmas. This can streamline and streamline the distribution and control of drug stocks. The history of drug transactions is presented in the form of frequent item sets, where drugs as frequently occurring items can be recommended for the provision of medicinal stock from the Health Office to the Puskesmas. It is said to often arise when it meets the minimum value of support or the minimum value of occurrence of all transactions.

This research uses the FP-Growth algorithm as an algorithm for frequent pattern mining that generates frequent item set without the need to generate feature candidates [3] [4] [5] [6] [7]. FP-Growth is an Apriori algorithm development, where there is a difference in database scanning [5] [8]. FP-Growth only does one scanning database, while Apriori needs to do scanning database repeatedly. This is because the Apriori algorithm requires the generation of feature candidates to get frequent item sets [9] [3]. However, the FP-Growth algorithm does not generate feature candidates because FP-Growth uses the concept of tree development in the search for frequent item sets [5] [8]. This is what causes the FP-Growth algorithm faster than Apriori algorithm.

In the upcoming section, we describe the concept of Supply Chain Management for Medicine Distribution in Section 2, Frequent Pattern Growth Algorithm for Extracting Frequent Medicine Sets in Section 3. Experiment and Result in Section 4, then Evaluation and Discussion in Section 5. Finally, conclusion of this research stated in Section 6.

**Figure 1. Use Case Diagram for SCM of Drug Distribution using FP-Growth Algorithm**

### 2. Supply Chain Management for Medicine Distribution

SCM is wider than its logistics oriented product flow only [10]. SCM connects and coordinates every business process between both the supplier and the consumer or the organization itself. In this study, the Health Office is the supplier that distributes the drug, while the Puskesmas is the consumer receiving the drug. To meet the needs of drugs in Puskesmas as well as vice versa, both Health Office and Health Center must manage the flow of drug transactions and drug stock needs well. Based on the SCM process framework, SCM implementation in this study is modeled with the use case diagrams in Figure 1.
3. Frequent Pattern Growth Algorithm for Extracting Frequent Medicine Sets

Frequent pattern mining is a process or technique that generates frequent itemsets that do not pay attention to the order of occurrence of items on a transaction, in contrast to the sequential pattern that takes into account the order of occurrence of items on transactions [9] [3] [7]. FP-Growth is one of the algorithms that generate frequent itemset by compressing the database into a tree structure called FP-Tree [11] [12] [5]. There are several major steps in generating frequent itemsets without generating feature candidates, among others [6] [11] [13] [14] [15] [16]: Perform a database scan and find support from each item; Eliminates all items that are not frequent; Sort frequent items based on their support. Starting from the highest to lowest support; Any transactions that are read and saved in the tree; The first root is labeled with null, then the items of each transaction are stored in a node that contains item-name, count, and node-link; FP-Growth extracts frequent itemsets from FP-Tree; Read from leaves to root (bottom-up algorithm); Devide and conquer.

| ID | Item sets DB                                      |
|----|--------------------------------------------------|
| 1  | Klorfeniramin Maleat (CTM) Tablet 4 mg, Paracetamol Tablet 500 mg |
| 2  | Piridoksin HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg |
| 3  | Kotirimoksazol Syr, Klorfeniramin Maleat (CTM) Tablet 4 mg |
| 4  | Kotirimoksazol Syr, Parasetamol Tablet 500 mg   |

Table 1. An Example of Medicine Transactions

Figure 2 illustrates the FP-Tree and FP-Growth algorithms for generating frequent itemsets of drug transactions in Table 1. Figure 2 is a FP-Tree with a minimum value of 50% support (from 4 transactions appearing at least 2 times). Where, the frequency of occurrence of Klorfeniramin Maleat (CTM) Tablet 4 mg is 3, Paracetamol Tablet 500 mg is 3, and Srimrimazasazole Syr is 2, whereas Piridoksin HCl Tablet will be eliminated because it does not meet the minimum support with frekuendi occurrence once. After each item is sorted by the largest frequency, then the FP-Tree is established in Figure 2. Furthermore, from FP-Tree generated frequent itemset with conditional pattern base and conditional FP-Tree [5] [11] which is described in Table 2.

The emergence of frequent itemset drugs indicates that the drug is the most widely used and needed, of course every puskesmas has different needs. In addition, often concurrent rug ems indicate that these drugs are often needed imultaneously. For example, for a sample transaction on Table 1 and generating a frequent pattern <Paracetamol Tablet 500 mg, Klorfeniramine Maleat (CTM) Tablet 4 mg> indicates that people tend to use Paracetamol always together with CTM. So when it will distribute Paracetamol, it is necessary to check also the stock CTM, vice versa. This is because based on the history of Paracetamol transactions often appear simultaneously with CTM.
### Table 2. Extraction Processes of Frequent Itemset or Frequent Pattern from FP-Tree on Figure 2

| Item                        | Freq. | Conditional Pattern Base | Conditional FP-Tree | Frequent Pattern |
|-----------------------------|-------|---------------------------|--------------------|-----------------|
| Kotrimoksazol Syr           | 2     | {}                        | <>                 | [< Kotrimoksazol Syr >]: 2 |
| Paracetamol Tablet 500 mg   | 3     | { Klorfeniramin Maleat (CTM) Tablet 4 mg } : 2; { Kotrimoksazol Syr } : 1 | < Klorfeniramin Maleat (CTM) Tablet 4 mg > : 2 | [< Paracetamol Tablet 500 mg ] : 3; [< Paracetamol Tablet 500 mg , Klorfeniramin Maleat (CTM) Tablet 4 mg >] : 2 |
| Klorfeniramin Maleat (CTM) Tablet 4 mg | 3 | { Kotrimoksazol Syr } : 1 | <> | [< Klorfeniramin Maleat (CTM) Tablet 4 mg >] : 3 |

### 4. The experiment and result

#### Table 3. Scenarios and Black Box Test Result

| Scenario                                           | Result | Expected Output                                      | Real Output                                      |
|----------------------------------------------------|--------|------------------------------------------------------|--------------------------------------------------|
| Clear the form then submit                         | Success| The system gives an error message. Column can not be empty. | The system gives an error message. Column can not be empty. |
| Displays the highest frequency of drug use          | Success| The system successfully displays the highest drug frequency data. | The system successfully displays the highest drug frequency data. |
| Displays the highest frequency of drug use with minimum support | Success| The system successfully displays the highest drug frequency data with support ≥ minimum support that has been specified. | The system successfully displays the highest drug frequency data with support ≥ minimum support that has been specified. |
| Displays the highest frequency of drug use with minimum length | Success| The system succeeded in displaying the highest drug frequency data with a length of ≥ minimum length that has been specified. | The system succeeded in displaying the highest drug frequency data with a length of ≥ minimum length that has been specified. |

The test is done in black box and white box. Black box testing is done to test whether the SCM system is built in accordance with the specification of functional system requirement specification that is modeled on the use case diagrams in Figure 1. Table 3 shows the black box testing scenario and the result of testing each system functionality. While white box testing is done to see the accuracy and accuracy of FP-Growth algorithm in generating frequent item set from drug transactions. White box testing performed a test scenario to compare the results of the number of frequent item sets obtained by entering the minimum support varies with the same number of datasets. Testing is done by using 100 drug transaction data from Health Office Purwakarta regency of Indonesia which oversees 20 Puskesmas with minimum support and minimum length or minimum length of varied items. Table 4 to Table 7 show the result of frequent itemset produced by FP-Growth algorithm with variation of input parameter.

#### Table 4. Test Result on the Amount of frequent Itemset with Minimum Support ≥ 20% and Length = 1, 2, and 3

| No  | Frequent Itemsets                          | Support | Length |
|-----|-------------------------------------------|---------|--------|
| 1   | Paracetamol Tablet 500 mg                 | 69%     | 1      |
| 2   | Klorfeniramin Maleat (CTM) Tablet 4 mg    | 51%     | 1      |
| 3   | Klorfeniramin Maleat (CTM) Tablet 4 mg, Paracetamol Tablet 500 mg | 41% | 1 |
| 4   | Piridoksin HCl. Tablet                    | 37%     | 1      |
|     | Piridoksin HCl. Tablet, Paracetamol Tablet 500 mg | 36% | 1 |
|     | Giserial Guaikolat Tablet 100 mg          | 27%     | 1      |
| 7   | Giserial Guaikolat Tablet 100 mg, Paracetamol Tablet 500 mg | 25% | 1 |
| 8   | Piridoksin HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg | 25% | 1 |
| 9   | Piridoksin HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg, Paracetamol Tablet 500 mg | 25% | 1 |
| 10  | Giserial Guaikolat Tablet 100 mg, Klorfeniramin Maleat (CTM) Tablet 4 mg | 23% | 1 |
| No | Frequent Itemsets | Support | Length |
|----|------------------|---------|--------|
| 11 | Kotrimoksazol Syr | 23%     | 1      |
| 12 | Kotrimoksazol Syr, Parasetamol Tablet 500 mg | 21% | 1 |
| 13 | Kotrimoksazol Syr, Klorfeniramin Maleat (CTM) Tablet 4 mg | 20% | 1 |
| 14 | Antasida tab | 20% | 1 |
| 15 | Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg | 41% | 2 |
| 16 | Piridoks in HCl. Tablet, Parasetamol Tablet 500 mg | 36% | 2 |
| 17 | Gliseril Guaikolat Tablet 100 mg, Parasetamol Tablet 500 mg | 25% | 2 |
| 18 | Piridoks in HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg | 25% | 2 |
| 19 | Piridoks in HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg | 25% | 2 |
| 20 | Gliseril Guaikolat Tablet 100 mg, Klorfeniramin Maleat (CTM) Tablet 4 mg | 23% | 2 |
| 21 | Kotrimoksazol Syr | 23% | 2 |
| 22 | Kotrimoksazol Syr, Parasetamol Tablet 500 mg | 21% | 2 |
| 23 | Kotrimoksazol Syr, Klorfeniramin Maleat (CTM) Tablet 4 mg | 20% | 2 |
| 24 | Piridoks in HCl. Tablet, Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg | 25% | 3 |

Table 6. Test Result on the Amount of frequent Itemset with Minimum Support ≥ 40% and Length = 1 and 2

| No | Frequent Itemsets | Support | Length |
|----|------------------|---------|--------|
| 1  | Parasetamol Tablet 500 mg | 69%     | 1      |
| 2  | Klorfeniramin Maleat (CTM) Tablet 4 mg | 51% | 1 |
| 3  | Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg | 41% | 1 |
| 4  | Piridoks HCl. Tablet | 37% | 1 |
| 5  | Piridoks HCl. Tablet, Parasetamol Tablet 500 mg | 36% | 2 |
| 6  | Klorfeniramin Maleat (CTM) Tablet 4 mg, Parasetamol Tablet 500 mg | 41% | 2 |
| 7  | Piridoks HCl. Tablet, Parasetamol Tablet 500 mg | 36% | 2 |

Table 7. Test Result on the Amount of frequent Itemset with Minimum Support ≥ 50% and Minimum Length ≥ 1

| No | Frequent Itemsets | Support | Length |
|----|------------------|---------|--------|
| 1  | Parasetamol Tablet 500 mg | 69% | 1 |
| 2  | Klorfeniramin Maleat (CTM) Tablet 4 mg | 51% | 1 |

5. Evaluation and Discussion

Based on the SCM framework [2][10], each process is largely represented by each use case in this study. Customer Relationship Management, Supplier Relationship Management and Customer Service Management even though it is technically more done outside the system, but through the system automatically also woke up the communication between the Department of Health as a supplier and health center as customers were represented by the use case "Show Report of medicine Transaction" and "Create Report of Medicine Transaction". Especially the Demand Management, Order Fulfillment and Manufacturing Flow Management processes that are directly met by all use cases. Especially for Demand Management and Order Fulfillment facilitated by appropriate drug experts on the needs of each health center with implementing the FP-Growth algorithm that is modeled in use case "See Recommendation from the Highest Frequency" and "Running FP-Growth Algorithm and Generate frequent itemset". For Product Development and commercialization process, and Return Management is not directly represented by the use case is there, but the use case "Manage Medicine Transaction" there are controls on drugs that enter and exit, including the expired drugs and conditions can be restored if there is a mismatch.
Furthermore the FP-Growth algorithm has been implemented well in accordance with the literature [5] [6] [11]. Evident from the test results with several scenarios variation input parameters, including minimum support and minimum length, frequent item set of drugs produced in accordance with existing transaction data. In addition, the test results show the consistency of frequent item set produced. Seen in Table 4 to Table 7 although the minimum support and minimum length values entered are different (with the same transaction data), as long as the item is said to be frequent it will still produce the same frequent pattern. It also proves that function and structure of FP-Growth algorithm has been applied correctly.

6. Conclusion

The conclusion that can be drawn from this research is SCM can be applied to manage the whole process of distribution and management of drugs between the Health Office and Puskesmas. Much of this research fills the SCM framework by modeling and implementing it into system functionality. In addition, based on the test results indicate that the FP-Growth Algorithm can be used for the recommendation of drug needs planning in the Department of Health so that the drug distribution can be done effectively and efficiently according to the needs of each Puskesmas. For future research can be developed the process of SCM framework that has not been implemented, among others Product Development and Commercialization. Furthermore, in testing the algorithm needs more data because it can affect the accuracy of the test. In addition, it can also use a frequent pattern algorithm other than FP-Growth so that the formation process of frequent item set can be optimized in terms of time and effort required.

References

[1] Lambert D M and Cooper M C 2000 Issues in Supply Chain Management Industrial Marketing Management 29 pp 65-83.
[2] Lambert D M and Enz M G 2017 Issues in Supply Chain Management: Progress and potential Industrial Marketing Management 62 pp 1-16.
[3] Han J and Kamber M 2006 Data Mining: Concepts and Techniques (USA: Morgan Kaufmann Publisher).
[4] Han J, Cheng H, Xin D and Yan X 2007 Frequent Pattern Mining: Current Status and Future Data Mining Knowledge Discovery 15 pp 55-86.
[5] Jiawei H, Pei J and Yin Y 2000 Mining Frequent Patterns without Candidate Generation Proc. of the 2000 ACM SIGMOD Int. Conf. on Management of data.
[6] Maylawati D S 2015 Development of Pre-processing Library with Set of Frequent Word Item Set Representation for Text Mining Application Case Study: Indonesian’s Slang Language Thesis Report Bandung Institute of Technology.
[7] Maylawati and Saptawati G A P 2016 Set of Frequent Word Item sets as Feature Representation for Text with Indonesian Slang 1st Int. Conf. on Computing and Applied Informatics (ICCAI).
[8] Lin KC, Liao IE and C ZS 2011 an Improved Frequent Pattern Growth Method For Mining Association Rules Expert System with Applications 38 pp 5154-5161.
[9] Agrawal R and Srikant R 1995 Mining Sequential Pattern IBM Research Center.
[10] Christoper M 2011 Logistics and Supply Chain Management Fourth Edition (Edinburg: Prentice Hall).
[11] Pawar D 2013 The Comparative Study of Apriori and FP-Growth Algorithm for Association Rule Mining [Online] source: https://www.slideshare.net/deepti92pawar/the-comparative-study-of-apriori-and-fpgrowth-algorithm.
[12] Miraldi R N, Rahmat A and Susanto B 2014 Implementasi Algoritma FP-Growth untuk Sistem Rekomendasi Buku di Perpustakaan UKD Informatika, 10.
[13] Ardani N R and Fitriansi N 2016 Sistem Rekomendasi Pemesanan Sparepart dengan Algoritma FP-Growth (Studi Kasus PT Rosalia Surakarta) (Yogyakarta: STMIK AMIKOM Yogyakarta).
[14] Rtrianti 2014 Implementasi Algoritma FP-Growth pada Aplikasi Prediksi Persediaan Sepeda Motor (Studi Kasus PT Pilar Deli Labumas) Petita Informatika Budi Darmawan VI.
[15] Arfin R N 2014 Implementasi Algoritma Frequent Pattern Growth (FP-Growth) Menentukan Asosiasi Antar Produk (Studi Kasus Nadiamart) (Yogyakarta: STMIK AMIKOM Yogyakarta).
[16] Sari E W, Rianto A and Andarawarif S D 2014 Decision Support on Supply Chain Management System Using Apriori Data Mining Algorithm College of Informatics and Computer Engineering Indonesia (STIKI).