Fuzzy Clustering based Anomaly Detection for Distributed Multi-view Data

Hongmei Wang*, Tianhua Chen†, Hongtao Wang‡, Xuqiang Shao‡ and Pan Su‡,

*Science and Technology College, North China Electric Power University, China
Email: wanghm@ncepu.edu.cn
†Department of Computer Science, School of Computing and Engineering, University of Huddersfield, UK
Email: T.Chen@hud.ac.uk
‡School of Control and Computer Engineering, North China Electric Power University, China
Email: wanght@ncepu.edu.cn, shaoxuqiang@163.com, supan@ncepu.edu.cn

***Abstract*—Anomaly detection aims to identify the abnormal instances, whose behavior deviates significantly from the others. Nowadays owing to the existence of diverse data generation sources, different attributes of the same instances may be located on distributed parties forming a multi-view dataset. Thus multi-view anomaly detection has become a key task to discover outliers across various views. Traditionally, to perform multi-view anomaly detection, one needs to centralize data instances from all views into a single machine. However, in many real-world scenarios, it is impractical to send data from diverse views to a master machine due to the privacy issues. Inspired by this, we propose a fuzzy clustering based distributed approach for multi-view anomaly detection that simultaneously learns a membership degree matrix for each view and then detects anomalies for all parties. Specifically, we first introduce a combined fuzzy c-means clustering method for multi-view data and then design an anomaly measurement criterion to quantify the abnormal score from membership degree matrix. To solve the proposed model, a protocol is provided to unify all parties performing a well-designed optimization in an iterative way. Experiments on three datasets with different anomaly settings demonstrate the effectiveness of our approach.**

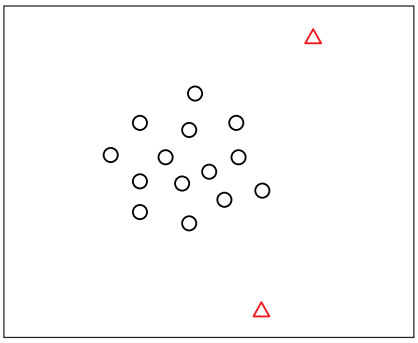
I. Introduction

Recently, the burst of big data encourages the development of many new technologies which concentrates on understanding these data for knowledge discovery. Anomaly detection, being one of the key issues in both data science and security management, aims to detect suspicious individuals, patterns, events and groups which do not conform to the expected modality. Anomaly detection has been applied to data from a wide spectrum of domains, such as detecting network intrusions [1], spams [2] and road traffic events [3].

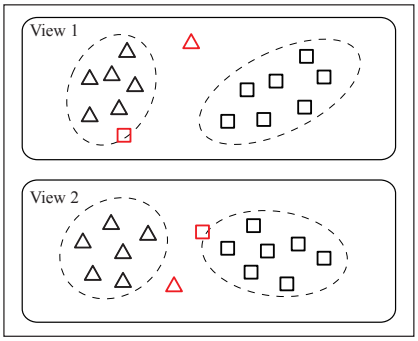
Most existing works on anomaly detection are centralized anomaly detection methods, in which data must be collected to a centralized machine before implementing the detection algorithms. There are two categories on centralized anomaly detection methods. One category is *single-view anomaly detection*, which aims to discover abnormal data instances as outliers from one single view (as shown in Figure 1(a)). Several effective methods in this category have been proposed, for example, one class SVM [4], one-class CRF [5] and isolation forest [6], to name a few. However, the fast accumulation and collection of diverse data sources in recent years leads to the generation of large amounts of multi-view data such as social media, sensor readings, and scientific publication networks. These data expedites the need of the other category of centralized anomaly detection: *multi-view anomaly detection*. The goal of multi-view anomaly detection is not only to identify instances that deviate from normal clusters, but also to identify instances with inconsistent behavior among multiple views. As it is shown in Fig. 1(b), red triangle is a typical multi-view anomaly since it is abnormal in both views, and it can be easily detected by common single-view anomaly detectors by checking each view. However, red rectangle represents a multi-view anomaly belonging to different clusters under two views. Performing single-view anomaly detection cannot directly differentiate the anomalies under this circumstance [7].

In many practical scenarios, centralizing the multi-view data (for anomaly detection) is not available in terms of the privacy issues. For example, different hospitals may hold different medical records that refer to the same set of patients. It is very important for these hospitals to protect the patients' private data from leaking to others while they working collaboratively to detect anomalies for treatment. There exists distributed algorithms for single-view anomaly detection [8], [9], [10], in which data instances of all attributes are distributed on different parties. However, for privacy-preserving multi-view anomaly detection, it still remains a challenge due to the complicated organizations or distributions of anomalies in each views.

To address the above-mentioned problem, we propose a distributed multi-view anomaly detection approach, denoted by DOMAD, which allows all parties to operate on their individual data views iteratively in a distributed manner. In the proposed approach, a united fuzzy c-means technique is utilized to simultaneously learn clustering structures on each view as well as the membership degree matrix. After that, we estimate the anomaly score for each sample by devising a criterion from the membership degree matrices, considering the different kinds of anomalies in multiple views.



(a) single-view anomaly



(b) multi-view anomaly

Fig. 1. Two categories of centralized anomaly detection. (a) Red triangles are single-view anomalies. (b) Red triangle and rectangle are multi-view anomalies.

To solve the proposed model, we develop a protocol to unify all parties performing a well-designed optimization algorithm in an iterative way. Finally, experiments on several public benchmark datasets are reported to show the effectiveness of the method.

The main sections of this paper are organized as follows: In Section III, we propose a novel multi-view anomaly detection approach based on unified fuzzy c-means technique. Our approach can give a fuzzy membership degree for multi-view instances and anomalies are identified from the perspective of membership degree matrix. In Section IV, we present an efficient optimization algorithm to solve the objective of anomaly detection approach, and devise a protocol to supervise multi-parties to perform a well-designed distributed optimization algorithm in an iterative way, while protecting the privacy of their own data. To evaluate the proposed model, we conduct extensive experiments on three UCI datasets and demonstrate the effectiveness of our approach in Section V.

II. Related Works

The growing popularity of data acquisition techniques from diverse data sources has made it easy to collect multiple views for the same entity. Many data sets in real world are naturally comprised of different views, which are often called *Multi-view Data* [11]. Studying multi-view data is with great challenges compared to single-view data due to the diverse generation mechanisms under different views.

One of the popular research on multi-view data is multi-view clustering or co-clustering, which aims to divide instances into clusters based on multiple views. One class of solution is to utilize traditional single-view clustering algorithms and then fuses them based on consensus [12]. [13] projects multi-view data into a common lower dimensional subspace and applies k-means to get the partitions. Another class of methods is based on fuzzy set theory that gives soft partitions on multi-view data. For example, [14] proposed a FCM-type fuzzy co-clustering algorithm for categorical multivariate data. [15] devised a fuzzy co-clustering algorithm on distributed co-occurrence data, where vertically partitioned co-occurrence information is stored in multiple parties. However, these methods could not identify anomalies or outliers hidden in multi-view datasets.

To tackle the problem of multi-view anomaly detection, horizontal anomaly detection (HOAD)[16] is presented to detect sample whose behavior is inconsistent among different sources. Affinity propagation (AP) [17] based multi-view anomaly detection algorithm is proposed by measuring the differences between each sample and its neighborhoods in different views. Recently, dual-regularized multi-view outlier detection method (DMOD) [7] and multi-view low-Rank analysis (MLRA) method [18] are proposed respectively. However, these methods can only be performed on a single master machine, which comes with limitations on distributed computation due to the privacy issues. Although several privacy-preserving multi-party clustering algorithms have been proposed [19], [20], distributed privacy-preserving multi-view anomaly detection approach has not been addressed. In this paper, we fill this gap by utilizing distributed multi-view fuzzy clustering with privacy concerns.

III. Multi-view Fuzzy Anomaly Detection Approach

In this section, we first introduce the preliminary knowledge of fuzzy c-means technique. After that we propose our unified multi-view fuzzy anomaly detection approach.

A. Preliminary Knowledge

Many tools and methods [21], [22], [23], [24], [25], [26], [27] have been successfully developed in the framework of fuzzy set theory, owing to its capability to deal with vague and imprecise information. Among these, fuzzy c-means(FCM) [28] extends conventional clustering algorithms that produces individual clusters separated by sharp boundaries, and allows an object belonging to differen clusters at the same time to various degrees, overcoming boolean boundaries that are often not natural or even counterintuitive. Let $\{x_i \in \mathbb{R}^d | i = 1, \cdots, N\}$ be a set of N data instances. The objective of FCM using the representation of membership degree can be written as:

$$\begin{aligned} &\min_{U,B} \sum_{c=1}^{C} \sum_{i=1}^{N} u_{ci}^m \|x_i - b_c\|^2 \\ &s.t.\ u_{ci} \in [0,1], \sum_{c=1}^{C} u_{ci} = 1, \forall i = 1, \cdots, N \end{aligned} \tag{1}$$

where $b_c \in \mathbb{R}^d$ is the centroid vector for cluster c ($c = 1, \cdots, C$), u_{ci} denotes the membership degree of i th instances to the c th fuzzy cluster, and m denotes the weighting exponent that controls the extent of membership sharing between fuzzy clusters. Let $U = (u_{ci}) \in \mathbb{R}^{C\times N}$, $B = (b_1^T, b_2^T, \cdots, b_C^T) \in \mathbb{R}^{C\times d}$ be the membership degree matrix and cluster centroid matrix in latent space, respectively. Note that the sum of each column of U should equal one. The objective is to learn the above two matrices by minimizing the sum of distances between the data instances and the corresponding cluster centers. Like traditional k-means, FCM also suffers from the vulnerability to outliers, especially for distributed multi-view data [11].

B. The Proposed Approach

Suppose there are V parties for the same data instances, and each party has a collection of features to form a view. To detect the anomalies, although it is straightforward to perform common algorithms view by view separately, the performance is not prominent largely because an multi-view anomaly may not behave anomalous in every single view (see red rectangle in Figure 1(b)). Therefore, optimizing objective functions on their individual view's data could not sufficiently utilize other view's useful information and lead to bias on multi-view anomaly detection.

To address this issue, we formalize a united Multi-view Fuzzy c-means approach, denoted as MFCM, to learn a common membership degree matrix U by utilizing data from all views. We denote the dataset $X = \{X^{(1)}, \cdots, X^{(k)}, \cdots, X^{(V)}\}$, where $X^{(k)} = (x_1^{(k)}, \cdots, x_N^{(k)}) \in \mathbb{R}^{d_k\times N}$ is the input data with N instances in the k th view. Then the objective of our approach can be formulated as:

$$\begin{aligned} &\min_{U,B^{(1)},\cdots,B^{(V)}} \sum_{k=1}^{V}\sum_{c=1}^{C}\sum_{i=1}^{N} u_{ci}^m \|x_i^{(k)} - b_c^{(k)}\|^2 \\ &s.t.\ u_{ci} \in [0,1], \sum_{c=1}^{C} u_{ci} = 1, \forall i = 1, \cdots, N \end{aligned} \tag{2}$$

where $B^{(k)} = (b_1^{(k)T}, \cdots, b_C^{(k)T}) \in \mathbb{R}^{C\times d_k}$ is the cluster centroid matrix for the k th view.

Remark Unlike running the single-view FCM method in each view, our MFCM approach formalize a united objective by simultaneously summarizing the objectives of all views. This method takes all view's data involved to a single objective function and shares the same membership degree matrix U for each view, which can guarantee that more latent anomalies with inconsistent cluster memberships could be discovered.

C. Anomaly Measurement Criterion

We have discussed the procedure of our approach in clustering all multi-view instances into groups with fuzzy membership degree. In order to make a quantitative estimation of anomaly, we propose a novel entropy-based anomaly score function $f(x_i)$ for sample x_i as

$$f(x_i) = -\sum_{c=1}^{C} u_{ci} \log u_{ci} + \alpha \sum_{c=1}^{C}\sum_{k=1}^{V} \|x_i^{(k)} - b_c^{(k)}\|^2 \tag{3}$$

where α is a trade-off parameter.

Eqn. (3) could help up to identify anomalies in two fold. The first term is the well-known information entropy measuring the confidence of x_i attached to cluster c. If x_i is an anomaly, the membership degree for each cluster could be almost equally low and thus lead to a high entropy. On the contrary, the entropy could be low if the membership degree for one cluster c is extremely large.

The second term is a trade-off term to consider the following scenario. When two instances x_i and x_j have similar membership degree vector, their distances to cluster centroid could become an important factor measuring their deviation degree. Traditionally, α is small to trade-off since the entropy is the main factor of anomaly score while the deviation degree only impact some special cases.

IV. A Distributed Multi-view Anomaly Detection Scheme

To realize the approach proposed, in this section we firstly illustrate the optimization solution to problem (2). Then we design a distributed multi-party architecture and protocol to run the solution in a privacy-preserving way.

A. Model Optimization

Similar to FCM, we introduce the Lagrange multiplier $\lambda_i (i = 1, \cdots, N)$ for each instances, and then the augmented Lagrange function for (2) is re-written as:

$$\begin{aligned} \mathcal{L} = &\sum_{k=1}^{V}\sum_{c=1}^{C}\sum_{i=1}^{N} u_{ci}^m \|x_i^{(k)} - b_c^{(k)}\|^2 \\ &+ \lambda_1 (\sum_{c=1}^{C} u_{c1} - 1) + \cdots + \lambda_N (\sum_{c=1}^{C} u_{cN} - 1) \end{aligned} \tag{4}$$

By computing the derivative of (4) with respective to the variables independently, we get the updated formula as follows:

$$u_{ci} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\sum_{k=1}^{V} \|x_i^{(k)} - b_c^{(k)}\|^2}{\sum_{k=1}^{V} \|x_i^{(k)} - b_j^{(k)}\|^2} \right)^{\frac{1}{m-1}}} \tag{5}$$

and

$$b_c^{(k)} = \frac{\sum_{i=1}^{N} (x_i^{(k)} u_{ci}^m)}{\sum_{i=1}^{N} u_{ci}^m} \tag{6}$$

After updating u_{ci} and $b_c^{(k)}$ for each view iteratively, (4) can be optimized finally.

Algorithm 1 Distributed Optimization for Multi-view Anomaly Detection (DOMAD)

Input: Data matrices $X = \{X^{(1)}, \cdots, X^{(k)}, \cdots, X^{(V)}\}$, each party with $X^{(k)}$.

Output: Binary anomaly label vector Y.

Initialize U, λ_i, m, α for all parties.
For each party with $X^{(k)}, k = 1, \cdots, V$, run the following steps:
while not converged in each iteration **do**
  for cluster $c = 1, \cdots, C$ **do**
    Calculate $b_c^{(k)}$ by Eqn. (6);
    for instance $i = 1, \cdots, N$ **do**
      Calculate and store the distance $\|x_i^{(k)} - b_c^{(k)}\|^2$ into a matrix $D^{(k)}$;
  Broadcast $D^{(k)}$ to other parties;
  Collect other parties' distance matrix;
  for cluster $c = 1, \cdots, C$ **do**
    for instance $i = 1, \cdots, N$ **do**
      Calculate u_{ci} by Eqn. (5);
Calculate the anomaly scores for all instances by (3);
Generate the binary anomaly label Y, if $f(i) > \tau, Y(i) = 0$; otherwise, $Y(i) = 1$;

B. Multi-party Protocol

Note that for the multi-party scenario, the calculation of $b_c^{(k)}$ in (6) can be performed in each party once the acquisition of the value of U. However, to calculate u_{ci} in (5), one need to collect all the distances between $x_i^{(k)}$ and $b_j^{(k)}$ for all clusters in each view k. For the privacy issues, every party can not share their data to others, but only share the resultant distances $\|x_i^{(k)} - b_c^{(k)}\|^2$ to other parties. This is secure since even through a malicious party know other parties' centroid vector they could not calculate their exact values.

There are two ways to calculate and share u_{ci}. The first one is to introduce an independent third-party in charge of collecting all distances shared. After calculating the membership degree matrix U, it distributes U to all parties. Another way is to utilize the broadcast mechanism among all parties to share their distances to others at each iteration. We adopt the latter scheme and summarize the protocol into an algorithm as shown in in Algorithm 1.

V. Performance Evaluation

In order to illustrate the effectiveness of the proposed method for anomaly detection in distributed multi-view data, three benchmark data sets from the UCI Machine Learning Repository [29] have been selected, i.e, *iris*, *breast-wdbc* and *ionosphere*. These data sets are not multi-view data sets. Thus we simulate multi-view data following the approach in [16], [17], [7]. To generate multi-view data sets, the collection of features in each data set is divided into two subsets, where each subset of features and the corresponding feature values are considered as one view of the data. A summary of the resulting data sets is shown in Table I.

TABLE I
A Summery of The Data Sets Used

	Iris	Breast-wdbc	Ionosphere
Number of Features	4	32	34
Number of Classes	3	2	2
Number of Instances	150	569	351
Number of Views	2	2	2
Number of Features in Each View	2	16	17

Two types of outliers, i.e., class-outlier and attribute-outlier are considered in the experiments. Class-outlier is an outlier that exhibits inconsistent characteristics (e.g., cluster membership) across different views. Attribute-outlier is an outlier that exhibits consistent abnormal behaviors in each view. In order to generate these two types of outliers, the following two steps of data-processing have been adopted:

1) For class-outliers, take two objects from two different classes and swap the subsets in one view but not in the other [16];
2) For attribute-outliers, a sample is randomly selected, and its features in all views are replaced by random values [7].

This is the first work on distributed multi-view anomaly detection, to the best of our knowledge. To validate the performance of the proposed DOMAD method, we compare it with four centralized multi-view anomaly detection baselines as follows:

- Direct Robust Matrix Factorization (DRMF) [30] projects multiple views into a matrix and follows a single-view anomaly detection approach.
- Low-Rank Representation (LRR) [31] is also a representative anomaly detection method for single-view data. We merge multiple views into a single view to detect anomalies.
- HOrizontal Anomaly Detection (HOAD) [16] is a cluster-based multi-view anomaly detection method finding the inconsistency instances among multiple views.
- Anomaly detection using Affinity Propagation (AP) [17] proposed two affinity matrices and presented four anomaly measurement strategies for multi-view anomaly detection.

Receiver Operating Characteristic (ROC) curve is commonly used to show the trade-off between the True Positive Rate (TPR) and the False Positive Rate of an algorithm. The TPR and FPR are defined as $TPR = TP/(TP + FN)$ and $FPR = FP/(TP + TN)$, where TP, FN, TN and FP represent true positives, false negatives, true negatives, and false positives, respectively. In this experiment, the positives are either class-outliers or attribute-outliers generated by the two steps mentioned above. Furthermore, the Area Under ROC Curve (AUC) is employed in the experiment to evaluate the general performance of the proposed methods.

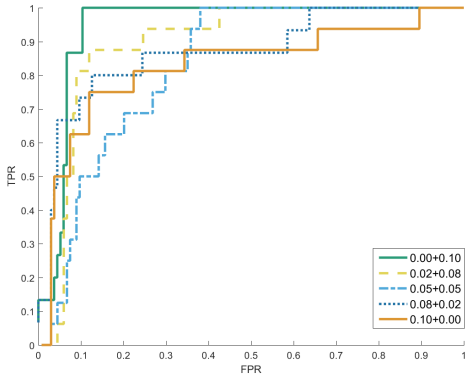
In order to compare the results with [17], [16], the outliers are generated randomly for 50 times for each data set.

Therefore, each entry of averaged result and standard deviation in Table II is reported based on 50 times of assessments. To simulate the real-world applications in different circumstances, we conduct three settings by mixing both outliers with different ratios: (1) 2% class-outlier of the total sample number + 8% attribute-outlier of the total sample number, represented in format "DatasetName-2-8"; (2) 5% class-outlier + 5% attribute-outlier in format DatasetName-5-5; (3) 8% class-outlier + 2% attribute-outlier in format DatasetName-8-2. The number of clusters in DOMAD is set to the number of classes in each data sets and the value of α is set to 0.5×10^{-3}.

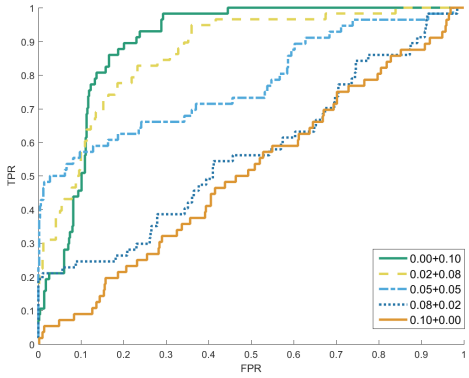
Table II reports the AUC values on 3×3 data sets with different outlier settings. It can be seen from Table II that the proposed method DOMAD generally outperforms the other methods when the class-attribute outlier rate is low. Specifically, in the experiments with "-2-8" configuration, DOMAD performs better than other methods significantly. However, in the experiments with "-8-2" configuration, the method proposed in [17] achieved the best performance on two of the three tested data sets.

As it can be seen from Table II and Fig. 2 that the proposed DOMAD achieves better results when dealing with attribute-outliers than that on class-outliers. This is mainly because the proposed algorithm is not regularized neither for individual views nor individual clusters. This problem opens up an avenue for significant further investigation in fuzzy clustering approaches for multi-view anomaly detection.

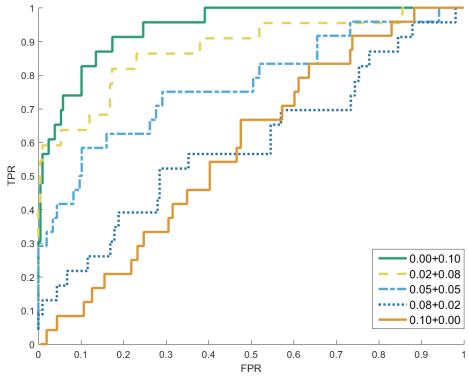
Since the number of data views k is usually pre-defined in practical problems, the only major parameter in the proposed DOMAD algorithm is α (in Eqn. (3)). In order to evaluate the robustness and stability of DOMAD against the change of α and k, Fig. 3(a) is provided to show the average AUC values (of 50 runs) with respect to the changing values of α with a step of $10^{-n}, n = \{-4, -3, \cdots, 3\}$. Figure 3(b) shows the ROC curves of DOMAD on the Breast-wdbc-5-5 data set with different numbers of data views. It can be seen from them that the performance of DOMAD in these settings are relatively steady. Therefore, we chose $\alpha = 0.5 \times 10^{-3}$ in the experiments of this paper and recommend it as the default value in practice. It is worth noticing that since the trends of average AUC values(and ROC curves) with respect to different values α (and k, respectively) are similar on the 3 testing data sets, only the results on the breast data set are presented.



(a) Iris



(b) Breast-wdbc



(c) Ionosphere

Fig. 2. Results of ROC

VI. Conclusion

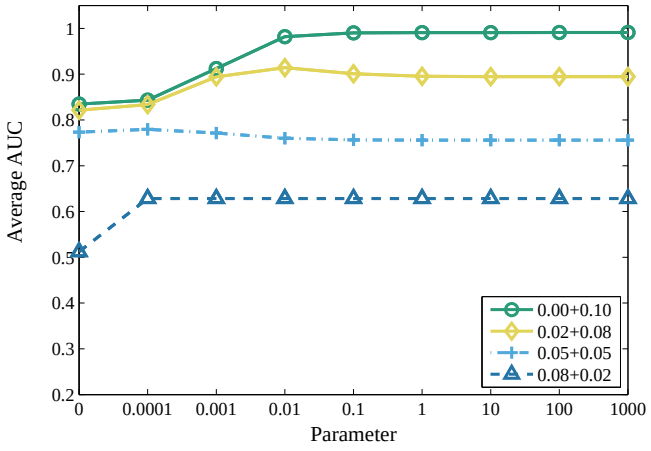
In this paper, we propose a novel fuzzy anomaly detection approach for distributed multi-view data , named DOMAD. With this approach, we utilize a fuzzy clustering method to simultaneously learn a membership degree matrix for each view and then detects anomalies for all parties. DOMAD can be performed independently on each parties under a multi-party protocol and thus can prevent information leakage from their private data. Finally we devise an novel anomaly detection criterion to quantify the abnormal score from the shared membership degree matrix. Extensive experiments on real world datasets demonstrate the effectiveness of our approach on detecting multi-view anomalies privately and collaboratively.

VII. Acknowledgement

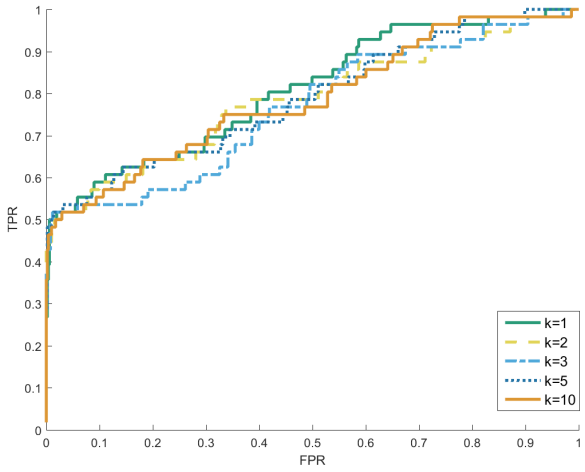
This work was supported in part by the Natural Science Foundation of Hebei Province under Grant F2016502069, and in part by the National Natural Science Foundation of China under Grant 61502168.

TABLE II
RESULTS OF AUC

Data Set	DRMF[30]	LRR[31]	HOAD[16]	AP[17]	DOMAD(this paper)
Iris-2-8	0.749 ± 0.044	0.779 ± 0.062	0.167 ± 0.057	0.326 ± 0.027	**0.904 ± 0.032**
Iris-5-5	0.714 ± 0.038	0.762 ± 0.107	0.309 ± 0.062	0.630 ± 0.021	**0.894 ± 0.040**
Iris-8-2	0.651 ± 0.037	0.740 ± 0.100	0.430 ± 0.055	0.840 ± 0.021	**0.864 ± 0.051**
Breast-wdbc-2-8	0.764 ± 0.013	0.586 ± 0.037	0.555 ± 0.072	0.293 ± 0.012	**0.874 ± 0.015**
Breast-wdbc-5-5	0.708 ± 0.034	0.493 ± 0.017	0.586 ± 0.061	0.532 ± 0.024	**0.772 ± 0.031**
Breast-wdbc-8-2	0.684 ± 0.024	0.508 ± 0.043	0.634 ± 0.046	**0.693 ± 0.023**	0.636 ± 0.040
Ionosphere-2-8	0.705 ± 0.029	0.699 ± 0.025	0.446 ± 0.074	0.632 ± 0.033	**0.865 ± 0.028**
Ionosphere-5-5	0.676 ± 0.040	0.627 ± 0.029	0.422 ± 0.051	**0.761 ± 0.025**	0.736 ± 0.043
Ionosphere-8-2	0.634 ± 0.023	0.511 ± 0.014	0.448 ± 0.041	**0.822 ± 0.030**	0.620 ± 0.053



(a) Average AUC Values with α Changing



(b) ROC of Breast-wdbc-5-5 with Number of View k Changing

Fig. 3. Test of Robustness on Breast-wdbc Data Set.

REFERENCES

[1] C. Tsang, S. Kwong, and H. Wang, "Anomaly intrusion detection using multi-objective genetic fuzzy system and agent-based evolutionary computation framework," in *Proceedings of the 5th IEEE International Conference on Data Mining*, 2005, pp. 789–792.

[2] S. Fakhraei, J. Foulds, M. Shashanka, and L. Getoor, "Collective spammer detection in evolving multi-relational social networks," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1769–1778.

[3] H. Wang, H. Wen, F. Yi, H. Zhu, and L. Sun, "Road traffic anomaly detection via collaborative path inference from GPS snippets," *Sensors*, vol. 17, no. 3, p. 550, 2017.

[4] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[5] Y. Song, Z. Wen, C.-Y. Lin, and R. Davis, "One-class conditional random fields for sequential anomaly detection," in *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, 2013, pp. 1685–1691.

[6] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Transactions on Knowledge Discovery from Data*, vol. 6, no. 1, p. 3, 2012.

[7] H. Zhao and Y. Fu, "Dual-regularized multi-view outlier detection," in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015, pp. 4077–4083.

[8] K. Das, K. Bhaduri, and P. Votava, "Distributed anomaly detection using 1-class SVM for vertically partitioned data," *Statistical Analysis and Data Mining*, vol. 4, no. 4, pp. 393–406, 2011.

[9] M. Stolpe, K. Bhaduri, K. Das, and K. Morik, "Anomaly detection in vertically partitioned data by distributed core vector machines," in *Proceedings of the Machine Learning and Knowledge Discovery in Databases - European Conference*, 2013, pp. 321–336.

[10] C. O'Reilly, A. Gluhak, and M. A. Imran, "Distributed anomaly detection using minimum volume elliptical principal component analysis," *IEEE Transactions on Knowledge Data Engineering*, vol. 28, no. 9, pp. 2320–2333, 2016.

[11] S. Bickel and T. Scheffer, "Multi-view clustering," in *Proceedings of the 4th IEEE International Conference on Data Mining, 1-4 November 2004, Brighton, UK*, 2004, pp. 19–26.

[12] E. Bruno and S. Marchand-Maillet, "Multiview clustering: a late fusion approach using latent models," in *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2009, pp. 736–737.

[13] K. Chaudhuri, S. M. Kakade, K. Livescu, and K. Sridharan, "Multi-view clustering via canonical correlation analysis," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 129–136.

[14] Y. Kanzawa, "Fuzzy clustering based on α-divergence for spherical data and for categorical multivariate data," in *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*, 2015, pp. 1–8.

[15] K. Honda, T. Oda, and A. Notsu, "Fuzzy co-clustering of vertically partitioned cooccurrence data with privacy consideration," in *Proceedings of the 2014 IEEE International Conference on Fuzzy Systems*, 2014, pp. 2500–2504.

[16] J. Gao, W. Fan, D. S. Turaga, S. Parthasarathy, and J. Han, "A spectral framework for detecting inconsistency across multi-source object relationships," in *Proceedings of the 11th IEEE International Conference on Data Mining*, 2011, pp. 1050–1055.

[17] A. M. Alvarez, M. Yamada, A. Kimura, and T. Iwata, "Clustering-based anomaly detection in multi-view data," in *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management*, 2013, pp. 1545–1548.
[18] S. Li, M. Shao, and Y. Fu, "Multi-view low-rank analysis for outlier detection," in *Proceedings of the 2015 SIAM International Conference on Data Mining*, 2015, pp. 748–756.
[19] J. Vaidya and C. Clifton, "Privacy-preserving *k*-means clustering over vertically partitioned data," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, pp. 206–215.
[20] T. Yu, D. T. Lee, S. Chang, and J. Zhan, "Multi-party k-means clustering with privacy consideration," in *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing with Applications*, 2010, pp. 200–207.
[21] P. Su, C. Shang, T. Chen, and Q. Shen, "Exploiting data reliability and fuzzy clustering for journal ranking," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 5, pp. 1306–1319, 2017.
[22] P. Su, Q. Shen, T. Chen, and C. Shang, "Ordered weighted aggregation of fuzzy similarity relations and its application to detecting water treatment plant malfunction," *Engineering Applications of Artificial Intelligence*, vol. 66, pp. 17–29, 2017.
[23] P. Su, C. Shang, and Q. Shen, "A hierarchical fuzzy cluster ensemble approach and its application to big data clustering," *Journal of Intelligent & Fuzzy Systems*, vol. 28, no. 6, pp. 2409–2421, 2015.
[24] T. Chen, Q. Shen, P. Su, and C. Shang, "Fuzzy rule weight modification with particle swarm optimisation," *Soft Computing*, vol. 20, no. 8, pp. 2923–2937, 2016.
[25] P. Su, C. Shang, and Q. Shen, "Link-based approach for bibliometric journal ranking," *Soft Computing*, vol. 17, no. 12, pp. 2399–2410, 2013.
[26] T. Chen, C. Shang, P. Su, and Q. Shen, "Induction of accurate and interpretable fuzzy rules from preliminary crisp representation," *Knowledge-Based Systems*, vol. 146, pp. 152–166, 2018.
[27] H. Wang, H. Wang, F. Yi, H. Wen, G. Li, and L. Sun, "Context-aware personalized path inference from large-scale GPS snippets," *Expert Systems with Applications*, vol. 91, pp. 78–88, 2018.
[28] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms," *Advanced Applications in Pattern Recognition*, vol. 22, no. 1171, pp. 203–239, 1999.
[29] M. Lichman, "UCI machine learning repository," 2013. [Online]. Available: http://archive.ics.uci.edu/ml
[30] L. Xiong, X. Chen, and J. G. Schneider, "Direct robust matrix factorizatoin for anomaly detection," in *Proceedings of the 11th IEEE International Conference on Data Mining*, 2011, pp. 844–853.
[31] A. Liu and D. N. Lam, "Using consensus clustering for multi-view anomaly detection," in *Proceedings of the 2012 IEEE Symposium on Security and Privacy Workshops*, 2012, pp. 117–124.