Anomaly Detection Algorithms for Smart Metering using Swarm Intelligence

Pradeep Subhash Paikrao and Ranjan Bose
Department of Electrical Engineering
Indian Institute of Technology Delhi, India

ABSTRACT

Advancement in the information and communication technology has introduced Advanced Metering Infrastructure (AMI) in the electricity metering system, which has replaced old mechanical meters with smart electric meters. This modernization also introduced a lot of scope for the different anomalies and attacks on smart meters. Hence to tackle these challenges, we have proposed three anomaly detection algorithms (VBA, HBA, KBA) which are truly based on the principles of Swarm Intelligence (SI). The swarm intelligence is the emerging subbranch of artificial intelligence which studies the collective intelligence of groups of simple agents. The theory is corroborated by its performance in terms of probability of detection and probability of false alarm. The proposed algorithms entrust the probability of detection and probability of false alarm close to 1.00 and 0.17 respectively.

KEYWORDS

Advanced Metering Infrastructure (AMI); smart meter; anomaly detection; Swarm Intelligence (SI)

ACM Reference Format:
Pradeep Subhash Paikrao and Ranjan Bose. 2018. Anomaly Detection Algorithms for Smart Metering using Swarm Intelligence. In *FICN '18: 1st Int'l Workshop on Future Industrial Communication Networks, Nov. 2, 2018, New Delhi, India.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3243318.3243319

1 INTRODUCTION

In today's world electricity is the most valuable thing for the betterment of human life. Its generation and optimum use is a noteworthy concern and for this lot of efforts have been put. Still, instability of a grid, theft of energy, different types of technical and non-technical losses have made present power grid unreliable. Hence conventional power grid is not reasonable for the present time [1]. With a considerable measure of innovations and growth of technology has driven us from traditional power grid to smart grid that includes Advanced Metering Infrastructure (AMI) which is featured with two-way communication, high reliability, real-time demand response, self-healing and security [2].

These meters in AMI are called as smart meters because, in addition to record customer's energy consumption, smart meters can have a duplex communication path between utility centers and customer [3]. But these smart meters are more vulnerable to different types of attack which is in favor of attackers. Numerous techniques as of now have been developed to detect such attacks [2],[4]. Our objective is to design anomaly detection techniques that can identify anomalous smart meters in the AMI network using swarm intelligence.

The rest of this paper is organized as follows. Section 2 introduces swarm intelligence. Section 3 presents the system model. Section 4 details proposed anomaly detection algorithms and its simulation results are shown in section 5. Finally, Section 6 presents the conclusion of this paper.

2 SWARM INTELLIGENCE

The concept of swarm intelligence is taken from social conduct that can be noticed in nature, such as ant colonies, flock of birds, fish schools and bee hives, where a number of individuals with limited capabilities are able to come to intelligent solutions for complex problems [5]. This swarming practices of gatherings of creatures empowers them to solve problems that are beyond the capabilities of single individuals. Swarm Intelligence (SI) is artificial intelligence based on the collective behavior of decentralized, self-organized systems. SI has numerous advantages over the centralized system as follows [6]:

- Self-organization: swarm does not have any leader.
- Reduced workload: complete work is divided among the agents in the swarm so the workload is reduced.
- Long life: due to decentralized nature if any agent is failed then also work in the swarm will not stop.
- Low cost: agents in the swarm are dumped or individually cannot take the decision, requires less hardware.

FICN'18, November 2, 2018, New Delhi, India

ACM ISBN 978-1-4503-5929-0/18/11…$15.00
https://doi.org/10.1145/3243318.3243319

The swarm intelligence algorithms include Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Glowworm Swarm Optimization (GSO), and Cuckoo Search Algorithm (CSA) [7],[8], [9]. One of the famous algorithms in swarm intelligence, Ant Colony Optimization (ACO) which is inspired by frogging behavior of ants [10]. In ACO, ants find the shortest path between their nest and food with the deposition of pheromone. The swarm accomplishes its objectives by means of the collaborations of the whole gathering. The organisms use simple local rules to administer their activities. They typically do not follow commands from a leader, or any global plan. With this swarm intelligence guarantees the capacity to oversee complex systems of interacting individuals through minimal communication with only local neighbors to deliver a global emergent action [11].

Till now, swarm intelligence algorithms have been broadly applied to robotics, DNA computing, vehicle routing problem, data-intensive Internet of Things (IoT) operations and services [12]. Hence we are encouraged to apply this idea in the anomaly detection area. In this paper, we have proposed three algorithms in the view of swarm principles to detect anomalous or compromised smart meters in the network.

3 SYSTEM MODEL

Formation of swarm should be completely random and the number of smart meters in each swarm is a random choice and we have considered this in our implementation of algorithms. Swarm formation of smart meters in AMI is as shown in Fig.1. In this, blue nodes are representing normal smart meters and red nodes are anomalous smart meters.

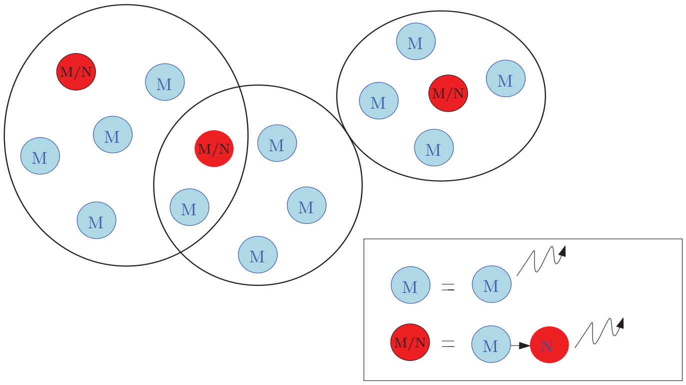

Figure 1: Formation of swarms in AMI.

$$M_i = \begin{cases} N \times X_i, & \text{if } N \neq 1 \cdots \textit{Anomalous} \\ X_i, & \text{otherwise} \cdots \textit{Normal} \end{cases} \tag{1}$$

In above equation, $N \times X_i$ is the i^{th} anomalous smart meter and with case $N = 1$, M_i is the normal smart meter. N is the attack vector and throughout this paper considered attacks are constant, uniform and gaussian distributed.

4 ANOMALY DETECTION ALGORITHMS

In this section, we introduce three SI based anomaly detection algorithms.

4.1 Vector Based Algorithm (VBA)

In this algorithm, all the smart meters in the randomly formed swarm are indexed from 1 to n. Then as shown in Fig.2, smart meter 1 securely shares the last T consumption readings $(M_1 + Y)$ with smart meter 2 in the swarm where Y is the random vector only known to smart meter 1. Similarly smart meter 2 shares $M_1 + M_2 + Y$ readings to smart meter 3 and so on. Finally, smart meter 1 subtracts Y from the summation of consumption readings and calculates the average consumption of the swarm. Then this average consumption is shared with every meter in the same way.

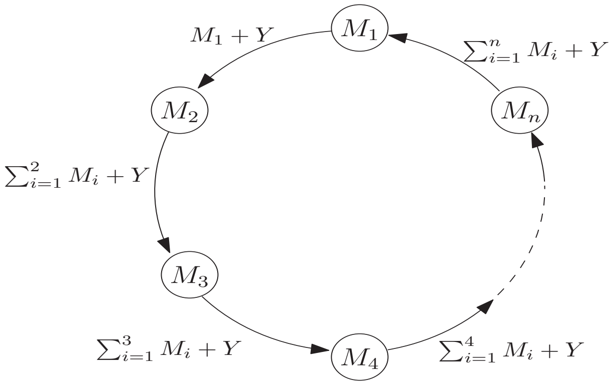

Figure 2: Sharing of consumption readings in swarm

Subsequently, every smart meter finds its probability distribution $(p(m_i))$ from its histogram and calculates mean (μ_{ij}) and entropy (H_{ij}) of these T readings as shown in (2) and (3) where i represents index of the smart meter in the swarm and j represents the swarm index.

$$\mu_{ij} = \frac{1}{T} \sum_{t=1}^{T} m_{it} \tag{2}$$

$$H_{ij} = -\sum p(m_i) log(p(m_i)) \tag{3}$$

Also, all meters calculate mean (μ_{cj}) and entropy (H_{cj}) of average consumption of the swarm as shown in (4) and (5) where c is used for center and N_m represents total number of smart meters in the swarm.

$$\mu_{cj} = \frac{1}{N_m T} \sum_{t=1}^{T} \sum_{i=1}^{N_m} m_{it} = \frac{1}{Nm} \sum_{i=1}^{N_m} \mu_{ij} \tag{4}$$

$$H_{cj} = -\sum p(m_{ci}) log(p(m_{ci})) \tag{5}$$

$$\delta_{ij} = \sqrt{(\mu_{ci} - \mu_{ij})^2 + (H_{cj} - H_{ij})^2} \tag{6}$$

In (6), δ_{ij} is calculated which represents the deviation of i^{th} smart meter from center point of the j^{th} swarm. Based on these, the smart meter who deviates more is flagged in the swarm depending on the threshold (see (7)).

$$\delta_{ij} \underset{F_{ij}=0}{\overset{F_{ij}=1}{\gtrless}} \max_{1\leq i\leq N_m} \delta_{ij} C_1 \tag{7}$$

Finally, after N_s different swarm realizations smart meters are categorized as anomalous smart meter who's flagging status is above the threshold. Flagging the meters in every swarm calculation is fundamentally the same as pheromone deposition seen in ACO [10].

$$z_i = \sum_{j=1}^{N_s} F_{ij} \underset{i \neq Anomalous}{\overset{i = Anomalous}{\gtrless}} \max_{1\leq i\leq N_m} \sum_{j=1}^{N_s} F_{ij} C_2 = z_{max} C_2 \tag{8}$$

In (8), Z_i defines the flagging status of i^{th} smart meter and C_2 is the threshold parameter. The algorithm is explained using flowchart in Fig.3.

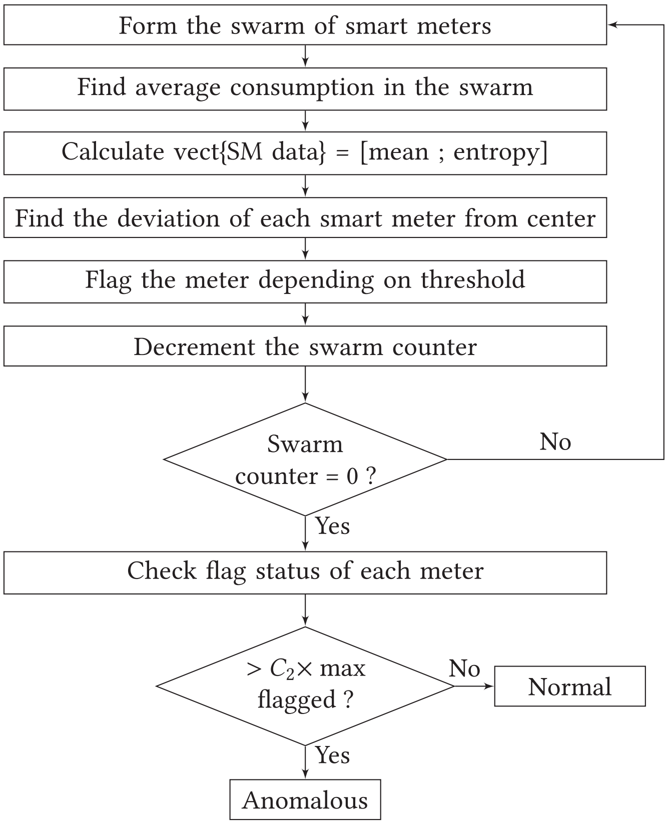
Form the swarm of smart meters
Find average consumption in the swarm
Calculate vect{SM data} = [mean ; entropy]
Find the deviation of each smart meter from center
Flag the meter depending on threshold
Decrement the swarm counter
Swarm counter = 0 ?
No
Yes
Check flag status of each meter
> C_2× max flagged ?
No
Normal
Yes
Anomalous

Figure 3: Flowchart for Vector Based Algorithm (VBA)

4.2 Honesty coefficient Based Algorithm (HBA)

This algorithm is modified from LUD (LU decomposition) algorithm presented by Sergio Salinas [13] to fulfill the properties of swarm intelligence in anomaly detection. According to swarm principles, the formation of swarm should be completely random and must be a decentralized process which is not possible in LUD algorithm as the collector in AMI acts as a leader. The collector's reading $(\overline{P_{t_i}})$ is used to solve (9).

$$k_1 p_{t_i,1} + k_2 p_{t_i,2} + \cdots + k_n p_{t_i,n} = \overline{P_{t_i}} \tag{9}$$

For a randomly formed swarm (see Fig.4, here randomly formed swarm is circled red), the collector's readings cannot be used directly. Also, the inclusion of the collector in every swarm to have actual total consumption values is not practical.

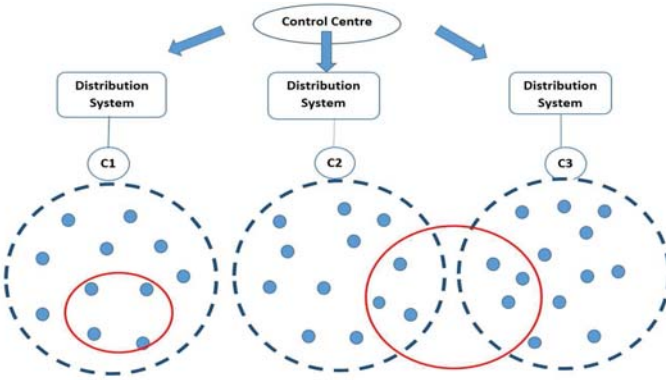

Figure 4: Randomly formed swarms in AMI

Hence, we have proposed the concept of the virtual collector in which the actual total consumption in each swarm is approximated as shown in (10).

$$\widetilde{P_{t_i}} = \sum_{l=1}^{n} \frac{\overline{P_{t_i l}}}{N_l} N_{ml} \tag{10}$$

where $\widetilde{P_{t_i}}$ is virtual collector reading at i^{th} time instant, N_l is the total number of smart meters belongs to l^{th} collector, N_{ml} is the number of smart meters belongs to a randomly formed swarm and l^{th} collector. So $\overline{P_{t_i}}$ is replaced with $\widetilde{P_{t_i}}$ for our algorithm. Hence (9) for time instant 1 to n becomes,

$$\begin{gathered} k_1 p_{t_1,1} + k_2 p_{t_1,2} + \cdots + k_n p_{t_1,n} = \widetilde{P_{t_1}} \\ \cdots \\ k_1 p_{t_n,1} + k_2 p_{t_n,2} + \cdots + k_n p_{t_n,n} = \widetilde{P_{t_n}} \end{gathered} \tag{11}$$

which can be represented in matrix form:

$$Pk = \widetilde{P} \tag{12}$$

The energy consumption matrix P is decomposed in the lower triangular matrix L and upper triangular matrix U, i.e., $P = LU$. After calculating the elements of L and U, following equations are solved in order to find out honesty coefficient of every smart meter.

$$Ly = \widetilde{P} \tag{13}$$

$$Uk = y \tag{14}$$

Rest of the algorithm is same as LUD algorithm to find honesty coefficient of every meter [13]. Depending upon the value of honesty coefficient, smart meters are flagged and based on the flagging status of a smart meter, the final anomaly decision is made. The algorithm is explained using flowchart in Fig.5.

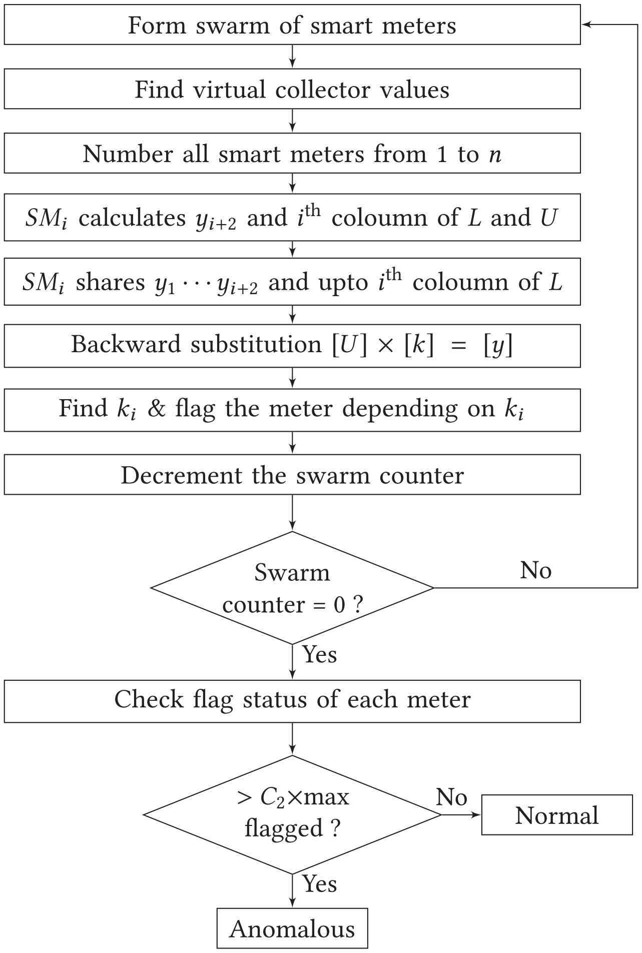
Form swarm of smart meters
Find virtual collector values
Number all smart meters from 1 to n
SM_i calculates y_{i+2} and i^{th} coloumn of L and U
SM_i shares $y_1 \cdots y_{i+2}$ and upto i^{th} coloumn of L
Backward substitution $[U] \times [k] = [y]$
Find k_i & flag the meter depending on k_i
Decrement the swarm counter
Swarm counter = 0 ?
No
Yes
Check flag status of each meter
> C_2×max flagged ?
No
Normal
Yes
Anomalous

Figure 5: Flowchart for Honesty coefficient Based Algorithm (HBA)

4.3 KLD Based Algorithm (KBA)

In this algorithm, every meter securely shares last T consumption readings with other meters in the swarm (refer Fig.2) and the average consumption readings of the swarm is calculated and shared in the same way. After that, every smart meter finds histogram of its own consumption readings and of average consumption readings. Then from the histogram, meter finds the probability distribution of its own and average consumption readings. Now, each meter finds Kullback–Leibler Distance (KLD) [14] between the distribution of its own consumption and the average consumption in the swarm as shown in (15).

$$D_{KL}(p(M_i) \parallel p(M_{avg})) = \sum p(M_i) log(\frac{p(M_i)}{p(M_{avg})}) \tag{15}$$

where D_{KL} is Kullback–Leibler distance, $p(M_i)$ is the probability of i^{th} smart meter and $p(M_{avg})$ is the probability of average consumption in the swarm.

Based on this KL distance meters are flagged and the final decision of anomaly is made depending upon the threshold. The algorithm is explained using flowchart in Fig.6.

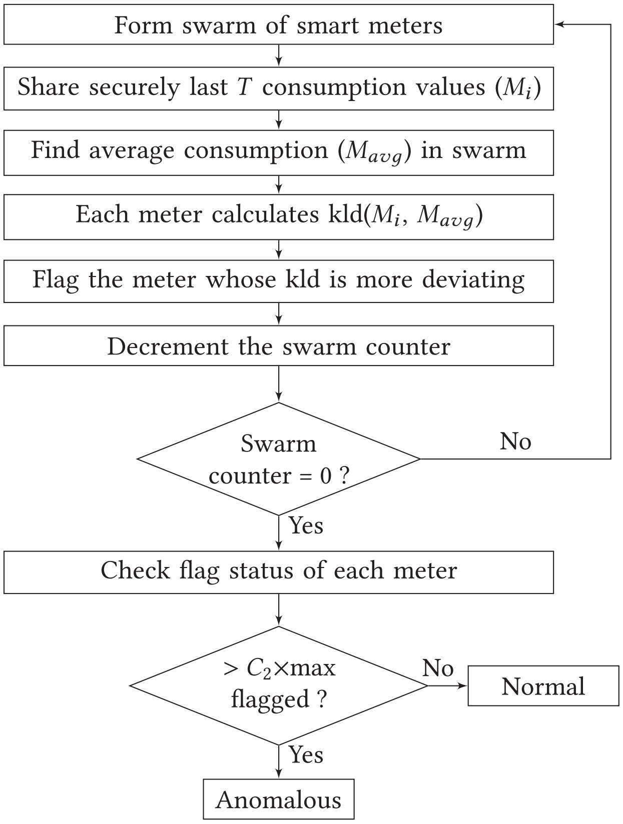
Form swarm of smart meters
Share securely last T consumption values (M_i)
Find average consumption (M_{avg}) in swarm
Each meter calculates kld(M_i, M_{avg})
Flag the meter whose kld is more deviating
Decrement the swarm counter
Swarm counter = 0 ?
No
Yes
Check flag status of each meter
> C_2×max flagged ?
No
Normal
Yes
Anomalous

Figure 6: Flowchart for KLD Based Algorithm (KBA)

5 RESULTS

For the performance analysis of these proposed algorithms, we have considered the dataset having consumption record of 25 smart meters from NYISO which is obtained from [15]. Proposed algorithms are tested for 500 simulation rounds on Matlab, under false data injection attack in which we have suppressed meters reading [16]. The types of attack we consider are fixed attack, uniform and Gaussian distributed attack with the single and multiple cases. The metric that we considered for the performance analysis is the probability of detection (P_D) and probability of false alarm (P_F) as shown in (16) and (17) respectively [17].

$$P_D = prob(Z_i \geqslant Z_{max}C_2|M_i = NX_i) \tag{16}$$

$$P_F = prob(Z_i \geqslant Z_{max}C_2|M_i = X_i) \tag{17}$$

From above equations, it is clear that the optimal value of threshold parameter C_2 is highly desirable in order to get a high probability of detection and low probability of false alarm. Fig.7 shows the plot of P_D and P_F for the different values of the threshold parameter (C_2) for VBA. It points out that by selecting C_2 to 0.6 the probability of detection will be lies around 0.93 and probability of false alarm is 0.22. A similar sort of behavior is additionally observed for HBA and KBA henceforth we have kept C_2 equivalent to 0.6.

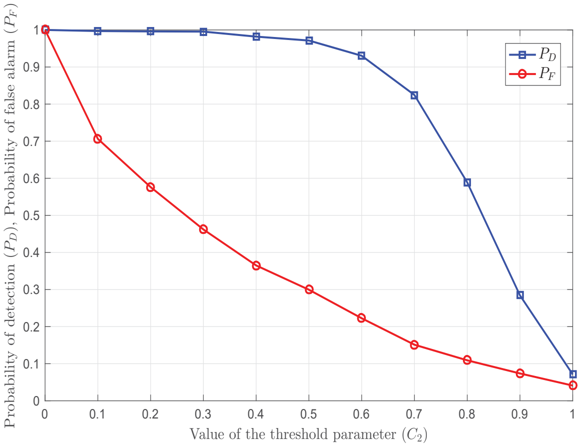

Figure 7: Plot of P_D and P_F vs threshold parameter C_2 (%) in VBA

For VBA, Fig.8 shows that after 400 different swarm realizations the probability of detection and false alarm converges to 1 and 0.20 respectively. This indicates that Vector Based Algorithm gives high performance after 400 swarm realizations.

For HBA, Fig.9 shows that after 200 randomly formed swarms the probability of detection and false alarm converges to 1 and 0.17 respectively. Hence HBA performs better after 200 swarm formations.

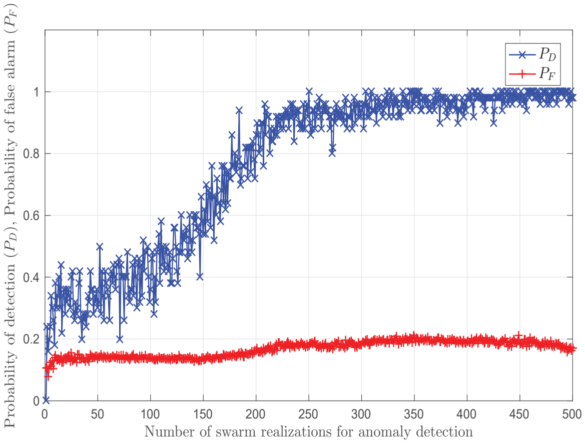

Figure 8: Plot of P_D and P_F vs Number of swarm realizations for VBA

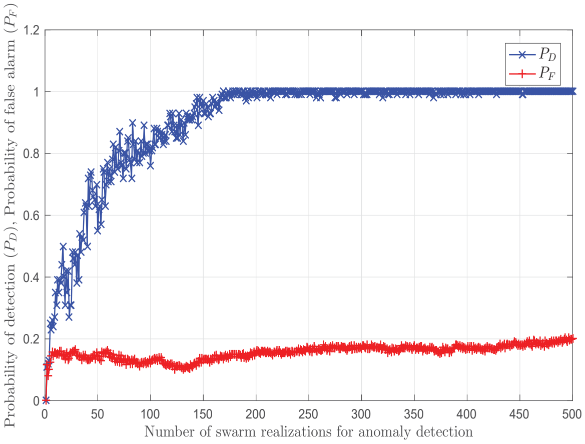

Figure 9: Plot of P_D and P_F vs Number of swarm realizations for HBA

For KBA, Fig.10 shows that after 100 different realizations of swarm the probability of detection and false alarm converges close to 1 and 0.30 respectively. Hence with this algorithm decision can be taken after 100 swarm realizations.

The above results show that as the number of iterations increases, the performance of all algorithms is improved and getting more closer to the solution ie. detection of anomalous smart meters. The list of parameters on which the performance of the designed anomaly detection algorithms depends is :

- Number of swarm formations required to take decision (around 100 to 400 requires from above results).
- Number of meters in each swarm.

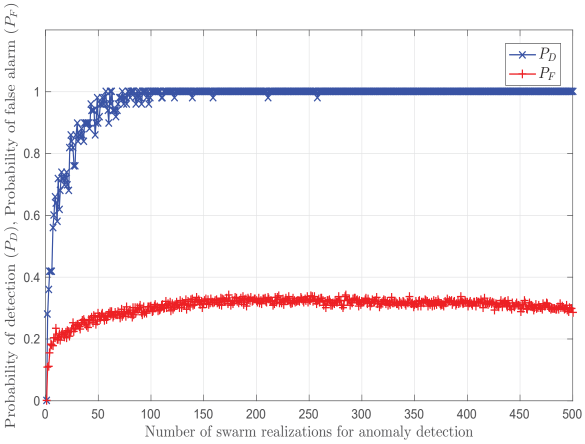

Figure 10: Plot of P_D and P_F vs Number of swarm realizations for KBA

- Level of attack.
- Number of meters attacked (Single meter and multiple meter attack).
- Point of threshold (value of C_2 in (8)).

The comparison of these three anomaly detection algorithms is given in Table1. From Table1, it is observed that probability of detection is close to 1.00 and the probability of false alarm lies between 0.17 to 0.30.

Table 1: State of the art

Algorithms	VBA	HBA	KBA
Number of swarm realizations	400	200	100
Probabilty of detection	0.95-1	1	1
probabilty of false alarm	0.20	0.17	0.30
Single & multiple meter attack	✓	✓	✓
Constant attack	✓	✓	✓
Uniform attack	✓	✓	✓
Gaussian attack	✓	✓	✓
Privacy preservation	✓	✓	✓
Used dataset	NYISO	NYISO	NYISO

6 CONCLUSION

In this paper, we have presented three anomaly detection algorithms to identify compromised meters in AMI with the concept of Swarm Intelligence. Based on the iterative process, in every randomly formed swarm suspicious meters are flagged and after the sufficient number of iterations decision of anomaly detection is made. For the performance analysis of proposed algorithms, we have used NYISO data set under various attack cases. Also, we observe that the probability of detection is close to 1.00 but the probability of false alarm is approximately 0.17. The thresholds C_1 and C_2 can be further optimized to reduce the probability of false alarm.

REFERENCES

[1] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen. EPPA: An Efficient and Privacy-Preserving Aggregation Scheme for Secure Smart Grid Communications. *IEEE Transactions on Parallel and Distributed Systems*, 23(9):1621–1631, Sept 2012.

[2] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. S. Shen. Energy-theft detection issues for advanced metering infrastructure in smart grid. *Tsinghua Science and Technology*, 19(2):105–120, April 2014.

[3] R. Berthier, W. H. Sanders, and H. Khurana. Intrusion Detection for Advanced Metering Infrastructures: Requirements and Architectural Directions. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 350–355, Oct 2010.

[4] D. M. Menon and N. Radhika. Anomaly detection in smart grid traffic data for home area network. In *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pages 1–4, March 2016.

[5] E.A. Mishra, M.N. Das, and T.C. Panda. Swarm intelligence optimization editorial survey. *International Journal of Emerging Technology and Advanced Engineering*, 3:217–230, 01 2013.

[6] Manju and Chander Kant. Ant Colony Optimization: A Swarm Intelligence based Technique. *International Journal of Computer Applications*, 73, issue 10:30–33, 07 2013.

[7] Vandana Jagtap. Survey of different swarm intelligence algorithms. *International Journal of Advance Engineering and Research Development* 1, 12 2014.

[8] A. Enache, V. SgÃćrciu, and M. Togan. Comparative Study on Feature Selection Methods Rooted in Swarm Intelligence for Intrusion Detection. In *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, pages 239–244, May 2017.

[9] Ab Wahab, Mohd Nadhir, Samia Nefti-Meziani, and Adham Atyabi. A Comprehensive Review of Swarm Optimization Algorithms. *PLOS ONE*, 10(5):1–36, 05 2015.

[10] M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):29–41, Feb 1996.

[11] Yan fei Zhu and Xiong min Tang. Overview of swarm intelligence. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, volume 9, pages V9–400–V9–403, Oct 2010.

[12] T. Chakraborty and S. K. Datta. Application of swarm intelligence in Internet of Things. In *2017 IEEE International Symposium on Consumer Electronics (ISCE)*, pages 67–68, Nov 2017.

[13] S. Salinas, M. Li, and P. Li. Privacy-Preserving Energy Theft Detection in Smart Grids: A P2P Computing Approach. *IEEE Journal on Selected Areas in Communications*, 31(9):257–267, September 2013.

[14] Ranjan Bose. *Information theory, coding and cryptography*. Tata McGraw-Hill Education, 2008.

[15] Load Data: Market and Operational Data (NYISO). [online] Available: http://www.nyiso.com/public/markets_operations/index.jsp.

[16] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi. Joint-Transformation-Based Detection of False Data Injection Attacks in Smart Grid. *IEEE Transactions on Industrial Informatics*, 14(1):89–97, Jan 2018.

[17] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory Part I*. John Wiley & Sons, Inc, 2002.