International Conference on Information Technology Systems and Innovation (ICITSI) 2014
Bandung-Bali, 24-27 November 2014
ISBN: 978-1-4799-6526-7

Improvement of Fuzzy Geographically Weighted Clustering using Particle Swarm Optimization

Arie Wahyu Wijayanto
School of Electrical Engineering and Informatics
Insitut Teknologi Bandung
Bandung, Indonesia
ariewahyu@students.itb.ac.id

Ayu Purwarianti
School of Electrical Engineering and Informatics
Insitut Teknologi Bandung
Bandung, Indonesia
ayu@stei.itb.ac.id

Abstract — **One of the most interesting interdisciplinary research topics between Geographic Information Systems and Data Mining are known as Geo-demographic analysis (GDA). Among some methods in this field, Fuzzy Geographically Weighted Clustering (FGWC) is considered as the state-of-the-art for Geo-demographic segmentation. In other hand, recent studies show that its quality of clustering achievement can be improved. One of its limitations is in cluster initialization phase, which performed randomly. In this paper, we introduce a hybrid approach of FGWC based on Particle Swarm Optimization (PSO), namely FGWC-PSO to overcome the limitation. The experimental evaluation through public data set studies shows that the proposed method obtains better clustering quality.**

Keywords—clustering; fuzzy geographically weighted clustering; geo-demographic analysis; particle swarm optimization

I. INTRODUCTION

Geo-demographic analysis (GDA) which was characterized as "the investigation of spatially referenced geo-demographic and lifestyle information", is one of the most interesting inter-disciplinary research topics between Geographic Information Systems and Data Mining, and is generally utilized as a part of the public and private sector for planning and provision of products and services [1], [2]. An accurate fuzzy geo-demographic clustering task needs to be performed to handle the classification and genuine geographical neighborhood effect problems [3].

Some of well known fuzzy clustering methods in this field are Fuzzy C-Means (FCM) [3], [4], Gustafson-Kessel algorithm [4], the Neighborhood Effect [3], and Fuzzy Geographically Weighted Clustering (FGWC) [5]. Many current research defines that FGWC is the existing state-of-the-art method in this specific field [1], [6], [7]. It is because FGWC provides a better approach to handle the geographical neighborhood effect among elements in spatial data set to perform a qualified fuzzy geo-demographic clustering [5], [6].

In other hand, FGWC method has some limitations when performing a clustering task [1], [6]. FGWC similar to the classical FCM, determine the cluster centers (centroids) during iterative process, randomly [5]. This limitation in the random selection of center points during initialization phase, as reported in other similar clustering problem, makes iterative process failed to reach a global optimum solution [8]. This problem could affect the quality of cluster that is produced by FGWC. Previously we also proposed simplified design to handle this limitation in [9], but there is still not any experimental proof about its effectiveness.

Particle Swarm Optimization (PSO) is one of metaheuristic methods, which can be applied to differing optimisation problems with little change necessary as a global optimization tools [10], [11]. PSO as an optimization method is an iterative procedure, starting from an initial condition into a certain number of iterations, which could converge towards to a stable and optimal solution [12]. PSO is also an effective option when used on optimization of fuzzy clustering [13]–[15]. In this research we aim to propose an integration of PSO based optimization and FGWC, namely FGWC-PSO to improve the geo-demographic clustering quality by overcoming the limitation of FGWC in its initialization phase.

II. RELATED WORKS

Geo-demographic Analysis widely uses clustering methods to classify the Geo-demographic data into groups, making the data more valuable and manageable for analysis purposes [6]. The use of classical fuzzy clustering methods, e.g. FCM, in geo-demographic analysis has been widely used to perform a clustering task. Grekousis & Thomas performed segmentation analysis of various real geo-demographic data set using FCM and Gustaffson-Kessel algorithm in the prefecture Attica, Greece [4].

In other research, based on the results reported in [10], Runkler & Katz proposed a hybrid Fuzzy C-Means and PSO to minimize the objective function of FCM using two different methods, which are using the cluster centers and membership matrix as an independent variable separatedly [10]. In [16], the authors proposed an integration of clustering methods using Kmeans and PSO. A recent results of modification in segmentation task using FCM and Mahalanobis distance is also reported in [17], which is successful integrating the PSO into the clustering algorithms. Those various methods was proven experimentally in performing efficient clustering process. However, FCM and Kmeans on those methods misses geographical factor in their design [1].

Table 1 provide list of current most related research with their own limitation. It shows a clear research gap in this field which motivated us to propose a new initiative to overcome the fuzzy geo-demographic clustering problem. As a summary, the differences of FGWC-PSO with those other clustering methods listed in Table 1 is two fold: Firstly, the FGWC-PSO, as derivative of FGWC, is specially designed for the geo-demographic analysis problem which requires the modification of geographical spatial effects to the methods itself and not covered by classical fuzzy clustering such as Fuzzy C-Means; secondly, it is focused on overcoming the limitation of FGWC in initialization phase by performing PSO to select cluster centers or membership matrix under constraint of objective function minimization.

TABLE I. LIST OF RELATED RESEARCH

Authors	*Limitation*
Son et.al. (2012)	- Proposed algorithms have overcome the limitation of FGWC in terms of running time and computational process, but not considered to deal with initialization problems of FGWC [6].
Brouwer & Groenwold (2010)	- Proposed methods, hybrid PSO in classical fuzzy clustering, which is FCM, have not considered the aspect of geo-demographic neighborhood effect or spatial data handling [14].
Wang et.al. (2012)	- Proposed algorithms only deals with standard FCM and PSO, which still have not considered the aspect of spatial data handling and geo-demographic neighborhood effect [18].
Liu et.al. (2009)	- Proposed methods only modify the distance function of FCM using Mahalanobis and integrate it with PSO. It has not used the Fuzzy Geographically Weighted Clustering (FGWC) [17].
Runkler & Katz (2006)	- Proposed algorithms only deals with FCM and FCM-AO using PSO, and have not considered the aspect of spatial data handling and geographically aware clustering [10].
Niu & Huang (2011)	- Proposed algorithms, only reformulated and improve the methods in [10], and still have not considered the aspect of spatial data handling and geo-demographic neighborhood effect [13].
Niknam & Amiri (2010)	- Proposed solutions only compare the effectivity of PSO and Ant Colony Optimization (ACO) in standard FCM. It still has not considered the aspect of spatial data handling and geo-demographic neighborhood effect [19].

III. THEORETICAL BACKGROUND

A. Fuzzy Geographically Weighted Clustering (FGWC) Algorithm

FGWC Algorithm, which firstly proposed by Mason and Jacobson in 2007, provides a geographically aware alternative to a standard Fuzzy C-Means algorithm by providing the capability to apply population and distance effects for analyzing a geo-demographic segmentation [5]. The influence of one area upon another is considered by FGWC as the product of the populations of the areas. The divisor implements a distance decay effect through the weighting factor [5].

FGWC improves the previous research, a proposed method using neighbourhood effect (NE) by Feng and Flowerdew (1998) [3], which gives *ex post facto* adjustment of the cluster memberships after original fuzzy clustering [3], [5]. FGWC incorporating geography into geo-demographic analysis, so the cluster being sensitive to neighbourhood effects and will have an influence on the cluster centre values to create "geographically aware" clusters [5].

Feng and Flowerdew incorporated neighbourhood effects (NE) after the process of fuzzy clustering, which gives better result for fuzzy clustering [3], [5]. Mason and Jacobson integrated those two process in FGWC, where the cluster membership and characteristics evolve throughout the process of fuzzy clustering [5]. The adjusted cluster membership for the fuzzy geographically weighted clustering algorithm, which is calculated in each iteration of the fuzzy clustering algorithm,is shown in equation below [5]:

$$\mu'_i = \alpha\mu_i + \beta \frac{1}{A}\sum_j^n w_{ij}\ \mu_j \qquad (1)$$

Where μ'_i is the new cluster membership of area i and μ_i is the old cluster membership of area i. The w_{ij} is the weight measuring the amount of interaction between a pair of areas. The weight is decided by distance between area centroids or the length of common boundary between areas, or both. The A parameter is determined to ensure that the average of weighted membership values is still in the range of 0 and 1. The α and β are respectively weights to old membership and the mean of membership values of surrounding geographical areas and they observe [5]:

$$\alpha + \beta = 1 \qquad (2)$$

Where α and β are scalling variables to affect the proportion of the original membership vs the weighted (calculated) membership [5].

$$w_{ij} = \frac{(m_i\, m_j)^b}{d_{ij}^a} \qquad (3)$$

Where $m_i\, m_j$ are the population of areas i and j respectively, d_{ij} is the distance between i and j, then a and b are user definable parameters [5].

B. Particle Swarm Optimization (PSO)

Particle Swarm Optimization, which firstly proposed by James Kennedy and Russel Eberhart in 1995 [20], exploits a population of individuals to probe promising regions of the search space [21], [22]. In analogy with evolutionary computation methods, a swarm is similar to population and a particle is similar to an individual. PSO follows a stochastic optimization method based on Swarm Intelligence (SI). The fundamental idea is that each particle represents a potential solution which it updates according to its own experience and that of neighbours. The PSO algorithm searches in parallel using a group of individuals. Individuals or particles in a swarm, approach to the optimum through its present velocity, previous experience and the experience of its neighbors [21].

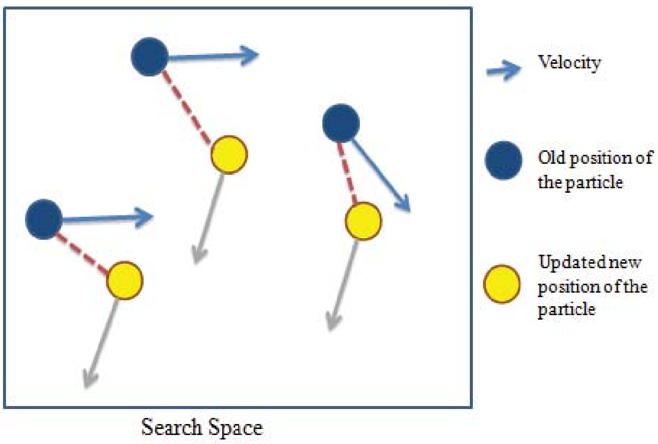

Figure 1. Conceptual overview of particle swarm with their associated positions and velocities [23]

PSO searches the problem domain by adjusting the trajectories of moving points in a multidimensional space. The motion of individual particles for the optimal solution is governed through the interactions of the position and velocity of each individual, their own previous best performance and the best performance of their neighbors. For a swarm of n particles the i-th particle is represented by a position denoted as $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ where n is the number of particles. Except the position, each particle of a swarm is represented in dimensional space with a velocity $v_i = (v_{i1}, v_{i2}, \ldots, v_{id})$ [21].

Suppose that we have a minimization problem that is defined over a continuous domain of d dimensions. We also have a population of N candidate solutions, denoted as $\{x_i\}, i \in [1, N]$. Furthermore, suppose that each individual x_i is moving with some velocity v_i through the search space. A basic particle swarm optimization algorithm for minimizing the n-dimensional function $f(x)$, where x, is the i-th candidate solution and v_i is its velocity vector. The notation $a \circ b$ means element-by-element multiplication of the vectors a and b [21].

A conceptual overview of particle swarm with their associated positions and velocities is presented in Figure 1. Initial conditions of PSO consists of initial number of particle swarm in certain positions and velocities. During iterative process, the methods performing some update towards new positions and velocities of whole particle. The complete pseudo code representation of Particle Swarm Optimization method can be seen in Figure 2. It is shown the iterative process until termination criteria reached that updates positions of all particles in decision area and calculate their associated velocity to find best solution in minimizing the fitness function [22], [23].

Initialize a random population of individuals $\{x_i\}, i \in [1, N]$
Initialize each individual's n-element velocity vector $v_i, i \in [1, N]$
Initialize the best-so-far position of each individual: $b_i \leftarrow x_i, i \in [1, N]$
Define the neighborhood size $\sigma < N$
Define the maximum influence values $\phi_{1,max}$ and $\phi_{2,max}$
Define the maximum velocity v_{max}
While not (termination criterion)
    For each individual $x_i, i \in [1, N]$
        $H_i \leftarrow \{\sigma$ nearest neighbors of $x_i\}$
        $h_i \leftarrow \arg\min_x \{f(x) : x \in H_i\}$
        Generate a random vector ϕ_1 with $\phi_1(k) \sim U[0, \phi_{1,max}]$ for $k \in [1, n]$
        Generate a random vector ϕ_2 with $\phi_2(k) \sim U[0, \phi_{2,max}]$ for $k \in [1, n]$
        $v_i \leftarrow v_i + \phi_1 \circ (b_i - x_i) + \phi_2 \circ (h_i - x_i)$
        If $|v_i| > v_{max}$ then
            $v_i \leftarrow \frac{v_i v_{max}}{|v_i|}$
        End if
        $x_i \leftarrow x_i + v_i$
        $b_i \leftarrow \arg\min \{f(x_i), f(b_i)\}$
    Next individual
Next generation

Figure 2. Particle Swarm Optimization Pseudo Code [24]

IV. THE PROPOSED METHODS

FGWC method has some limitations when performing a clustering task. The center point of each cluster (centroids) are defined randomly, which makes iterative process falling into the local optimal solution easily. The basic ideas is using PSO algorithm to determine the cluster centers or membership matrix in the initialization phase of FGWC clustering.

The basic objective function which will be minimized is:

$$J_{FGWC}(U, V; X) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{m} |v_i - x_k|^2 \rightarrow min \quad (4)$$

where m is a weigting exponent which determines the fuzziness of the clusters, u_{ij} is an element of partition matrix, v_i is a cluster center and x_k is a data point.

The cluster center itself can be determined using the following formulas:

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} \quad (5)$$

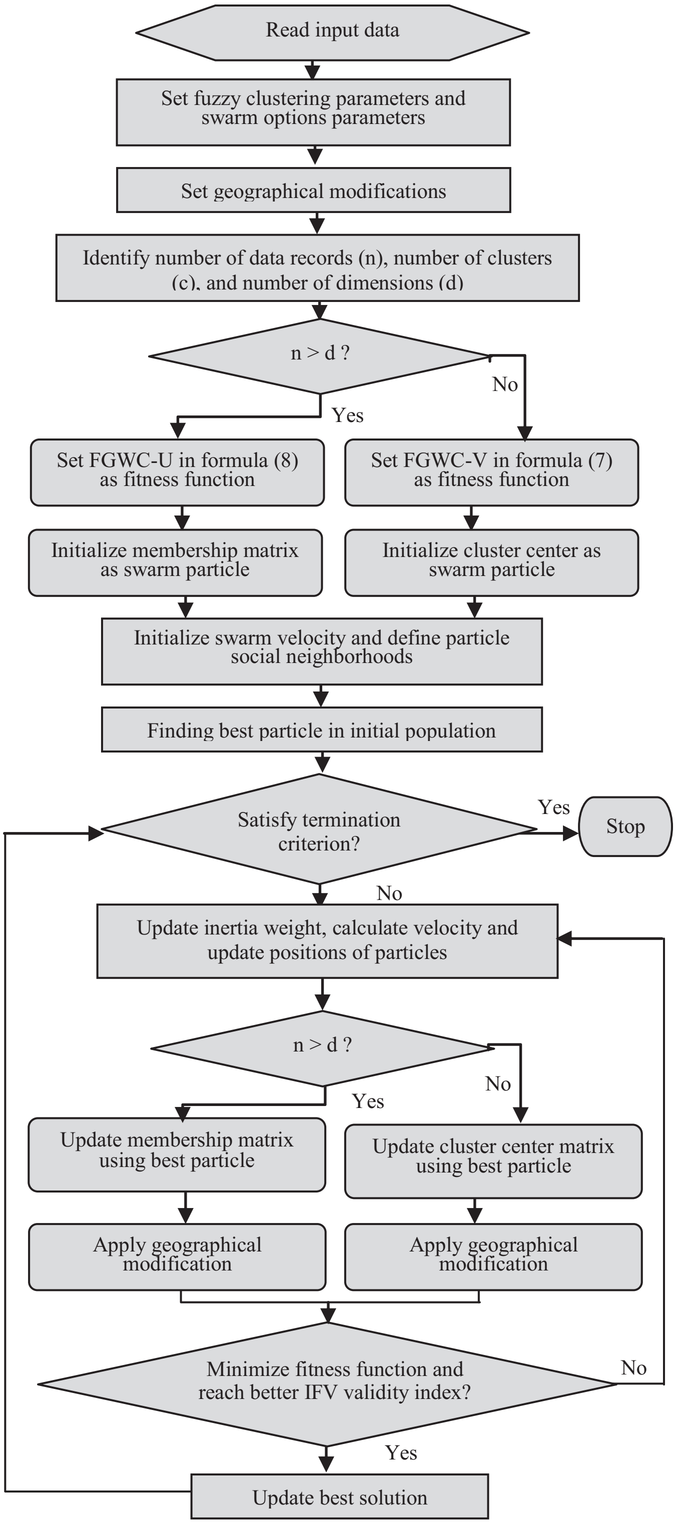

Figure 3. Flowchart of improved FGWC using PSO algorithm

In other hand, the membership matrix of fuzzy cluster before geographical modification can be calculated as follows:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{|v_i - x_k|}{|v_j - x_k|}\right)^{\frac{2}{m-1}}} \quad (6)$$

After being calculated, we modify the membership matrix using formulas (1). Inspired by popular method to optimize objective function in classical fuzzy clustering as reported by Runkler and Katz [10], we can reformulate the subsequent computation of cluster centers in formulas (5) and the membership matrix in formulas (6) and (1) into two different objective functions, respectively:

$$J_{FGWC}(V;X) = \sum_{i=1}^{c}\sum_{k=1}^{n} \frac{|v_i - x_k|^2}{\left(\sum_{j=1}^{c}\left(\frac{|v_i - x_k|}{|v_j - x_k|}\right)^{\frac{2}{m-1}}\right)^{m}} \rightarrow min \quad (7)$$

$$J_{FGWC}(U;X) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} \left|\frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} - x_k\right|^2 \rightarrow min \quad (8)$$

Those reformulation process is widely known as alternating optimization (AO) [10], [25], which in this case it is used to optimize the FGWC model through the necessary conditions extrema of J_{FGWC}. In this research, we called these reformulated function as FGWC-U and FGWC-V respectively. Where u_{ik} is the geographically modified cluster membership. The detailed flowchart of the proposed method are shown in Figure 3.

V. Results

The proposed method has implemented in Matlab R2013a and has executed under the environment of Intel Core i5-3210M CPU @2.50GHz, 4GB RAM and Windows 7 64bit operating system. We use public data set from Indonesian Population Census 2010 [26] which consist of 110 sociodemographic variables and 33 regions/provinces to evaluate the proposed method. The improved algorithm was tested against existing FCM [27], FGWC [5], and NE [3]. Parameters of those algoritms are set up using threshold $\varepsilon = 10^{-4}$, a = 1, b = 1, α = 0.5 and β = 0.5. All methods were run 100 times to evaluate the robustness.

The evaluation of geo-demographic clustering process use 4 kind of validity indices which are proven in many research as better tool to measure the clustering quality, which are: PC index [28], [29], CE index [29], SC index [29], and IFV index[30]. The IFV index is highly recommended as better validity index for measuring clustering process on spatial data [1], [30]. The better clustering quality gives maximum value of PC index and IFV index, also give minimum value of CE index and SC index.

A. PC Validity Index Measurement

After performing simulation, FGWC-PSO is proven as a best methods measured in PC validity index. Using various number of clusters in table 2 and figure 4, FGWC-PSO always gives greater PC value than others, which shows better clustering quality. The FCM, NE and FGWC gives very similar result after rounded into 3 decimal number.

TABLE II. Comparison of PC Index Value

C	PC Index (for m = 2.0)				PC Index (for m = 3.0)			
	FCM	*NE*	*FGWC*	*FGWC-PSO*	*FCM*	*NE*	*FGWC*	*FGWC-PSO*
2	0.556	0.500	0.500	**0.999**	0.500	0.500	0.500	**0.949**

C	PC Index (for m = 2.0)				PC Index (for m = 3.0)			
	FCM	*NE*	*FGWC*	*FGWC-PSO*	*FCM*	*NE*	*FGWC*	*FGWC-PSO*
3	0.387	0.333	0.333	**0.573**	0.333	0.333	0.333	**0.876**
4	0.299	0.250	0.250	**0.975**	0.250	0.250	0.250	**0.525**
5	0.271	0.200	0.200	**0.834**	0.200	0.200	0.200	**0.592**
6	0.238	0.167	0.167	**0.614**	0.167	0.167	0.167	**0.546**
7	0.200	0.125	0.125	**0.747**	0.143	0.143	0.143	**0.560**
8	0.200	0.125	0.125	**0.873**	0.125	0.125	0.125	**0.519**
9	0.181	0.111	0.111	**0.582**	0.111	0.111	0.111	**0.480**

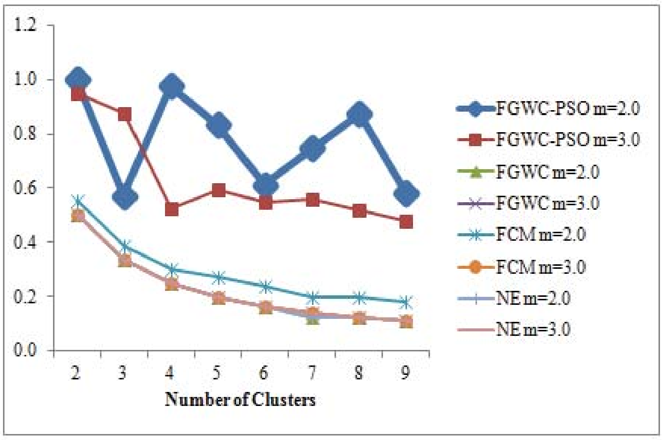

Figure 4. Comparison of PC Validity Index Evaluation

B. *CE Validity Index Measurement*

Measured using CE index as shown in table 3, FGWC-PSO is able to reach better clustering quality than FCM, NE, and FGWC for various fuzziness exponent value and number of clusters. The FCM, NE and FGWC gives different result, but not significant after 3 decimal rounding.

TABLE III. COMPARISON OF CE INDEX VALUE

C	CE Index (for m = 2.0)				CE Index (for m = 3.0)			
	FCM	*NE*	*FGWC*	*FGWC-PSO*	*FCM*	*NE*	*FGWC*	*FGWC-PSO*
2	0.636	0.693	0.693	**0.005**	0.693	0.693	0.693	**0.120**
3	1.018	1.099	1.099	**0.619**	1.099	1.099	1.099	**0.272**
4	1.290	1.386	1.386	**0.072**	1.386	1.386	1.386	**0.739**
5	1.451	1.609	1.609	**0.367**	1.609	1.609	1.609	**0.836**
6	1.614	1.792	1.792	**0.575**	1.791	1.792	1.792	**0.742**
7	1.869	2.079	2.079	**0.444**	1.946	1.946	1.946	**0.743**
8	1.869	2.079	2.079	**0.279**	2.079	2.079	2.079	**0.964**
9	1.994	2.197	2.197	**0.622**	2.197	2.197	2.197	**0.948**

Figure 5 shows that FGWC-PSO for m = 2.0 and m = 3.0, despite its fluctuation, always give less CE value than others, which shows better clustering quality. Overall the use of parameter fuzziness exponent m = 2.0 gives better result than m = 3.0 in FGWC-PSO.

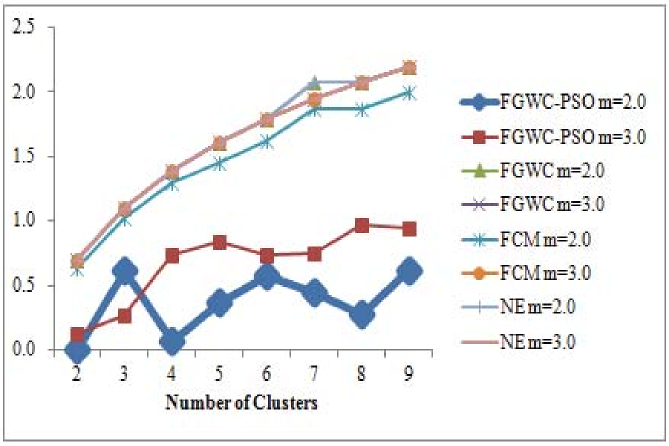

Figure 5. Comparison of CE Validity Index Evaluation

C. *SC Validity Index Measurement*

SC validity index of FGWC-PSO for various fuzziness exponent and number of clusters is less than FGWC and NE and not too worse than FCM. It is proven that the proposed FGWC-PSO can gives better clustering quality as can be seen in table 4 and 5 for m = 2.0 and m = 3.0 respectively.

TABLE IV. COMPARISON OF SC INDEX VALUE FOR M = 3,0

C	Algorithm			
	FCM	*NE*	*FGWC*	*FGWC-PSO*
2	7.08E-06	6.44E+03	3.97E+08	**4.01E-04**
3	4.82E-07	2.04E+03	7.73E+07	**2.46E-04**
4	2.90E-07	1.31E+03	3.88E+07	**1.22E-04**
5	8.24E-08	5.09E+02	1.38E+07	**7.92E-05**
6	7.51E-08	1.22E+03	7.38E+06	**3.70E-05**
7	7.48E-08	1.76E+02	1.72E+06	**2.40E-05**
8	1.94E-07	2.79E+02	1.85E+07	**6.17E-05**
9	9.53E-08	1.63E+02	2.22E+06	**5.78E-05**

TABLE V. COMPARISON OF SC INDEX VALUE FOR M = 2,0

C	Algorithm			
	FCM	*NE*	*FGWC*	*FGWC-PSO*
2	2.25E-12	3.57E+02	7.01E+07	**2.70E-04**
3	1.50E-12	1.96E+02	1.64E+07	**1.34E-04**
4	9.90E-13	7.97E+01	1.50E+07	**4.32E-05**
5	1.21E-12	3.53E+01	5.44E+06	**3.03E-05**
6	1.82E-12	6.17E+01	5.49E+06	**1.15E-05**
7	1.07E-11	1.96E+01	4.65E+06	**3.57E-05**
8	1.07E-11	1.96E+01	4.65E+06	**2.31E-05**
9	1.64E-11	1.13E+01	6.00E+05	**1.52E-05**

D. *IFV Validity Index Measurement*

As can be seen in table 6 and 7, for various number of clusters and fuzziness exponent, FGWC-PSO gives greater IFV value, which shows better clustering quality. Despite not better than FCM, FGWC-PSO is succesfully improved FGWC to reach better spatial clustering quality, which is measured by IFV index.

TABLE VI. COMPARISON OF IFV INDEX VALUE FOR M = 3,0

C	Algorithm			
	FCM	***NE***	***FGWC***	***FGWC-PSO***
2	7.30E+01	1.17E-05	7.84E-07	**1.41E+02**
3	1.43E+03	3.84E-05	5.37E-06	**2.42E+02**
4	2.37E+03	5.76E-05	1.07E-05	**5.33E+02**
5	4.25E+03	1.13E-04	1.54E-05	**9.22E+02**
6	5.02E+03	4.71E-05	3.08E-05	**9.69E+02**
7	2.72E+03	1.76E-04	7.13E-05	**1.26E+03**
8	1.74E+03	1.45E-04	1.12E-05	**6.49E+02**
9	2.02E+03	2.30E-04	5.04E-05	**4.84E+02**

TABLE VII. COMPARISON OF IFV INDEX VALUE FOR M = 2,0

C	Algorithm			
	FCM	***NE***	***FGWC***	***FGWC-PSO***
2	1.02E+07	2.12E-04	4.44E-06	**1.47E+02**
3	2.33E+07	4.02E-04	2.50E-05	**7.43E+02**
4	3.44E+07	9.57E-04	2.75E-05	**7.50E+02**
5	4.95E+07	1.60E-03	4.04E-05	**2.43E+03**
6	5.22E+07	9.28E-04	4.37E-05	**2.06E+03**
7	5.03E+07	2.00E-03	4.40E-05	**1.16E+03**
8	5.03E+07	2.00E-03	4.40E-05	**2.51E+03**
9	4.22E+07	3.40E-03	1.98E-04	**1.26E+03**

VI. CONCLUSIONS

This paper aims to propose the improvement of the limitations in fuzzy geo-demographic clustering algorithm by proposing an integration of Particle Swarm Optimization(PSO) based optimization and Fuzzy Geographically Weighted Clustering (FGWC), namely FGWC-PSO algorithm to reach a better geo-demographic clustering. The design are using PSO algorithm to determine the cluster centers or membership matrix in the initialization phase of FGWC. We evaluate the proposed methods using 4 different popular fuzzy clustering validity index. The experimental results shows that FGWC-PSO is effective to optimize the quality of clustering result. Using various fuzziness exponent and number of clusters, FGWC-PSO is succesfully give better geo-demographic clustering quality than the existing FGWC and other mature methods e.g. Fuzzy C-Means and Neighborhood Effect.

REFERENCES

[1] L. H. Son, B. C. Cuong, P. L. Lanzi, and N. T. Thong, "A novel intuitionistic fuzzy clustering method for geo-demographic analysis," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9848–9859, Aug. 2012.

[2] L. H. Son, "Enhancing clustering quality of geo-demographic analysis using context fuzzy clustering type-2 and particle swarm optimization," *Appl. Soft Comput.*, vol. 22, pp. 566–584, Sep. 2014.

[3] Z. Q. Feng and R. Flowerdew, "Fuzzy geodemographics: a contribution from fuzzy clustering methods," in *Innovations in GIS 5 : Selected Papers from the Fifth National Conference on GIS Research UK (GISRUK)*, S. Carver, Ed. London: Taylor & Francis Group, 1998, pp. 112–120.

[4] G. Grekousis and H. Thomas, "Comparison of two fuzzy algorithms in geodemographic segmentation analysis: The Fuzzy C-Means and Gustafson–Kessel methods," *Appl. Geogr.*, vol. 34, pp. 125–136, May 2012.

[5] G. A. Mason and R. D. Jacobson, "Fuzzy Geographically Weighted Clustering," in *Proceedings of the 9th International Conference on Geocomputation*, 2007, no. 1998, pp. 1–7.

[6] L. H. Son, P. L. Lanzi, B. C. Cuong, and H. A. Hung, "Data Mining in GIS : A Novel Context-Based Fuzzy Geographically Weighted Clustering Algorithm," *Int. J. Mach. Learn. Comput.*, vol. 2, no. June, pp. 1–4, 2012.

[7] L. H. Son, B. C. Cuong, and H. V. Long, "Spatial interaction – modification model and applications to geo-demographic analysis," *Knowledge-Based Syst.*, vol. 49, no. 2013, pp. 152–170, Sep. 2013.

[8] H. Izakian and A. Abraham, "Fuzzy C-means and fuzzy swarm for fuzzy clustering problem," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 1835–1838, Mar. 2011.

[9] A. W. Wijayanto and A. Purwarianti, "Improvement Design of Fuzzy Geodemographic Clustering using Artificial Bee Colony Optimization," in *The 3rd International Conference on Information Technology for Cyber and IT Service Management*, 2014, vol. 1, no. 2, pp. D13–D16.

[10] T. a. Runkler and C. Katz, "Fuzzy Clustering by Particle Swarm Optimization," in *2006 IEEE International Conference on Fuzzy Systems*, 2006, no. 3, pp. 601–608.

[11] R. J. Mullen, D. Monekosso, S. Barman, and P. Remagnino, "A review of ant algorithms," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9608–9617, Aug. 2009.

[12] X.-S. Yang, "Swarm intelligence based algorithms: a critical analysis," *Evol. Intell.*, vol. 7, no. 1, pp. 17–28, Dec. 2013.

[13] Q. Niu and X. Huang, "An improved fuzzy C-means clustering algorithm based on PSO," *J. Softw.*, vol. 6, no. 5, pp. 873–879, May 2011.

[14] R. K. Brouwer and A. Groenwold, "Modified fuzzy c-means for ordinal valued attributes with particle swarm for optimization," *Fuzzy Sets Syst.*, vol. 161, no. 13, pp. 1774–1789, Jul. 2010.

[15] A. W. Wijayanto and Takdir, "Fighting Cyber Crime in Email Spamming: An Evaluation of Fuzzy Clustering Approach to Classify Spam Messages," in *International Conference on Information Technology Systems and Innovation (ICITSI) 2014*, 2014, p. 2014.

[16] D. W. van der Merwe and A. P. Engelbrecht, "Data clustering using particle swarm optimization," in *The 2003 Congress on Evolutionary Computation, 2003. CEC '03.*, 2003, pp. 215–220.

[17] H. Liu, J. Yih, W. Lin, and T. Liu, "Fuzzy C-Means Algorithm Based on PSO and Mahalanobis Distance," *Int. J. Innov. Comput. Inf. Control*, vol. 5, no. December 2009, pp. 5033–5040, 2009.

[18] D. Wang, B. Han, and M. Huang, "Application of Fuzzy C-Means Clustering Algorithm Based on Particle Swarm Optimization in Computer Forensics," *Phys. Procedia*, vol. 24, pp. 1186–1191, Jan. 2012.

[19] T. Niknam and B. Amiri, "An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis," *Appl. Soft Comput.*, vol. 10, no. 1, pp. 183–197, Jan. 2010.

[20] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948.

[21] S. Rana, S. Jasola, and R. Kumar, "A review on particle swarm optimization algorithms and their applications to data clustering," *Artif. Intell. Rev.*, vol. 35, no. 3, pp. 211–222, Nov. 2010.

[22] X. Yang and L. Press, *Nature-Inspired Metaheuristic Algorithms Second Edition*, 2nd ed. Luniver Press, 2010.

[23] E.-G. Talbi, *Metaheuristics: From Design to Implementation*. Wiley Publishing, 2009.

[24] D. Simon, *Evolutionary Optimization Algorithms : Biologically Inspired and Population-Based Approaches to Computer Intelligence*. Wiley, 2013.

[25] T. A. Runkler, "Ant colony optimization of clustering models," *Int. J. Intell. Syst.*, vol. 20, no. 12, pp. 1233–1251, Dec. 2005.

[26] BPS, "Statistics Indonesia - 2010 Population Census," 2014. [Online]. Available: sp2010.bps.go.id. [Accessed: 01-Sep-2014].

[27] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell USA: Kluwer Academic Publishers, 1981, p. 256.

[28] A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murtagh, "Validity-guided (re)clustering with applications to image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 2, pp. 112–123, May 1996.

[29] B. Balasko, J. Abonyi, and B. Feil, "Fuzzy Clustering and Data Analysis Toolbox: For Use with Matlab," Veszprem, Hungary, 2005.

[30] C. Hu, L. Meng, and W. Shi, "Fuzzy clustering validity for spatial data," *Geo-spatial Inf. Sci.*, vol. 11, no. 3, pp. 191–196, Jan. 2008.