Postprint: Research on V-MDAV Algorithm Based on Grey Relational Analysis

Authors: Zhang Qishan, Zheng Lijun

Date: 2018-11-29T00:00:00+00:00

Abstract

Distance metrics influence the clustering effectiveness of microaggregation algorithms. To improve the algorithm' s privacy protection capability, the equilibrium proximity from grey relational analysis is employed to replace the Euclidean distance for measuring inter-record distances in the V-MDAV algorithm, proposing a V-MDAV algorithm based on grey relational analysis, namely the V-GRAV (variable-size grey relation to average vector) algorithm. As equilibrium proximity encompasses both the grey relational degree' s measurement of overall proximity and the equilibrium degree's measurement of sequence balance, it overcomes the issue of Euclidean distance being heavily influenced by local singular values. Consequently, the V-GRAV algorithm substantially reduces privacy leakage risk while maintaining information loss comparable to V-MDAV, with experimental results demonstrating the algorithm' s effectiveness.

Full Text

Preamble

Vol. 37 No. 1
Application Research of Computers
ChinaXiv Cooperative Journal

Research on V-MDAV Algorithm Based on Grey Relational Analysis

Zhang Qishan, Zheng Lijun
(School of Economics & Management, Fuzhou University, Fuzhou 350108, China)

Abstract: Distance metrics significantly influence the clustering effectiveness of microaggregation algorithms. To enhance privacy preservation capabilities, this paper proposes the V-GRAV (variable-size grey relation to average vector) algorithm, which replaces the Euclidean distance in V-MDAV with the balanced

adjacent degree from grey relational analysis to measure inter-record distances. The balanced adjacent degree incorporates both the grey relational degree' s measurement of overall proximity and the balance degree' s assessment of sequence equilibrium, thereby overcoming the susceptibility of Euclidean distance to local singular values. Consequently, V-GRAV substantially reduces privacy disclosure risk while maintaining information loss comparable to V-MDAV. Experimental results demonstrate the algorithm' s effectiveness.

Keywords: privacy protection; V_GRAV algorithm; balanced adjacent degree; information loss; privacy disclosure risk

0 Introduction

With the maturation of data mining technologies, leveraging data to extract useful information or knowledge has garnered increasing attention from academia and industry. However, personal information such as habits and preferences that individuals wish to keep private is being inferred and consumed at an alarming rate, making privacy protection an increasingly prominent concern.

To address privacy protection issues, Samarati et al. [1] proposed k-anonymity technology in 1998. This technique requires at least k indistinguishable records in published data, preventing attackers from tracing individuals through released information and thereby protecting personal privacy. During implementation, most privacy protection approaches employ generalization/suppression techniques [2–5], but determining the generalization domain for attribute values remains challenging. Moreover, generalizing numerical data often leads to semantic loss and unnecessary information loss.

Many scholars have turned to microaggregation algorithms from Statistical Disclosure Control (SDC) technology for data table k-anonymization to overcome the limitations of generalization techniques in numerical data applications. Among these, the MDAV algorithm is a crucial and high-performance microaggregation algorithm for numerical data anonymization, proven to perform best in terms of homogeneity within equivalence groups. However, MDAV (maximum distance to average vector), as a fixed-size heuristic algorithm, can produce k-partitions far from optimal in certain cases.

To eliminate the impact of unnatural k-partitions, Solanas et al. [6] proposed V-MDAV (variable-size maximum distance to average vector), a new heuristic multivariate microaggregation method. This approach forms variable-size equivalence groups that adapt more naturally to dataset distributions, increasing intra-group homogeneity. Nevertheless, V-MDAV employs Euclidean distance to measure inter-record distances, which is significantly influenced by singular values and consequently affects privacy protection effectiveness.

Building upon the idea that introducing the balanced adjacent degree into the MDAV algorithm can enhance privacy protection capabilities [7], this paper

proposes the V-GRAV algorithm. This algorithm replaces Euclidean distance with the balanced adjacent degree for k-aggregation in the V-MDAV framework. As an improved method of grey relational analysis, the balanced adjacent degree differs from regression analysis in mathematical statistics: it imposes no restrictions on sample size or regularity, involves small computational overhead, and produces quantitative results consistent with qualitative analysis. It is particularly suitable for small-sample, irregular data studies [8]. The algorithm inherits grey relational analysis' s independence from data distribution patterns while overcoming the point correlation tendency present in Euclidean distance and traditional grey relational analysis. Therefore, using the balanced adjacent degree to measure inter-record distances for k-anonymity can substantially reduce privacy disclosure risk while maintaining information loss comparable to V-MDAV.

1.1 V-MDAV Algorithm

The optimal multivariate microaggregation problem cannot be solved exactly in polynomial time, making it NP-hard [9]. Consequently, heuristic methods represent the only feasible approach. The MDAV algorithm [10, 11] is a well-known fixed-size heuristic that generates equivalence groups with a fixed cardinality of k. However, as a fixed-size heuristic, MDAV can produce k-partitions far from optimal in certain scenarios.

To address this limitation, Solanas et al. [6] proposed V-MDAV (Variable-size maximum distance to average vector), which adapts to dataset distributions to produce more natural groupings and increase intra-group homogeneity. Additionally, Solanas et al. proposed using genetic algorithms [12] for microaggregation of small multivariate datasets with up to 100 records, introducing a new N-ary encoding to handle the multivariate nature of microaggregation and conducting comprehensive experiments to determine optimal values for genetic algorithm parameters such as population size, crossover rate, and mutation rate. However, this method only applies to small datasets.

To overcome this limitation, literature [13] combined genetic algorithms with V-MDAV, using V-MDAV to partition large original datasets into smaller subsets manageable by genetic algorithms, then applying genetic algorithms to obtain microaggregated datasets. Huang et al. [14] integrated the advantages of microaggregation and surface subdivision, proposing the Hybrid-VMDAV algorithm to address location privacy issues. Under the guidance of l-diversity principles, they also proposed an l-diverse VMDAV (LD-VMDAV) as an improvement to effectively prevent temporal and spatial privacy attribute leakage.

Compared to V-MDAV, the V-GRAV algorithm can produce more natural k-partitions. This is illustrated in the following example:

Example 1. Let S be a dataset consisting of 9 records with 2 attributes:
S = {(2.4, 3), (1.68, 4.9), (3.18, 5.54), (5.32, 3.6), (18.68, 11.49), (20.14, 9.56), (19.85, 10.33), (21.28, 10.9), (23, 11.5)}.

Figure 1 [Figure 1: see original paper] depicts the output when MDAV is applied to dataset S with k = 3. Circles represent records in S, while triangles represent the global centroid. The red-marked group appears highly dispersed, degrading the overall 3-partition quality. This example demonstrates that MDAV' s fixed-size characteristic may inadequately adapt k-partitions to specific datasets. In contrast, Figure 2 [Figure 2: see original paper] shows V-MDAV's output, which clusters according to data distribution, producing two equivalence groups that are clearly more reasonable than those in Figure 1. Thus, V-MDAV improves upon MDAV by adapting to the natural distribution of records in microdatasets.

1.2 Grey Relational Analysis

Grey relational analysis, proposed by Professor Deng Julong as part of grey system theory, determines correlation by assessing the geometric similarity of sequence curves—the closer the curves, the greater the association between sequences. This method imposes no requirements on data volume or distribution patterns, produces quantitative results consistent with qualitative analysis, and can address uncertainty and nonlinear problems.

Traditional grey relational analysis suffers from local point correlation tendency. Professor Zhang Qishan introduced the balance degree into grey relational degree to propose the concept of balanced adjacent degree [15], which effectively eliminates point correlation tendency and overcomes traditional limitations. The balanced adjacent degree has been successfully applied as a similarity measure in clustering algorithms. Literature [16] utilized it to measure data similarity, overcoming parameter sensitivity issues and improving traditional spectral clustering performance. Li Liqiong et al. [17] proposed a novel algorithm integrating balanced adjacent degree into FCM, enabling global assessment of data similarity while mitigating the impact of strong local correlations.

Definition 1. Grey Relational Degree. Let X be the grey relational factor set, X = {x (k) | k K} be the reference sequence, and X = {x (k) | k K} be the comparison sequences, where i = {1, 2, 3, ···, h} and K = {1, 2, 3, ···, n}. The grey relational coefficient is:

$$r(x_0(k), x_i(k)) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|}$$

where is the distinguishing coefficient. The grey relational degree between reference sequence X and comparison sequence X is:

$$r(X_0, X_i) = \frac{1}{n} \sum_{k=1}^{n} r(x_0(k), x_i(k))$$

Definition 2. Balance Degree. Given R as the set of grey relational coefficient sequences, where R = {r(x (k), x (k)) | k K}, i = {1, 2, 3, ···, h}, and K = {1, 2, 3, ···, n}, the balance degree B(R) between reference sequence X and comparison sequence X is:

$$B(R_i) = -\frac{1}{\ln n}\sum_{k=1}^{n} p_i(k)\ln p_i(k)$$

where $p_i(k) = \frac{r(x_0(k),x_i(k))}{\sum_{k=1}^{n} r(x_0(k),x_i(k))}$. Here, $\ln n$ represents the maximum value of grey entropy, and from equation (3), $p_i(k)$ ranges between (0, 1).

Definition 3. Balanced Adjacent Degree. Let X be the grey relational factor set, X X be the reference sequence, X X be the comparison sequences, $r(X_0, X_i)$ be the grey relational degree, and $B(R_i)$ be the balance degree. The balanced adjacent degree $B(X_0, X_i)$ is:

$$B(X_0, X_i) = B(R_i) \times r(X_0, X_i)$$

A larger balanced adjacent degree indicates stronger correlation between the comparison and reference sequences.

2.1 Algorithm Description

The V-GRAV algorithm employs the balanced adjacent degree as the distance metric between records. Since this metric incorporates both the grey relational degree' s assessment of overall proximity and the balance degree' s evaluation of sequence equilibrium, introducing it into microaggregation yields the V-GRAV (variable-size grey relation to average vector) algorithm, providing a novel implementation approach for V-MDAV microaggregation.

The algorithm differs from V-MDAV in several key aspects:

a) V-GRAV measures inter-record similarity using balanced adjacent degree, whereas V-MDAV uses Euclidean distance. The balanced adjacent degree comprehensively considers both entropy correlation and point correlation, encompassing measurements of both point-wise distance proximity and overall non-differential proximity between sequences.

b) In V-GRAV, larger balanced adjacent degree values indicate greater similarity and closer distance between records, while in V-MDAV, larger Euclidean distance values indicate greater separation—the opposite relationship.

2.1.1 Group Generation

Input: Original dataset S, anonymity parameter k, distinguishing coefficient , gain factor .
Output: Anonymized dataset S' .

1. Calculate the balanced adjacent degree D between all records in dataset S.
2. Compute the centroid C of dataset S.
3. Identify the record r with the smallest balanced adjacent degree to centroid C among unassigned records.
4. Form an equivalence group centered at r with the k-1 records having the largest balanced adjacent degree to r.
5. Expand the group (see Section 2.1.2).
6. Repeat steps 3-5 until fewer than 2k records remain unassigned.
7. If fewer than k records remain, assign them to their nearest subset; otherwise, form a final subset with the remaining records.

2.1.2 Group Expansion

The expansion step enables V-GRAV to adapt to the natural distribution of records. After generating a subset of k records, this step identifies candidate records for potential inclusion and adds them if they are closer to the subset than to other unassigned records. The expansion process works as follows:

For a given group g with m records, let e_{max} be the closest unassigned record to g, b_{in} be the maximum balanced adjacent degree between e_{max} and g, and b_{out} be the maximum balanced adjacent degree between e_{max} and other unassigned records:

$$b_{in} = \max_{i \in [1,m]} B(e_{max}, e_g^i)$$

$$b_{out} = \max_{j \in [1, N_{un}], j \neq max} B(e_{max}, e_j)$$

where e_g^i represents the i-th record in group g, e_j represents the j-th record in the unassigned dataset, and N_{un} is the number of unassigned records. If multiple records satisfy the condition for e_{max}, one is randomly selected.

Finally, to determine whether to add e_{max} to group g, we compare b_{in} and b_{out} using the following criterion:

$$Add\_record = \begin{cases} YES & \text{if } b_{in} > \gamma \cdot b_{out} \\ NO & \text{otherwise} \end{cases}$$

To enhance V-GRAV's adaptability, the gain factor must be adjusted. However, determining its optimal value is non-trivial and beyond the scope of this paper

due to length constraints. The expansion process repeats until the group size reaches 2k - 1 or the condition in equation (9) is no longer satisfied, as literature [11] indicates that optimal k-partitions contain between k and 2k - 1 records per group.

2.2 Algorithm Evaluation

This paper proposes the V-GRAV algorithm for privacy protection applications. Data publishing privacy protection requires algorithms to achieve privacy goals while ensuring data utility. V-GRAV implements anonymization technology, but replacing records within groups with group centroids introduces data distortion that reduces utility. Therefore, evaluating V-GRAV' s performance requires assessing both information loss and privacy disclosure risk to determine microaggregation algorithm effectiveness.

2.2.1 Information Loss of Anonymous Tables

Various methods exist for calculating information loss. This paper adopts the IL (information loss) metric from the literature as a common approach for continuous data. The process for calculating information loss in an anonymized table is as follows:

1) **Calculate the total homogeneity measure SST of the original data table.**
   For an original data table with n records partitioned into h equivalence groups after anonymization, where each group contains m_i records, SST is computed as:

$$SST = \sum_{i=1}^{h} \sum_{j=1}^{m_i} d(W_{ij}, \overline{W})^2$$

where W_l represents the l-th record in the original data table and $\overline{W}$ is the overall mean of the original data table.

2) **Calculate the homogeneity measure SSE of anonymized equivalence classes.**
   First compute the mean of each equivalence group:

$$\overline{W}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} W_{ij}$$

Then calculate SSE by summing the squared differences between each record in the anonymized table and its group mean:

$$SSE = \sum_{i=1}^{h} \sum_{j=1}^{m_i} d(W_{ij}, \overline{W}_i)^2$$

3) **Calculate the total information loss IL after anonymization:**

$$IL = \frac{SSE}{SST}$$

For a given static dataset, the total homogeneity SST is fixed, while SSE varies depending on the k-partition. Thus, SSE dominates the evaluation of algorithm quality: greater intra-group homogeneity yields smaller SSE and consequently lower information loss.

2.2.2 Privacy Disclosure Risk of Anonymous Tables

Probabilistic record linkage and the distance-based record linkage privacy risk evaluation model DLD (distance-linked disclosure risk) are the most widely used methods for assessing privacy disclosure risk [19]. Due to its easier implementation and operation, this paper employs the DLD model to evaluate V-GRAV' s privacy disclosure risk.

The DLD model calculates distances between original and protected records to reflect the re-identifiability of anonymized records. Given an original dataset S = {s , s , ···, s } and its anonymized version S' = {s ' , s ' , ···, s ' }, for each record s ' in S' , we identify the two closest records in the original table S. If either of these records corresponds to the true original record s of s ' , then tuple s ' is considered successfully linked.

Assuming the number of successfully linked records in the anonymized table is *linked_records* and the total number of records is *total_records*, the privacy disclosure risk is measured as:

$$DLD = \frac{linked\_records}{total\_records}$$

3.1.1 Datasets

This study employs three classic datasets: Tarragona, Census, and EIA. The Tarragona dataset contains information from 834 companies in the Tarragona region in 1995; the Census dataset provides demographic census information from the 2000 U.S. Census; and the EIA dataset contains 1996 U.S. energy information from the Energy Information Administration. These datasets comprise 12, 12, and 9 numerical attributes respectively, plus one sensitive attribute, with 834, 1080, and 4092 records.

3.1.2 Experimental Software/Hardware Environment

Hardware: Intel Core 3.30 GHz CPU, 4096 MB RAM, 32-bit Windows 7 operating system.
Programming Environment: Mathworks MATLAB R2014a.

3.2 Information Loss Analysis

Using equation (14), we calculated the information loss for both V-MDAV and V-GRAV algorithms under different k values, with results shown in Figure 3 [Figure 3: see original paper]. V-MDAV results are represented by solid lines, while the proposed V-GRAV algorithm uses dashed lines. The gain factor is set to 0.2 and the distinguishing coefficient to 1.8; determining optimal values for these parameters is complex and beyond the scope of this paper due to length constraints.

Figure 3 shows that information loss increases with k, as larger equivalence groups reduce intra-group homogeneity. Among the three datasets, Tarragona exhibits the highest information loss, followed by Census, with EIA showing the lowest. The relationship between dataset sizes is Tarragona < Census < EIA. With the same k value, larger datasets offer more candidate records for partitioning equivalence groups, enabling selection of more homogeneous records and thus lower information loss after anonymization.

Comparing the two algorithms reveals that V-GRAV's information loss is slightly higher than V-MDAV' s. Since information loss calculation uses Euclidean distance—consistent with V-MDAV' s partitioning criterion—V-MDAV achieves lower information loss. However, the difference between the two algorithms does not exceed 5%, ensuring that V-GRAV maintains acceptable data utility.

As all three datasets show similar trends, Figure 4 [Figure 4: see original paper] selects the EIA dataset for comparative analysis of four algorithms. Variable-size microaggregation algorithms demonstrate lower information loss than fixed-size algorithms because they adapt to the natural distribution of records in micro-datasets, producing more reasonable clustering results. Both GRAV algorithms show higher information loss than MDAV algorithms due to the Euclidean distance metric used in calculations aligning better with MDAV' s partitioning approach.

3.3 Privacy Disclosure Risk Analysis

Using equation (15), we calculated the privacy disclosure risk for V-MDAV and V-GRAV, with results presented in Figure 5 [Figure 5: see original paper]. Figures 5(a)-(c) show privacy disclosure risk across the Census, Tarragona, and EIA datasets, with solid lines representing V-MDAV and dashed lines representing V-GRAV.

All three datasets exhibit similar trends: privacy disclosure risk decreases as k in-

creases. Larger equivalence classes reduce intra-class homogeneity and increase information distortion, making it more difficult for attackers to re-identify users. The overall privacy disclosure risk follows the pattern: Tarragona < Census < EIA, reflecting the inverse relationship between data utility and privacy risk—higher utility corresponds to greater disclosure risk.

Crucially, V-GRAV achieves significantly lower privacy disclosure risk than V-MDAV. This occurs because the DLD evaluation model directly employs Euclidean distance formulas, which are inconsistent with V-GRAV' s equivalence class partitioning criteria. Consequently, anonymized datasets produced by V-GRAV often fail to link to true records under the DLD model, resulting in lower privacy disclosure risk.

Figure 6 [Figure 6: see original paper] analyzes the EIA dataset across four algorithms, showing that both GRAV algorithms achieve lower privacy disclosure risk than MDAV algorithms for the same reason of metric inconsistency with the DLD model. Combined with Figure 4, these results demonstrate that variable-size microaggregation algorithms achieve lower information loss than fixed-size algorithms with comparable privacy disclosure risk, validating their effectiveness.

3.4 Comprehensive Evaluation of V-GRAV Algorithm

Figure 7 [Figure 7: see original paper] compares the four algorithms by calculating absolute differences in information loss and privacy disclosure risk between algorithm pairs. Blue bars represent differences between fixed-size microaggregation algorithms, while orange bars represent differences between variable-size algorithms.

Comparing MDAV and GRAV reveals that their information loss difference does not exceed 1.5%, while their privacy disclosure risk difference generally exceeds 1.5%. Since the information loss gap is smaller than the privacy risk gap, GRAV algorithms significantly enhance privacy protection while sacrificing minimal data utility, demonstrating the effectiveness of introducing balanced adjacent degree into microaggregation.

Similarly, comparing V-MDAV and V-GRAV shows that V-GRAV achieves substantially lower privacy disclosure risk. The information loss difference between V-MDAV and V-GRAV remains below 1.5%, while their privacy disclosure risk difference exceeds 1.5%. This confirms that the proposed V-GRAV algorithm effectively reduces privacy disclosure risk.

In summary, V-GRAV, which employs balanced adjacent degree as its distance metric, proves to be an effective approach.

3.5 Experimental Summary

Based on experimental results, we conclude:

a) The effectiveness of introducing balanced adjacent degree into microaggregation algorithms is validated through comparisons between MDAV and GRAV algorithms.

b) V-GRAV reduces privacy disclosure risk while maintaining acceptable information loss, demonstrating the viability of balanced adjacent degree as a distance metric in microaggregation.

c) As k increases, V-GRAV produces anonymized tables with higher information loss but lower privacy disclosure risk.

d) Consistent with V-MDAV, V-GRAV shows decreasing information loss trends as dataset size increases.

e) Information loss and privacy disclosure risk exhibit an inverse relationship: higher information loss corresponds to lower privacy disclosure risk.

Overall, experimental results confirm that V-GRAV effectively implements k-anonymity models and produces anonymized tables with strong privacy protection capabilities.

4 Conclusion

Distance measurement is a critical issue in microaggregation algorithms for k-clustering. To address Euclidean distance' s vulnerability to singular values and point correlation tendency, this paper proposes the V-GRAV algorithm, a novel multivariate microaggregation method that introduces balanced adjacent degree to enhance privacy protection.

V-GRAV replaces Euclidean distance in V-MDAV with balanced adjacent degree for inter-record distance measurement. By incorporating both overall proximity measurement and sequence equilibrium assessment, balanced adjacent degree eliminates point correlation tendency, enabling V-GRAV to reduce privacy disclosure risk while maintaining information loss comparable to V-MDAV.

However, this research has limitations. Future work will focus on:
a) Detailed analysis of determining optimal gain factor values for given datasets.
b) Extending V-GRAV to handle categorical attributes, as the current balanced adjacent degree approach is designed for numerical data. Future research will explore grey relational methods for anonymizing mixed-type data.

References

[1] Samarati P, Sweeney L. Generalizing data to provide anonymity when disclosing information (abstract) [C]// Proc of the 17th ACM-SIGMOD-SIGACT-SIGART Symposium on the Principles of Database Systems. New York: ACM Press, 1998: 188.

[2] Valls A. Semantic adaptive microaggregation of categorical microdata [J]. Computers & Security, 2012, 31(5): 653-672.

[3] Li Jiuyong, Baig M M, Sattar S A H M, et al. A hybrid approach to prevent composition attacks for independent data releases [J]. Information Sciences, 2016, 367-368: 324-336.

[4] Amiri F, Yazdani N, Shakery A. Bottom-up sequential anonymization in the presence of adversary knowledge [J]. Information Sciences, 2018, 450: 316-334.

[5] Li Boyu, Liu Yanheng, Han Xu, et al. Cross-Bucket generalization for information and privacy preservation [J]. IEEE Trans on Knowledge & Data Engineering, 2018, 30(3): 449-459.

[6] Solanas A, Martínez-Ballesté A. V-MDAV: a multivariate microaggregation with variable group size [J]. Seventh Compstat Symposium of the Iasc, 2006.

[7] Zheng Qunhua. Research on k-anonymity Model and Algorithm for Privacy Preserving based on Grey Relational Analysis [D]. Fuzhou: Fuzhou University, 2012.

[8] Liu Sifeng, Yang Yingjie, Wu Lifeng. Grey system theory and its application [M]. 7th ed. Beijing: Science Press, 2014: 63-108.

[9] Laszlo M, Mukherjee S. Iterated local search for microaggregation [J]. Journal of Systems & Software, 2015, 100: 15-26.

[10] Rebollo-Monedero D, Forné J, Pallarès E, et al. A modification of the lloyd algorithm for k-anonymous quantization [J]. Information Sciences, 2013, 222(2): 185-202.

[11] Soria-Comas J, Domingo-Ferrer J. Differentially private data publishing via optimal univariate microaggregation and record perturbation [J]. Knowledge-Based Systems, 2018, 153: 78-90.

[12] Solanas A, González-Nicolás U, Martinez-Balleste A. A variable-MDAV-based partitioning strategy to continuous multivariate microaggregation with genetic algorithms [C]// Proc of International Joint Conference on Neural Networks. Piscataway, NJ: IEEE Press, 2010: 1-7.

[13] Solanas A, González-Nicolás Ú, Martínez-Ballesté A. Mixing genetic algorithms and V-MDAV to protect microdata [J]. Computational Intelligence for Privacy and Security, 2012, 394: 115-133.

[14] Huang Kuanlun, Kanhere Salil S, Hu Wen. Preserving privacy in participatory sensing systems [J]. Computer Communications, 2010, 33(11): 1268-1280.

[15] Zhang Qishan, Deng Julong, Shao Yong. A grey correlational analysis by the method of balance and approach [J]. Journal of Huazhong University of Science and Technology, 1995, 23(11): 94-98.

[16] Guo Kun, Zhang Qishan. Spectral clustering based on grey relational analysis [J]. System Engineering Theory and Practice, 2010, 30(7): 1260-1265.

[17] Li Liqiong, Liu Zhanghui, Guo Kun. Fuzzy C-means based on grey relational analysis [J]. Journal of Fuzhou University: Natural Science Edition, 2016, 44(2): 170-175.

[18] Reza Mortazavi, Saeed Jalili. Preference-based anonymization of numerical datasets by multi-objective microaggregation [J]. Information Fusion, 2015, 25(C): 85-104.

[19] Mortazavi R, Jalili S. Enhancing aggregation phase of microaggregation methods for interval disclosure risk minimization [J]. Data Mining and Knowledge Discovery, 2016, 30(3): 605-639.

Note: Figure translations are in progress. See original paper for figures.