Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2016) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria

A DECISION SUPPORT SYSTEM FOR STAFF WORKLOAD BALANCING

Siamak Naderi[a,*], Kemal Kilic[a]

[a]Sabancı University, Orta Mahalle, Üniversite Cddesi, No:27 Tuzla, İstanbul, 34956, TURKEY

Abstract

This research arises from a real life problem that a third party solution partner of health care insurance companies in Turkey faces. The company develops Electronic Claim Processing System (ECPS) that private health care insurance companies utilize in order to check and validate whether an expenditure (i.e., claims) of their customers abides by the rules of the contracts between the insurance companies and the customers as well as the insurance companies and the health care service providers. As part of their business, the company also trains and assists the health care service providers regarding to the ECPS and the regulation of the insurance companies. Every year their training staffs visit 400+ health care service providers in Istanbul and in order to reduce the total travelling cost and balance the workload among them is a problem that they face. A decision support system that utilizes fuzzy zoning as a solution for the planning problem that the company faces is being developed. In this paper the fuzzy zoning algorithm utilized for this purpose will be presented and its performance will be evaluated.

Peer-review under responsibility of the Organizing Committee of ICAFS 2016.

Keywords: Zoning, Districting, Decision Support System, Fuzzy Zoning

1. Introduction

A third party solution partner for private health care insurance companies (will be referred to as "the Company" hereinafter) faces a staff planning problem at the beginning of each planning period where the objective is

* Corresponding author. Tel.: +90-538-236-8631;
E-mail address: siamak@sabanciuniv.edu

Peer-review under responsibility of the Organizing Committee of ICAFS 2016.

minimizing the traveling cost of their staff and at the same time balancing their workload in order to complete the tasks concurrently and maintain the fairness of the working environment. The Company develops and maintains the Electronic Claim Processing System (ECPS) for private health care insurance companies. Every transaction made by the customers of the insurance companies, flows through the ECPS and whether it is appropriate to make the payment is monitored by the ECPS. The transactions that violate the terms of the contracts both signed between the insurance companies and their customers as well as the insurance companies and the health care service providers are denied payment and those that abide are approved by the ECPS.

Since the health insurance companies that has the majority of the market share in Turkey use the ECPS (and associated services) of the Company, more or less most of the health care service providers (hospitals, health clinics, laboratories, medical imaging facilities, etc.) in the country have the Company's ECPS installed. As part of their obligations, training personnel of the Company visit quite a number of health care service providers (~400 alone in Istanbul) regularly and train their staff and inform the regulations of the insurance companies.

As of today, four of their employees are used as training personnel by the Company and the management should distribute the workload among them as part of the planning process. One of the major objectives of the Company is distributing the health care service providers to the training staff so that the total traveling cost is minimized. In a typical day, a training staff visits ~3 health care service providers due to the time required for the training session and the traveling time between two consecutive visits. Considering that Istanbul is a city with a population of ~15 Million which is located in two continents that are separated by the Bosporus (i.e., the strait that connects Black Sea and the Sea of Marmara) and connected by two bridges, traveling between two points (particularly during the day time traffic) usually requires significant amount of time. Hence, the time spent idle during traveling between two consecutive visits constitute the majority of the operating cost and should be minimized. On the other hand, the workload should also be equally distributed among the staff in order to maintain the fairness (or the perceived fairness) in the workplace. Furthermore, the concurrent completion of the tasks is desired in the company so that the second phase of the training period can start (which requires traveling to other cities as teams). Here, "equal distribution" does not necessarily imply that each employee should be assigned to same number of health care service provider, but the total time required during a planning period to conclude the tasks (the total time required for training at the facility and the traveling between the facilities) should be balanced.

Without the balance constraint, minimizing the total traveling cost could be addressed with conventional clustering approaches that groups the health care facilities with respect to their distance to each other (i.e., that minimizes the total within cluster distances). However, the resulting clusters would not necessarily be homogeneous in terms of the total workload required in the obtained clusters since the spatial distribution of the health care service providers is not homogeneous. Since balanced clusters are required, the problem on hand resembles zoning (i.e., districting) problem as opposed to a general clustering problem. In this paper, we will present a novel fuzzy zoning algorithm, which is proposed as the engine of the decision support tool that is developed for the Company which will be used in their training staff planning process. The algorithm aims to minimize the total within distance and balance the workload of the employees at the same time and yields the list of the facilities that each employee should visit during a planning period (typically two months).

The rest of the paper is organized as follows. In the next section, the relevant literature will be reviewed. The proposed fuzzy zoning algorithm will be presented in the third section. The fourth section will discuss the results of the computational experiments that are conducted in order to evaluate the performance of the proposed methodology for proof of concept purposes. The paper will be finalized with some concluding remarks and future research agenda.

2. Relevant Literature

Zoning/Districting problem usually refers to geographical design of an area with respect to various criteria and have wide range of applications in distribution (collection) supply chain subsystems, political districting, public services (e.g., police, health services) districting, design of sales territories, transportation planning, etc. Various design criteria could be considered during the districting process such as the minimum variation in spatial characters (i.e., size, shape, etc.), contiguity, compactness and homogeneity of the zones. Depending on the application, some of these design criteria might be more important, e.g., for political districting contiguity is a hard constraint in order to avoid gerrymandering, but is not that crucial for design of sales territories.

In the literature various researchers address different zoning/districting problems. Among them political districting is one of the most extensively studied problem which deals with dividing a given area into c districts such that each district has almost the same population of voters, is contiguous and compact (Bozkaya, Erkut, and Laporte, 2003; Garfinkel and Nemhauser, 1970; Hess et al., 1965; Hojati, 1996).

On the other hand, balancing the workload of the sales representatives, distribution vehicles or service providers also received some attention from the researchers. For example, Salazar-Aguilar, Rios-Mercado, and Gonzalez-Velarde (2011) address a real life problem arose in Mexico in a bottled beverage distribution company. Given customers information, they were interested in clustering the customers in such a way that number of customers in each cluster is equal (i.e., balanced). In Pavone et al. (2011) the problem is dividing a region to a specified number of sub regions, and then assigning a responsible employee to each sub region such that the workload of each responsible employee is equal. Nikolakopoulou et al. (2004) consider a problem where routing is also an important issue to be considered. The objective is minimizing the total travel distance by vehicles and balancing the workload. Rios-Mercado and Fernandez (2009) work on a real life problem in a beverage distribution firm where minimization of a measure of territory dispersion, balancing the different node activity measures among territories and contiguity is considered.

Zhou, Min, and Gen (2002) address equality in a location-allocation problem since in a typical location-allocation problem, customer demand data are often aggregated according to some arbitrary spatial points and such points do not represent true sources of customer demands. Hence allocation of aggregated customers to distribution centers can lead to underutilization of distribution centers and deterioration of customer services. In some of the applications the objective is grouping the objects so that maximum load is minimized rather than balancing the load. For example, locating cellular phone towers (e.g. facilities) in a given region is such an application. Baron et al. (2007) models the problem of locating c facilities on the unit square to minimize the maximal demand faced by each facility such that assignment to the closest facility and coverage constraints are satisfied.

Note that, all of the above mentioned research assume crisp zones. That is to say, at the end of the day an object (e.g., population unit, customer, voter, etc.) is assigned to a specific zone. However, such a dichotomous approach limits the underlying mathematical model since a hard balanced constraint would virtually always lead to infeasible solution. In order to overcome this problem, the existing literature incorporates an arbitrary tolerance level which softens the balance constraint. Another approach could be softening the zones themselves (i.e., fuzzy zones) as opposed to softening the balance constraint arbitrarily. Such an approach would also be useful in certain applications since it can be utilized for backup planning (for instance if a server is not available the alternatives can be easily determined).

In this paper a novel fuzzy zoning approach will be presented. To the best of our knowledge an algorithm which provides totally balanced clusters or an algorithm which assigns data points to not a single cluster but to the several clusters is not available in the literature. The proposed fuzzy zoning algorithm will address this gap in the literature.

3. Proposed Fuzzy Zoning Algorithm

Fuzzy set theory has received extensive attention from research community since it is introduced by Zadeh (1965) and has been successfully applied in various fields such as control theory (Takagi and Sugeno, 1985), health care (Kilic et al., 2004), system modelling (Uncu, Turksen, and Kilic, 2003; Uncu, Kilic, and Turksen, 2004), etc.

Fuzzy clustering is also proven to be a fertile extension of the fuzzy set theory and widely used in data mining applications. In conventional fuzzy clustering, given a set of data points (n data points) and their associated coordinates, one is interested in grouping them into a specific number of clusters (c clusters). The objective is usually maximizing the compactness of the clusters, which is translated as minimizing the total within cluster distances. Bezdek's infamous Fuzzy C-Means (FCM) algorithm (Bezdek, 1980) is still among the most frequently used fuzzy clustering algorithm**.**

The objective function of FCM can be written mathematically as $\sum_{k=1}^{n}\sum_{j=1}^{c}\bar{u}_{jk}^{m}\, d_{jk}^{2}$ where d_{jk}^{2} is simply the square p-norm distance between data point k and centroid of the j^{th} cluster. A *degree of fuzziness* (weighing exponent) (*m*) is also included to the objective function that controls the level of fuzziness. The smaller the m (close to one) is, the less fuzziness is obtained and clustering becomes a non-fuzzy clustering, i.e., crisp clustering, while a high degree of fuzziness forces all membership to be equal to $\frac{1}{c}$ (total or extreme fuzziness). Note that FCM does not necessarily yield balanced clusters, as the rest of the conventional clustering algorithms since equality constraint is not part of the underlying mathematical model. Adding an "equality constraint" yields the following mathematical model:

$$Minimize \quad \sum_{k=1}^{n}\sum_{j=1}^{c} u_{jk}^{m} d_{jk}^{2} \tag{1}$$

$$Subject\ to:$$

$$\sum_{j=1}^{c} u_{jk} = 1 \qquad \forall k \in \{1,\dots,n\}, \tag{2}$$

$$\sum_{k=1}^{n} p_k u_{jk} = \frac{p}{c} \qquad \forall j \in \{1,\dots,c\}, \tag{3}$$

$$0 \le u_{jk} \le 1 \qquad \forall k \in \{1,\dots,n\}, \forall j \in \{1,\dots,c\} \tag{4}$$

where (1), (2) and (4) are fuzzy c-means algorithm's mathematical model and p_k is population of each data point and p is total population. Objective 1 minimizes total distance square errors to the cluster centers. Constraint 2 forces total membership of each data point to all cluster centers sum up to one. Note that in the context of zoning applications summation of membership degrees to 1 is desired for practical purposes. Constraint 3, which is balance constraint, guarantees that total population with respect to membership degrees within clusters will be equal (hence balanced). Note that the objective function (1) is non-linear since u_{jk} and $d_{jk}$are decision variables. In this paper, in order to solve this mathematical model, a Lagrangean relaxation approach is adopted. Constraint sets (2) and (3) are relaxed and added to objective function. By solving the relaxed problem (5) and (6) are obtained.

$$u_{jk}^{\frac{1}{m}} = \frac{d_{jk}^{\frac{2}{1-m}}}{\left(\frac{\lambda_k}{m} + \frac{\gamma_j p_k}{m}\right)^{\frac{1}{1-m}}} \tag{5}$$

$$\frac{\sum_{k=1}^{n} x_k u_{jk}}{\sum_{k=1}^{n} u_{jk}} = v_j \tag{6}$$

By solving (5) and (6) iteratively the fuzzy zones can be obtained. Algorithm 1 demonstrates the steps of this iterative process.

Algorithm 1 Fuzzy Zoning Algorithm

1: Define c and ε

2: Initialize Cluster Centers

3: **while** $||u_{jk}^{r+1}$-$u_{jk}^{r} \le \varepsilon||$ **do**

4: Calculate Euclidean Distance

5: Solve the system of equations and determine λ_k and γ_j

6: Use equation *5* and find u_{jk}

7: Use equation *6* and find v_j

8: **end while**

4. Experimental Analysis

In order to test the performance of the proposed algorithm, since there is no other method available which is applied to the discussed problem, a similar problem (with readily available tools that can be used for comparison) is considered as part of the proof of concept stage. Hence, the proposed fuzzy zoning algorithm is applied to a similar problem in the context of Wireless Sensor Networks (WSN). WSN consists of hundreds of thousands of autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, humidity, light, vibration, etc., equipped with data processing and communication units to pass data to a main location, i.e., Base Station (BS). WSN are used in many applications such as environmental monitoring, acoustic detection, seismic detection, inventory tracking, medical monitoring, smart spaces, military applications, etc. Note that in both problems (i.e., WSN and staff balancing discussed above) has an objective in common which is the total number of the objects in each cluster (sensors in WSN and health centers in staff balancing) which needs to be balanced.

In the literature there exist a number of cluster-based protocols proposed by various researchers. One of the most popular among them is Low Energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman, Chandrakasan, and Balakrishnan, 2000), which is a typical cluster-based protocol using a distributed cluster formation algorithm. According to this protocol, the CHs are selected with a predefined probability, other nodes select the closest cluster to join, which is identified based on the signal strength of the advertisement message they receive from the potential CHs. The CHs change over time among all the nodes in the network to save energy of the CHs because of high-traffic load in CHs. In order to compare the performance of the proposed fuzzy zoning algorithm, a protocol is developed as follows:

Algorithm 2 Fuzzy Zoning Based Protocol

1: Initialize energy of the network
2: **while** Number of rounds is less than maximum number of rounds **do**
3: Use Algorithm 1 to form clusters
4: Find ψ closest sensors to the centroids and choose one with the most residual energy
5: Update energy of network (both CHs and sensors)
6: **end while**

We have tested the performance of the fuzzy zoning based protocol in a simulated environment in MATLAB. An experimental setup is constructed with different number of data points, $n = 150$, $n = 200$ and $n = 300$, and their associated coordinates. Also different sized areas, namely, 300×300, 500×500 and 1000×1000 are considered. Two measures are considered in order to compare the performance of the algorithm with the commonly used LEACH protocol. These were namely, total number of alive sensors and total remained energy of network throughout a period. For each experimental condition *10* random replications were generated in the analysis and *15* rounds are assumed in each replication. The performance of the proposed algorithm with different levels of ψ (the parameter used in the proposed algorithm) was also analyzed during the experiments.

Table 1 illustrates the total number of alive sensors related with different data sets ($N = 150$, *200* and *300*). Table 2 illustrates total remained energy of network. Results in the tables are the averages of the rounds. The results suggest that the fuzzy zoning based protocol outperforms the LEACH in the experiments for different levels of ψ. On the other hand, as ψ increases total number of alive sensor also increases, but same relation is not true between ψ and total remained energy which indicates that the energy of the remaining sensors is not identical.

Figure 1, depicts the fate of the network with $N=150$ for the total remained energy for the proposed algorithm with ψ=4 and LEACH. The results suggest that the fuzzy zoning algorithm yields a smoother reduction throughout the rounds.

Table 1: Average number of alive sensors.

N	LEACH	Zoning			
		ψ=2	ψ=3	ψ=4	ψ=5
150	0	6	11	13	16
200	0	19	27	32	46
300	0	34	36	39	44

Table 2: Average total remained energy of the network.

N	LEACH	Zoning			
		ψ=2	ψ=3	ψ=4	ψ=5
150	0	0.946887	1.51753	1.932424	2.222152
200	0	8.093512	6.610239	7.037405	7.660732
250	0	7.309429	8.434438	9.13621	6.188448

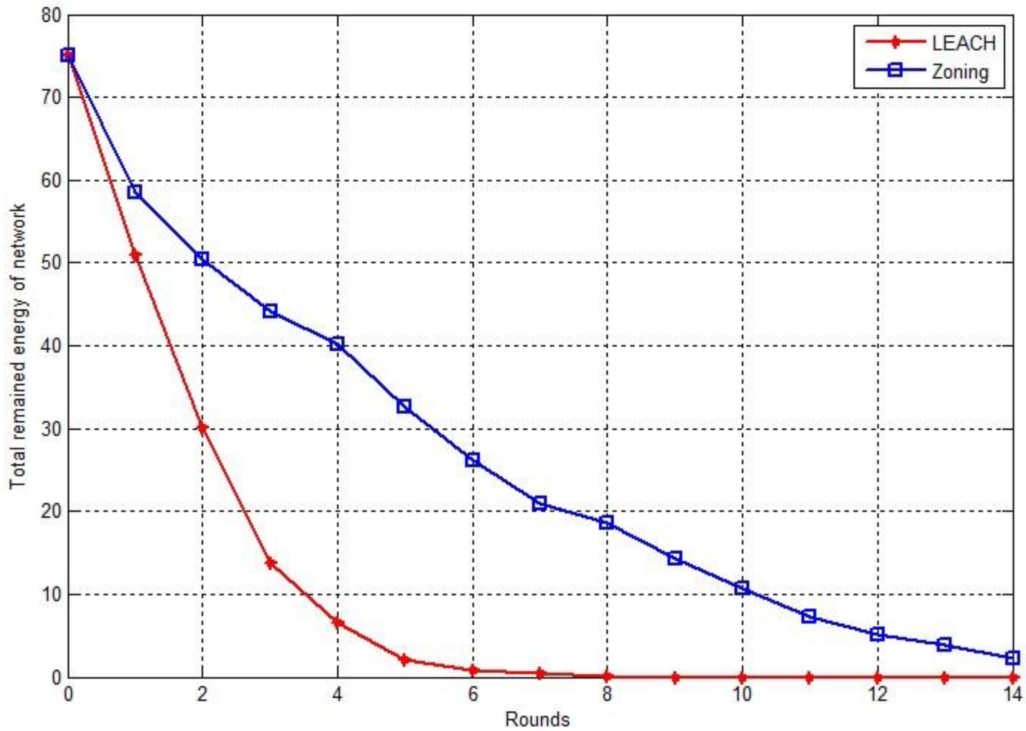

Fig. 1. Total remained energy of network, with *n = 150* and ψ=4.

5. Future Research Agenda

The results of the experimental analysis revealed that the developed fuzzy zoning algorithm works well for a problem where the total traveling within cluster distance and the balance among the clusters is important at the same time. Next thing in our research agenda is applying the proposed approach to the real problem that the Company is facing and comparing its performance with other zoning approaches available in the literature. In order to make such a comparison actual day time travel information is required. Even though it is always possible to assume that the traveling time between the facilities (hard to determine) is proportional to the distance (easier to determine), however, this is not the case particularly in cities with heterogeneous traffic conditions like Istanbul. Note that, the very same algorithm can also be used for the daily task assignment of each employee. The only difference is rather than determining the zones for each employee, each day can be treated as a zone. This is also left as part of the future research.

References

1. Baron, O., O. Berman, D. Krass, and Q. Wang. *The equitable location problem on the plane*. European Journal of Operational Research 183 (2), 2007, 578–590.
2. Bezdek, J. C. *A convergence theorem for the fuzzy ISODATA clustering algorithms*. IEEE Transactions on Pattern Analysis & Machine Intelligence (1), 1980, 1–8.
3. Bozkaya, B., E. Erkut, and G. Laporte.*A tabu search heuristic and adaptive memory procedure for political districting*. European Journal of Operational Research 144 (1), 2003, 12–26.
4. Garfinkel, R. S. and G. L. Nemhauser. *Optimal political districting by implicit enumeration techniques.* Management Science 16 (8), 1970, B–495.
5. Hess, S. W., J. Weaver, H. Siegfeldt, J. Whelan, and P. Zitlau. *Nonpartisan political redistricting by computer.* Operations Research 13 (6), 1965, 998–1006.
6. Hojati, M. "Optimal political districting." Computers & Operations Research 23 (12), 1996, 1147–1161.
7. Kilic, K., B. A. Sproule, I. B. Türksen, and C. A. Naranjo. *Pharmacokinetic application of fuzzy structure identification and reasoning.* Information Sciences 162 (2), 2004, 121–137.
7. Nikolakopoulou, G., S. Kortesis, A. Synefaki, and R. Kalfakakou. *Solving a vehicle routing problem by balancing the vehicles time utilization.* European Journal of Operational Research 152 (2), 2004, 520–527.
8. Pavone, M., A. Arsie, E. Frazzoli, and F. Bullo. *Distributed algorithms for environment partitioning in mobile robotic networks.* Automatic Control, IEEE Transactions on 56 (8), 2011, 1834–1848.
9. Rios-Mercado, R. Z. and E. Fernandez. *A reactive GRASP for a commercial territory design problem with multiple balancing requirements.* Computers & Operations Research 36 (3), 2009, 755–776.
10. Salazar-Aguilar, M. A., R. Z. Rios-Mercado, and J. L. Gonzalez-Velarde. *A bi-objective programming model for designing compact and balanced territories in commercial districting.* Transportation Research Part C: Emerging Technologies 19 (5), 2011, 885–895.
11. Takagi, T. and M. Sugeno. *Fuzzy identification of systems and its applications to modelling and control.* Systems, Man and Cybernetics, IEEE Transactions on (1), 1985, 116–132.
12. Uncu, O., I. Turksen, and K. Kilic. *Localm-fsm: A new fuzzy system modeling approach using a two-step fuzzy inference mechanism based on local fuzziness level.* In: Proceedings of international fuzzy systems association world congress, 2003, pp. 191–194.
13. Uncu, O., K. Kilic, and I. Turksen. *A new fuzzy inference approach based on Mamdani inference using discrete type 2 fuzzy sets.* In: Systems, Man and Cybernetics, 2004 IEEE International Conference on. Vol. 3. IEEE, pp. 2272–2277.
14. Zadeh, L.. *Fuzzy sets.* Information and Control 8 (3), 1965, 338–353.
15. Zhou, G., H. Min, and M. Gen. *The balanced allocation of customers to multiple distribution centers in the supply chain network: a genetic algorithm approach.* Computers & Industrial Engineering 43 (1), 2002, 251–261.