# Fuzzy Zoning: A Lagrangean Relaxation Approach

Siamak Naderi
Faculty of Engineering and
Natural Sciences
Sabancı University
Istanbul, Turkey
Email: Siamak@sabanciuniv.edu

Kemal Kılıç
Faculty of Engineering and
Natural Sciences
Sabancı University
Istanbul, Turkey
Email: kkilic@sabanciuniv.edu

*Abstract*—This research arises from the need of equality in real life problems. Clustering algorithms are being used in many applications where equality is an interest, such as districting (either zonal or political) and industry (distribution companies). One of the well known clustering algorithms is Fuzzy clustering. We add an equality constraint to the existing model. We call the new problem "Zoning" problem. One of the application where equality can play a critical role is Wireless Sensor Network. A Lagrangean relaxation based approach is developed to solve Zonnig problem. The proposed algorithm is simulated and the results show robust performance regarding the equality of the clusters.

*Index Terms*—Zoning, Lagrangean Relaxation, Fuzzy Clustering, Optimization

## I. INTRODUCTION

Clustering is the identification of natural groups (i.e., clusters) in an environment. "Natural grouping" has different definitions in various contexts. In demographics, graph theory and Data Mining (DM), clusters usually refer to groups of objects that are similar to each other and different from others in terms of a similarity metric.

One of the most important issues that in the recent years has attracted a great deal of attention is equality (i.e; balance) within clusters. Equality can be defined in various contexts. For example, governments aim to assign equal resources to different groups of people (budget equality). Another example is in industrial companies seek to assign equal workloads among their employees or assign equal resources to the different regions they cover. Consider a distribution company which serves a city through the multiple vehicles. A big concern in such companies is assigning an equal workload to each vehicle. In some cases, equality is targeted for the purpose of fairness as mentioned in either governmental applications or industries. On the other hand, in some applications, equality provides optimization advantages. Wireless sensor network is one of the applications where equality must be taken into account because of its effect on optimizing the total cost.

## II. RELEVANT LITERATURE

Fuzzy set theory has received extensive attention from research community since it is introduced by Zadeh [1] and has been successfully applied in various fields such as control theory (Takagi and Sugeno [2]), health care (Kilic *et.al* [3]), system modeling (Uncu, Turksen, and Kilic [4]; Uncu, Kilic, and Turksen [5]), etc. Clustering is a field that uses the the basis of fuzzy sets. In the literature such clustering methods are called fuzzy clustering algorithms.

The most famous fuzzy clustering algorithm is *Fuzzy C-Means* (FCM) algorithm [6]. Bill Wee is the first researcher who worked on fuzzy pattern recognition in his Ph.D. thesis. Woodbury and Clive [7] combined fuzziness and probability in hybrid fuzzy clustering and at the same time, Dunn [8] and Bezdek [9] published their work on FCM model.

In the classical FCM clustering problem, given a set of data points ($n$ data points) and their associated coordinates, one is interested in grouping them in a specific number of clusters ($c$ clusters). Objective is maximizing the compactness of the clusters which leads us to minimize the total within distances of the data points in clusters. This objective function is considered as *least squares model* based on [10]. However, FCM does not guarantee that the clusters created have equal number of objects.

*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 *et al.* [11]; Garfinkel and Nemhauser [12]; Hess *et al.* [13]; Hojati [14]).

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 *et al.* [15] 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.* [16] 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.* [17] 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 [18] 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 [19] 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 under utilization 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.* [20] 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.

## III. PROPOSED FUZZY ZONING ALGORITHM

Fuzzy clustering is 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 [9]) is still among the most frequently used fuzzy clustering algorithm (Xu and Wunsch [21]).

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 ($d_{jk} = (v_j - x_k)^2$, where $x_k$ is the coordinates of data points and $v_j$ is the centroid of each cluster). A *degree of fuzziness* (weighting 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:

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

subject to:

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

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

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

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. Solving constrained non-linear problems are hard and time consuming however, Lagrangean relaxation works well in the case of unconstrained non-linear problems. 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. The relaxed problem is an unconstrained problem with continuous objective function. By taking derivative with respect to decision variables and lagrangean multipliers and make them equal to zero equations (5) and (6) are obtained. Note that taking derivative with respect to lagrangean multipliers let us obtain constraints (2) and (3).

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

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

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

Note that $\lambda_k$ and $\gamma_j$ are lagrangean multipliers associated to constraints (2) and (3) respectively.

## IV. EXPERIMENTAL ANALYSIS

In order to test the performance of the proposed algorithm 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). WSNs are used in many applications such as environmental monitoring, acoustic detection, seismic detection, inventory tracking, medical monitoring, smart spaces, military applications, etc [22].

Since sensor nodes are powered by limited energy source like battery, *energy conservation* is considered to be the most important feature in order to keep the connectivity and operation of the network. Grouping sensors to the clusters and assigning a cluster head (CH) to each cluster can save energy as each single sensor communicates only with the associated CH and the CH transmits the information to Base Station (BS) after processing the data. Note that consumption of energy has a direct relation with distance, that is to say, the farther the CH is to the sensor, the more energy sensor must consume to communicate with it. Since transmitting

---

**Algorithm 1** Fuzzy Zoning Algorithm

1: Define $c$ and $\epsilon$
2: Initialize Cluster Centers
3: **while** $||u_{jk}^{r+1} - u_{jk}^{r} \leq \epsilon||$ **do**
4:     Calculate Euclidean Distances
5:     Solve the system of equations and determine $\lambda_k$ and $\gamma_j$.
6:     Use equation 5 and find $u_{jk}$
7:     Use equation 6 and find $v_j$
8: **end while**

---

**Algorithm 2** Fuzzy Zoning Based Protocol

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

---

data to BS by each single sensor is energy costly, clustering helps energy conservation.

In the literature there exist a number of cluster-based protocols proposed by various researchers. One of the most popular protocols among them is *Low Energy Adaptive Clustering Hierarchy* (LEACH) (Heinzelman *et al.* [23]), 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 in Algorithm 2. According to proposed protocol, in each round, clusters are created based on fuzzy zoning algorithm, developed in this research. Furthermore, CHs are selecting among $\psi$ closest sensors to the centroid of clusters ($v_j$). Among these sensors the one with the highest residual energy is selected. This idea helps finding CHs with high level of energy that have minimum distances to the other sensors in the cluster.

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, $300 \times 300$, $500 \times 500$ and $1000 \times 1000$ 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. Note that, a round is an iteration where clusters are made, CHs are selected and signals are sent from sensors to the CHs and from there to BS. The performance of the proposed algorithm with different levels of $\psi$ was also analyzed during the experiments.

Table I illustrates the total number of alive sensors related with different data sets ($N = 150, 200$ and $300$). Table II illustrates total remained energy of network. The results
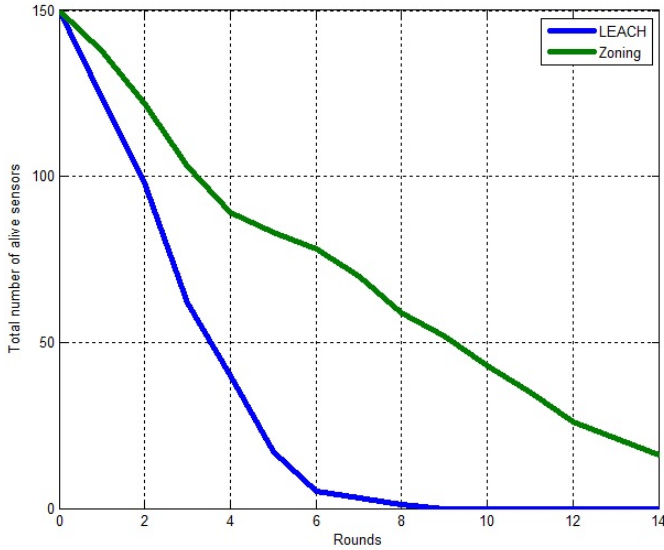
**Fig. 1:** Number of alive sensor: LEACH vs Zoning based protocol, $N = 150$, and $\psi = 5$



**Fig. 2:** Remained energy of network: LEACH vs Zoning based protocol, $N = 150$, and $\psi = 5$

suggests that the fuzzy zoning based protocol outperforms the LEACH in the experiments for different levels of $\psi$. On the other hand, as $\psi$ increases, total number of alive sensors also increases since rather than considering only the distance to the centroid or only the level of residual energy, both are considered. Conversely, by increasing $\psi$, total remaining energy of network does not increase necessarily. In all cases, LEACH protocol failed after 6 rounds of simulation.

Figure 1 and 2, depict the fate of the network with $N = 150$ for two measures (i.e., number of alive sensors and total remained energy) and compare it with LEACH. The results shows that the fuzzy zoning algorithm yields a smoother reduction of energy throughout the rounds.

**TABLE I:** Average number of alive sensors.

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

**TABLE II:** Average total remained energy of the network.

| N | LEACH | Zoning | | | |
|---|---|---|---|---|---|
| | | $\psi=2$ | $\psi=3$ | $\psi=4$ | $\psi=5$ |
| 150 | 0 | 0.946887 | 1.51753 | 1.932424 | 2.222152 |
| 200 | 0 | 8.093512 | 6.610239 | 7.037405 | 7.660732 |
| 300 | 0 | 7.309429 | 8.434438 | 9.13621 | 6.188448 |

## V. FUTURE RESEARCH AGENDA

The results of the experimental analysis revealed that the developed fuzzy zoning algorithm works well for a problem where equality is a target. In WSNs, equality lets the network last more since the reduction of energy happens smoothly.
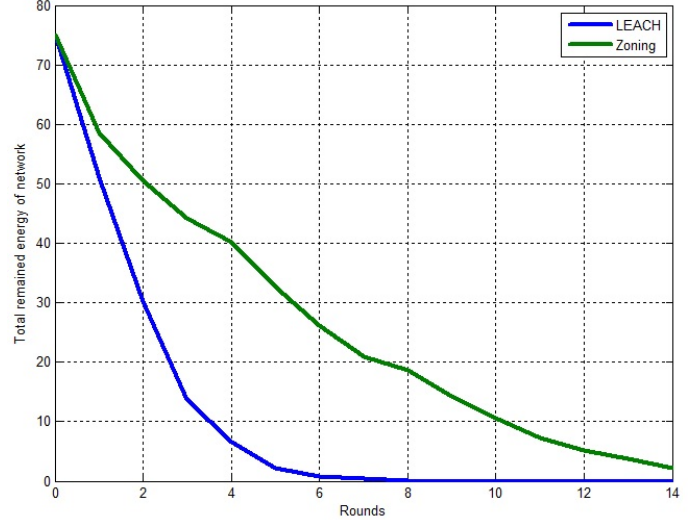
Additionally, in the developed algorithm, we allow each sensor to communicate with more than a single cluster head. This causes lets cluster heads to save more energy. The assignment of sensors to more than one cluster head is a natural result of using fuzzyness.

Fuzzy Zonnig is compared with LEACH algorithm, in a simulated environment. Fuzzy zonnig outperformed LEACH with respect to two criteria, total remained energy of network and number of remained nosed.

For the future research, we would be interested to apply fuzzy zonnig to other application such as staff balancing problem.

## REFERENCES

[1] L. Zadeh, *Fuzzy sets*, Information and Control, 8, (3), 338-353, 1965.

[2] T. Takagi, and M. Sugeno *Fuzzy identification of systems and its applications to modelling and control.* Systems, Man and Cybernetics, IEEE Transactions on (1), 116-132, 1985.

[3] K. Kilic, B. A. Sproule, I. B. Trksen, and C. A. Naranjo. *Pharmacokinetic application of fuzzy structure identification and reasoning.* Information Sciences, 162, (2), 121-137, 2004

[4] O. Uncu, I. B. 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. 191-194, 2003.

[5] O. Uncu, K. Kilic, and I. B. Turksen. *"A new fuzzy inference approach based on Mamdani inference using discrete type 2 fuzzy sets.* In: Systems, Man and Cybernetics, IEEE International Conference on, 3, 2272-2277, 2004

[6] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms.* Kluwer Academic Publishers, 1981.

[7] M. A. Woodbury, and J. Clive. *Clinical pure types as a fuzzy partition.* 111-121, 1974.

[8] J. A. Dunn. *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.* 32-57, 1973.

[9] J. C. Bezdek, *A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2, (1), 1-8, 1980.

[10] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms.* Kluwer Academic Publishers, 1981.

[11] B. Bozkaya, E. Erkut, and G. Laporte *A tabu search heuristic and adaptive memory procedure for political districting*, European Journal of Operational Research, 144, (1), 12-26, 2003.

[12] R. S. Garfinkel, and G. L. Nemhauser *Optimal political districting by implicit enumeration techniques.* Management Science, 16, (8), 1970.

[13] S. W. Hess, J. Weaver, H. Siegfeldt, J. Whelan, and P. Zitlau. *Nonpartisan political redistricting by computer.* Operations Research, 13, (6), 998-1006, 1965.

[14] M. Hojati. *Optimal political districting.* Computers & Operations Research, 23, (12), 1147-1161, 1996.

[15] M. A. Salazar-Aguilar, 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), 885-895, 2011.

[16] M. A. Pavone, Arsie Frazzoli, and F. Bullo. *Distributed algorithms for environment partitioning in mobile robotic networks.* Automatic Control, IEEE Transactions on, 56, (8), 1834-1848, 2011.

[17] G. Nikolakopoulou, 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), 520-527, 2004.

[18] R. Z. Rios-Mercado, and E. Fernandez. *A reactive GRASP for a commercial territory design problem with multiple balancing requirements.* Computers & Operations Research, 36, (3), 755-776, 2009.

[19] G. H. M. Zhou 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), 251-261, 2002.

[20] O. Baron, O. Berman, D. Krass, and Q. Wang. *The equitable location problem on the plane.* European Journal of Operational Research, 183, (2), 578-590, 2007.

[21] Rui Xu, D. Wunsch. *Survey of clustering algorithms.* Neural Networks, IEEE Transactions on, 16, (3), 645-678, 2005.

[22] S. Bandyopadhyay and E. J. Coyle *An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks.* INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies,(3), 2003.

[23] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. *Energy-Efficient Communication Protocol for Wireless Microsensor Networks.* Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000.