

A New Approach for Event Detection using k -means Clustering and Neural Networks

Muyiwa O. Oladimeji, Mikdam Turkey, Mohammad Ghavami (SMIEE), Sandra Dudley (MIEEE)

School of Engineering, London South Bank University, London, UK.

{oladimm2, turkeym, ghavamim, dudleymy}@lsbu.ac.uk

Abstract— In composite event detection systems such as fire alarms, the two foremost goals are speed and accuracy. One way to achieve these goals is by performing data aggregation at central nodes. This helps reduce energy consumption and redundancy. In this paper we present a new hybrid approach that involves the use of k -means algorithm with neural networks, an efficient supervised learning algorithm that extracts patterns and detects trends that are hidden in complex data. Previous research on event detection concentrates majorly on the use of feed forward neural network and other classifiers such as naive Bayes and decision tree alone for modern fire detection applications. In our approach presented here, we combine k -means with neural networks and other classifiers in order to improve the detection rate of event detection applications.

To demonstrate our approach, we perform data aggregation on normalized multi-dimensional fire datasets in order to remove redundant data. The aggregated data forms two clusters which represent the two class labels (actual outputs) with the aid of k -means clustering. The resulting data outputs are trained by the Feed Forward Neural Network, Naive Bayes, and Decision Trees. This approach was found to significantly improve fire detection performance.

Keywords—Data Aggregation, Clustering, Classification, Machine learning.

I. INTRODUCTION

There is a growing need for prompt and accurate fire detection in any extensive indoor environment, and this can only be achieved by employing appropriate wireless sensor network (WSN) deployment. WSN ad-hoc networks are made of tiny wireless nodes called sensor nodes that measure any event or exceptional change in environmental conditions. The sensors organise themselves into a single multi-hop or hierarchical network structures with several clusters and cluster heads. Each sensor node is capable of sensing, processing, and transmitting data to the base station [1].

A composite event is the combination of different observation of attributes. For an event fire alarm application to make a decision of normal or abnormal situation, there may be the need to combine several attributes based on large number of sensor types (temperature, carbon monoxide (CO), smoke) which are spatially distributed over a wide area [2]. Data

obtained from a composite event are multidimensional in nature. One of the key measures of enhancing accurate fire detection decisions is to perform data aggregation at intermediate nodes or at the clustered head. Data aggregation usually involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the base station (sink). Data aggregation helps to remove redundant and highly correlated data generated from neighboring sensors at the intermediate node before transmission to the base station [3]. Data aggregation techniques are also very effective in reducing communication overhead by collecting the most critical data from the sensors and making it available to the sink in an energy efficient manner with minimum data latency. Data latency is a crucial requirement in most event detection application such as fire detection applications.

The aggregated data at the clustered head are sent to the decision center (base station) for a final decision using an appropriate machine learning algorithm e.g. to detect the event status. This machine learning algorithm is divided into supervised and unsupervised learning approaches. The former approach relies on the presence of labeled data and a training phase whereas the latter approach does not require labeled data, training and prior knowledge of the event patterns [4,5]. The accuracy of supervised learning is often negatively affected when many attributes depend on one another. Learning large and complex models has increased difficulty with supervised learning than with unsupervised learning.

In section II of this paper, we briefly review contributions for fire detection mechanism using WSN. In section III we discuss on data aggregation in cluster-based WSN. In section IV, we propose our own approach and also discuss different techniques involved such as data aggregation, clustering and classification. In the final section, we perform an experiment using some test data to show the performance of our system approach.

II. LITERATURE REVIEW

According to the National Fire Danger Rating System (NFDRS) for forest fire detection, four sensor types (temperature, humidity, smoke and wind speed) were used to generate a fire-likelihood index [6]. The contribution of this study is the function of a feed-forward neural network (FFNN) and data aggregation for reducing overhead communication. A system approach was proposed in [7] for forest fire detection

using sensor nodes, gateway(s) and task manager(s). The sensor types used were temperature and humidity. Data obtained from different sensor nodes were fused together at the gateways. The data analysis and decisions are taken at the task manager.

Early fire detection in open spaces such as forests and urban areas was proposed using a sensor network approach [8]. For early fire detection, the authors used temperature sensor and maximum likelihood (ML) to fuse sensor data. Their system architecture is made up of the sensing, computing and localized alerting unit. According to previous works on fire detection using WSN, it is found that accurate and early fire detection can be approached from two perspectives [9]. Firstly, sensor data from several nodes of the same sensor type can be aggregated into one. Secondly, an artificial intelligence (AI) can be incorporated to recognize patterns in that data. Generally, the selection of sensors is based on random process or assumption. Researcher have discovered that the use of a single sensor type such as temperature sensor cannot guarantee accurate fire detection and so, there is need to employ multi-sensor type fire detectors that are capable of monitoring the environment against any changes in the amount of carbon monoxide (CO), carbon dioxide (CO₂) and oxidized gas. The use of multi-sensor will help to provide more accurate fire detection decision and discrimination between fire and noise [9, 10].

In our hybrid approach, we select four optimal sensor types, which are the temperature, ionization, photoelectric and CO sensors. The flaming fire and smoldering fire are detected by the ionization and photoelectric sensors respectively. A two-storey building example is used. It is assumed that every node in the WSN contains all the required sensors. At intermediate stages, data aggregation was performed on the continuous data obtained from different sensor points of the same type. Data aggregation helped to avoid communication overhead between neighboring nodes. *k*-means clustering was subsequently performed to divide the aggregated data of the selected optimal sensor into two clusters. And finally, we demonstrate the use of efficient and cheap detection algorithm such as the feed forward neural network (FFNN), Naïve Bayes and Decision Tree to show that the performance accuracy increases significantly.

III. DATA AGGREGATION IN CLUSTER-BASED WSN

In some WSNs where sensor nodes are densely deployed, each sensor node senses similar data from the physical environment due to closeness of sensor nodes. This type of sensor network will result in transmitting redundant data and this has the potentials to degrade the overall network. To solve this issue, there is a need to perform some grouping of sensor nodes and also combining or compressing data and transmitting only the compact data. [11].

In cluster-based WSN, sensors are grouped in clusters and in-network data aggregation is done locally within the clusters. A cluster head plays the role of aggregator which aggregate data received from cluster members locally and then transmits the result to base station (sink). In heterogeneous networks, the clustered head or aggregator node has a higher energy capability compare with the member sensor nodes

within the same cluster network. A data aggregation scheme is energy efficient if it maximizes the functionality of the WSN in the sense that sensor nodes should spend the same amount of energy in every data gathering round. Figure 1 below shows a cluster-based sensor network organization that involves a long range transmissions or multi hopping through other cluster heads to the sink or base station.

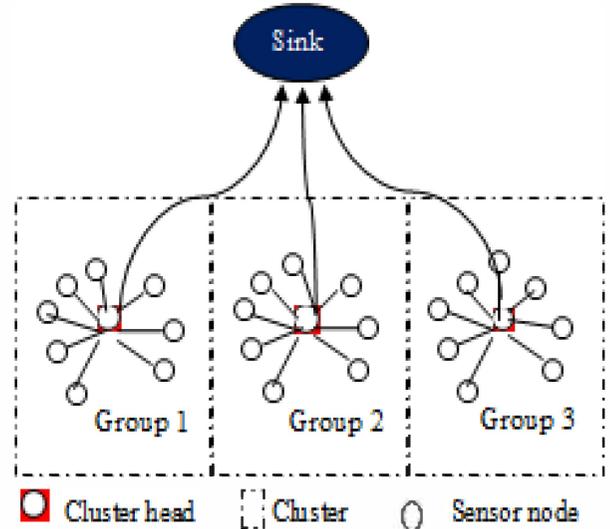


Figure 1 Cluster based sensor network.

The performance measures of data aggregation in cluster base WSN for event detection application are discussed below:

- Network lifetime: Network lifetime is defined as the number of data aggregation rounds till certain percentage of the sensors die. In applications such as fire detection where all sensor nodes are vital, the lifetime is define as the number of rounds until the first sensor drains off its battery energy or dies completely
- Data Accuracy: This is the evaluation of the ratio of the total number of reading received at the base station to the actual total number of data generated.
- Latency: Latency is defined as the delay involved in data transmission, routing and data aggregation. It can be measured as the time delay between the data packets received at the sink and the data generated at the source nodes.

IV. PROPOSED HYBRID APPROACH

In this hybrid approach, the *k*-means clustering algorithm is used to assign labels to a set of features from four sensor types (temperature, ionization, photoelectric and CO). *k*-means clustering was used to partitioned the data into two clusters (i.e. $K=2$), and assigned labels to the data. The labeled data output is fed into the classifier for training and testing purposes. The FFNN, Naïve Bayes and Decision tree are used as classifier and the prediction accuracy was presented in section V.

A. k-MEANS CLUSTERING ALGORITHM

k-means clustering aims to divide a set of n observations (x_1, x_2, \dots, x_n) into k ($\leq n$) sets disjoint $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the sum-of-squares criterion

$$J = \sum_{j=1}^k \sum_{X \in S_j} \|X - \mu_j\|^2 \quad (6)$$

Where μ_j is the mean of points in S_j , X is a vector representing the n observations. In general, the algorithm does not achieve a global minimum of J over the assignments. The limitation of this algorithm is that the reachable minimum is not a proper local minimum because the algorithm uses discrete assignment rather than a set of continuous parameters. However, the algorithm is often used due to its ease of implementation [12].

According to [13], the main k-means clustering algorithm steps are as follows:

1. Select an initial partition with K clusters; repeat steps 2 and 3 until cluster membership stabilizes.
2. Generate a new partition by assigning each pattern to its closest cluster centre.
3. Compute new cluster centres.

In this work, k-means clustering will be combined with FFNN, Naïve Bayes and Decision Tree for accurate fire detection.

B. CLASSIFICATION TECHNIQUES

Supervised learning techniques strongly rely on the presence of labeled data and require a training phase.

1. Feed Forward Neural Network

Feed forward neural network (FFNN) is a type of the neural network, in which each layer is fed by its back layer [14]. FFNN consists of one input layer, one or more hidden layers and one output layer. Figure 1 shows the FFNN's architecture.

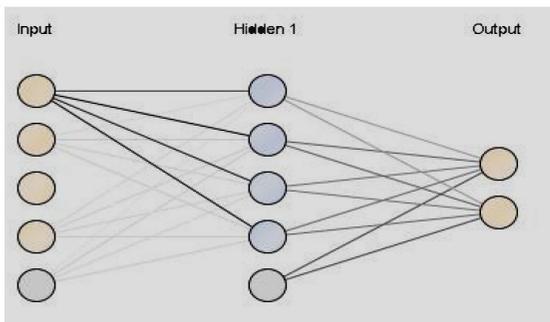


Figure 2 Feed forward Neural Network

FFNN is trained by learning iteratively processing a training set and comparing the predicted output with the known target output using back-propagation algorithm. During the training process, the mean square error between the predicted and actual target value is minimize by adjusting the weights in a backward direction. One major challenge of FFNN is finding the optimal weight and one of the ways of

finding the weights is through gradient descent (GD) approach [15].

2. Naive Bayes classifier

A Naïve Bayes classifier is a simple probabilistic classifier that obtains the posterior probability of each class, C_i using Bayes rule. The Naive Bayes classifier (NBC) makes the simplifying assumption that the attributes A, are independent given the class, so the likelihood can be obtained by the product of the individual conditional probabilities of each attribute given the class [16]. Thus, the posterior probability $P(C_i | A_1, \dots, A_n)$ is given by:

$$P(C_i | A_1, \dots, A_n) = \frac{P(C_i)P(A_1 | C_i) \dots P(A_n | C_i)}{P(A)} \quad (2)$$

In simple terms, NBC assumes that the presence (absence) of a particular feature does not relate to the presence (absence) of any other feature. The merit of Naive Bayes is that it requires a small number of training data to compute the means and variances that is used for classification. However, independent variables are assumed because only the variances for each labels is required and not the total covariance matrix [17].

3. Decision tree

A decision tree is an inverted tree-like model because of its top root and bottom branches structure. The goal of this model is to predict the value of a target attribute called class or labels based on several input attributes of the datasets. In Rapid Miner™ an attribute with *label* role is predicted by the Decision Tree operator.

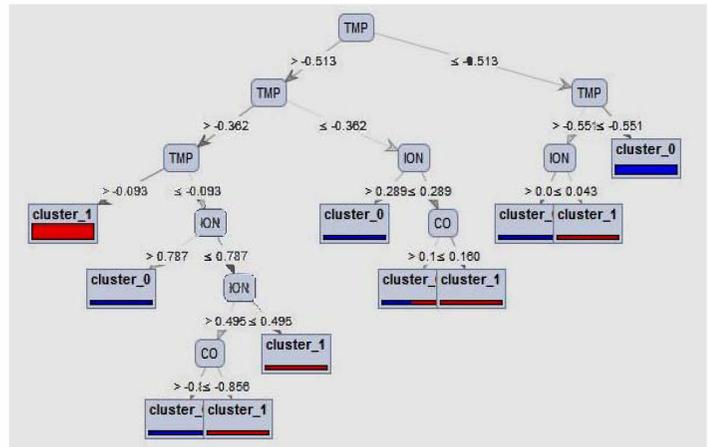


Figure 3 Diagram of a Decision Tree in Rapid Miner

Note: In figure 3 above, the temperature, carbon monoxide and ionization sensors are represented as TMP, CO, ION respectively.

Each interior node of tree is matched to input attributes. The number of possible values of the input attribute is equal to the number of edges of nominal interior node. Disjoint ranges label is assigned to outgoing edges of numerical attributes. Each leaf node describes the value of *label* attribute given the values of the input attributes

represented by the path from the root to the leaf. The advantage of decision tree is that data representation is meaningful and easy to interpret compared with other approaches [17].

V. EMPIRICAL RESULTS

To evaluate the prediction accuracy of the hybrid combination of k -means clustering with FFNN, Naive Bayes and Decision trees against the use of the classifiers alone, a set of data were obtained and a number of experiments were conducted. To evaluate the performance of the hybrid approach, we carried out experiments on six fire datasets obtained from NIST website (<http://smokealarm.nist.gov/>) namely two soldering fire dataset, two flaming fire dataset and two nuisance resource dataset. This dataset are merged together and pre-processed into a total of 1400 data instances with hundreds of attributes (based on sensor types and range), all having same units. All data from the same sensor type were fused together using the average operator. The aggregated data was grouped into two clusters and finally passed to the classifier. The goal of the classifier is to accurately separate the data and classify them into their respective class, i.e., fire and non-fire (noise).

The aggregated data obtained from the four sensor nodes (CO, temperature, photoelectric and ionization). After labelling, the data were passed to the classifiers. To perform a cross validation, 1400 data instances were divided to a 1000 training data and a 400 test data. All data were randomly mixed and given to the classifiers. Each test was repeated 10 times and the average prediction accuracy by changing the classifiers is shown in Table 1 below.

Table 1 Empirical Results for all Classifiers

Hybrid Approach	Prediction Accuracy (Percentage)	
	Fire (Cluster 0)	Noise (Cluster 1)
FFNN	100.00%	99.79%
Naive Bayes	100.00%	99.26%
Decision Tree	98.26%	99.68%

Table 2 Distribution table for the four sensor types

Attribute	Parameter	Cluster 0	Cluster 1
Temperature	Mean	-1.271	0.614
Temperature	Standard deviation	0.388	0.502
CO	Mean	0.216	-0.104
CO	Standard deviation	0.880	1.038
ION	Mean	1.113	-0.538
ION	Standard deviation	0.817	0.523
Photo	Mean	-0.003	0.001
Photo	Standard deviation	0.150	0.166

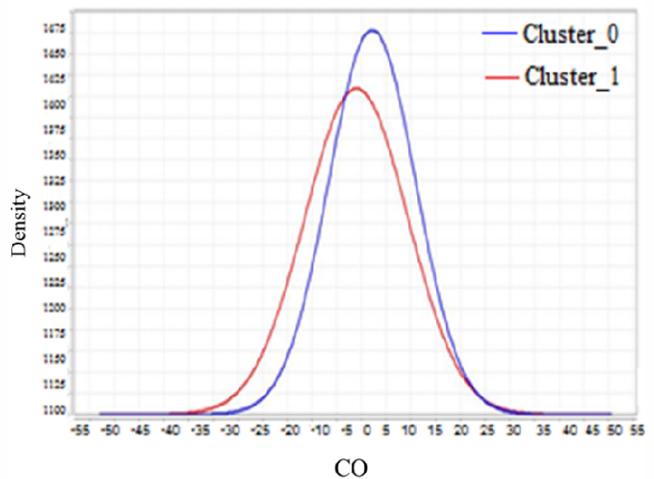


Figure 4a Density curves of CO Sensor Data

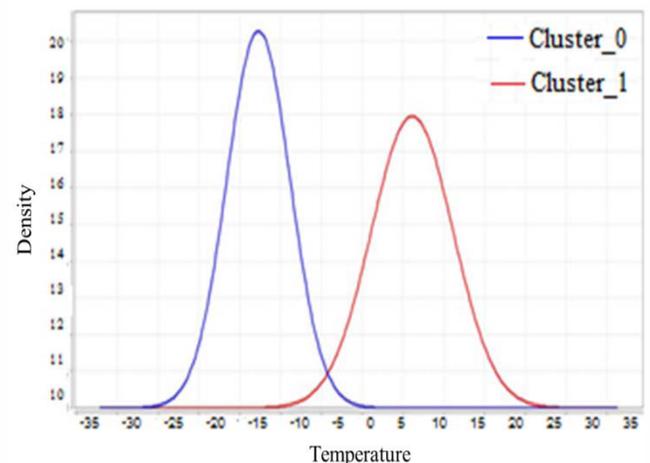


Figure 4b Density curves of Temperature Sensor Data

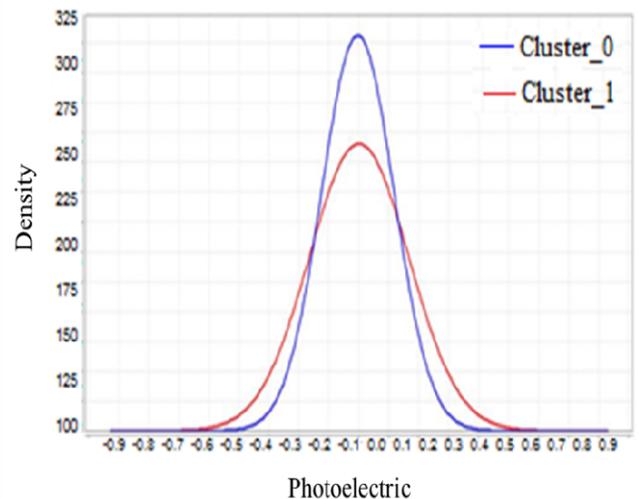


Figure 4c. Density curves of Photoelectric Sensor Data

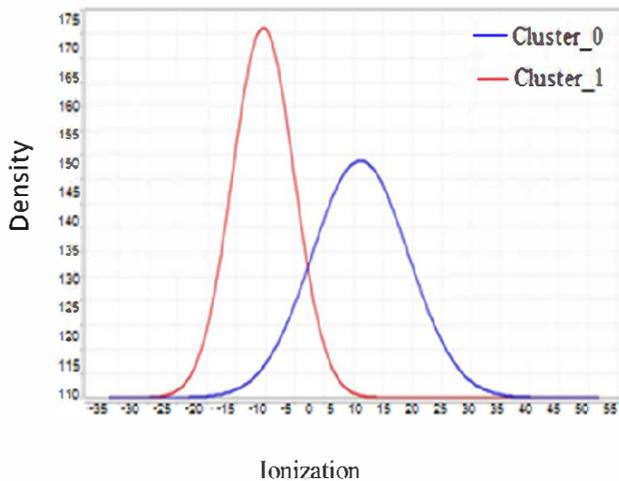


Figure 4d Density curves of Ionization Sensor Data

Note: The blue curve denotes fire event and the red curve denotes a non-fire event.

Table 1 shows the prediction accuracy of the classifiers can be improved by using this hybrid approach. The prediction accuracy of FFNN was shown to be 100%, which is a significant improvement compared to 97.49% prediction accuracy obtained in [15].

Also, the symmetrical density curves for CO, Temperature, photoelectric and ionization sensors are displayed in Figure 4 (a-d). For the density curve of temperature sensor data, it can be observed that the fire (cluster 0) curve slightly overlaps the noise (cluster 1) curve compared with the other sensors. This shows that the temperature is the most contributing attribute for fire prediction while the photoelectric sensor is the least contributor due to the obvious overlap between the fire and noise density curves. In Table 2, the photoelectric data attributes has the same approximate standard deviation of ~ 0.2 for cluster 0 and cluster 1.

VI. CONCLUSION

In this paper, we have proposed and demonstrated the use of data aggregation to combine multidimensional data obtained from sensor nodes to reduce the data complexity and communication overhead. The hybrid combination of k -means clustering and the three popular classification approaches such as the FFNN, Naive Bayes and decision tree has enable us to generate a better fire prediction accuracy against the use of only the classifiers.

REFERENCES

- [1] C. Shen, C. Srisathapornphat, and C. Jaikaeo, "Sensor information networking architecture and applications," *IEEE Personnel Communications*, Aug. 2001, pp.52-59.
- [2] I. Memon, T Muntean, "Cluster-based Energy-efficient Composite Event Detection for Wireless Sensor Networks", Sixth International Conference on Sensor Technologies and Applications, pp.241-247, 2012
- [3] R. Rajagopalan, P.K. Varshney, Data aggregation techniques in sensor networks: a survey, *IEEE Communications Surveys Tutorials* 8 (4) (2006) 48–63
- [4] S. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, no. 3, pp. 249–268, 2007.
- [5] A. Nagpal, A. Jatain, D. Gaur, "Review based on data clustering algorithms," *IEEE Conference on Information and Communication Technologies (ICT)*, pp 298-303, 2013.
- [6] Yu, L., N. Wang, and X. Meng. Real-time forest fire detection with wireless sensor networks. *Wireless Communications, Networking and Mobile Computing*. 2005.
- [7] Zhiping, L., et al."The Design of Wireless Sensor Networks for Forest Fire Monitoring System" 2006, White Paper: School of Electronics and Information, Hangzhou Dianzi University.
- [8] Zervas, E., et al. Fire Detection in the Urban Rural Interface through Fusion Techniques. In *Mobile Adhoc and Sensor Systems (MASS 2007)*. 2007.
- [9] Cestari, L.A., C. Worrell, and J.A. Milke, *Advanced Fire Detection Algorithms Using Data from the Home Smoke Detector Project*. *Fire Safety Journal*, 2005. **40**: p. 1-28.
- [10] Milke, J.A. *Using Multiple Sensors for Discriminating Fire Detection*. In *Fire Suppression and Detection Research Application Symposium*. 1999: National Fire Protection Research Foundation
- [11] K. Maraiya, K. Kant, & N. Gupta, "Efficient cluster head selection scheme for data aggregation in wireless sensor network." *International Journal of Computer Applications* 23, no. 9, pp10-18, 2011.
- [12] S. Na, L. Xumin and G. Yong, "Research on k-means clustering algorithm: An improved k-means clustering algorithm", *Proc. 3rd IEEE Int. Symp. Intell. Inf. Technol. Security Informat*, pp.63 -67 2010.
- [13] A. K. Jain "clustering: 50 years beyond K-means", *Pattern recognition letters*, Vol 31, Issue 8, pp 651–666, 2010.
- [14] Mehrotra, K., C.K. Mohan, and S. Ranka, *Elements of Artificial Neural Networks*. 1996, MIT Press.
- [15] Bahrepour, Majid, Nirvana Meratnia, and Paul JM Havinga. "Use of AI techniques for residential fire detection in wireless sensor networks." (2009): 311-321.
- [16] Zhang, H. *The Optimality of Naive Bayes*. in *Seventeenth Florida Artificial Intelligence Research Society Conference*. 2004: AAAI Press.
- [17] Fareed Akthar and Caroline Hahne, "RapidMiner 5: Operator Reference" *Rapid-I* (2012), pp 651-660 <http://www.csce.uark.edu/~xintaowu/4013/RapidMine>