Date of publication July 31, 2022, date of current version Aug. 12, 2022.

Improving Customer Experience in Supermarkets: A New Approach based on Travel Path

RSN Dilrukshi [1#], HA Caldera [2]

[1,2]The University of Colombo, School of Computing, Sri Lanka
[1#] rsnirma@gmail.com

ABSTRACT In today's competitive market, understanding its customers is a key to the success of any business. The market contains various customer subgroups that can be distinguished based on purchasing habits, time spent, product selection, and travel path. To identify the pattern hidden inside these subgroups, is needed to use real data as it reflects the ordinary behaviour of the c ustomers. Analysis of the travel path data that customers make inside the shopping mall enables retailers to understand and predict customer behaviour, which has become a critical point in effective decision making for increasing sales with more customer comfort. Introducing the right discount for the right products acts as an important mediating factor in the customer relationship. Traditional methods of determining the discount and layout have dealt only with customer transactions, which have missed other important characteristics of customers' purchasing behaviour. This paper addresses the problem of sales increase based on personalized discount schemas and improved store layout using customers' shopping travel paths. It uses the Frequent Pattern Growth (FP Growth) algorithm to improve the sales and the RFM (Recency, Frequency, and Monitory value) analysis to identify the customer segments based on the dataset of Instacart from the Kaggle website. An FP growth algorithm has been used to identify the frequent locations and frequent products of a customer's purchases. An improved version of the supermarket layout has been suggested based on the frequent travel paths of customers. The findings of this approach can be used by retailers to improve the in-store shopping experience of customers.

INDEX TERMS: personalized discount, shopping path, travel path, supermarket layout

I INTRODUCTION

Shopping is one of the main investments of energy that people make for their benefits. As technology gets updated with new trends, retailers try to make their customers' shopping experiences more and more interesting and easy. In order to identify and understand the different needs of customers, customer segmentation techniques, which are commonly based on purchase behaviour, time spent on purchase, and purchase history, have been used. The findings have allowed to provide better services and preferable products to customers. These segmentation techniques allow to create profitable segments based on competitive advantages [8]. But identifying the segments based on suitable measures and determining the right marketing campaigns are challenging tasks for markets.

The term "competition" in retail refers to the rivalry among retailers who are keen to retain their customers and attract new ones. The most common and well-known method of improving sales and revenue for retailers with customer satisfaction is by providing discounts to customers. Discounts are typically assigned to items that are not frequently purchased by many customers, regardless of whether they are interested in them. Enabling a discount alone for non-frequent items and keeping them in the original location of the layout do not affect the improvement of the sales of that item. Product placement and layout also play a major role in improving sales in supermarkets because a typical customer decides their next return to the store based on their experience and impression of the current layout of the store. From the perspective of the retailer, the effectiveness of the layout of the store determines the level of the customer's exposure to the goods and affects the chance of the item being bought. The layout design has a direct impact on the travel paths of customers who are searching for their needs inside the supermarket. Therefore, the customer travel paths can be used as a measure to identify the customer's behaviour when buying products, which opens up a new way of increasing sales by analyzing the related data.

This paper proposes a novel approach to improve the sales of the supermarket based on the customer's travel path. A dynamic discount will be calculated for each customer based on their travel path, and a personalized supermarket layout will be suggested by placing the non-frequent sale items in frequent customer travel locations in order to improve the sales. The paper suggests the possibility of improvement in sales and design a personalized supermarket layout by applying the FP growth algorithm to the customer travel path data.

The remainder of the paper is organized as follows: section II provides discussions on shopping analysis of past literature, the proposed methodology is presented in section III , section IV bring out the result and generates a discussion on the findings and lastly, section V concludes the paper with conclusion and future work.

II LITERATURE REVIEW

Innovating marketing strategies are needed in order to create a long lasting and profitable relationship with customers in today's business world. The most salient issue in designing a marketing strategy is that the variation of customer needs from one to another is significantly different. As a result, customer segmentation based on their features and behaviours play a key role in business viability.

A *Customer Segmentation*

Customer segmentation has led to a deeper understanding of the customer's buying pattern [3]. It costs five times more to acquire a new customer than it does to keep an existing one, and ten times more to re-engage a dissatisfied customer [14]. As many companies focus more on improving their marketing strategies to enhance their market share, they primarily focus on customer segmentation. Dividing the customers into clusters based on their behaviour parameters can lead to a significant growth in their revenue [3]. Segmentation is a key area for retailers as it allows them to identify profitable groups and to organize marketing campaigns accordingly [8].

Considering the data of 250 bank customers, research found five different clusters, which are different based on factors such as loan amount, degree of loyalty, account balance, default risk, and profitability for the bank [3]. Findings suggest that customer clustering can help the financial sector because it augments their competitiveness to improve their marketing methods to target segment-based marketing approaches.

The RFM model has been widely used by researchers for customer segmentation. RFM analyses the behaviour of the customer based on Recency (R) – how recently a customer has made a purchase, Frequency (F) – How frequently a customer buys and Monetary (M) – How much money a customer spends on purchases [5] [7] [10].

A case study defines RFM as a three-dimensional way that is used for classifying or ranking customers, which is based on the 80/20 principle that 20% of customers bring 80% of the revenue of a company [1]. They used RFM to study the scoring of the active e-banking users. In this study, it used clustering as a technique for data mining and organized the findings into cluster groups based on the pyramid model. They used a two- step clustering method to identify the most important customers. The pyramid model has been useful for different businesses because it improves issues such as decision making, future revenue forecast, simulation of inactive customers, and prediction of alteration of the customer position in the pyramid [1].

B *Shopping path Analysis*

Retailers foremost concern is the shopping layout as it defines the way of customers expose to the products. According to a study, retailers try to expose many products through layout as it enables higher exposer rate, sales and return on investment [12]. Customers want store to create the layout in a way that minimizes this unwanted steps and motion in the shopping process [4]. Common layout design is based on the product category approach where products that share the same functional characteristics or origins are placed nearby. But according to this approach, it failed to respond the needs of the time pressured consumers of the shop. To satisfy the consumer, store sections need to be redesigned based on the consumer desire. Consumer's desire can be identified by analysing their demographic and transactional data. These data only provide the basic information such as quantity and purchase amount and how often they visited and etc. However, to understand consumers' behaviour completely, it needs more data and data mining techniques.

Based on the regional analysis on customer purchase history or recommending products through customer segmentation will not provide sufficient information to understand the customer shopping behaviour in the physical store [11]. To overcome these problems, they used longest common subsequence (LCS) which provide the capability of identifying hotspots where most of the customers visited and dead spots with few visits.

As a result of technological advancements in recent years; RFID, Bluetooth beacons, video cameras and Wi-Fi location tracking have been used for analysis of the customer behaviours inside the super markets. A research was done by Sorensen Associates based on a dataset that collected by tracking the travel paths of customers inside the actual supermarket in western United States [13]. For path tracking they used grocery carts which attached a RFID tag to the bottom. They categorized the path travel by each shopper using clustering algorithm and identifying 14 different canonical paths of it. Through the research they have identified that area at the entrance, end cap of aisles and checkout areas get more attention of the customer. Identification of the pattern in customer behaviour inside the supermarket will result to classify the frequent paths or locations travelled. Pattern discovery from a sequence of data is one of the main tasks in data mining research area. Extract hidden information from large database and then generate association between the items in it defined as association rule mining [2]. Market basket analysis is the common implementation in this method. Through this method it produces association rules which measures the dependency of each item in a dataset [9]. Association rule contain with mainly two parts: an antecedent(if) and a consequent (then). In order to identify the strength of an association rule mainly two measures 'Support' and 'Confidence' are used. Association rules were used to extract the knowledge from a transactional dataset which a shopping layout based on the association among the product categories is suggested [4]. They used Apriori algorithm with dimensionally reduction for identification of the association rules of transactional dataset. They allowed retailers to cluster the products based on the consumers buying habit and to create a strong appeal for the consumers' needs. As an example, rather than placing the coffee in the beverage section, ham in the meat section they suggested to place these products in breakfast food section which allocated for foods related to breakfast. In this paper the authors proposed a method of clustering the different product categories into same section based on the products that customers bought together.

The Frequent Pattern Growth algorithm is another algorithm that is used for frequent pattern mining [6]. By applying the FP growth algorithm in a step-by-step process, it removes unnecessary data and improves the performance of the overall process [2]. The generation of rules in FP Growth must be accompanied by a validation process to ensure that they are applicable and authentic.

By combining the advantages of FP growth with Apriori, an algorithm called Search Space Reduction (SSR) was introduced [15]. In SSR, it first scans the transactions once to count the support and identify frequently occurring items that are higher than the minimum support threshold. Then it generates an FP Tree using the identified frequent item sets. As a result, there is only one item-prefix tree in the memory at a time. In SSR, it uses the function Item-prefix-tree-construction for constructing item-prefix trees and it uses the function Frequent-pattern-generation for candidate generation and frequent pattern generation.

III METHODOLOGY

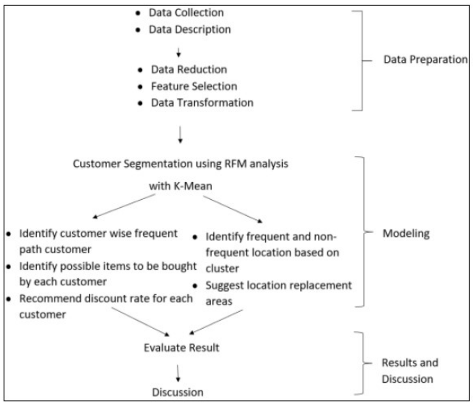

Figure 1. Work flow of the proposed approach

This paper proposed a solution to the above identified problem in the introduction section through RFM analysis and a frequent pattern growth algorithm. The proposed methodology includes three major phases: Data preparation, Modeling and Result and Discussion. Figure 1 shows a general model of how the data is acquired and data analysis is conducted in order to construct the proposed solution.

A Data Preparation

This study incorporates data regarding the customers' transactions and synthetically generated travel path data related to a dataset of Instacart from the Kaggle website. The dataset contained around 1 million grocery orders, which were placed by around 3500 users. It includes details based on the customers' transactions, products with assigned departments, aisles of products, and the shopping path of customers. The initial layout of the shopping mall is synthetically designed, and the layout is divided into sub partials with an assigned location code. Even though the original dataset contained around 3 million records, because of the hardware limitation, for the analysis, around 3000 users with their transitions were selected.

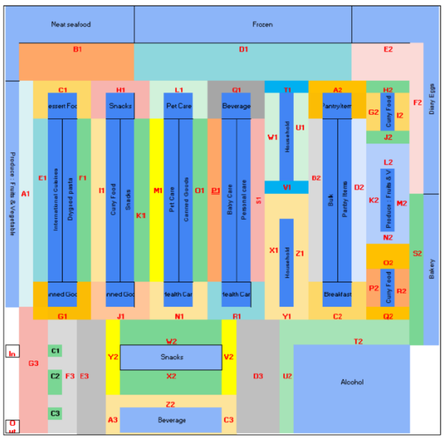

Figure. 2: Layout of the shopping mall

Figure 2 illustrate the layout of the shopping mall with each area assigned with a particular code. This code was used to track the path of the user in traveling inside the mall.

B Modelling

Clustering is the process of grouping objects based on their similarity. RFM analysis is used for customer segmentation. Segmentation is done based on the total orders per customer, average days between orders per customer, and average size of the orders per customer (number of products in an order) aligned with recency (R), frequency (F), and monitory value(M) respectively.

Based on these three measures, the selected dataset is clustered using the K-means clustering algorithm. Identification of the best number of clusters or k value in the k-mean algorithm is important as it leads to minimizing the effect of outliers and the best number of clusters (k) are being evaluated based on distortion score and silhouette score. The dataset was inserted into the k-mean algorithm and model performance was calculated using silhouette score first.

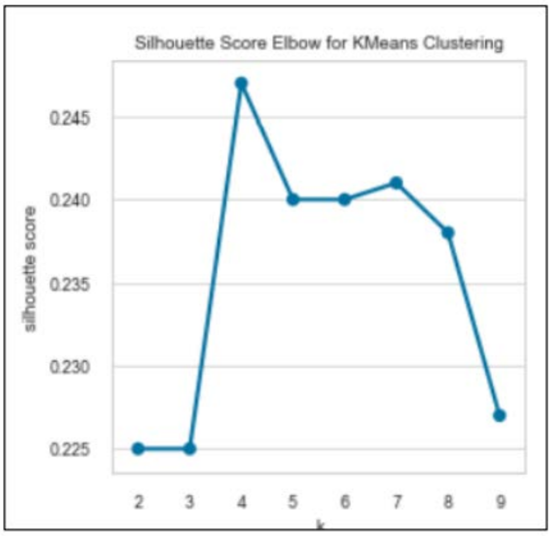

Figure 3: Silhouette score graph

The silhouette score measures how close the point lies to its nearest neighbour points across the clusters by considering the variables such as variance, high-low difference, and skewness. The resultant silhouette score for the number of clusters two to nine was visualized in Figure 3. The Silhouette score reached its' global maximum at k = 4, where it contains an ideal peak value. The best k value for the cluster is four based on the definition of the silhouette score, but for further clarification, the distortion score is also considered.

The distortion score was used in order to clarify the best number of clusters of customers which were identified under the silhouette score. In this method, the K-elbow visualizer is implemented based on the 'elbow' method in k-mean clustering. In this scoring method, the user must first specify in advance the range of clusters, and then the elbow method computes the average score for each cluster. According to the scores plotted in Figure 4, elbow point k = 4 was identified as the best k value.

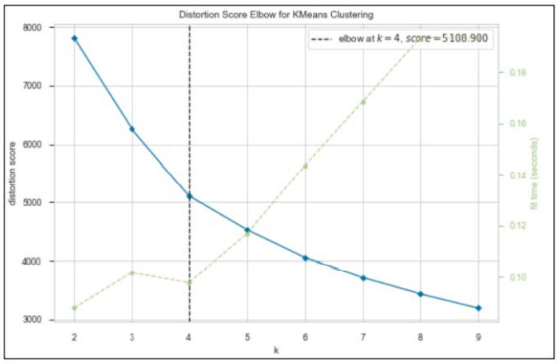

Figure 4: Distortion score graph

K-mean clustering was then concluded with an optimum cluster size of k = 4 for customer segmentation for further analysis.

Figure 5 illustrates the snake plot graph to visualize the average value of the main three features, which are identified in R, F, and M for each cluster. Compared with other graphical data analysis representation techniques, Snake plot graph perform well for customer perception analysis. The labels 'Orders,' "Lag," and "Products" used in Figure 5's X-axis represent the size of the order, the number of days between each order, and the number of products in each order, respectively, and correspond to the recency, frequency, and monitory of RFM analysis.

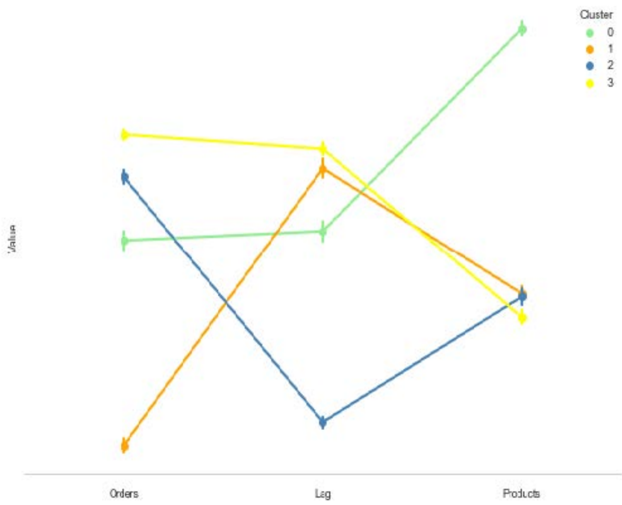

Figure 5: Snake Plot Graph for 4 clusters

Each line represents behavior of each cluster with appropriate scale endpoints for above criteria. Following points discuss each cluster behavior separately.

Cluster 1: Customers who place lowest order rate but not visit the shop often and once visit placed average number of products in an order.

Cluster 2: Customers who place more orders and visit the shop often with each order it contain average number of products.

Cluster 3: Customers who place order most but not frequent with least number of products in the orders when comparing to the other clusters.

1 *Frequent path for a particular customer:*

By applying the FP growth algorithm to the data of a particular By applying the FP growth algorithm to the data of a particular customer's travelled path inside the mall, his/her frequent path was generated. Figure 6 visualized a heatmap based on the travel frequency for each location by considering the travel path data of customer '516'.

In the heatmap in Figure 6, the variation of the colour areas defined the travel frequency for each location by the customer. Dark-coloured areas defined high travel frequency, while light-coloured areas visualized less travel frequency. By applying the FP growth algorithm to the path data of a customer, frequent itemset are identified for the customer, as illustrated in Figure 7 for customer '516'.

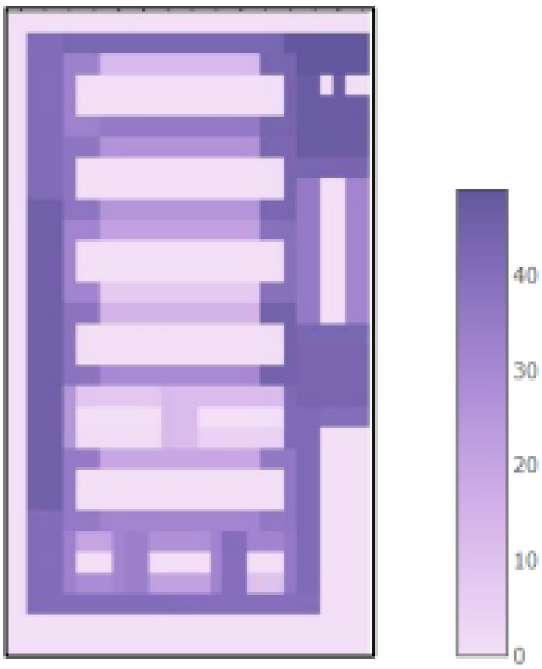

Figure. 6: Heatmap for traveled path of the customer '516'

Item List
'Organic Lemon'
'Bag of Organic Bananas'
'Organic Large Grade AA Brown Eggs'
'Strawberries'
'Organic Blueberries'

Figure 7: Frequent itemset of customer '516'

Customer Purchase Quantity < Average purchase for product	**Customer Purchase Quantity >= Average purchase for product**	**If suggested item not bought by customer previously**
Calculate discount to move the purchase quantity to average level ↓ Item Discount Rate = Calculate Discount using eq (1)	Item Discount Rate = Fixed rate	Item Discount Rate = Fixed rate
For example, consider the customer '516'		
➢ Average purchase per customer for product 'Organic Avocado' = 7490 / 3049 = 2.42 ➢ Average purchase for '516' for product 'Organic Avocado' = 2 ➢ Based on eq (1), customer '516' discount for item 'Organic Avocado' (if t =1 year): = $(2.42/2)^{1/1} - 1$ = 0.21%	➢ Average purchase per customer for product 'Organic Yellow Onion' = 4290 / 2449 = 1.75 ➢ Average purchase of customer '516' for product 'Organic Yellow Onion' = 4 ➢ Discount for 'Organic Yellow Onion' is a fixed rate.	➢ Product 'Organic Gala Apples' is never bought by customer '516'. ➢ Discount for 'Organic Gala Apples' is a fixed rate.

Figure 8: Discount recommendation

2 *Frequent path for selected group of customers:*

Based on the clusters that were identified previously, cluster two was identified as the most effective customer for the revenue of the shop. A sample of 100 customers were selected to apply the FP growth algorithm. By applying the FP growth algorithm to a group of selected customers, their frequent locations were identified.

Frequent Path for 100 customers in cluster 02: K2, E3, V2, A1, N2, D3, C2, E2, G3, Y2, D1, G1, F3, B1, D2, F2, H1, M2

3 *Recommendation of the discount:*

Discount value for products based on the travel path of user was calculated using following eq(1)[16] and eq(2).

$$Discount\ Rate = \left(\frac{Expected\ Rate}{Current\ Rate}\right)^{1/t} - 1 \quad (1)$$

t – Number of Years

$$Average\ Purchase\ per\ customer\ for\ Product\ 'ItemName' = \frac{Total\ bought\ Quantity\ for\ the\ item}{Number\ of\ items\ bought\ by\ a\ customer} \quad (2)$$

Figure 8 illustrate the three scenarios which are considered in discount recommendation.

‘Customer purchase quantity’ and ‘Average purchase for product’ is considers as main criteria in discount recommendation.

‘Customer purchase quantity’: Purchase quantity for particular product by a customer

‘Average purchase for product’: Average purchase quantity for a product

4 *Layout Recommendation:*

RFM analysis has resulted in four clusters, and among them, RFM analysis has resulted in four clusters, and among them, cluster two was identified as the most important segment for the shop because it includes the high order placement rate, the minimum days between order placement, and each order containing an average size. As this cluster represents the regular visitors of this shop, it was considered for the layout recommendation. In the frequent path, it only includes most of the traveling areas, but not all traveling areas of the customer. If the frequent travel areas include the most travelled areas, customers may tend to buy new products as they travel through the shop. This may result in improved sales and the expansion of the customer’s travel areas.

For the layout updating process, a sample of 100 customers from cluster two was selected as it implies the behaviour of all the customers in the cluster and to overcome the technical limitations that will be faced when executing the full dataset at once. By applying FP growth to these customers, the most travelled areas within them were identified. Through this result, the items that are in the most frequent traveling path, infrequent locations, and items that are bought by other customers from frequent locations are identified. Aisles related to infrequent locations can be relocated to the frequent path of the customers. This has allowed to identify the items that are in the relocated aisles and bought by the customers in parallel. By considering the above item lists, it is clear that there are new items in the frequent path that are already bought by customers. When items are in the frequent path of customers, they will tend to buy more products than before. This scenario concludes that relocating aisles will result in increase of sales and enhanced shopping experience for customers.

IV RESULT AND DISCUSSION

A *Improve sales based on discount schema*

Using the FP growth algorithm, it is possible to select the frequent path of a particular customer. This path represents the location codes through which the customer travelled the most in the store during the purchase(s).

In this frequent path, there are locations that are visited by the customer, but items may not be bought from some locations. This research, has introduced a discount schema for items in order to solve this issue. The introduction of discount schemas, may encourage customers to buy products that they have never bought or bought less frequently.

Item List
'Bag of Organic Bananas'
'Asparagus'
'Organic Hass Avocado'
'Organic Avocado'
'Pineapple Chunks'
'Strawberries'
'Organic Blueberries'
'Bag of Organic Bananas','Organic Hass Avocado'

Figure. 9: Frequent Items of customer ‘626’

When comparing the frequent item list of the customer ‘626’ as depicted in Figure 9 to the customer ’516’ in Figure 7, the ‘Organic Avocado’ item is not a frequent item of customer ‘516’. By introducing a discount for non-frequent items, customers may be more inclined to buy them on their next shopping trip. This led to an increase in product sales and an increase in the revenue of the shop.

B *Effectiveness of current layout over previous layout*

Based on the Figure 2 location codes, Figure 10 illustrates the overall behaviour of the clusters based on the travel frequency of each location inside the supermarket. According to the above Figure 10, each customer group’s travel frequency for some of the aisle locations, such as ‘W1’, ’U1’, and ’O2’, were lower compared to others.

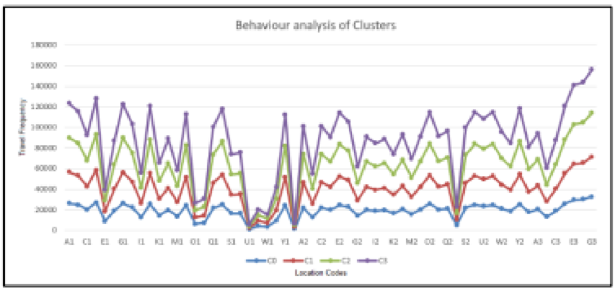

Figure. 10: Behavior analysis of cluster

Locations such as ‘A1’, ‘G1’ and ‘Y1’ contain high travel frequency. Figure 10 concludes that each cluster group has similar behaviour compared to each other for travel frequency inside the shopping mall. Because of that, only a subset of customers from a cluster is considered for layout generation. Based on the previous findings, cluster two was identified as the customer set which is most affected by the layout because of that, hundreds of customers from cluster two were selected for further analysis in the layout for suggesting a process.

Frequent Path for 100 customers in cluster 02: K2, E3, V2, A1, N2, D3, C2, E2, G3, Y2, D1, G1, F3, B1, D2, F2, H1, M2

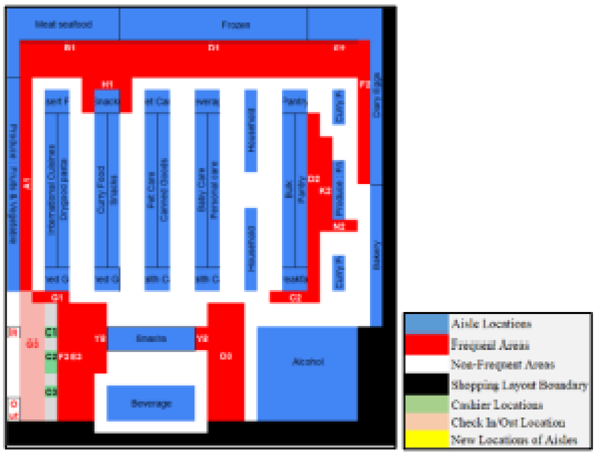

Figure 11: Previous Layout

Figure 11 represents the most travelled areas of the hundred customers who selected previously. Frequent visit areas are represented by the red colour, and white-coloured areas represent less frequent travel areas of customers. According to Figure 11, most of the customers have travelled to the aisles that are placed in the corners of the shop but not to the aisles in the middle of the shop.

According to the location path codes in Figure 2, Figure 12 illustrates the sales of the products by considering 30 path codes. According to Figure 12, a significant number of products which are purchased by customers are located in the frequent path but some of the frequented path locations do not contain a considerable number of purchases. One reason for this could be that even though customers have to travel to some locations, they may not be interested in the items placed there, such as products near the cashier counter, entrance or alcohol area.

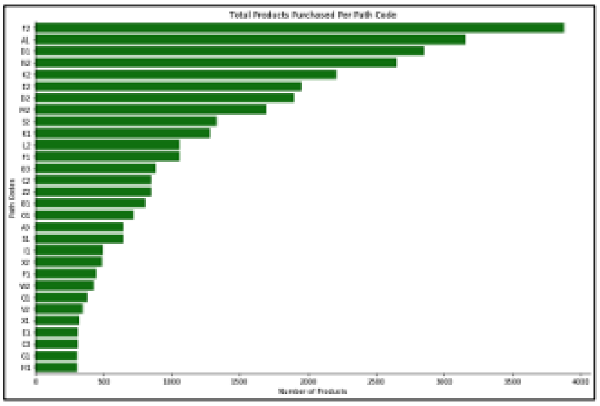

Figure 12: Total products purchased per Path

Relocating the non-frequent sale items to a frequent travel path will result in an increase in sales. Figure 13 illustrates the current layout of the shop, where some of the aisle locations were changed in order to increase sales. Even though some products are bought by the customers, they are not located in their frequent path. However, according to the current layout, they are placed in frequent paths, which may assist customers to easily locating the products and thereby leads to an increase in sales. The layout update mostly affects the customers who visit more often but introducing a discount schema will improve the sales of irregular customers, and their visits.

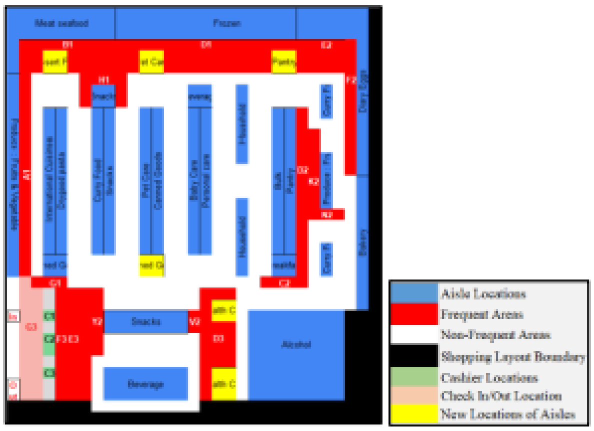

Figure 13: Current Layout

V CONCLUSION AND FUTURE WORK

This project proposed a new discount schema and a new supermarket layout by applying the FP growth algorithm and RFM analysis to travelled path data of customers inside the supermarket. Results are computed based on the supermarket transaction data, which indicates the actual behaviour of the customer. An FP growth algorithm was applied to identify the frequent paths of each customer, from which the frequent and non-frequent areas of customer travel were identified. Consequently, the most frequent items and non-frequent items of the customer were identified. As the suggested discountschemas are based on the customers' travel patterns and are personalized for each customer, they will have a positive impact on the revenue of the retailers. Relocating non-frequent items to the frequent path with discounts will possibly impact the revenue of the shop positively by directing customers to buy more products.

A new shopping layout was introduced by relocating the aisles from non-frequented areas to frequented travel areas, which will allow the customers to easily find their products and also will increase sales simultaneously. Thus, providing a personalized discount based on a customer's travel path improves the sales more than a static discount schema for each and every customer. Our research has finally concluded that recommending discount schemas and shopping layouts based on the travel paths of customers increases sales and improves customer satisfaction at the shop.

Deployment of the research will be a future task. For future research, implementing an advanced pattern mining algorithm in parallel to the FP growth algorithm will improve the performance of result generation. For accurate result generation in path tracking, indoor location tracking techniques can be used. As the suggested discount schemas are for the product categories, they can be further personalized item-wise for a customer and used as an effective discount schema for discount calculation.

REFERENCES

[1] V. Aggelis and D. Christodoulakis, "Customer Clustering using RFM analysis", ICCOMP'05 Proceedings of the 9th WSEAS International Conference on Computers, 2nd ed, 2005, pp. 1-5, [online]: Available:https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.7091&rep=rep1&type=pdf

[2] K. H. Alyoubi, "Association Rule Mining on Customer's Data using Frequent Pattern Algorithm", IJCSNS International Journal of Computer Science and Network Security, vol. 20, issue 5, 2020, 103-110, [online]: Available:http://paper.ijcsns.org/07_book/202005/20200513.pdf

[3] A. Ansari and A. Riasi, "Taxonomy of Marketing Strategies Using Bank Customers' Clustering". International Journal of Business and Management, vol. 11, issue 7, 2016, pp 106-119, doi:10.5539/ijbm.v11n7p106

[4] I. Cil, "Consumption universes-based supermarket layout through association rule mining and multidimension scaling. Expert Systems with Applications", vol. 39, issue 10, 2012, pp 8611-8625, https://doi.org/10.1016/j.eswa.2012.01.192

[5] Y. L. Chen, M. Kuo, S. Wu and K. Tang, "Discovering recency, frequency, and monetary (RFM) sequential patterns", Electronic Commerce Research and Applications, vol. 8, issue 5, 2009, pp 241-251, https://doi.org/10.1016/j.elerap.2009.03.002

[6] C. H. Chee, J. Jaafar, I. A. Aziz, M.H. Hasan and W. Yeoh, "Algorithms for frequent itemset mining: a literature". Artif Intell Rev, vol. 8, issue 5, 2018, pp 2603-2621, doi:10.1007/s10462-018-9629-z

[7] K. Coussement,F. Van den Bossche and K. De Bock, "Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees", Journal of Business Research, vol. 67, issue 1, 2014, pp 2751-2758, doi: 10.1016/j.jbusres.2012.09.024

[8] O. Dogan, E. Aycin and Z. A. Bulut, "Customer Segmentation by using RFM Model and clustering methods: A Case study in Retail Industry", International Journal of Contemporary Economics and Administrative Science, vol. 8, issue 1, 2018, pp 1-19, [online]: Available: http://www.ijceas.com/index.php/ijceas/article/view/174

[9] S. Gurudath, "Market Basket Analysis & Recommendation System Using Association Rules", Thesis, Master of Science in Big data Management and Data Analytics, Griffith College, Dublin, 2020, [online]: Available:

[10] Y. Hu and T. W. Yeh, "Discovering valuable frequent patterns based on RFM analysis without customer identification information. Knowledge-Based Systems", vol. 61, 2014, pp. 76-88, doi:10.1016/j.knosys.2014.02.009

[11] I. Jung and Y. S. Kwon, "Grocery Customer Behavior Analysis using RFID based shopping paths data", World Academy of Science, Engineering and Technology, vol. 5, issue 11, 2011, pp. 1508-1512, doi.org/10.5281/zenodo.1081169

[12] O. A. Kasapoglu, "Data Analysis for Retail Layout Decision", International Journal of Advanced Research and Review, vol. 1, issue 6, 2016, pp. 115-126.

[13] J. S. Larson, E. Bradlow and P. S. Fader, "An exploratory look at supermarket shopping paths", International Journal of Research in Marketing, vol. 22, issue 4, 2005, pp. 395-414, https://doi.org/10.1016/j.ijresmar.2005.09.005

[14] R. Soudagar, "Customer Segmentation and Strategy Definition in Segments: An Case study: An Internet Service Provider in Iran", Master Thesis, Master of Arts, Business Administration, Lulea University of Technology, Sweden, 2012, doi:10.1109/ICEBE.2018.00027

[15] S. W. Yen, C. H. Wang and L. Y. Ouyang, "A Search Space Reduced Algorithm for Mining Frequent Patterns", Journal of Information Science and Engineering, vol. 28, issue 1, 2012, pp. 177-191, doi:10.6688/JISE.2012.28.1.12

[16] M. Thakur, "Discount Rate Formula ",EDUCBA, Accessed: Aug 20,2021[online], Available https://www.educba.com/discount-rate-formula/

ACKNOWLEDGMENT

Researchers would like to acknowledge all the staff members of the University of Colombo School of Computing for the guidance and support they have given throughout the research work.

AUTHOR BIOGRAPHIES

RSN Dilrukshi received BSc (Hons) in IT degree from General Sir John Kotelawala Defence University and currently pursuing her master's degree at University of Colombo School of Computing.

HA Caldera received hid PhD degree in computer science from University of Western Sydney, Australia in 2005. Since then, he has been a Senior Lecturer at the University of Colombo School of Computing. His current interests include Data Mining, Web Mining, Business Intelligence and Natural Language Processing. He has published and reviewed many international conference and journal papers.