Clustering Approach Using Artificial Bee Colony Algorithm for Healthcare Waste Disposal Facility Location Problem

Zeynep Gergin, PhD
Assistant Professor, Istanbul Kültür University, Department of Industrial Engineering, E5 Yanyol, Yenibosna, 34156, Istanbul, Turkey
z.gergin@iku.edu.tr; +902124984456

Nükhet Tunçbilek
PhD Candidate, Istanbul University, Department of Industrial Engineering, Faculty of Engineering, A Blok, Kat:5, Avcilar, Istanbul, Turkey
nukhettuncbilek@gmail.com

Şakir Esnaf , PhD
Professor, Istanbul University, Department of Industrial Engineering, Faculty of Engineering, A Blok, Kat:5, Avcilar, Istanbul, Turkey
sesnaf@istanbul.edu.tr; +902124737070

ABSTRACT

In this study, Artificial Bee Colony (ABC) based clustering algorithm is proposed for solving continuous multiple facility location problems. Unlike original version applied to multivariate data clustering, the ABC based clustering here solves the two-dimensional clustering. On the other hand, the multiple facility location problem the proposed clustering algorithm deals with is aimed to find site locations for healthcare wastes. After applying ABC based clustering algorithm on test data, a real world facility location problem is solved for identifying healthcare waste disposal facility locations for Istanbul Municipality. Geographical coordinates and healthcare waste amounts of Istanbul hospitals are used to decide the locations of sterilization facilities to be established for reducing the medical waste generated. ABC based clustering is performed for different number of clusters predefined by Istanbul Metropolitan Municipality, and the total cost - the amount of healthcare waste produced by a hospital, multiplied by its distance to the sterilization facility - is calculated to decide the number of facilities to be opened. Benchmark results with four algorithms for test data and with two algorithms for real world problem reveal the superior performance of the proposed methodology.

Keywords: Artificial Bee Colony, Clustering Algorithms, Healthcare Waste Disposal Facility Location, Real World Problem

INTRODUCTION

Facility Location Problem (FLP) is defined by Tavakkoli and Shayan (1998) as "locating *n* facility to *m* locations ($n<m$) so as to minimize the transportation costs". These problems consider identifying the places of the facilities to satisfy customers' demand and, assigning every customer to a specific facility location under defined constraints. These constraints define the specific characteristic of the FLP (Daskin, 1995);

- Whether the facility locations are selected from a finite/infinite set of possible locations; continous/discrete problems
- Whether the facility capacities are limited; capacited/uncapacited problems
- Whether the facilities to be opened are singular or plural; multi-facility problems
- Whether the facility demands are static or dynamic; problems with time periods
- Whether the facility demands are stable or subject to change; deterministic/probabilistic problems
- Whether the product is singular or plural; multi-product problems
- Whether the problem has one/more objectives; multi-objective problems
- Whether the problem is hierarchic; multi-stage problems

The use of clustering methods, that aim to separate and group elements in a data space depending on similarity, is seen for location selection problems in literature due to the four components that characterize the location problems (ReVelle & Eiselt, 2005); (1) customers located on points, (2) facilities to be located, (3) a metric that indicated distances between customers and facilities, and (4) a space in which customers and facilities are located. The facilities that are to be located are assumed to be desirable in the sense that the closer they are to the customers, the better the value of the objective function. Yet the location of healthcare waste disposal facilities may have different conditions based on the treatment system that are thermal processes, chemical processes, biological processes, mechanical processes, and irradiation technologies respectively. The choice of treatment system involves various factors such as, waste characteristics, quantity of wastes for treatment and disposal, technology capabilities, environmental and safety factors, public acceptability, regulatory requirements and costs – many of which depend on local conditions and consequent decisions (World Health Organization, 2014). The most established waste treatment technologies focus on disinfection that is realized with thermal processing in autoclaves. Autoclaves have the advantage of being designed in various capacities that are suitable to fit every area preferred, and every source generating medical waste can sterilize its own load. On the other hand, since the handling of large amounts of medical waste require special precautions, in big cities municipalities manage the collection and treatment of healthcare waste, and local healthcare waste disposal facility locations are identified and facilities with required capacities are established for treatment.

In this study, Artificial Bee Colony (ABC) clustering algorithm is exploited for an uncapacitated continuous multiple facility location problem for a healthcare waste disposal facility. The contributions of our paper are as follows. In this paper, Artificial Bee Colony (ABC) based clustering algorithm is used for solving continuous multiple facility location problems for the first time. This is the first contribution. Second contribution is that the proposed algorithm is applied for medical waste disposal facility location. According to our knowledge this is the first study in the healthcare management literature.

The paper is organized as follows; in Section 2 the mathematical definition of the uncapacitated continuous multiple facility location problem (MFLP) subject to this study is given. Literature review on the application of clustering algorithms for FLP is given in Section 3. Section 4 describes the general information about the ABC algorithm and its application for clustering. In Section 5, initially the results of the benchmark tests applied with library data to evaluate the performance of ABC clustering are reported. This section then covers the case study for a real world problem; alternative healthcare waste disposal facility locations are evaluated for Istanbul and the results are explained and summarized with tables. The experimental results are both giving the performance of clustering method using ABC algorithm, and the alternatives of sterilization facility location and capacities for the company of Istanbul Municipality that is responsible for healthcare waste collection and disposal. Four benchmark algorithms are also applied to the real world data. Finally, the last section is for concluding the findings of the research.

PROBLEM DEFINITION

In this research a continuous uncapacitated multiple facility location problem (MFLP) is studied. Continuous MFLPs are concerned with determining the location or coordinates of *c* facilities in a plane to serve *n* customers having fixed locations. The solution process of the MFLP is to find an optimal solution that satisfies all customer demand and minimizes the total cost. There is no limit of capacities for any facility, and whole demand of each customer has to be assigned to one of the facilities.

This problem can be formulated as a mathematical model with the following equations (Esnaf & Küçükdeniz, 2013);

$$\min_{\overline{v_1},\overline{v_2},\dots,\overline{v_m}} \sum_{i=1}^{m} \sum_{x_k \in V_i} w_k d(\bar{x}_k, \bar{v}_i)\, z_{ik} \tag{1}$$

$$s.t. \sum_{j=1}^{n} z_{ik} = 1 \qquad \textit{for } i = 1, \dots, m \tag{2}$$

$$z_{ik} = \{0, 1\} \qquad \textit{for } i = 1, \dots, m \textit{ and } k = 1, \dots, n \tag{3}$$

$$x, y, p, q \in R \tag{4}$$

where,

$\bar{x}_k = (x_k, y_k)$=The location of customer *k* in a plane, *k=1,2...n*
w_k =The demand of customer *k*, $w_k > 0$, *k=1,2...n*
$\bar{v}_i = (p_i, q_i)$ = The location or center of facility *i*
V_i = Cluster of the customer that is assigned to the i^{th} facility
z_{ik} : the binary variable used for the road from customer *k* to facility *i*
$d_k(\bar{x}_k, \bar{v}_i)$= Distance between the facility *i* and customer *k*, which is Euclidean distance and formulated as follows:
$d_k(\bar{x}_k, \bar{v}_i) = \sqrt{(x_k - p_i)^2 + (y_k - q_i)^2}$

LITERATURE REVIEW

Tan et al. (2005) defines the basic aim of clustering as the separation and grouping of elements in a data space depending on similarity. Clustering methods try to maximize within-cluster resemblance and between cluster dissemblance. Due to the features of facility location problems that are listed below, clustering algorithms can be applied for facility location problems (ReVelle & Eiselt, 2005;);

(1) Customers located on various locations,
(2) Candidate facilities to be located on various locations,
(3) A metric that indicated distances between customers and facilities, and
(4) A space in which customers and facilities are located, and that each customer should be assigned to a facility.

Similarly Esnaf & Küçükdeniz (2013) suggests modelling continuous multiple facility location problems by portraying the following characteristics;

(1) The facilitied can be located to any coordinate, and the optimum location is found iteratively.
(2) There is no transportation (arc) allowed between the facilities.
(3) Each customer is served by only one facility, and the demand cannot be splitted to more than one facility.
(4) Transportation cost is assumed to be proportional with Euclidean distance.

(5) Each customer is assigned to the nearest facility.
(6) Setup costs are not considered.
(7) The customers' coordinates on the continuous space and their demands are constant.

Clustering Applications for Facility Location Problems

The application of clustering methods for MFLP is initially seen with the study of Franca et al. (1999). The writers partitioned a given set of customers with distinct demands into p clusters with limited capacities. Hsieh & Tien (2004), proposed a heuristic method using Kohonen's feature maps to solve uncapacitated location problems. Sheu et al. (2005), used fuzzy α-cut clustering algorithm for grouping 50 pre-determinded disaster areas into five clusters to identify the locations of disaster-relief points. The work of Zalik (2006) initially introduced Fuzzy C-Means (FCM) algorithm for location problems, and compared its performance with classic c-means. Zalik suggested the applicability of the fuzzy algorithm for FLP. Esnaf & Küçükdeniz (2009) solved an uncapacitated MFLP with a fuzzy logic clustering approach combined with Center-of-Gravity (COG) method. In this study, initially the demand points are clustered, and then facility locations in the clusters are calculated with center of gravity algorithm. They applied their two-step methodology for a real World problem. Kashan et al. (2012) used a modified ABC algorithm, named DisABC for an uncapacited FLP. Küçükdeniz et al. (2012) integrated FCM and convex programming for solving a capacitated multi-facility location problem. The writers applied the proposed method to both OR library and real world data, and concluded that the proposed method outperforms original FCM in terms of transportation costs. Tunçbilek et al. (2012) solved a discrete uncapacited FLP with ABC algorithm, and concluded the superior performance of the method over Particle Swarm Optimization (PSO) algorithm for a network of 768 demand points and 11 facilities. The work of Esnaf et al. (2014) proposed a single-iteration version of FCM algorithm for solving uncapacitated facility location problems (UFLPs). The researchers tested and compared the results with PSO and ABC algorithms for various UFLPs such as discrete, continuous, discrete with local search and continuous with local search. They declared that the proposed algorithm's performance surpassed the PSO-based and ABC-based algorithms based on the results obtained from real life application. Mohrechi & Hatamlou (2015) applied ABC algorithm to a discrete FLP for locating emergency medical centers. They also applied benchmarks tests with the case data, and concluded the better performance of the ABC algorithm compared with the results of genetic algorithm and PSO. Li et al. (2017), proposes a hybrid discrete artificial bee colony (HDABC) algorithm for solving the large scale location allocation problem in reverse logistics network system. The benchmarks implemented with GA, MFOA and classical ABC algorithms.

Facility Location Problems for Hazardous Waste Disposal

With the increase on environmental concerns, studies on locating disposal or recycling facilities are seen in the literature, however few are on healthcare waste. Cappanera et.al. (2004) have proposed a discrete combined location-routing problem for all obnoxious waste. Alumur & Kara (2005), has considered not medical but all other types of hazardous waste disposal by proposing a model for selecting treatment technologies and disposal locations. Among the rare studies on healthcare waste disposal, Alagöz & Kocasoy (2008) have evaluated different scenarios on the medical waste collection problem of Istanbul. They proposed one central sterilization system, and routed the transportation of waste to this disposal facility. Li-hong (2009) presents a mathematical model for optimizing the medical waste reverse logistic network. This study discusses the location of collecting and processing centers as well as transportation costs. Shanmugasundaram et.al. (2016) have developed a Geographic Information System (GIS) based model for locating centralized healthcare waste treatment facility and a route optimization in Lao. Hariz, Dönmez and Sennaroğlu (2017) employed a two stage analysis to identify a suitable location or handling, and disposal of health care waste in Kenya. Initially they used Geographic Information System (GIS) to eliminate unsuitable land, and then, Multi-Criteria Decision Analysis (MCDA) methods are used to analyse and rank the potential sites. Thakur and Ramesh (2017) have discussed and proposed a model for healthcare waste disposal strategy selection by using grey theory based AHP approach. The authors have also implemented the proposed model for an Indian case. Yılmaz, Kara and Yetiş (2017), have conceptualized a multi-objective mixed integer location/routing model for hazardous waste and implemented for a case study in Turkey. Wichapa and Khokhajaikiat (2018) solved a facility location problem model which hybridizes fuzzy AHP and goal programming (GP) to select new suitable locations for infectious waste disposal in Upper-Northeastern Thailand among six candidate municipalities.

Clustering Algorithms for Waste Disposal Facility Location

Studies on using clustering algorithms for waste disposal facility location are also quite scarce, and none of them concern healthcare waste; Negresios & Palhano (2006) studied a two-phase clustering algorithm to be applied to the design of garbage collection zones. Gomes et al. (2007) proposed a solution for recycling facility location with Self Organizing Maps (SOM) and Fuzzy C-Means (FCM) algorithms. Ayoub et al. (2007) used a simulation model for deciding the locations of biomass collection points. The data used in the simulation model are derived by running a FCM clustering algorithm. Büyüksaatçi et al. (2008), identified the optimum locations of the recycling facilities for an asphalt company under capacity, demand and geographical location constraints. They used a hybrid model of Gustaffson-Kessel Fuzzy clustering algorithm and convex programming. Zhang & Lee (2013), proposed the use of ABC algorithm for FLP of collection centers aiming to minimize the total logistics costs, and stated that the performance of ABC algorithm proved it is efficiency for design of reverse logistics network. The study of Gergin & Esnaf (2013) compared the performance of four clustering algorithms by solving a real by solving a real multi-facility location problem. This study integrated center of gravity (COG) method with Self Organizing Maps (SOM), and concluded that the integration of COG method to FCM and SOM algorithms improved the performance of the clustering approach of the FLP studied. Esnaf et al. (2014) proposed Single-Iteration Fuzzy C-Means (SIFCM) clustering algorithm in order to assign the demands points to facilities when locations were already defined. This study also showed the superior performance of the proposed algorithm on linear programming and ABC algorithm with real world data. There is only one study uses fuzzy clustering for medical waste disposal location by Gergin and Esnaf (2013).

METHODOLOGY

The novel approach of this study is the application of Artificial Bee Colony based clustering algorithm for healthcare waste disposal facility location. The following sections introduce the original ABC method and its adaptation for clustering.

Artificial Bee Colony Algorithm

Artificial Bee Colony (ABC) algorithm is one of most recently introduced swarm intelligence based meta-heuristic method by Derviş Karaboğa in 2005. The algorithm has been motivated by the intelligent behaviour of honeybees. This algorithm is as simple as Particle Swarm Optimization (PSO) and Differential Evolution (DE) algorithms. (Karaboğa, 2005)

Karaboğa (2005) proposed the artificial bee colony (ABC) algorithm that is inspired by foraging behaviour of honey bees. In the ABC algorithm, a problem is solved by exploring good solutions, which are represented as food sources. The quality of the solution is represented by the nectar amount of that food source. In this algorithm, the first half of the bee colony are employed bees, the second half are the onlooker bees. The number of food sources is same as the number of employed bees. It is assumed that there is only one employed bee for every food source. Each employed bee is placed on a food source, and starts extracting nectar. The employed bee becomes a scout when the food source has no more nectar and moves away to look for another food source. As soon as a scout bee finds a new food source it again becomes an employed bee. The ABC algorithm initially places all employed bees on randomly generated food sources (solutions). Then iteratively, every employed bee determines a food source nearby their currently associated food source and evaluates its nectar amount (fitness). If it has more nectar than that of its current food source, then that employed bee moves to this new food source, otherwise it stays on its current food source.

The search can be materialized with following steps:

- Employed bees locate a food source close to current food source, which is in their memory.

- The other half of the colony, onlooker bees waits in the hive and get information about rich food sources from employed bees, which returned into the hive. Then the onlookers decide to go to one of the food sources and locate to a food source, which is close to this food source.
- After some period, food source may be exhausted. An employed bee on such a food source becomes a scout and starts to search a new food source randomly.

The ABC algorithm is presented below:

Initialize food sources at random positions and locate employed bees
Repeat
Move the employed bees around their food sources and determine their nectar amounts.
Move the onlookers towards rich food sources and determine their nectar amounts.
Determine exhausted food sources and assign employed bees as scout bees for searching new food sources.
Memorize the best food source found so far.
until total number of cycles executed

This cycle is repeated up to predefined number of iterations or predefined limit on CPU time. A food source can be interpreted as a possible solution to the optimization problem. The nectar amount of a food source represents the quality of the solution represented by that food source. Scout bees move to new directions so that colony can explore new food sources. While onlookers and employed bees exploiting good solutions in the search space, the scouts explore new unknown solutions (Karaboğa & Baştürk, 2007). The algorithm is exploited with the following four steps;

Step 1: Initial solutions x_{ij}, i=1,…,N, j=1,…,M are set with randomly generated values;

$$x_i^j = x_{min}^j + rand[0,1]\left(x_{max}^j - x_{min}^j\right) \quad (6)$$

Step 2: Employed bees examines their neighbourhood by calculating new food source positions.
Movement of an employed bee around its current position is probabilistically formulated. Scanning food sources in the neighbourhood of a particular food source is done by altering the value of one randomly chosen solution parameter (dimension j) and keeping other parameters unchanged. The value of the chosen parameter is changed by using the following formula:

$$v_{ij} = x_{ij} + \varphi_{ij}(x_{ij} - x_{kj}) \quad (7)$$

where

j : 1,…,D randomly chosen dimension
v_i : new candidate location
x_i : current location
ϕ_{ij} : random factor between -1 and +1 which is generated by uniform probability distribution
k : a randomly chosen neighbour, where k is not equal to i
x_k : location of a randomly chosen neighbour (k) at chosen dimension j

If the calculated value v_{ij} exceeds the acceptable range for dimension j, it is set to the corresponding extreme value in that range.
Fitness values for x_i and v_i are compared and the one with better fitness value is chosen as the new position for i^{th} employed bee. This is a greedy selection process. If nectar amount of the candidate location is better than the present one, the bee forgets the present location and memorizes the candidate location. Otherwise, the bee keeps its present location in the memory.

Step 3: Onlooker bees follow the information given by returning employed bees.
Onlooker bees move according to the information taken from an employed bee that is returned back to hive. When employed bees have finished collecting nectar, they come back to their hive and share information with the onlooker bees by dancing longer or shorter, according to nectar amount of the last visited food source. Onlooker bees select a food source according to a probability which is proportional to the nectar amount of that food source. The probability p_i of selecting a food source i is determined using the following expression:

$$P_i = \frac{fitness(i)}{\sum_{n=1}^{SN} fitness(n)} \tag{8}$$

where
fitness(i) : fitness value of i^{th} solution which represents nectar amount at food source at i^{th} position.
SN : total number of employed bees (also number of food sources).

Step 4: Employed bees become scout bees if their fitness values cannot be improved.
If a particular employed bee does not improve the solution in a predefined number of iterations called "food limit", then that employed bee becomes a scout bee by leaving its current position and looking for a new food source at a randomly set position.
Step 2, 3 and 4 are repeated until predefined total number of cycles are executed.

Application of ABC Algorithm to Clustering Problem

In this study, ABC algorithm is used to solve uncapacitated continuous MFLPs as a clustering problem. The aim is to minimize the total distances between candidate cluster centers and member points. Cluster centers are represented by X_i are moved by ABC algorithm and fitness value is calculated by the following expression:

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} z_{ij} d(X_j, P_i) \tag{9}$$

$$d(X_j, P_i) = \sqrt{(x_j - a_i)^2 + (y_j - b_i)^2} \tag{10}$$

$$z_{ij} = \begin{cases} 1 & \text{if } j = argmin\, d(X_k, P_i) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where;

m	: number of cluster centers
n	: number of points
a_i	: horizontal position of point i $(i=1,\dots,n)$
b_i	: vertical position of point i $(i=1,\dots,n)$
P_i	: static position of point i (a_i, b_i)
z_{ij}	: assignment of point i to cluster center j $\{0,1\}$
x_j	: horizontal position of cluster center j $(j=1,\dots,m)$
y_j	: vertical position of cluster center j $(j=1,\dots,m)$
X_j	: variable position of cluster center j (x_j, y_j)
$d(X_j, P_i)$	: Euclidean distance between point i and cluster center j

The following steps are taken to look for best clustering solution to given points.
Step 1: ABC algorithm is initialized by setting parameters “number of food sources”, “food limit”, “number of iterations” and “number of dimensions” (Number of dimensions = number of clusters × 2).
Step 2: Food source locations are initially set with random values.
Step 3: The following steps are taken until the total number of iterations is reached;

- Employed bee phase is taken and fitness values are calculated after each movement.
- Onlooker bee phase is taken and fitness values are calculated after each movement.
- Scout bee phase is taken and fitness values are calculated after each movement.

Step 4: Cluster centers are extracted from best known solution vector of ABC algorithm.
Step 5: Each point is assigned to the closest cluster center.

Conversion of the model inputs and variables for the proposed ABC-based clustering algorithm can be seen in Table 1.

Table 1. Inputs and variables for model and the proposed ABC-based clustering algorithm

Definition	**Problem**	**ABC Clustering**
{0,1} assignment of customers to facilities	z_{ik}	z_{ij}
Euclidean distance from facilities to the customers	$d_k(\bar{x}_k, \bar{v}_i)$	$d(X_j, P_i)$
Location of facilities	$\bar{v}_i = (p_i, q_i)$	$X_j = (x_j, y_j)$
Location of customers	$\bar{x}_k = (x_k, y_k)$	$P_i = (a_i, b_i)$

RESULTS OF BENCHMARKS AND CASE STUDY

In order to evaluate the performance of the proposed algorithm same data is clustered with different algorithms for benchmarking. Benchmark clustering algorithms used to compare with the performance of ABC are Fuzzy C-Means (FCM), Center of Gravity integrated Fuzzy C-Means (FCM-COG), Self Organizing Maps (SOM), and Center of Gravity integrated Self Organizing Maps (SOM-COG).
FCM algorithm: It was initially proposed in 1973 by Dunn, and later improved by Bezdek in 1981. This algorithm lets a point to be a member of two or more clusters due to its fuzzy logic. As the fuzzy logic principle declares, an element of the data space can be a member of many clusters with differing membership values between [0,1] (Höppner et al, 2000). Higher membership values indicate the closer distance to the related cluster center.
SOM : SOM is a special type of neural networks where high dimension inputs are represented with lower dimension outputs. It is also named as Kohonen Maps, since initially proposed by Teuvo Kohonen. SOM operates, like other neural networks, in two phases. Initially system trains itself, and then maps the new input. This algorithm uses competitive learning during that phase. Haykin (1999) proves that SOM behaves very similar to C-Means algorithm for small number of neurons.
FCM-COG and SOM-COG algorithms: These algorithms are the integrated versions of FCM and SOM with Center of Gravity method to recalculate the cluster centers (Esnaf and Küçükdeniz, 2009). With the integration, the cluster center coordinates are revised so as to be closer to the nodes that are producing larger amounts of waste.

Benchmark Test Results

FCM clustering is applied using the program developed by Balasko et al. (2005) for MATLAB, and SOM clustering is established with SOM Toolbox of MATLAB developed by Alhoneimi et al. (2000). A code is developed in MATLAB for FCM-COG clustering based on the program of Balasko et al. (2005) and another for SOM-COG based on the codes of Alhoneimi et al. (2000). The developed model initially clusters the hospitals with FCM and SOM algorithms separately. FCM Algorithm is run for once due to its convergence feature; however SOM algorithm is run until it gives the minimum cost since it is a learning algorithm and gives better results with each new run. In the second step, the cluster centers are recalculated with integrated COG algorithm considering the weights of member points, and new clusters are formed. The objective is minimizing the "total cost" which is the amount of waste produced by a hospital, multiplied by its distance to the cluster center. The parameters used for the algorithms are "*m=2*" for FCM, and "*epoch=1000*" for training phase, for SOM. ABC clustering is done by the program developed as a Microsoft Windows application with Microsoft Visual C# 2012 Express Edition. The test runs were executed on a computer with Intel Core i5-430M processor at 2.26 GHz and 4 GB RAM.
The set of instances comprising real data were collected using the Geographical Information System ArcView, and report the central area of São José dos Campos city. Six instances (*100x10), (200x15), (300x25), (300x30), (402x30)* and *(402x40*) are created, containing 100, 200, 300 and 402 nodes. Each point is located on a block, which presents a demand node and is also a possible place to locate medians. Demand was estimated considering the number of houses (apartments) at each block. An empty block received value 1 (Lorena and Senne, 2004). Total costs which are calculated as sum of Euclidean distances from cluster center to every demand point multiplied by demand quantity of each demand point, are given in Table 2 for all benchmark algorithms.

Table 2. Benchmark results for total costs

	Points	Clusters	SOM	SOM-COG	FCM	FCM-COG	ABC
SJC1	100	10	985,328	759,113	1,010,600	843,925	708,430
SJC2	200	15	1,538,210	1,283,330	1,642,290	1,399,010	1,238,925
SJC3a	300	20	2,005,560	1,812,150	2,064,670	1,772,300	1,362,355
SJC3b	300	25	1,784,600	1,534,740	1,896,040	1,590,720	1,177,645
SJC4a	402	30	2,515,870	2,236,100	2,725,700	2,362,690	1,977,504
SJC4b	402	40	2,140,300	1,821,580	2,387,160	1,977,530	1,568,640

Table 3 displays the comparative results of the different cluster algorithms. In this table percent transportation cost differences of the ABC algorithm from the benchmark algorithms are given. The calculation of the percent differences for each data set is realized with the following formula:

$$\Delta = \left(\frac{H - M}{H}\right) \times 100 \quad (12)$$

Where *H* represents the objective function value, i.e. transportation cost, generated by benchmark algorithms for each data set, *M* is the transportation cost generated by the ABC algorithm for the corresponding data set.

Table 3. Percentage of transportation cost improvement of ABC algorithm from benchmark algorithms

	FCM	FCM-COG	SOM	SOM-COG
SJC1	28.10 %	6.68 %	29.90 %	16.06 %
SJC2	19.46 %	3.46 %	24.56 %	11.44 %
SJC3a	32.07 %	24.82 %	34.016 %	23.13 %
SJC3b	**34.01 %**	23.27 %	**37.89 %**	25.97 %
SJC4a	21.40 %	11.56 %	27.45 %	16.30 %
SJC4b	26.71 %	13.89 %	34.29 %	20.68 %
Average	26.96 %	13.95 %	31.350 %	18.93 %
Standard Deviations	5.73 %	8.64 %	4.938 %	5.32 %

The comparative results reveal that ABC algorithm outperforms all benchmark algorithms for all tests. Specifically, with problem set SJC3b, total cost results obtained by ABC are 34.01 % and 37.89 % better than FCM algorithm and SOM results respectively. The FCM-COG algorithm produced close results only for two instances, which are for the problems coded SJC1 and SJC2 respectively.

Case Study

In this study a real world MFLP is solved for identifying the healthcare waste disposal facility locations for Istanbul Municipality. The data space is the coordinates and waste disposal data of the hospitals with more than 20 bed capacity in Asian and European sides of Istanbul.

Due to the geographical location of the city, the medical wastes of the hospitals are collected on both sides of the Bosphorus using separate truck fleets. The treatment and disposal activities of healthcare waste are provided by Istanbul Metropolitan Municipality Environmental Management Industrial and Trade Inc. (ISTAÇ), which is the company established to assist the metropolitan municipality. The waste collection from hospitals with bed capacities over twenty is under the responsibility of ISTAÇ.

The medical waste collected is transferred to the sterilization facility on the European side. Table 4 shows the total amount of waste collected from the hospitals weekly.

Table 4. Total medical waste collected weekly

	Waste collected
Asian Side	109,093 kg/wk
European Side	179,315 kg/wk

ISTAÇ is planning to build sterilization facilities close to the hospitals to decrease the transportation costs as well as for environmental purposes. The reduction in the transferred distance will result in less fuel consumption. Consequently the emissions generated by the transfer processes will decrease, and the risk of transporting infectious waste through the city will also be mitigated.
ISTAÇ serves 99 hospitals on the Asian side and 158 hospitals on the European side of the city. The geographic coordinates of the hospitals in consideration are shown in Figure 1 for Asian side hospitals and Figure 2 for European Side hospitals.

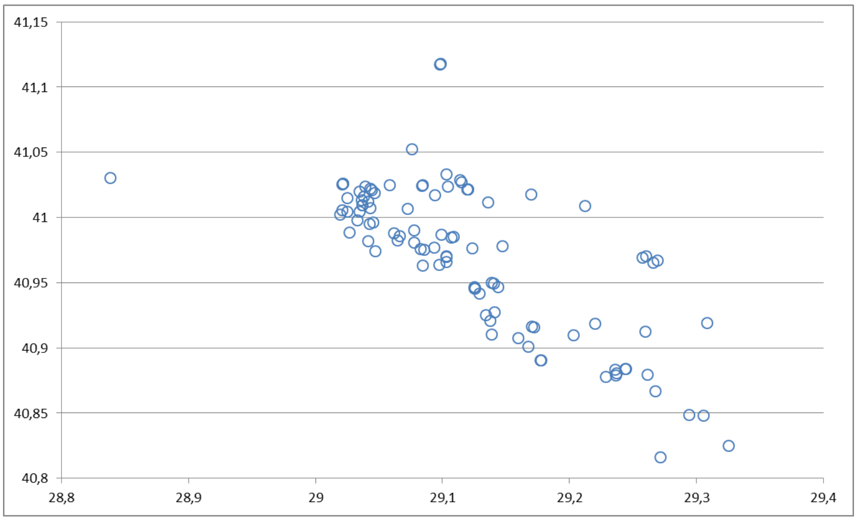

Figure 1. Asian side hospitals' coordinates

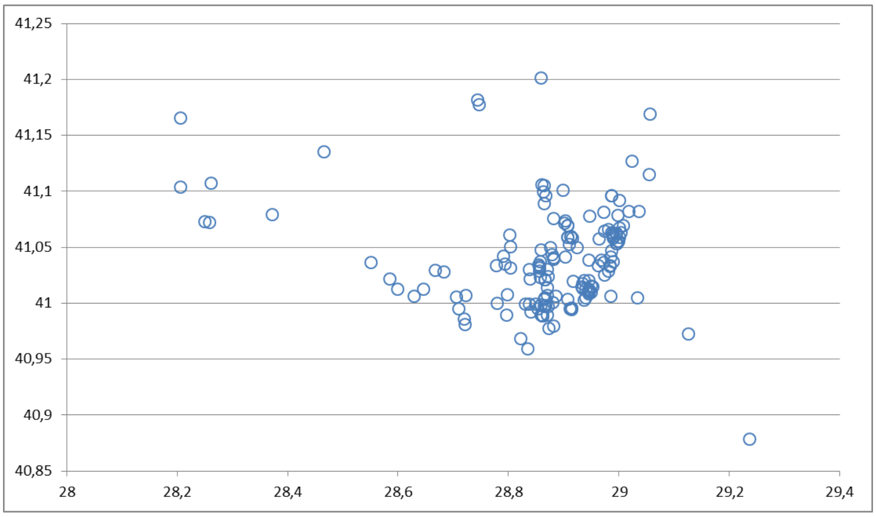

Figure 2. European side hospitals' coordinates

The aim of the study is to analyze and compare different clustering alternatives with ABC algorithm and to observe the one resulting the minimum transportation costs. ABC clustering is exploited for a continuous facility location problem. In that sense, facility locations are selected from an infinite set of possible locations. The algorithm is run with the following parameters to cluster the hospitals for each part of the city on two different continents;

Number of bees	= 20
Number of scout bees	= 1
Food limit	= 200
Iteration limit	= 2500

The fitness is defined as the total distance between the cluster center and the hospitals assigned to that cluster. ABC clustering algorithm is applied for various numbers of facilities to be opened, such as two, three and four facilities. These number of clusters are predefined due to the decisions of ISTAÇ supervisors, since the maximum number of facility to be built is planned to be four based on the land prices and possible public oppositions. The solutions obtained in seconds are given in Table 5 for Asian side, and in Table 6 for European side. These results also show the positive effect of cluster numbers on the costs. In other words, if the numbers of sterilization facilities to be opened are four, the total transportation costs of the medical waste will be minimum. Total transportation costs are calculated as weekly waste amounts in kilograms times the distance travelled in kilometers.

The real world problem data is also clustered with the benchmark algorithms. Consequently, another finding is the close results obtained from FCM-COG and SOM-COG. However, it is observed that the performance of ABC for total transportation costs is better than the benchmark algorithms as also observed in the tests with library data. These benchmark results are given in Table 7 and Table 8 for the Asian side, and Table 9 and Table 10 for the European side. The comparative results of the real world problem are also discussed under the following subsection.

The results of ABC clustering for the Asian side displays the minimum total cost for the four cluster scenario, and as 2757.15 km(kg/wk). Similarly, the best scenario is also for the four cluster scenario for European side. The results displayed below in Table 6 reveals that the minimum total cost is 6206.27 km(kg/wk).

Table 5. Asian side ABC algorithm clustering results

COORDINATES of CLUSTER CENTERS (X,Y)	**NUMBER OF CLUSTER MEMBERS**	**WASTE AMOUNT kg/wk**	**TOTAL COST km.(kg/wk)**
2 CLUSTERS			4436.79
29.1720 40.9160	41	27,193	
29.0710 41.0021	58	81,900	
3 CLUSTERS			3288.72
29.1720 40.9160	39	28,210	
29.0072 41.0070	26	42,695	
29.0780 40.9900	34	38,188	
4 CLUSTERS			2757.15
29.0750 40.9850	19	26,578	
29.0270 41.0070	26	41,140	
29.1030 41.0033	15	16,135	
29.1720 40.9160	39	25,240	

Table 6. European side ABC algorithm clustering results

COORDINATES of CLUSTER CENTERS (X,Y)	NUMBER OF CLUSTER MEMBERS	WASTE AMOUNT kg/wk	TOTAL COST km.(kg/wk)
2 CLUSTERS			10,762.55
28.8540 41.0000	79	79,380	
28.9450 41.0230	79	99,935	
3 CLUSTERS			8208.41
28.6460 41.0120	21	15,580	
28.8590 40.9990	58	63,800	
28.9450 41.0230	79	99,935	
4 CLUSTERS			6206.27
28.8590 40.9990	57	63,450	
28.9340 41.0170	38	59,505	
28.6460 41.0120	21	15,580	
28.9900 41.0580	42	40,780	

When FCM-COG and SOM-COG algorithm clustering results given in Table 7 and Table 8 are analysed for Asian side, it is observed that the total costs are 2997.83 km(kg/wk) and 2873.91 km(kg/wk) for four cluster scenarios respectively. Other scenarios of two and three cluster options produce considerably higher total costs. It is concluded that all costs are higher than the results obtained by ABC.

Table 7. Asian side FCM-COG algorithm clustering results

COORDINATES of CLUSTER CENTERS (X,Y)	NUMBER OF CLUSTER MEMBERS	WASTE AMOUNT kg/wk	TOTAL COST km.(kg/wk)
2 CLUSTERS			4450.41
29.1723 40.9159	41	27,193	
29.0502 41.0037	58	81,900	
3 CLUSTERS			4120.65
29.2454 40.9103	22	9800	
29.0376 41.0086	42	67,680	
29.1292 40.9504	35	31,913	
4 CLUSTERS			2997.83
29.2601 40.9123	19	8100	
29.0341 41.0068	28	42,560	
29.0952 41.0003	33	40,240	
29.1704 40.9164	19	18,193	

Table 8. Asian side SOM-COG algorithm clustering results

COORDINATES of CLUSTER CENTERS (X,Y)	NUMBER OF CLUSTER MEMBERS	WASTE AMOUNT kg/wk	TOTAL COST km.(kg/wk)
2 CLUSTERS			4443.40
29.1723 40.9158	41	27,193	
29.0485 41.0045	58	81,900	
3 CLUSTERS			3510.04
29.1726 40.9157	33	24,128	
29.0369 41.0091	32	46,410	
29.1011 40.9789	34	38,555	
4 CLUSTERS			2873.91
29.2601 40.9122	19	8100	
29.0277 41.0068	26	41,110	
29.0799 40.9909	34	41,590	
29.1704 40.9163	20	18,293	

The results for European side also reveal the better performance of ABC clustering. As given in Table 9 and Table 10, FCM-COG and SOM-COG clustering results for four cluster options give the lowest values among the scenarios. The total costs for four cluster scenario of FCM-COG and SOM-COG are 7192.32 km(kg/wk). However, the total costs are considerably higher than the total cost of 6206.27 km(kg/wk) that is obtained by ABC clustering.

Table 9. European side FCM-COG algorithm clustering result

COORDINATES of CLUSTER CENTERS (X,Y)	NUMBER OF CLUSTER MEMBERS	WASTE AMOUNT kg/wk	TOTAL COST km.(kg/wk)
2 CLUSTERS			11.869.90
28.9238 41.0194	137	163,735	
28.6003 41.0219	21	15,580	
3 CLUSTERS			8794.33
28.9464 41.0245	79	99,935	
28.2618 41.0974	8	4510	
28.8586 40.9987	71	74,870	
4 CLUSTERS			7129.32
28.6460 41.0125	14	11,870	
28.2539 41.0987	7	3710	
28.9555 41.0273	74	95,535	
28.8662 41.0050	63	68,200	

Table 10. European side SOM-COG algorithm clustering results

COORDINATES of CLUSTER CENTERS (X,Y)	NUMBER OF CLUSTER MEMBERS	WASTE AMOUNT kg/wk	TOTAL COST km.(kg/wk)
2 CLUSTERS			11,869.90
28.9238 41.0194	137	163,735	
28.6003 41.0219	21	15,580	
3 CLUSTERS			8791.96
28.9453 41.0233	79	99,935	
28.2618 41.0973	8	4510	
28.8574 40.9991	71	74,870	
4 CLUSTERS			7129.32
28.6460 41.0125	14	11,870	
28.2539 41.0987	7	3710	
28.9555 41.0273	74	95,535	
28.8662 41.0050	63	68,200	

Comparative Results

The tables given in this section show the comparative results of the different cluster algorithms applied to the real world data. Table 11 and Table 12 display the total costs [km.kg/wk] and the percent transportation cost differences of the ABC algorithm from the FCM-COG, and SOM-COG obtained from the different algorithms for various cluster numbers. The calculation of the percent differences for each data set is realized with the same formula used in the comparison of benchmark tests.
The comparative results declare that ABC algorithm outperforms for every scenario, and both for Asian and European side clusters.

Table 11. Asian side transportation cost differences of the ABC algorithm from FCM-COG and SOM-COG

	FCM-COG	SOM-COG
2 CLUSTERS	0.31%	0.15%
3 CLUSTERS	20.19%	6.31%
4 CLUSTERS	8.03%	4.06%

Table 12. European side transportation cost differences of the ABC algorithm from FCM-COG and SOM-COG

	FCM-COG	SOM-COG
2 CLUSTERS	9.33 %	9.33 %
3 CLUSTERS	6.66 %	6.64 %
4 CLUSTERS	12.95 %	12.95 %

CONCLUSION

In this study, a real world facility location problem is solved for identifying the healthcare waste disposal facility locations for Istanbul Municipality using Artificial Bee Colony (ABC) clustering algorithm. The algorithm is applied for clustering geographical coordinates and healthcare waste data of the Istanbul hospitals with more than twenty beds capacity. Clustering is also performed with SOM-COG and FCM-COG algorithms for benchmarking, and all algorithms are run for different number of clusters, and the total costs - the amount of healthcare waste produced by a hospital, multiplied by its distance to the sterilization facility - are calculated for comparing the results.

Benchmark tests applied with library data before the case study showed the better performance of ABC clustering algorithm over four algorithms, FCM, FCM-COG, SOM and SOM-COG respectively. Specifically, ABC clustering algorithm produced 31.65 % better than SOM, and 26.96 % better than FCM algorithms on average. SOM-COG and FCM-COG approached closer, however ABC clustering still outperforms FCM-COG 13.95 % on average, with a standard deviation of 8.64 %. The average for SOM-COG is 18.93 % with 5.32 % standard deviation. Consequently, FCM-COG and SOM-COG are selected for the case study.

The comparative results of the algorithms reveal the superior performance of the ABC based clustering algorithm. For the Asian side, total costs for two clusters are close. However for three clusters size, ABC clustering algorithm shows 20.19 % better performance than FCM-COG and 6.31 % better than SOM-COG. The results of ABC clustering for four clusters size are also outperforming the FCM-COG and SOM-COG benchmark algorithms as 8.03 %, and 4.06 % respectively. For European side, ABC algorithm constitutes 6.64 % better total cost value with three clusters, and the performance in four clusters case results with a 12.95 % better performance with respect to both benchmarks.

The total cost values will be used by ISTAÇ (Istanbul Metropolitan Municipality Environmental Management Industrial and Trade Inc.) for deciding the places of sterilization facilities for processing the medical waste generated in Istanbul hospitals. The results show the advantage of higher cluster sizes, since all the algorithms give better costs as the cluster size increase. On the other hand, it can be concluded that, the results obtained with the ABC algorithm is advised to be considered in decisions since it produces improved cost values for all cluster sizes.

Future research might be focused on hybridizing the ABC based clustering with another algorithm which will improve the performance of the proposed algorithm. This study which is in the part of ‘first group’ can be extended to the second part called ‘then route’. Hence routing of the vehicles inside the clusters considering not only the minimization of costs but also the CO_2 emissions. ABC based clustering algorithms are planned to apply capacitated continuous multiple waste disposal facility location problems and cases. These algorithms can also be used for the waste disposal facility location problem of other hazardous materials.

ACKNOWLEDGMENTS

The authors would like to thank the supervisors and staff of Istanbul Metropolitan Municipality Environmental Management Industrial and Trade Inc. (ISTAÇ), for their support throughout this study.

REFERENCES

Alagöz, A.Z. & Kocasoy, G. (2008). Improvement and modification of the routing system for the health-care waste collection and transportation in Istanbul, *Waste Management, 28,* 1461-1471.

Alhoneimi, E., Himberg, J. Parviainen, J., &Vesanto, J. (2000). SOM Toolbox 2.0, A software for MATLAB implementing SOM algorithm. Download available at: www.cis.hut.fi/projects/somtoolbox/

Alumur, S., & Kara, B.Y. (2005). A new model for the hazardous waste location-routing problem, *Computers & Operations Research, 34,* 1406-1423.

Ayoub, N., Martins, R., Wang, K., Seki, H., & Naka, Y. (2007). Two levels decision system for efficient planning and implementation of bioenergy production, *Energy Conversion and Management,48*(3), 709–723.

Balasko, B., Abonyi, J., & Feil, B. (2005). Fuzzy clustering and data analysis toolbox for MATLAB, Department of Process Engineering University of Veszprem, Hungary. Download available at: www.sunfinedata.com/wp-content/uploads/2009/10/FuzzyClustering Toolbox.pdf

Ballou, R.H. (1999). *Business Logistics Management*, 4th edition, NY: Prentice Hall.

Büyüksaatçi, S. (2009) *Green supply chain management, and an application*, Unpublished master's dissertation, Istanbul University, Industrial Engineering, (In Turkish).

Cappanera, P., Gallo, G., & Maffioli, F. (2004). Discrete facility location and routing of obnoxious activities, *Discrete Applied Mathematics*, *133*, 3-28.

Correia, I., & Captivo, M.E. (2006). Bounds for the single source modular capacitated plant location problem, *Computers & Operations Research, 33*, 2991-3003.

Daskin, M.S. (1995). *Network and discrete location: Models, algorithms, and applications*", USA: John Wiley and Sons.

Esnaf, Ş., & Küçükdeniz, T. (2009). A fuzzy clustering based hybrid method for a multi- facility location problem, *Journal of Intelligent Manufacturing, 20*(2), 259-265.

Esnaf Ş., & Küçükdeniz, T. (2013). Solving uncapacitated planar multi-facility location problems by a revised weighted fuzzy c-means clustering algorithm, *Journal Of Multiple-Valued Logic And Soft Computing*, *21,* 147-164.

Esnaf, Ş., Küçükdeniz, T., & Büyüksaatçi, S. (2008, June). *Fuzzy C-Means and center of gravity combined model for a capacitated planar multiple facility location problem,* Paper presented at International Conference on Multivariate Statistical Modeling & High Dimensional Data Mining, Kayseri, Turkey.

Esnaf, Ş., Küçükdeniz, T., & Tunçbilek, N. (2014). Fuzzy c-means algorithm with fixed cluster centers for uncapacited facility location problems: Turkish case study, In Kahraman C., & Öztayşi B., (Eds.), *Supply Chain Management Under Fuzziness, Studies in Fuziness and Soft Computing* (pp. 489-516), Heidelberg: Springer-Verlag,

Franca, P. M., Sosa, N. M., & Pureza, V. (1999). An adaptive tabu search algorithm for the capacitated clustering problem, *International Transactions in Operational Research*, *6*, 665-678.

Gergin, Z., & Esnaf, Ş. (2013, August) *Comparing the performance of different artificial intelligence based clustering algorithms in healthcare waste disposal location*, Paper presented at the 4th International Conference on Risk Analysis and Crisis Response, Istanbul, Turkey.

Gomes, H., Ribeiro, A. B., & Lobo, V. (2007). Locating model for CCA-treated wood waste remediation units using GIS and clustering methods, *Environmental Modelling and Software*, *22,* 1788-1795.

Hariz, H., Dönmez, C.Ç. & Sennaroglu, B. (2017) Siting of a central healthcare waste incinerator using GIS-based Multi-Criteria Decision Analysis, *Journal of Cleaner Production, 166,* 1031-1042.

Haykin, S., 1999, *Neural networks*, Pearson & Prentice Hall, USA.
Höppner, F., Klawonn, F., Kruse, R. & Runkler, T., 2000, *Fuzzy cluster analysis*, John Wiley&Sons, Chichester.

Hsieh, K., & Tien F. (2004). Self-organizing feature maps for solving location-allocation problems with rectilinear distances, *Computers and Operations Research, 31,* 1017-1031.

Karaboğa, D. (2005). *An idea based on honey bee swarm for numerical optimization,* Unpublished technical report (TR06), Erciyes University, Computer Engineering Department, Erciyes, Turkey.

Karaboğa, D., & Baştürk, B. (2007). A powerful and efficient algorithm for numerical function optimization: Artificial Bee Colony (ABC) algorithm, *Journal of Global Optimization*, *39*(3), 459-171.

Karabulut, M.,& İbrikçi, T. (2011). Assessment of clustering algorithms for unsupervised transcription factor binding site discovery, *Expert Systems with Applications, 38(9)*, 11160-11166.

Kashan, M.H., Nahavandi, N., & Kashan, A.H. (2012). DisABC: A new artificial bee colony algorithm for binary optimization, *Applied Soft Computing Journal*, *12*(1), 342-352.

Küçükdeniz, T., Baray, A., Ecerkale, K., & Esnaf, Ş. (2012). Integrated use of fuzzy-c means and convex programming for capacited multi-facility location problem, *Expert Systems with Applications, 39*(4), 4306-4314.

Li, J., Wang, J., Pan, Q. et al. (2017). A hybrid artificial bee colony for optimizing a reverse logistics network system, *Soft Computing, 21(20)*, 6001-6018.

Li-hong, S. (2009). A mixed integer linear programming model for medical waste reverse logistics network design, In preceedings of the *16th International Conference on Management Science & Engineering,* (pp. 1971-1975), September 14-16, Moscow, Russia.

Lorena, L.A.N., & Senne, E.L.F. (2004). A column generation approach to capacitated p-median problem, *Computers and Operations Research, 31,* 863–876.

Mohrechi, K., & Hatamlou, A. (2015). Locating optimal places for emergency medical centers using artificial bee colony algorithm, J*ournal of Advances in Computer Research Quarterly*, *6*(1), 115-124.

Negresios, M., & Palhano, A. (2006). The capacitated centered clustering problem, *Computers and Operations Research, 33*, 1639–1663

Re Velle, C. S., & Eiselt, H. A. (2005). Location analysis : A synthesis and survey, *European Journal of Operational Research, 165*(1), 1-19.

Scheuerer, S., & Wendolsky, R. (2006). A scatter search heuristic for the capacitated clustering problem, *European Journal of Operational Research*, *169*(2), 533-547.

Shanmugasundaram, J., Soulalay, V., & Chettiyappan, V. (2016). Geographic information system-based healthcare waste management planning for treatment site location and optimal transportation routeing, *Waste Management and Research*, *30*(6), 587-595.

Sheu, J-B., Chen, Y-H., & Lan, L.W. (2005). A novel model for quick response to disaster relief distribution, *Eastern Asia Society for Transportation Studies*, *5,* 2454 – 2462.

Sun, M. (2006). Solving the uncapacitated facility location problem using tabu search, *Computers and Operations Research, 33*, 2563–2589.

Tan, P-N., Steinbach, M., & Kumar, V. (2005). *Introduction to data mining*, USA: Pearson & Addison Wesley.

Thakur, V. & Ramesh, A. (2017). Healthcare waste disposal strategy selection using grey-AHP approach, *Benchmarking: An International Journal, 24(3),* 735-749.

Tunçbilek, N., Taşgetiren, F., & Esnaf, Ş. (2012). Artificial Bee Colony optimization algorithm for uncapacitated facility location problems, *Journal of Economic and Social Research*, *14*, 1-24.

Wichapa, N & Khokhajaikiat, P. (2018). Solving a multi-objective location routing problem for infectious waste disposal using hybrid goal programming and hybrid genetic algorithm.*International Journal of Industrial Engineering Computation ,* 9(1), 75-98.

World Health Organization. (2014). *Safe management of wastes from health-care activities*, 2nd Ed., Y.Chartier, J.Emmanuel, U.Pieper, A.Prüss, P.Rushbrook, R.Stringer, W.Townend, S.Wilburn, R.Zghondi (Eds.), Malta: WHO publications.

Yılmaz, Ö., Kara, B. Y. & Yetis, Ü. (2017) Hazardous waste management system design under population and environmental impact considerations, *Journal of Environmental Management, 203(2),* 720-731.

Zalik, K. R. (2006). Fuzzy C-means clustering and facility location problems, In proceedings of the *10th IASTED Conference Artificial Intelligence and Soft Computing*, August 28–30, Palma de Mallorca, Spain.

Zhang, S. Z., & Lee, C.K.M. (2013). Optimization of facility location problem in reverse logistics network using artificial bee colony algorithm, In proceedings of the *2013* IEEE, April 14-19, Turin, Italy.