

Customer segmentation for churn management by using ant colony algorithm

Batuhan Gulluoglu*, Department of Business, Bahcesehir University, Istanbul, Turkey

Evren Arifoglu, Department of Business, Bahcesehir University, Istanbul, Turkey

Adem Karahoca, Department of Business, Bahcesehir University, Istanbul, Turkey <https://orcid.org/0000-0003-4654-6351>

Dilek Karahoca, Department of Business, Bahcesehir University, Istanbul, Turkey <https://orcid.org/0000-0002-5017-3204>

Suggested Citation:

Gulluoglu, B., Arifoglu, E., Karahoca, A. & Karahoca, D. (2020). Customer segmentation for churn management by using ant colony algorithm. *Global Journal of Computer Sciences: Theory and Research*. 10(1), 018-026. <https://doi.org/10.18844/gjcs.v10i1.4749>

Received October 15, 2019; revised January 2, 2020; accepted March 22, 2020.

Selection and peer review under responsibility of Prof. Dr. Dogan Ibrahim, Near East University, Cyprus.

©2020 United World Center of Research Innovation and Publication. All rights reserved.

Abstract

It is extremely important for companies to set customer priorities and act in line with these priorities. The ant colony algorithm is used to perform customer segmentation. To do this, the shortest path approach was chosen. Besides, clustering is done by the Euclidean distance formula in the ant colony algorithm. The customer segmentation attributes are mostly related to the satisfaction factors, but some of them were eliminated by using ranker. These results are mostly related to the customer's income, tenure, equip call card and reside. These attributes are the most important satisfaction factors not to lose customers as expected. There are many reasons in changing GSM operator for subscribers, and it is very important for companies to predict if subscriber will change GSM operator or not. For this reason, companies that give GSM services have to monitor subscribers' behaviour and predict one step forward. In this study, changing subscribers' GSM operator will be predicted by using data mining techniques.

Keywords: Ant colony, churn management, customer segmentation, data mining.

Categories: I.2.1, I.2

* ADDRESS FOR CORRESPONDENCE: Batuhan Gulluoglu, Department of Business, Bahcesehir University, Istanbul, Turkey, Email: batuhanguoglu@gmail.com

1. Introduction

Customer priorities are very important for companies. Moreover, customer priorities must be determined, and campaigns must be ordered according to these priorities. Customer segmentation was done with the ant colony algorithm. The shortest path approach is used in the ant colony algorithm. Moreover, clustering is done by the Euclidean distance formula in the ant colony algorithm.

The customer segmentation attributes are mostly related to the satisfaction factors, but some of them were eliminated by using ranker. These results are mostly related to the customer's income, tenure, equip call card and reside. These attributes are the most important satisfaction factors not to lose customers as expected. There are many reasons in changing GSM operator for subscribers, and it is very important for companies to predict if subscriber will change GSM operator or not. For this reason, companies that give GSM services have to monitor subscribers' behaviour and predict one step forward. In this study, changing subscribers' GSM operator will be predicted by using data mining techniques.

Customer management and satisfaction are the most important issues for companies. Companies have to know their customers. Moreover, companies have to know their customers' characteristics, and they should determine the campaign according to these characteristics. Companies have to influence their customers due to priorities. The most important profit is customer satisfaction, and this will be an advantage or disadvantage according to the approach of companies (Rygielski, Wang & Yen; 2002).

The goal of clustering is to reduce the amount of data by categorising or grouping similar data items together.

Such grouping is pervasive in the way that humans process information, and one of the motivations for using clustering algorithms is to provide the automated tools to help in constructing categories or taxonomies (Jardine & Sibson; 1971)

We can divide clustering methods into two basic types: hierarchical and partitional clustering. Each type has a wealth of subtypes and there are different algorithms to the clusters. In this thesis, ant colony optimisation is used for clustering the customer data sets. The ant colony is one of the most important techniques to make a selection from the data sets (Zhao, Karypis & Fayyad, 2005).

Keeping the customers that we require more source than gaining new customers. Churn models are developed to detect the customers about the attrite and take action about them (Zeng & Cheung, 2009)

There is some hypothesis that is indicated in this study. One of the hypotheses is about customer's age and income. If the customer is under 18-year old and has no income, then he or she would not spend more than 500 Turkish liras in a month. Moreover, if the customer is over 18-year old woman and has some income, then she has expense more than her salary. Therefore, she has to have a credit card which has a limit more than the customer's salary. Besides, if the customer is an employee and has a constant salary, he or she will spend more than his or her salary (Al-Ani, 2005).

1.1. Churn management

Churn management is one of the most important things that provide differences in the system according to the necessities of customers. According to the customer's satisfaction, the bank changes its strategy to be more successful. The strategy is usually about the products and services that the customers use (Wei & Chiu, 2002).

Churn management is to revise the customer data and analyse the results. The results depend on customers' satisfaction. Churn management makes the customers' satisfaction in a stable position.

There are a lot of methods for clustering of customer data. Filter and wrapper approaches are used for unsupervised property selection. Ant colony optimisation clustering is used to solve the problems of clustering. Customer detention or customer churn is applied by some methods.

Filter clustering is related to filter the data according to the tool. There are a lot of attributes in churn management data. However, clustering is related to the accuracy, so the most important attributes have to be determined according to the accuracy. These attributes are related to data, but if the clustering is not accurate, then it should be abated amongst these attributes, and this filtering should be done unless the most accurate parameters are obtained (Tsai & Lu, 2009)

2. Previous Studies

There are a lot of studies related to data mining techniques with the ant colony. This study aimed to find the documents related to customer segmentation with ant colony optimisation to determine the customers, which will be useful for the campaign. There are a total of 122 topics related to these topics.

There are a lot of techniques that are used in ant colony optimisation. First, Deneubourg used ACO in 1991, and the study is related to robots. Deneubourg used this algorithm to provide some property to robots. Robots decide to take up or drop the objects according to these algorithm results. This is a result that is similar to the result of this study. In this study, the data are clustered according to the ACO algorithm, and then, classification is performed and there exist some rules according to classification such as robots' operations (Xiao-bin, Feng & Hui, 2009). Lumer and Faieta updated the Gutowitz algorithm in 1995. This study is related to tasks. Tasks are done according to the distances. These distances are obtained due to the clustered data as a result of ACO algorithm (Zeng & Cheung, 2009).

De Castro improved Lumer and Faieta's study. This study is related to robots' learning. He provided a determining matrix related to the distances. According to this matrix, robots do operations, respectively, in the fallacy of ACO algorithm. Moreover, a pheromone is used in his study, and similar operations are seen as a result of the pheromone level (Chow, Wang & Ma, 2008).

Gutowitz changed this algorithm in 1993. It is the same algorithm, but robots schedule the tasks related to the operations of robots and, respectively, do these operations.

Weili improved the study related to entropy. Entropy is related to the operations of robots. In this study, Faieta's algorithm is improved related to memory. Moreover, there exists a counter in this study, and this counter occupies the errors in a distance matrix. Shelokar's study is nearly the same as this study. Pheromone level is updated on each iteration, and most of the similar data are clustered according to the distances and pheromone level. This is determined according to the updated pheromone level. Zeng and Cheung have an algorithm related to property selection. This algorithm provides an index related to the similar properties. This index should not be pre-determined (Chow et al., 2008). All of these methods and algorithms have the same solution as this study mentioned. The most important topic is clustering data and determines the methodology related to data set. The operations are made to determine the similarities and then provide some properties according to these similarities.

3. Method

3.1. Ant colony optimisation clustering

Ant colony optimisation is one of the clustering algorithms, which is used in this study depending on Shelokar, Jayaraman and Kulkarni (2004) methodology.

3.2. Ant colony optimisation for clustering

The range of the objects is decreased to solve the problem of segmentation N objects to K clusters. Data set is specified by N . The result set contains defined clusters of the data set and is specified as S .

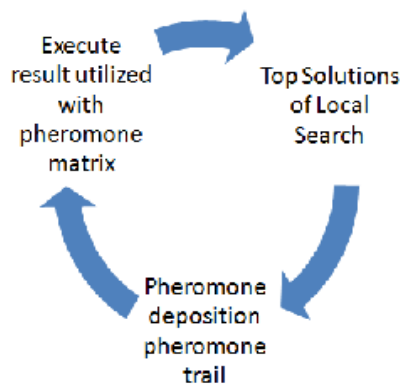
τ : $N \times K$ matrix of pheromone detention

$\tau_{i,j}$: Detention matrix of i ' th of j 'th form. Pheromone sign matrix is resettled to τ_0

The result structure process is ant's random behaviour. Ant selects the things depending on the sign pheromone matrix. Objects are become more desirable to select if pheromone invests on them. Ant's selection chances depend on the pheromone concentration of nominee objects to all sediment pheromone concentrations.

$$P_{i,j} = \frac{\tau_{i,j}}{\sum_{k=1}^K \tau_{i,k}}, j = 1..K$$

Building results depend on the agents that utilised the pheromone matrix info. The pheromone matrix is updated correspondingly.



According to this coincidental solution, the factors are searched. This method can execute for all the nominee results. Besides, the result is completed when the change is finished due to these results. In conclusion, the exchange process is executed between the solutions when an accurate value is bigger than the original result.

Ant colony optimisation's last step is about the exchange process. The accurate results are searched and compared with the original values, and the colony optimisation pheromone updates these values.

Pheromone matrix is affected by the best results.

Moreover, this method is executed to have the maximal repetitions of objects that are clustered.

3.3. SBS algorithm

Checking learning for the selection of properties provides a consecutive reflux search which is a heuristic method. The property subset values are eliminated according to the worst values and consecutive reflux selection that begins with a set which has all properties.

In conclusion, searching is finished if there is no evolution on the property set.

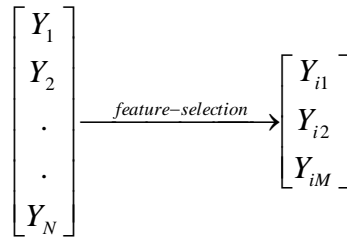


Figure 1. Feature selection

There is a procedure that is about to assess. This method begins on the set that has full properties. The solutions are determined according to this method. All the properties are taken, and new solutions are formed.

Using this method, there is always a feature selection on each stair of this method. One of the features is picked up on each step. Therefore, the accurate value is determined from the data that are property set.

The feature is selected due to the results related to the highest accuracy rate. This process is continued to find the last solution that has no innovation that can be determined.

3.4. Feature selection–ant colony clustering (FS-ACO CLUSTERING)

There are three parts of the wrapper’s algorithm. These parts are property search, unsupervised learning and assessment algorithm, respectively. The first algorithm is SBS in this thesis statement. The second one is ACO clustering algorithm, and the third one is about assessment algorithm. Besides, the assessment function uses unified criterion approach.

3.5. Unified evaluation criteria

The property choice and clustering process are associated with the Wrapper approach. This association process requires an evaluation procedure. There is only one assessment procedure that is used in this thesis statement for the property selection and clustering process.



Figure 2. FS-ACO clustering

If the number of parameters rises, then this index rises. This procedure is not related to property selection. There is a penalty glossary to control and check the value assessment.

The distance between objects to the centre is defined as evaluation functions. Nominal attributes are used in FS-ACO clustering algorithm in this thesis statement. There is a process. First, according to the format of data, these attributes are taken from the record.

Second, to decrement the computational supplementary load of the method, these values are updated as binary values.

The agents, respectively, execute results and this is a process. The clusters are picked to calculate the centroids of them, and then, an iteration performance is found by evaluation criteria.

Then, according to the results of that, there is an evaluation procedure to calculate the other results, and the top results are chosen by this procedure. In this way, the results are updated when the new results become more accurate with the oldest ones.

4. findings

4.1. Training the FS-ACO algorithm

This study solution is clearly expounded in this part. The offering FS-ACO aggregation method performance estimations on data are given.

The first section of the offered methods is the discretisation of the data. The property of the data is transformed into binned assets. Numerical amounts are transformed into nominal features with respect to their distribution.

Macros of the algorithm are optimised by the following different proofs. The education data set, which is explained in before part, sampled to decrease the calculate complexity and education time. For example, 280 samples are used for the education of the FS-ACO condensation algorithm.

They researched that some sampling factor will be skewed maximum, and first of all, the factors which are maximum skewed distribution are carried. Finally, the results have 20 features, which are used for the beginning education of the data.

If the data set increases, there are some overlap problems. These problems caused a suitability function to be zero, and it takes the local 0 (zero). This case is a solution by joining the term 1 to the suitability function.

The classification is being done by Weka, which is a data mining tool, with the Naive Bayesian method. The Naive Bayesian algorithm produces 10 rules for classification. There is a list below that includes the rules.

Rule 1: This rule contains the customers' age which is under 18 or over 18-year old. If the customers are under 18-year old, the company loses their customers easily. If the customers are older than 18-year old, the company does not lose their customers.

Rule 2: This rule contains the customers' age which is under 50-year old and over 30-year old. If the customers are over 30-year old and below 50-year old, the company loses their customers easily. If the customers over 50-year old, the company does not lose their customers.

Rule 3: This rule contains the customers' age and gender information which are between 30 and 50 years and male or female information. If the customer is under 50-year old and over 30-year old and gender is female, the company loses their customers when there is a limit under their income on their credit cards. This company has to give credit limit over their incomes.

Rule 4: This rule contains the customers’ income and monthly expenses. If the customer’s income is more than his or her monthly expense, the bank should prepare some campaigns to these group customers not to lose them.

Rule 5: This rule contains the customers’ region, age, gender and income. If the customer lives in the east, age is between 18 and 30 years old, gender is female and has the salary between 2,000 and 5,000 Turkish liras, then the company should provide her credit over her income. If the customer lives in the west, age is between 18 and 30 years old, gender is female and has the salary between 2,000 and 5,000 Turkish liras, the company does not have to provide credit over her income.

Rule 6: This rule contains information about customers’ marital status and gender. If the customer is married and male, the company does not have to provide some extra campaigns, but if the customer is married and female, the company has to provide extra campaign not to lose their customers.

Rule 7: This rule contains information about customers’ tenure. If the customer works below 5 years at the same company, the campaign does not need to provide credit campaigns. However, if the customer works between 2 and 5 years at the same company, the company has to provide extra campaigns to this group of customers. Hence, this type of customers is not reliable to all banks so they prefer the banks that provide them credit campaigns.

Rule 8: This rule contains information about customers’ age and retirement. If the customers are over 50-year old and retired, the company should provide some campaigns which include extra installments.

Rule 9: This rule contains information about customers’ credit card number and limits. If the customer has more than two credit cards and each of them has high limits, the company should provide extra bonus campaigns for their customers to assess this group of customers.

Rule 10: This rule contains information about customers’ bank information. If the customer is in the same bank below 5 years, the customers think to change their banks, so it should have some extra campaigns to customers which are in their banks below 5 years.

4.2. FS-ACO algorithm clustering vs. Weka comparison

The K-means algorithm is utilised to evaluate the wrapper algorithm with filtering methods. By using the K-means algorithm, then all the data set is aggregated to provide the properties. Clustering attribute chosen labels are utilised to get the data. Weka property chosen is utilised to get ahead of this situation.

Ant colony algorithm and Weka clustering results are shown in Figure 3. ‘WEKA_CLUSTER’ shows the Weka clustering algorithm, and the first ‘CLUSTER_COUNT’ shows the number of objects in the clusters. ‘ANT_CLUSTER’ shows an ant colony optimisation clustering algorithm, and the second ‘CLUSTER_COUNT’ shows the number of objects in the cluster ant colony.

ROWID	WEKA_CLUSTER	CLUSTER_COUNT	ANT_CLUSTER	CLUSTER_COUNT
ROW0	CLUSTER_0	10110	CLUSTER_0	10080
ROW1	CLUSTER_1	3020	CLUSTER_1	3050
ROW2	CLUSTER_2	1250	CLUSTER_2	1300
ROW3	CLUSTER_3	600	CLUSTER_3	500
ROW4	CLUSTER_4	20	CLUSTER_4	70

Figure 3. The Distribution of Clusters: Ant Colony and Weka

Figure 4 shows the accuracy of two different algorithms.

WINNER CLUSTER	CLUSTER_0	CLUSTER_1	CLUSTER_2	CLUSTER_3	CLUSTER_4
CLUSTER_0	10120	40	0	0	0
CLUSTER_1	0	1432	12	0	0
CLUSTER_2	0	0	2100	0	0
CLUSTER_3	0	0	0	0	0
CLUSTER_4	22	0	0	1286	0

Figure 4. Different algorithms' accuracy: ant colony and Weka

ROWID	TRUE POSITIVES	FALSE POSITIVES	TRUE NEGATIVES	FALSE NEGATIVES
CLUSTER_0	10120	22	4100	0
CLUSTER_1	1432	40	13410	21
CLUSTER_2	2100	12	12635	18
CLUSTER_3	1286	0	14321	25
CLUSTER_4	0	0		0

Figure 5. Ant colony and Weka clustering variables; The true positives, false positives, true negatives and false negatives

A linear correlation between the ant colony and Weka clustering results is shown in Figure 6.

ROWID	WEKA CLUSTER	ANT CLUSTER
WEKA CLUSTER	0.970	1
ANT CLUSTER	0.970	1

Figure 6. Linear correlation between ant colony and Weka clustering results

5. Conclusion and discussion

Clustering is one of the most important data mining problems that is studied in this thesis statement. There are a lot of techniques implemented for data sets to provide the most suitable parts of data.

Unsupervised property elimination is a technique used for these problems. Unsupervised property selection has two approaches, i.e., filter and wrapper. Moreover, there is better accuracy in the wrapper approach.

Ant colony optimisation clustering is used with SBS method for the elimination of properties of data in this thesis statement. Clustering algorithm and property elimination are joined, and the wrapper approach comes into being in this way.

Shelokar's ant colony clustering model contains an ant colony optimisation method that is utilised in this thesis statement. Manhattan distance is provided from the distance metric by using nominal attributes. Moreover, the features of mean values are utilised. The application of CRIT criteria is providing an innovation by the original function.

There are no improvements in ant colony optimisation and SBS wrapper approach before this thesis statement. The most accurate solutions are utilised to determine the features and separation of data sets.

There is an iteration according to the data sets. In all steps, the SBS algorithm is utilised. All the aims of this thesis statement are to determine the recoup of the customers. There is a process, and this process is related to choosing the data and converting them from numerical to nominal entities.

There are 280 items that are illustrated in this thesis statement. FS-ACO clustering uses this data. In conclusion, we can provide the customer's characteristics according to these results.

References

- Al-Ani, A. (2005). *Ant colony optimization for feature subset selection*. In WEC (2) (pp. 35-38).
- Anderson, E. T. (2002). Sharing the wealth: when should firms treat customers as partners?. *Management Science*, 48(8), 955–971. <https://doi.org/10.1287/mnsc.48.8.955.170>
- Chow, T. W. S., Wang, P. & Ma, E. W. M. (2008). A new feature selection scheme using a data distribution factor for unsupervised nominal data. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, 38(2), 499–509. <https://doi.org/10.1109/TSMCB.2007.914707>
- Cotter, S. F., Kreutz-Delgado, K. & Rao, B. D. (2001). Backward sequential elimination for sparse vector selection. *Signal Processing*, 81, 1849–1864. [https://doi.org/10.1016/S0165-1684\(01\)00064-0](https://doi.org/10.1016/S0165-1684(01)00064-0)
- Jardine, N. & Sibson, R. (1971). *Mathematical taxonomy*. London, UK: Wiley.
- Marco, D. (2004). *Ant colony optimization*. Cambridge, MA: Massachusetts Institute of Technology. ISBN: 0-262-04219-3.
- Marzuki, Z. & Ahmad, F. (2012). Data mining discretization methods and performances. *Lung*, 3(32), 57.
- Rahm, E. & Do, H. H. (2000). Data cleaning: problems and current approaches. *IEEE Data Engineering Bulletin*, 23(4), 3–13.
- Rygielski, C., Wang, J. C. & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, 24(4), 483–502. <https://doi.org/10.1088/1742-6596/910/1/012021>
- Shelokar, P. S., Jayaraman, V. K. & Kulkarni, B. D. (2004). An ant colony approach for clustering. *Analytica Chimica Acta*, 509(2), 187–195. <https://doi.org/10.1016/j.aca.2003.12.032>
- Tsai C.F. & Lu Y.H. (2009). Customer churn prediction by hybrid neural network. *Expert Systems with Applications*, 36(10), 12547–12553. <https://doi.org/10.1016/j.eswa.2009.05.032>
- Wei, C. P. & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: a data mining approach. *Expert Systems with Applications*, 23(2), 103–112. [https://doi.org/10.1016/S0957-4174\(02\)00030-1](https://doi.org/10.1016/S0957-4174(02)00030-1)
- Witten, I. H. & Frank, E. (2002). Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record*, 31(1), 76–77
- Xiao-bin Z., Feng G. & Hui H. (2009). *Customer-churn research based on customer segmentation* (pp. 443–446). Electronic Commerce and Business Intelligence, ECBI. International Conference on 2009.
- Zhao, Y., Karypis, G. & Fayyad, U. (2005). Hierarchical clustering algorithms for document datasets. *Data Mining and Knowledge Discovery*, 10(2), 141–168. <https://doi.org/10.1007/s10618-005-0361-3>
- Zeng, H. & Cheung, Y. M. (2009). A new feature selection method for Gaussian mixture clustering. *Pattern Recognition*, 42(2), 243–250. <https://doi.org/10.1016/j.patcog.2008.05.030>