Anomaly Detection in Smart Grid With Help of an Improved OPTICS Using Coefficient of Variation

Farid Fathnia
M.Sc. Student
Department of Electrical Engineering
Ferdowsi University
Mashhad, Iran
Farid.fathnia@mail.um.ac.ir

Mohammad Hossein Javidi Dasht Bayaz
Professor
Department of Electrical Engineering
Ferdowsi University
Mashhad, Iran
H-javidi@um.ac.ir

Abstract—**The vital and national infrastructure of a country includes a wide range of cyber-physical systems. In order to exploit critical services, it requires highly sophisticated controls and focused maintenance. Any kind of malfunctioning of these systems will cause widespread destruction in a country. In this paper, we try to cover the concept of protecting the confidentiality of customers' data in an intelligent network with the approach of using anomaly detection algorithm. For this purpose, OPTICS density-based method has been used. In order to increase the efficiency of this approach, the LOF index has been used alongside it. Then, by implementing the scattering criterion for the LOF indexes, we identify the malicious data. In order to validate the proposed method, we will help with the actual data of the London Smart Meter in January 2013. The results will be displayed in different scenarios.**

Keywords-Smart Grid; Anomaly Detection; OPTICS; LOF; Standard Deviation; Coefficient of Variation; Cyber Attack

I. INTRODUCTION

The concept of smart grid was introduced by Advanced Measurement Infrastructure (AMI) theory in order to improve the consumption management, energy efficiency and self-resilience of the electricity grid in order to improve reliability and respond appropriately to natural disasters [1]. Intelligent network is a fully automated network that has the ability to control and advanced operation of all network elements, in order to function safely and efficiently in production, transmission and distribution [2].

Today, the power industry faces huge challenges around the world, such as the diversification of power generation, the development and optimal expansion of expensive capital, demand response, energy storage and reduced production of carbon dioxide gas. It is clear that such critical issues can not be resolved within the existing power grid. The future grid, known as the smart grid, is expected to bridge a large part of the shortcomings and weaknesses of the current grid. In order to achieve comprehensive control and monitoring, the intelligent network is the result of convergence of information technology and communication with the power system. In order to successfully implement an intelligent network, we have to pay attention to a few things: the privacy and consumer choice must be respected, the consumer is the decision maker, the communication standards and intelligent network communication protocols must have free, flexible, secure and limited [3]. Here are some of the smart grid requirements, in order to take advantage of all of its extraordinary benefits. 1. Supervisory Control and Data Acquisition (SCADA) 2. Remote Terminal Units (RTUs) 3. Advanced Measurement Infrastructure (AMI) 4. Energy Management System (EMS) [4]. The SCADA system sends information to the energy management system measured by smart meters and a set of RTUs that are located at the strategic and important points of the smart grid. Then, the energy management system determines the necessary measures to provide the best conditions for network operation by means of state estimation and load and production forecasting algorithms.

So, as mentioned, maintaining the security of customer information and their personal data is one of the most important pillars of the emerging smart grid. For this reason, the purpose of this article is to provide a solution to find out the destructive effects that cyber attackers have on data transmitted over the smart grid intelligence infrastructure between customers' homes by smart meters and an energy management system. Security attackers are supposed to be attacking Modification mode due to a malfunction in security systems. Here, with the OPTICS approach to data density, we want to identify these malicious behaviors to prevent its bad effects. This method was introduced in 1999 in [5]. Then, by the LOF (Local Outlier Factor) index [6] and based on the dispersion index between the data, we are trying to improve the OPTICS method to identify abnormalities. The two previous articles by the authors of this paper [7-8], respectively, are about improving the performance of the OPTICS method by the LOF index and implement it in the Demand Dispatch Scheduling. Here we will have another innovation in this area by applying this scatter factor. For validation, the proposed method uses actual London data that relates to power consumption data transmitted through its smart meters. This information is accessible from reference [9].

A lot of research has been done in the field of detecting anomalies in smart grid that require telecommunication and computer infrastructure. For example, in [10], the authors proposed a network based Moving Target Defense Intrusion

978-1-5386-4916-9/18/$31.00 ©2018 IEEE

Detection System (MTDIDS) to monitor node traffic in smart grid IPv6 based AMI networks. MTDIDS analyzes incoming packets against dynamic signature arrays for anomaly detection. In [11], a new approach to create a tunable profile-based FDS was presented. This work is an important innovation by showing that it is possible to use only a small set of recent measures to define a consumption pattern. As we know Fraud Detection Systems (FDS) can be classified into two large categories: sensor-based and profile based. The FDS can be tuned, using an optimization procedure, by imposing constraints on the true or false alarm rates, or maximizing an objective function that represents the revenue of the utility. In [12] two models of auto-regressive moving average (ARMA) and Generalized Likelihood Ratio (GLR) have been used for this topic. By using the ARMA model, the probability distribution is shown by the customer's normal mode of consumption, and by the GLR index, the average amount of stealing power by the smart meters is estimated. In [13], a real-time anomaly detection framework, was developed which can be built based upon smart meter data collected at the consumers' premises. The model is designed to detect the occurrence of anomalous events and abnormal conditions at both lateral and customer levels. They propose a generative model for anomaly detection that takes into account the hierarchical structure of the network and the data collected from smart meters and also address three challenges existing in smart grid analytics: (i) large-scale multivariate count measurements, (ii) missing points, and (iii) variable selection. They present the effectiveness of their approach with numerical experiments. At last a review of the methods for analyzing consumption and intelligent data management is presented in [14].

As discussed in the articles, Nowadays, anomaly detection has become one of the issues in smart grids, which has focused more attention on itself. An approach that has not been highlighted in the above researches is the use of density and the amount of dispersion between the given data. Accordingly, this paper tries to make these concepts reach the goal. For the density between the data and improving the efficiency, as mentioned earlier, the OPTICS algorithm and the LOF index are used. Also, the statistical index used to model the error detection according to the scattering criterion between the data is similar to the Fano factor, with the difference that instead of using variance, the standard deviation is used which is further explained in the later sections.

The remaining of this paper is as follows. In Section II we will have a brief overview of the implications of the OPTICS algorithm, the LOF index, the desired dispersion criterion and the indicators by which we evaluate the validity of the proposed method. In Section III we will model abnormalities in smart meter data. In Section IV, the results of the simulations and their analysis will be presented, and finally, in Section V, we will summarize and follow up the work.

II. PROBLEM SOLVING METHOD

To discover clusters with arbitrary shape, density-based clustering methods have been developed. These typically regard clusters as dense regions of objects in the data space that are separated by regions of low density (representing noise). So here is a brief introduction to the OPTICS method, which is our approach in the data density space, and is not used in similar articles, and then we will mention the LOF index to improve the results and the scatter criterion to complete the proposed algorithm and its evaluation methods.

A. OPTICS ALGORITHM

By studying other density-based methods (DBSCAN and DENCLUE), one can find that in these algorithms, for all points, an imaginary radius is considered, and the number of points around that supposed radius (ε, for example) is determined. Then the user must define the minimum number of points (Minpts) to start the algorithm. The density of the distribution of data around these points is high. But more precisely in these methods, it turns out that for a constant value of the minimum points, higher-density clusters (that is, the smaller ε) are entirely within the lower-density clusters (more ε). So the selection of objects should be the one that must be considered first for the element that needs the lowest ε for cluster membership. OPTICS is a method that specifies this order, and for this purpose, it is necessary to calculate the two variables: core-distance and the reachability-distance [5].

- The core-distance of an object p is the smallest ε'value that makes {p} a core object. If p is not a core object, the core-distance of p is undefined.
- The reachability-distance of an object q with respect to another object p is the greater value of the core-distance of p and the Euclidean distance between p and q. If p is not a core object, the reachability-distance between p and q is undefined.

Figure 1 illustrates the concepts of core-distance and reachability-distance. Suppose that ε = 6 mm and Minpts = 5. The core-distance of p is the distance ε', between p and the fourth closet data object. The reachability-distance of q_1 with respect to p is the core-distance of p (i.e., ε'=3 mm), because this is greater than Euclidean distance from p to q_1. The reachability-distance of q_2 with respect to p is the Euclidean distance from p to q_2 because this is greater than the core-distance of p [6].

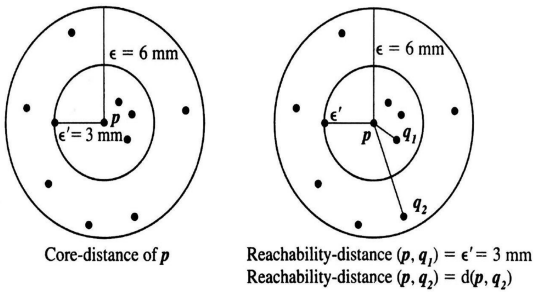

FIGURE 1 OPTICS TERMINOLOGY

The cluster ordering of a data set can be represented graphically, which helps in its understanding. For example, Figure 2 is the reachability plot for a simple two-dimensional data set, which presents a general overview of how the data are structured and clustered. The data objects are plotted in cluster order (horizontal axis) together with their respective reachability-distance (vertical axis). The three Gaussian bumps in the plot reflect three clusters in a data set [6].

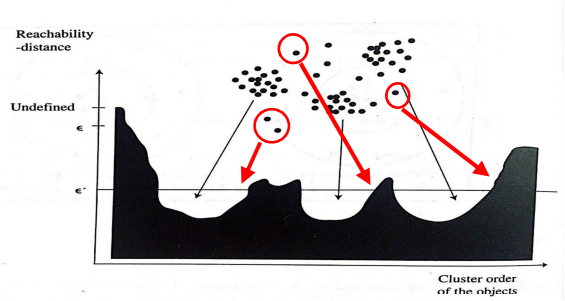

FIGURE 2 CLUSTER ORDERING IN OPTICS [5]

In Figure 2, also red arrows indicate that the points considered as noises have a higher reachability-distance because they are placed at a distance away from the points within a cluster with respect to Euclidean distance. The method of ordering points in the OPTICS algorithm is to transfer these points to the end of the cluster structure. For this reason, this method shows its performance against malicious data. Also, due to the presentation of the clustering structure, it is possible to understand the malicious cluster. The cluster is a malicious cluster, if all of which are malformed. Because finding the exact number of bad data in this method is difficult and there is no unique index, the LOF and data dispersion criteria help to identify the malicious data as well.

B. LOF INDEX

To define the local outlier factor of an object, we need to introduce the concepts of k-distance, k-distance neighborhood and local reachability density. The reachability-distance index is also required, which is explained in the previous section, which suggests a close relationship between these two concepts.

- The k-distance of an object p is the maximal distance that p gets from its k-nearest neighbors. This distance is denoted as k -distance (p). It is defined as the distance $d(p,o)$ between p and an object $o \in D$, such that for at least k objects, $\acute{o} \in D$ it holds that $d(p,\acute{o} \leq d(p,o))$.
- The k -distance neighborhood of an object p is denoted $N_k(p)$. By setting k to $MinPts$, we get $N_{MinPts}(p)$. It contains the $MinPts$ -nearest neighbors of p. That is, it contains every object whose distance is not greater than the $MinPts$-distance of p.
- The local reachability density (lrd) of p is the inverse of the average reachability density based on the $MinPts$-nearest neighbors of p.

LOF is the average of the ratio of the local reachability density of p and those of p's Minpts-nearest neighbors. It is easy to understand that the lower p's local reachability density is, and the higher the local reachability density of p's Minpts-nearest neighbors are, the higher LOF (p) is. The local reachability density and LOF equations are given below.

$$lrd_{MinPts}(p) = \frac{|N_{MinPts}(p)|}{\sum_{o \in N_{MinPts}(p)} reachability - distance_{MinPts}(p,o)} \quad (1)$$

$$LOF_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts}(p)} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{|N_{MinPts}(p)|} \quad (2)$$

C. COEFFICIENT OF VARIATION

In the theory of probability and statistics, the coefficient of variation (CV) is a normal criterion used to measure the distribution of statistical data. Which is obtained by dividing the standard deviation by the mean. In other words, the coefficient of variation expresses the amount of dispersion per unit of the mean. This value is defined when the mean is not zero and it has no dimension that makes it suitable for comparing statistical data of different units.

$$C_v = \frac{\sigma}{\mu} \quad (3)$$

In this paper, we will use this CV property to arrive at a better result, for which the value for two numbers always converges to a constant number if one of the numbers tends to infinity. Here's a proof of it.

Assume x_1 and x_2 will be our two numbers, then we have:

$$\bar{x} = \frac{x_1 + x_2}{2} \quad (4)$$

$$\begin{aligned} \sigma^2 &= (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 \\ &= x_1^2 + x_2^2 + 2\bar{x}^2 - 2\bar{x}(x_1 + x_2) \\ &= x_1^2 + x_2^2 + 2\bar{x}^2 - 4\bar{x}^2 \\ &= x_1^2 + x_2^2 - 2\bar{x}^2 \\ &= x_1^2 + x_2^2 - 2\left(\frac{x_1 + x_2}{2}\right)^2 \\ &= x_1^2 + x_2^2 - \frac{2}{4}(x_1^2 + x_2^2 + 2x_1x_2) \\ &= \frac{x_1^2}{2} + \frac{x_2^2}{2} - x_1x_2 \\ &= \frac{(x_1 - x_2)^2}{2} \end{aligned} \quad (5)$$

$$CV = \frac{\sigma}{\bar{x}} = \frac{|x_1 - x_2|/\sqrt{2}}{(x_1 + x_2)/2} = \frac{2}{\sqrt{2}} \frac{|x_1 - x_2|}{x_1 + x_2} \quad (6)$$

$$if\ x_1 > x_2 \xrightarrow{yields} \lim_{x_1 \to \infty} CV = \sqrt{2} \cong 1.4142 \quad (7)$$

So, as you can see, the coefficient of variation has such a feature that helps us find anomalies. It should be noted that in calculating the variance, we assumed the degree of freedom equal to one, so in the denominator, the number 1 is put. The reason is that if we have two observations, we have two independent observations for the mean of estimation, but there is only one independent observation to estimate the variance. Because both observations have a same distance from the mean.

In Figure 3, the proof of the hypothesis is also shown by MATLAB software.

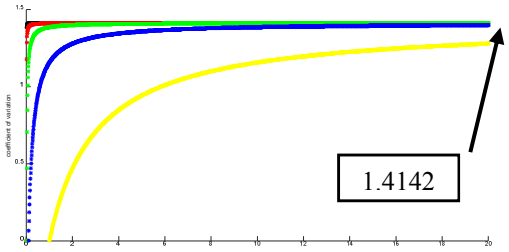

FIGURE 3 CV CURVES BY DIFFERENT SCALES

As shown in Figure 3, according to the first-number scale, various curves are obtained that all of them become radical 2 when the second number is infinitely. In this Figure, the yellow curve is for a non-decimal scale, a blue curve is for a tenth scale, a green curve is for a hundredth scale, a red curve is for a scale of one thousandth, and a black curve is for a scale of one ten thousand.

D. SOLVING PROCEDURE

Below, in Figure 4, the flowchart for the proposed method of finding anomalies is depicted. As can be seen, the reachability-distance (RD) and core-distance (CD) vectors are initially created by the OPTICS algorithm. Then, using the relationships of finding the local error equations, the LOF indexes are attributed to the dataset points in question. In the following, with care in the OPTICS algorithm, we conclude that in the reachability plot; if the input parameter (Minpts) is properly selected, the RD value is always lower than the CD value except in the case of the downside of the reachability plot. So the larger RD value from the CD value, is the greater the density of the points because their RD index is less. Additionally, we know that the noise and error at the end of this curve reveals itself, and the next important thing is that the reachability plot in its end section is ascending, so definitely one of the points which its RD value is larger from its CD value, become a non-anomaly point, and the smaller the RD index, it indicates its centrality in the desired cluster. Therefore, we select the least parameter among

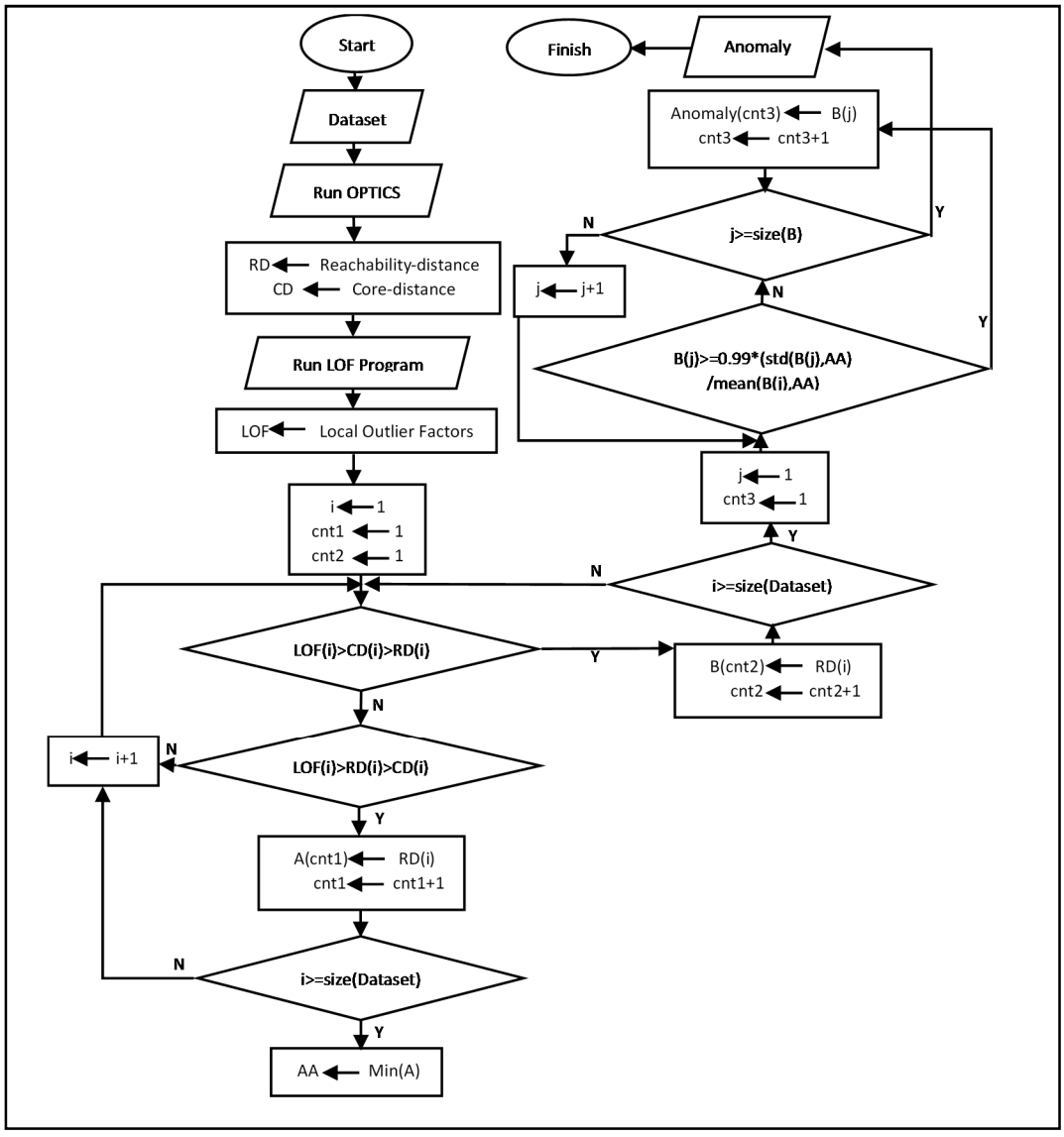

FIGURE 4 FLOWCHART OF PROPOSED METHOD

the points that have the same property and we will find anomalies in comparison with this point. According to the proven property of the coefficient of variation, all the data whose CD value is bigger than RD value, will be compared to the selected point with the coefficient of variation. Now, if this value is greater than 99% of radical 2 (Equation 7), we will consider it as an anomaly. In fact, we have set our own security margin or confidence interval by applying a 1% to the final limit value when one positive number reach infinity of CV value.

III. ABNORMALITIES MODELING

In order to test the proposed method and analyze the results, it is necessary to model the information errors in the computer network of smart grid. Therefore, the desired micro grid is shown in Figure 5. As you can see, customers' homes are equipped with smart meters, and sent and received information will be redirected to the controller's communications infrastructure, between the data center and smart meters in the micro grid. Here, it is assumed that the cyber-attacks somehow disturbs the information sent from the side of the smart meters. Here, the attackers carry out a Man In The Middle (MITM) attack to achieve their goals. This information includes their energy consumption and associated load factor. It is assumed that these anomalies are modeled by security invaders with normal distribution functions. In Figure 5, the effect of cybercrime is also depicted. Required information can be obtained from the reference [9].

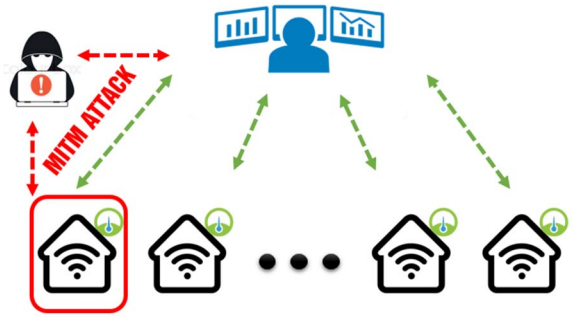

FIGURE 5 INFORMATION INFRASTRUCTURE WITH CYBER ATTACK

The power consumption data sent to the control center, at the end of each month, is collected in order to provide customers' electricity bill. Any change in this information, therefore, means that it is harmful both on the consumer's side and on the side of grid operator due to the less social welfare factor. For this reason, maintaining and protecting this information and securing the smart grid network is a very significant and important issue.

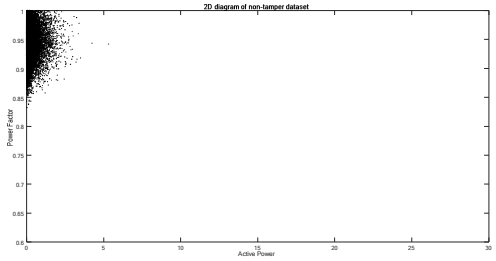

FIGURE 6 TWO DIMENSION OF NON-ANOMALY DATASET

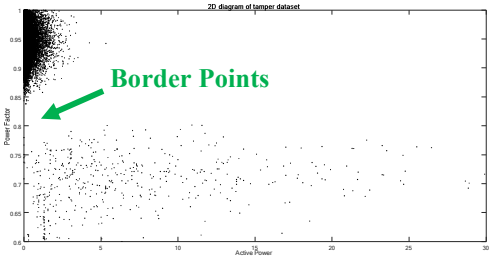

FIGURE 7 TWO DIMENSION OF ANOMALY DATASET

As shown in Figures 6 and 7, cyber attackers with help of MITM attack, carry on the control of the information sent to control center, and change them to implement their goals. Instead of reviewing consumer behavior as a time profile, all the information is plotted at the end of the month in the form of a multi-dimensional graph, and then, according to the algorithm presented, we identify the anomalies. Here, our issue has 2 dimensions. Power consumption and power factor. In fact, one of the advantages of the proposed method is to find abnormalities in high-dimensional information. Because in today's power grids, due to the advancement of technology, customer information is composed of various dimensions and should all of them be covered in analysis. The efficiency of the method is determined when it be able to identify points that are close to the conventional points. These points are named in Figure 7, as border points.

IV. SIMULATION AND RESULTS

In this section, we will simulate the proposed method on extracted data from the reference 9. Initially, by applying anomalies in the smart meter data of a particular customer, in the amount of energy consumed by it, according to the aforementioned modeling, we try to test the application of the LOF index.

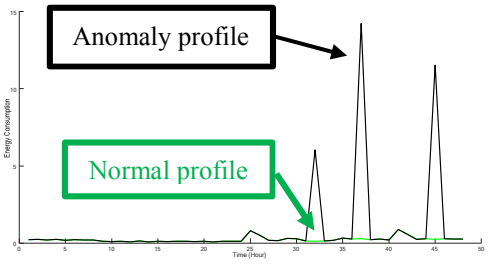

FIGURE 8 ENERGY CONSUMPTION PATTERN IN TWO MODE

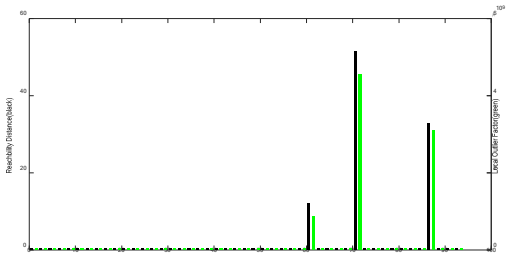

FIGURE 9 PERFORMANCE OF USING LOF INDEX

As shown in Figures 8 and 9, the LOF Index was able to detect the anomalies in the consumer smart meter data and take a large amount. Figure 8 shows these abnormalities in black and normal pattern in green color. In Fig. 9, the RD values for each data are also shown, these values also have large magnitudes at these points for the reasons stated. But 2 things are important here. 1. The use of this index for high productivity depends on the data that is far from the standard state and may not work well for border data (Fig. 7), and the exact boundary between anomalies and non-anomalies can not be detected. 2. Using the LOF and reachability-distance indicators alone is not possible, since the OPTICS algorithm for creating a cluster gives the start point a higher RD value, but that is not an anomaly point [8].To this end, we try to use the proposed method to improve the accuracy of performance. At first, we use six measures for comparing the performance [15]:

1) TPR or Sensitivity is the percentage of anomaly instances correctly detected.

2) FPR is the percentage of normal instances incorrectly classified as anomaly.

3) "Presicion" is the percentage if correctly detected anomaly instances over all the detected anomaly instances.

4) "Accuracy" is the percentage of all normal and anomaly instances that are correctly classified.

5) The "F-measure" is the equally-weighted (harmonic) mean of precision and sensitivity.

6) "Specificity" is the value of 1-FPR.

These measures determine how the proposed method perform in identifying anomaly instances. In Table 1, various scenarios for assessing the performance of the proposed method are shown.

TABLE 1 DATASET AND SCENARIOS

Scenario	Number of Data	Dimensions	Testing Instances	
			Normal	*Anomaly*
1	48	2	45	3
2	570	2	550	20
3	980	2	945	35
4	1489	2	1449	50

A. *Scenario 1*

TABLE 2 PROPOSED METHOD'S PERFORMANCE IN SCENARIO 1

FPR	Precision	Sensitivity	Specificity	Accuracy	Fmeasure
0	1	1	1	1	1

B. *Scenario 2*

TABLE 3 PROPOSED METHOD'S PERFORMANCE IN SCENARIO 2

FPR	Precision	Sensitivity	Specificity	Accuracy	Fmeasure
0.01	0.666	1	0.981	0.982	0.8

C. *Scenario 3*

TABLE 4 PROPOSED METHOD'S PERFORMANCE IN SCENARIO 3

FPR	Precision	Sensitivity	Specificity	Accuracy	Fmeasure
0.01	0.714	1	0.985	0.985	0.833

D. *Scenario 4*

TABLE 5 PROPOSED METHOD'S PERFORMANCE IN SCENARIO 4

FPR	Precision	Sensitivity	Specificity	Accuracy	Fmeasure
0.02	0.574	1	0.974	0.975	0.729

Figure 10 is a summary of tables 2 to 5. As can be seen from these tables and Figure 10, in scenario 1, where the number of data and the number of anomalies is low, the best possible results were obtained for an anomaly detection algorithm. In which the accuracy and amount of F-measure are as high as possible. In fact, the amount of F-measure is a measure of a test's accuracy. It considers both the precision and the sensitivity of the test to compute the score, where it reaches its best value at 1 (perfect precision and sensitivity) and worst at 0. So in this scenario we will have the best performance. As the number of anomalies, as well as the number of problem's data, increases, or in other words, we become closer to reality, the precision decreases in each scenario, which means that some non-anomaly data are known as anomalies .This amount of error in calculations is

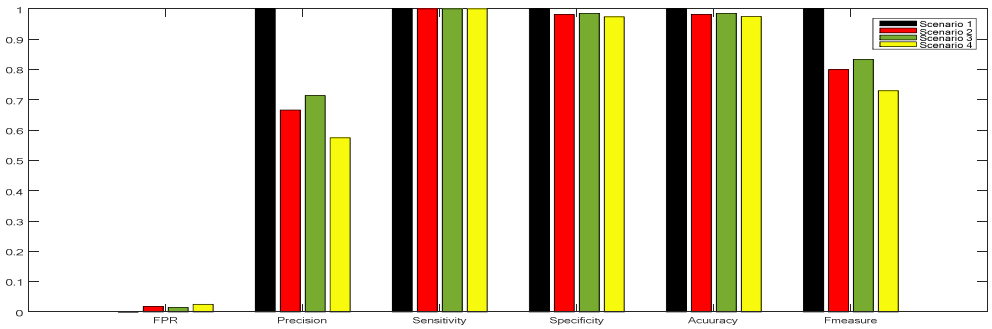

FIGURE 10 COMPARISON OF THE PROPOSED METHOD'S PERFORMANCE IN DIFFERENT SCENARIOS

integral to the current systems. But with a closer look at the results, we find that the accuracy of the algorithm is about 98% in four different scenarios, which is a very significant and widely accepted number of similar articles and highlights the advantage of the proposed method. The sensitivity index is in all cases equal to one, which indicates that the proposed algorithm detects all anomalies in all cases, and its error is just to maliciously know non-anomalous data, which despite 98% accuracy can be ignored. Because the precision index of a higher number of data is less than one, the F-measure also fades away from its ideal state, but maintains its ideal distance to the ideal state. This value is an average of 0.8, which results from the effectiveness of the proposed method. The FPR index is also small in all conditions and in scenario 1 it reaches zero, which, according to its concept, indicates the low error of the proposed method. Similarly, the specificity index, by its definition, adopts high values, and, in general, reports the performance of the proposed method very well.

V. CONCLUSION

In this paper, we tried to address the problem of detecting anomalies in the advanced intelligent network infrastructure by improving the OPTICS algorithm, which is a dense-axis approach to data mining studies. Because maintaining the security of data in the smart grid and protecting the privacy of customers is one of the most important issues in the operation of an intelligent network. It is worth noting the advantages of using the OPTICS algorithm, is to analyze the data in the multidimensional space. Initially, the diagnosis of abnormal data was done by the LOF index, due to the close relationship between the variables and the concepts of the two approaches mentioned. Then, with regard to the fact that all points are not clearly distinguishable from conventional and non-anomalous points, which also makes detection of anomalies more difficult, it was attempted to use the statistical concept of the coefficient of variation. With this concept and its used feature which is proved in the text of the paper, the relationship between the parameters of the OPTICS method and the LOF index was obtained and through it, we began to identify abnormalities. In order to test, real data from the smart meters in London was used, and we showed that the accuracy of the proposed method is high and it has just a slight error in various scenarios.

In order to continue this path, using more accurate error models, simulating in more scenarios and with big data, using other statistical models, and finally, testing methods in different values of input parameters, can improve the approach adopted.

REFERENCES

[1] F. Rahimi and A. Ipakchi, "Demand Response as a Market Resource Under the Smart Grid Paradigm," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 82-88, 2010.

[2] "Smart grids european technology platform smart grids," [Online]. Available: http://www.smartgrids.eu

[3] T. J. Lui, W. Stirling and H. O. Marcy, "Get Smart," *IEEE Power & Energy Magazine*, May/June, 2010.

[4] J. Aghaei and Mohammad-ImanAlizadeh, "Demand response in smart electricity grids equipped with renewable energy sources:A review," *Renewable and Sustainable Energy Reviews*, pp. 64-72, 2012.

[5] M. Ankerst, M. Breunig, H. Peter Kriegel, and J. Sander, "OPTICS: Ordering Points To Identify the Clustering Structure," *In Proceedings of ACM SIGMOD'99,* Institude of Cimputer Science, University of Munich, Germany, 1999.

[6] J. Han and M. Kamber, Data Mining: Concepts and Techniques, *ELSEVIER Inc*, 2nd Edition, 2006.

[7] F. Fathnia, F. Fathnia, and M. H. Javidi Dasht Bayaz, "Detection of Anomalies in Smart Meter Data: A Density-Based Approach," *Smart Grid Conference (SGC)*, December 2017.

[8] F. Fathnia, F. Daburi Farimani, F. Fathnia, and M. H. Javidi Dasht Bayaz, "The Effect of Cyber Attacks on the Demand Dispatch Application and Identify Them by OPTICS," *4TH International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, December 2017.

[9] https://data.london.gov.uk/dataset/

[10] B. Chatfield, R. J. Haddad, "Moving Target Defence Intrusion Detection System for IPv6 Based Smart Grid Advanced Metering Infrastructure," *SoutheastCon*, Charlotte, March 2017.

[11] M. Zanetti, E. Jamhour, M. Pellenz, M. Penna, V. Zambenetti, and I. Chueiri, "A Tunable Fraud Detection System for Advanced Metering Insfrastructure using Short-Lived Patterns," *IEEE Transactions on Smart Grid*, issue. 99, September 2017.

[12] Daisuke Mashima and Alvaro A. Cárdenas. *Research in Attacks, Intrusions, and Defenses: 15th International Symposium, RAID 2012, Amsterdam, The Netherlands, September 12-14, 2012. Proceedings*, chapter Evaluating Electricity Theft Detectors in Smart Grid Networks, pages 210–229. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[13] R. Moghaddass and J. Wang, "A Hierarchial Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data," *IEEE Transactions on Smart Grid*, April 2017.

[14] V. Ford, A. Siraj, and W. Eberle, "Smart grid energy fraud detection using artificial neural networks," *In IEEE Symposium on Computional Intelligence Applications in Smart Grid*, CIASG, January 2015.

[15] S. R. Gaddam, V. V. Phoha, and K. S. Balagani, "K-Means+ID3: A Novel Method for Supervised Anomaly Detection by Cascading K-Means Clustering and ID3 Decision Tree Learning Methods," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 3, March 2007.