

New Outlier Detection Method Based on Fuzzy Clustering

MOH'D BELAL AL-ZOUBI¹, ALI AL-DAHOUD², ABDEL FATAH A. YAHYA³

¹Department of Computer Information Systems
University Of Jordan
Amman
JORDAN
E-mail: mba@ju.edu.jo

²Faculty of Science and IT
Al- Zaytoonah University
Amman
JORDAN
E-mail: aldahoud@alzaytoonah.edu.jo

³Faculty of Science and IT
Al- Zaytoonah University
Amman
JORDAN
E-mail: science@alzaytoonah.edu.jo

Abstract: - In this paper, a new efficient method for outlier detection is proposed. The proposed method is based on fuzzy clustering techniques. The c-means algorithm is first performed, then small clusters are determined and considered as outlier clusters. Other outliers are then determined based on computing differences between objective function values when points are temporarily removed from the data set. If a noticeable change occurred on the objective function values, the points are considered outliers. Test results were performed on different well-known data sets in the data mining literature. The results showed that the proposed method gave good results.

Keywords: Outlier detection, Data mining, Clustering, Fuzzy clustering, FCM algorithm, Noise removal

1. Introduction

Outlier detection is an important pre-processing task [1]. It has many practical applications in several areas, including image processing [2], remote sensing [3], fraud detection [4], identifying computer network intrusions and bottlenecks [5], criminal activities in e-commerce and detecting suspicious activities [6]. Different approaches have been proposed to detect outliers, and a good survey can be found in [7, 8, 9].

Clustering is a popular technique used to group similar data points or objects in groups or clusters [10]. Clustering is an important tool for outlier analysis. Several clustering-based outlier detection techniques have been developed, most of which rely

on the key assumption that normal objects belong to large and dense clusters, while outliers form very small clusters [11, 12].

It has been argued by many researchers whether clustering algorithms are an appropriate choice for outlier detection. For example, in [13], the authors reported that clustering algorithms should not be considered as outlier detection methods. This might be true for some of the clustering algorithms, such as the k-means clustering algorithm [14]. This is because the cluster means produced by the k-means algorithm is sensitive to noise and outliers [15, 16].

The Fuzzy C-Means algorithm (FCM), as one of the best known and the most widely used fuzzy clustering algorithms, has been utilized in a wide variety of applications, such as medical imaging, remote sensing,

data mining and pattern recognition [17, 18, 19, 20, 21]. Its advantages include a straightforward implementation, fairly robust behavior, applicability to multidimensional data, and the ability to model uncertainty within the data.

FCM partitions a collection of n data points x_i , $i = 1, \dots, n$ into c fuzzy groups, and finds a cluster center in each group so that the objective function of the dissimilarity measure is minimized. FCM employs fuzzy partitioning so that a given data point can belong to several groups with the degree of belongingness specified by membership grades between 0 and 1. The membership matrix U is allowed to have elements with values between 0 and 1. However, imposing normalization stipulates that the summation of degrees of belongingness for a data set should always be equal to unity, as shown in Equation (1):

$$\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n \quad (1)$$

The objective function for FCM is:

$$J(U, c_1, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_j^n u_{ij}^m d_{ij}^2, \quad (2)$$

where u_{ij} values are between 0 and 1; c_i is the cluster center of fuzzy group i ; $d_{ij} = \|c_i - x_j\|$ is the Euclidean distance between i th cluster and j th data point; and $m \in [1, \infty)$ is the weighting exponent.

The necessary conditions for Equation (2) to reach its minimum are:

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}, \quad (3)$$

and

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (4)$$

The FCM algorithm is simply an iterated procedure through the preceding two necessary conditions. The algorithm determines the cluster centers c_i and the membership matrix U using the following steps [17]:

Step [1]: Initialize the membership matrix U with random values between 0 and 1 so that the constraints in Equation (1) are satisfied.

Step [2]: Calculate c fuzzy cluster centers

c_i , $i = 1, \dots, c$, using Equation (3).

Step [3]: Compute the objective function according to Equation (2). Stop if either it is below a certain tolerance value or its improvement over previous iteration is below a certain threshold, ε .

Step [4]: Compute a new U using Equation (4). Go to step 2.

In this paper, a new method based on FCM algorithm is proposed. Note that our method can be easily implemented on other clustering algorithms.

2. Related Work

As discussed in [11, 12], there is no single universally applicable or generic outlier detection approach. Therefore, many approaches have been proposed to detect outliers. These approaches can be classified into four major categories based on the techniques used [22] which are: distribution-based, distance-based, density-based and clustering-based approaches.

Distribution-based approaches [23, 24, 25] develop statistical models (typically for the normal behavior) from the given data and then apply a statistical test to determine if an object belongs to this model or not. Objects that have low probability to belong to the statistical model are declared as outliers. However, Distribution-based approaches cannot be applied in multidimensional scenarios because they are univariate in nature. In addition, a prior knowledge of the data distribution is required, making the distribution-based approaches difficult to be used in practical applications [22].

In the distance-based approach [8, 9, 26, 27], outliers are detected as follows. Given a distance measure on a feature space, a point q in a data set is an outlier with respect to the parameters M and d , if there are less than M points within the distance d from q , where the values of M and d are decided by the user. The problem with this approach is that it is difficult to determine the values of M and d .

Density-based approaches [28, 29] compute the density of regions in the data and declare the objects in low dense regions as outliers. In [28], the authors assign an outlier score to any given data point, which is known as the Local Outlier Factor (LOF), depending on its distance from its local neighborhood. A similar work is reported in [29].

Clustering-based approaches [11, 30, 31, 32, 33] consider clusters of small sizes as clustered outliers. In these approaches, small clusters (i.e., clusters containing

significantly less points than other clusters do) are considered outliers.

The advantage of the clustering-based approaches is that they do not have to be supervised. Moreover, clustering-based techniques are capable of being used in an incremental mode (i.e., after learning the clusters, new points can be inserted into the system and tested for outliers).

Custem and Gath [31] present a method based on fuzzy clustering. In order to test the absence or presence of outliers, two hypotheses are used. However, the hypotheses do not account for the possibility of multiple clusters of outliers.

Jiang et. al. [32] presented a two-phase method to detect outliers. In the first phase, the authors proposed a modified k-means algorithm to cluster the data, and then, in the second phase, an Outlier-Finding Process (OFP) is proposed. The small clusters are selected and regarded as outliers, using minimum spanning trees.

In [11], clustering methods have been applied. The key idea is to use the size of the resulting clusters as indicators of the presence of outliers. The authors use a hierarchical clustering technique. A similar approach was reported in [34].

Acuna and Rodriguez [33] performed the PAM algorithm [16] followed by the Separation Technique (henceforth, the method will be termed PAMST). The separation of a cluster A is defined as the smallest dissimilarity between two objects; one belongs to Cluster A and the other does not. If the separation is large enough, then all objects that belong to that cluster are considered outliers. In order to detect the clustered outliers, one must vary the number k of clusters until obtaining clusters of a small size with a large separation from other clusters.

In [35], the authors proposed a clustering-based approach to detect outliers. The K-means clustering algorithm is used. As mentioned in [15], the k-means is sensitive to outliers, and hence may not give accurate results.

In [36], the author proposed a clustering-based (will be termed CB) method to detect outliers. First, the PAM algorithm is performed, producing a set of clusters and a set of medoids (cluster centers). To detect the outliers, the Absolute Distances between the Medoid, μ , of the current cluster and each one of the Points, p_i , in the same cluster (i. e., $|p_i - \mu|$) are computed. The produced value is termed (*ADMP*). If the *ADMP* value is greater than a calculated

threshold, T , then the point is considered an outlier; otherwise, it is not. The value of T is calculated as the average of all *ADMP* values of the same cluster multiplied by (1.5).

The basic structure of the proposed method is as follows:

Step 1. Perform PAM clustering algorithm to produce a set of k clusters.

Step 2. Determine small clusters and consider the points (objects) that belong to these clusters as outliers.

For the rest of the clusters (not determined in Step 2)

Begin

Step 3. For each cluster j, compute the ADMPj and Tj values.

Step 4. For each point i in cluster j, if ADMPij > Tj then classify point i as an outlier; otherwise not.

End

The test results show that the CB method gave good results when applied to different data sets. In this paper, we compare our proposed method with the CB method.

3. Proposed Approach

In this paper, a new clustering-based approach for outliers detection is proposed. First, we execute the FCM algorithm, producing an objective function. Small clusters are then determined and considered as outlier clusters. We follow [9] to define small clusters. A small cluster is defined as a cluster with fewer points than half the average number of points in the c clusters.

To detect the outliers in the rest of clusters (if any), we (temporarily) remove a point from the data set and re-execute the FCM algorithm. If the removal of the point causes a noticeable decrease in the objective function value, the point is considered an outlier; otherwise, it is not.

The objective function represents the (Euclidean) sum of squared distances between the cluster centers and the points belonging to these clusters multiplied by the membership values (grades) of each cluster produced by the FCM.

Our idea is based on the following philosophy: the Objective Function (OF) produced by the FCM algorithm represents the distances between the cluster centers and the points belonging to these clusters. Removing a point from the data set will cause a decrease in the OF value because of the total sum of

distances between each point and the cluster center belonging to it. If this decrease is greater than a certain threshold, the point is then nominated to be an outlier.

To explain our proposed method in a more formal way, we will use the following parameter names:

- OF the objective function for the whole set.
- OF_i the objective function after removing point p_i from the set.
- Let DOF_i = OF - OF_i. and T = (1.5).

The basic structure of the proposed method is as follows:

Step 1. Execute the FCM algorithm to produce an Objective Function (OF)

Step 2. Determine small clusters and consider the points that belong to these clusters as outliers.

Step 3. For the rest of the points (not determined in Step 2):

Begin

SUM = 0

For each point p_i in the set

DO

remove p_i from the set

calculate OF_i

calculate DOF_i

SUM = SUM + DOF_i

return p_i back to the set;

End DO

AvgDOF = SUM / n

For each point, p_i

DO

if (DOF_i > T(AvgDOF)) then classify p_i as an outlier

End DO

End

4. Results and Discussion

In this section, we will investigate the effectiveness of our proposed approach when applied on three data sets: data set1, which is discussed in [16], 22 data points, and two dimensions. Data set2 is the Wood data set [37] and has been tested for outlier detection in [38, 39] with 20 points and six dimensions. Data set3 is the Bupa data set with six dimensions. Data set3 was obtained from the UCI repository of machine learning databases [40]. The fourth data set is the the well-known Iris data set

[40]. The Iris data set has three classes of Iris Plants: Iris- setosa, Iris-versicolor and Iris viginica and has four variables (dimensions). It is known that the Iris data have some outliers.

We will start our experimentations with data set1, which is shown in Fig. 1. It is clear that the set contains three natural clusters and two outliers, namely, points p₆ and p₁₃ (highlighted in Table 1).

When the defined number of clusters is three, then Step 3 will classify the points p₆ and p₁₃ as outliers. Table 1 shows the point's ID (ID) in the first column. The second column shows the OF_i values when the current point is removed from the set. The third column shows the DOF_i values. The value of the objective function for data set1, OF, is (57.688). The AvgDOF is (2.88). The value of T(AvgDOF) is (4.31), which is our threshold value. Since the DOF_i values for each of the points p₆ and p₁₃ were greater than T(AvgDOF), these two points were detected as outlier points.

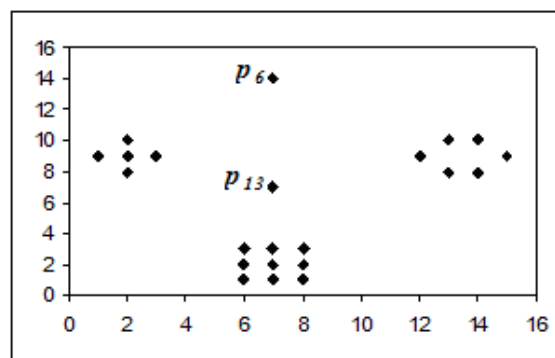


Fig. 1 Data set1 (contains 22 data points)

The value of the threshold (T) was obtained, following the try and error approach, when testing different data sets.

The second data set (data set2) is The Wood dataset with data points 4, 6, 8, and 19 being outliers [37]. The outliers are said not to be easily identified by observation [37, 38]. We have found that our method clearly identifies all outliers in the set.

Table 1: Data set1[16].

ID	OF _i	DOF _i
1	55.33	2.36
2	56.70	0.99
3	57.45	0.24
4	55.83	1.86
5	57.17	0.52
6	34.84	22.85
7	55.71	1.98
8	56.59	1.10
9	56.19	1.50
10	56.14	1.55
11	55.73	1.96
12	54.38	3.32
13	47.79	9.90
14	55.82	1.87
15	56.85	0.84
16	55.83	1.86
17	56.57	1.12
18	57.67	0.02
19	56.57	1.12
20	55.24	2.45
21	56.27	1.42
22	55.24	2.45

The Third data set is the Bupa data set [39]. This data set has two classes and 6 dimensions. Acuna and Rodriguez [29] have applied many different criteria with the help of a parallel coordinate plot to decide about the doubtful outliers in the Bupa data set. They found 22 doubtful outliers in class 1 and 26 doubtful outliers in class 2 (a total of 48 doubtful outliers).

Applying our proposed method, 16 outliers were identified in class 1 with the point's ID (ID) shown in Table 2, column 1, and 23 outliers were identified in class 2, as shown in Table 3.

Applying the CB method [36], 19 outliers were identified in class 1 with the point's ID (ID) shown in Table 4, column 1, and 17 outliers were identified in class 2, as shown in Table 5.

Although the CB method performed better in class1 of the Bupa data set, however our proposed method performed better in class 2 of the Bupa data set.

The fourth data set is the well-known Iris data set (Blake and Merz, 1998). The Iris data set has three classes of Iris Plants: Iris- setosa, Iris-versicolor and Iris virginica and has four variables (dimensions). It

is known that the Iris data have some outliers. Different experimentations have been conducted in [Acona] and showed that there are 10 outliers in class3 (Iris viginica) of the Iris data set as shown in Table 6 (column1). Applying our proposed method, nine outliers are detected as shown in Table 7 (column2). Applying the CB method, eight outliers are detected. Hence, our proposed method performed better in this case.

5. Conclusion

An efficient method for outlier detection is proposed in this paper. The proposed method is based on fuzzy clustering techniques. The FCM algorithm is first performed, then small clusters are determined and considered as outlier clusters. Other outliers are then determined based on computing differences between objective function values when points are temporally removed from the data set. If a noticeable change occurred on the objective function values, the points are considered outliers. The test results show that the proposed approach gave effective results when applied to different data sets.

However, our proposed method is very time consuming. This is because the FCM algorithm has to run n times, where n is the number of points in a set. This will be our focus in the future work.

Table 2: The data points detected by our proposed method from class1 of the Bupa data set (T = detected, and F = not detected).

ID	Detected
20	T
22	T
25	F
148	T
167	F
168	F
175	T
182	T
183	T
189	F
190	T
205	T
261	T
311	F
312	T
313	T
316	T
317	T
326	T
335	T
343	F
345	T

Table 3: The data points detected by our proposed method from class2 of the Bupa data set (T = detected, and F = not detected).

ID	Detected
2	T
36	T
77	T
85	T
111	T
115	T
123	T
133	F
134	T
139	T
157	T
179	T
186	T
187	T
224	T
233	T
252	T
278	F
286	T
294	T
300	T
307	F
323	T
331	T
337	T
342	T

Table 4: The data points detected by the CB method from class1 of the Bupa data set (T = detected, and F = not detected).

ID	Detected
20	F
22	F
25	T
148	T
167	T
168	T
175	T
182	T
183	T
189	T
190	T
205	T
261	T
311	T
312	T
313	F
316	T
317	T
326	T
335	T
343	T
345	T

Table 5: The data points detected by the CB method from class2 of the Bupa data set (T = detected, and F = not detected).

ID	Detected
2	F
36	T
77	T
85	T
111	F
115	T
123	T
133	T
134	T
139	T
157	F
179	T
186	T
187	T
224	F
233	F
252	F
278	T
286	T
294	F
300	T
307	F
323	T
331	T
337	F
342	T

Table 6: Outliers in the Iris dataset, Class 3 (column1), outliers detected by our proposed method (T = detected, and F = not detected).

ID	Detected
106	T
107	T
108	F
110	T
118	T
119	T
123	T
131	T
132	T
136	T

Table 7: Outliers in the Iris dataset, Class 3 (column1), outliers detected by the CB method (T = detected, and F = not detected).

ID	Detected
106	T
107	T
108	F
110	T
118	T
119	T
123	T
131	F
132	T
136	T

References

- [1] Han, J. and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2nd edition, 2006.
- [2] Al-Zoubi, M. B, A. M. Kamel and M. J. Radhy Techniques for Image Enhancement and Noise Removal in the Spatial Domain, *WSEAS Transactions on Computers*, Vol. 5, No. 5, 2006, pp. 1047-1052.
- [3] Yang J. and Z. Zhao, A Fast Geometric Rectification of Remote Sensing Imagery Based on Feature Ground Control Point Database, *WSEAS Transactions on Computers*, Vol. 8, No. 2, 2009, pp. 195-204.
- [4] Bolton, R. and D. J. Hand, Statistical Fraud Detection: A Review, *Statistical Science*, Vol. 17, No. 3, 2002, pp. 235-255.
- [5] Lane, T. and C. E. Brodley. Temporal Sequence Learning and Data Reduction for Anomaly Detection, 1999. *ACM Transactions on Information and System Security*, Vol. 2, No. 3, 1999, pp. 295-331.
- [6] Chiu, A. and A. Fu, Enhancement on Local Outlier Detection. *7th International Database Engineering and Application Symposium (IDEAS03)*, 2003, pp. 298-307.
- [7] Knorr, E. and R. Ng, Algorithms for Mining Distance-based Outliers in Large Data Sets, Proc. the 24th *International Conference on Very Large Databases (VLDB)*, 1998, pp. 392-403.
- [8] Knorr, E., R. Ng, and V. Tucakov, Distance-based Outliers: Algorithms and Applications. *VLDB Journal*, Vol. 8, No. 3-4, 2000, pp. 237-253.
- [9] Hodge, V. and J. Austin, A Survey of Outlier Detection Methodologies, *Artificial Intelligence Review*, Vol. 22, 2004, pp. 85–126.
- [10] Jain, A. and R. Dubes, *Algorithms for Clustering Data*. Prentice-Hal, 1988.
- [11] Loureiro, A., L. Torgo and C. Soares,. Outlier Detection Using Clustering Methods: a Data Cleaning Application, in *Proceedings of KNet Symposium on Knowledge-based Systems for the Public Sector*. Bonn, Germany, 2004.
- [12] Niu, K., C. Huang, S. Zhang, and J. Chen, ODDC: Outlier Detection Using Distance Distribution Clustering, T. Washio et al. (Eds.): *PAKDD 2007 Workshops, Lecture Notes in Artificial Intelligence (LNAI)* 4819, pp. 332–343, Springer-Verlag.
- [13] Zhang, J. and H. Wang, Detecting Outlying Subspaces for High-Dimensional Data: the New Task, *Algorithms, and Performance, Knowledge and Information Systems*, Vol 10, No. 3, 2006, pp.333-355.
- [14] MacQueen, J., Some Methods for Classification and Analysis of Multivariate Observations. *Proc. 5th Berkeley Symp. Math. Stat. and Prob*, 1967, pp. 281-297.
- [15] Laan, M., K. Pollard and J. Bryan, A New Partitioning Around Medoids Algorithms, *Journal of Statistical Computation and Simulation*, Vol. 73, No. 8, 2003, pp. 575-584.

- [16] Kaufman, L. and P. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, 1990.
- [17] Bezdek, J. L. Hall, and L. Clarke, Review of MR Image Segmentation Techniques Using Pattern Recognition, *Medical Physics*, Vol. 20, No. 4, 1993, pp. 1033–1048.
- [18] Pham, D, Spatial Models for Fuzzy Clustering, *Computer Vision and Image Understanding*, Vol. 84, No. 2, 2001, pp. 285–297.
- [19] Rignot, E, R. Chellappa, and P. Dubois, Unsupervised Segmentation of Polarimetric SAR Data Using the Covariance Matrix, *IEEE Trans. Geosci. Remote Sensing*, Vol. 30, No. 4, 1992, pp. 697–705.
- [20] Al- Zoubi, M. B., A. Hudaib and B Al-Shboul A Proposed Fast Fuzzy C-Means Algorithm, *WSEAS Transactions on Systems*, Vol. 6, No. 6, 2007, pp. 1191-1195.
- [21] Maragos E. K and D. K. Despotis, The Evaluation of the Efficiency with Data Envelopment Analysis in Case of Missing Values: A fuzzy approach, *WSEAS Transactions on Mathematics*, Vol. 3, No. 3, 2004, pp. 656-663.
- [22] Zhang, Q. and I. Couloigner, 2005. A New and Efficient K-Medoid Algorithm for Spatial Clustering, in O. Gervasi et al. (Eds.): *ICCSA 2005, Lecture Notes in Computer Science (LNCS)* 3482, pp. 181 – 189, Springer-Verlag, 2005.
- [23] Hawkins, D., *Identifications of Outliers*, Chapman and Hall, London, 1980.
- [24] Barnett, V. and T. Lewis, 1994. *Outliers in Statistical Data*. John Wiley.
- [25] Rousseeuw, P. and A. Leroy, *Robust Regression and Outlier Detection*, 3rd ed.. John Wiley & Sons, 1996.
- [26] Ramaswami, S., R. Rastogi and K. Shim, Efficient Algorithm for Mining Outliers from Large Data Sets. *Proc. ACM SIGMOD*, 2000, pp. 427-438.
- [27] Angiulli, F. and C. Pizzuti, Outlier Mining in Large High-Dimensional Data Sets, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 No. 2, 2005, pp. 203-215.
- [28] Breunig, M., H. Kriegel, R. Ng and J. Sander, 2000. LOF: Identifying Density-Based Local Outliers. In *Proceedings of ACM SIGMOD International Conference on Management of Data*. ACM Press, 2000, pp. 93–104.
- [29] Papadimitriou, S., H. Kitawaga, P. Gibbons, and C. Faloutsos, LOCI: Fast Outlier Detection Using the Local Correlation Integral. *Proc. of the International Conference on Data Engineering*, 2003, pp. 315-326.
- [30] Gath, I and A. Geva, Fuzzy Clustering for the Estimation of the Parameters of the Components of Mixtures of Normal Distribution, *Pattern Recognition Letters*, Vol. 9, 1989, pp. 77-86.
- [31] Cutsem, B and I. Gath, Detection of Outliers and Robust Estimation using Fuzzy Clustering, *Computational Statistics & Data Analyses*, Vol. 15, 1993, pp. 47-61.
- [32] Jiang, M., S. Tseng and C. Su, Two-phase Clustering Process for Outlier Detection, *Pattern Recognition Letters*, Vol. 22, 2001, pp. 691-700.
- [33] Acuna E. and Rodriguez C., A Meta Analysis Study of Outlier Detection Methods in Classification, Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez, available at academic.uprm.edu/~eacuna/paperout.pdf. In proceedings IPSI, 2004, Venice.
- [34] Almeida, J., L. Barbosa, A. Pais and S. Formosinho, Improving Hierarchical Cluster Analysis: A New Method with Outlier Detection and Automatic Clustering, *Chemometrics and Intelligent Laboratory Systems* Vol. 87, 2007, pp. 208–217.
- [35] Yoon, K., O. Kwon and D. Bae, 2007. An Approach to Outlier Detection of Software Measurement Data Using the K-means Clustering Method, First *International Symposium on Empirical Software Engineering and Measurement (ESEM 2007)*, Madrid, pp. 443-445.
- [36] Al- Zoubi, M. B., An Effective Clustering-Based Approach for Outlier Detection, *European Journal of Scientific Research*, Vol. 28, No. 2, 2009, pp. 310-316.
- [37] Draper , N. R. and H. Smith. *Applied Regression Analysis*, John Wiley and Sons, New York, 1966.

- [38] Rousseeuw, P. J. and A. M. Leroy. *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc., New York, 1987.
- [39] Williams, G., Baxter, R., He, H., Hawkins, S., and Gu, L., A Comparative Study of RNN for Outlier Detection in data Mining. In *Proceedings of the 2002 IEEE International Conference on Data Mining*. IEEE Computer Society, Washington, DC, USA, 2002, pp. 709 – 712.
- [40] Blake, C. L. & C. J. Merz, UCI Repository of Machine Learning Databases, <http://www.ics.uci.edu/mllearn/MLRepository.html>, University of California, Irvine, Department of Information and Computer Sciences, 1998.