International Journal of Digital Crime and Forensics
Volume 13 • Issue 5 • September-October 2021

Design and Development of Ternary-Based Anomaly Detection in Semantic Graphs Using Metaheuristic Algorithm

M. Sravan Kumar Reddy, Vellore Institute of Technology, India
Dharmendra Singh Rajput, Vellore Institute of Technology, India

ABSTRACT

At present, the field of homeland security faces many obstacles while determining abnormal or suspicious entities within the huge set of data. Several approaches have been adopted from social network analysis and data mining; however, it is challenging to identify the objective of abnormal instances within the huge complicated semantic graphs. The abnormal node is the one that takes an individual or abnormal semantic in the network. Hence, for defining this notion, a graph structure is implemented for generating the semantic profile of each node by numerous kinds of nodes and links that are associated to the node in a specific distance via edges. Once the graph structure is framed, the ternary list is formed on the basis of its adjacent nodes. The abnormalities in the nodes are detected by introducing a new optimization concept referred to as biogeography optimization with fitness sorted update (BO-FBU), which is the extended version of the standard biogeography optimization algorithm (BBO). The abnormal behavior in the network is identified by the similarities among the derived rule features. Further, the performance of the proposed model is compared to the other classical models in terms of certain performance measures. These techniques will be useful to detect digital crime and forensics.

KEYWORDS

Anomaly Detection, Digital Crime, Semantic Graph, Ternary List Formation

1. INTRODUCTION

In real-world applications, the social networks and the sensor networks are playing a crucial role from politics to healthcare and hence the computational analysis of graphs is a vital area of study (Wang et al., 2018) (Yao et al., 2016) (Lin & Chalupsky, 2008). The amount of graph data generated from diverse sources is in the exploration stage, it is a bit complex to analyze and understand the graph data. The ubiquitous presence of graphs includes social networks, citation networks, computer networks, biological networks, and the Web (Lhioui et al., 2017) (Etaiwi & Awajan, 2020) (Sun et al., 2020) (Rehman Javed et al., 2020) (Numan et al., 2020). The rich information these days is proliferating in real-world graphs and hence the attributes associated with the characteristics and properties of the information are described as the vertices and edges of the graph. In the similarity graphs, several lexical matching techniques were offered to detect the similarity between the node pairs (Lampropoulos et al., 2020) (Bounhas et al., 2019) (Antonello et al., 2020) (Chen et al., 2018). Among them, the semantic similarity approaches are more attractive, such that it has gained the attention of the current researchers. "A semantic graph is a graph where nodes represent objects of different types (for

DOI: 10.4018/IJDCF.20210901.oa3

example, persons, papers, organizations, etc.) and links represent binary relationships between those objects (for example, friend, citation, etc.)". Semantic graphs (Lugowski et al., 2015) (Guesmi et al., 2016) with various types of associations are known as MRNs. Further, with an increase in the web-scale graphs and high-frequency sensor data in the MRNs, the anomaly detection is of great focus. Typically, anomaly detection refers to the problem of identifying patterns in data that do not conform to an expected behavior (Assi et al., 2019) (Zhao et al., 2018) (Ahmad et al., 2017). A node is said to be suspicious and abnormal if the corresponding network encompasses the unique or abnormal semantics (Javed et al., 2020) (Mittal et al., 2020). In order to realize the concept of abnormal node detection, a semantic profile is generated for each node by means of summarizing a graph structure surrounding it. This is based on the different types of links and nodes connected to the node within a certain distance (Vlietstra et al., 2017) (Vela et al., 2017). The abnormal nodes here are identified as the nodes with abnormal semantic profiles. In the traditional approaches, random walks and SNA were developed as unsupervised network algorithms for detecting the nodes with abnormal semantic profiles. The major drawback of this approach is, they do not consider the semantics of links. Till now, the contribution of the optimization algorithms in the field of Anomaly Detection in Semantic Graphs is in the budding stage (Mittal et al., 2020) (Javed et al., n.d.).

The major contribution of the current research is described below:

- A novel optimization algorithm referred to BO-FBU is introduced for optimizing the node pairs, thereby detecting the anomaly nodes.
- The proposed method is optimizing the solution using Semantic Graphs using Metaheuristic Algorithm that will be enhanced to detect cyber-attack (Iwendi et al., 2020) (Ch et al., 2020). It will provide an optimal solution for collecting digital evidence, through to detection and classification APT attack and Study of propagation behavior. (Gupta & Sheng, 2019)
- The proposed meta-heuristic approach doesn't get struck into local optimal point of search.
- This approach has better exploration and exploitation rate when compare with other meta-heuristic approaches which is used for anomaly detection.
- The overall evaluation shows that the presented work is 45.9%, 66.6%, 64%, 25.9% and 2.1% better than the existing models like WOA, BBO, DA, FW-DA and T-LAU, respectively

1.1. Organisation of Paper

The rest of this paper is organized as: Section II portrays about the literature works undergone under this subject. Section III portrays the overall architecture of proposed anomaly detection in the semantic graph. Section IV describes the abnormal node detection using the proposed optimization concept. Further, Section V discusses the acquired results and Section V concludes the current research work.

2. LITERATURE REVIEW

2.1. Related Works

In 2019, Zhang *et al*. (2019) have developed a TG model for annotating the web image, which encompassed subgraphs like region data graph, image data graph, and label graph. The authors have connected these sub-graphs with tripartite graph that were induced from the label assignments and the resultants of the image segmentation. In addition, they have performed vertex-to-vertex, region-to-label, image-to- region and image-to-label relevance with multilevel Random Walk and Restart algorithm onto TG graphs. In the unannotated regions of the image, the labels were predicted by means of adding the unlabeled images with a semi-supervised learning approach. The resultant of the presented work had exhibited good performance on image region annotation.

In 2014, Vidal *et al*. (2014) have developed a novel approach for semantic annotation in document enrichment filed with linked data. The relevant terms were connected to graph on ontology and instance related to relevant terms are expanded to a depth limit that was defined in prior. Thus, the relevant term was fully connected to the graph. In the expansion process, they have neglected out the unrelated instances from the contextualized information contained in the document. They have evaluated the presented work in the e-Learning domain and have validated its efficiency.

In 2019, Albukhitan *et al*. (2019) have investigated the utilization of the word embeddings from deep learning algorithms for annotating the structure of Arabic words semantically. They have utilized the complicated morphological structure of the Arabic words to improve the semantic annotation performance. The proposed framework was evaluated with the group of domain ontology corresponding to the annotated document. The resultant had exhibited promising performance in terms of precision and recall.

In 2019, Cui *et al*. (2019) have developed a novel anomaly detection method on the basis of GGL matrix to make the spatiotemporal relationship corresponding to μPMU data visualized. The authors have segmented μPMU data into multiple segments. The CGL matrix was estimated with the Lagrangian function. The evaluation of the μPMU-based spatiotemporal was accomplished using the GGL matrix normalized diagonal elements.

In 2020, Cheng and Wang (2020) have developed a novel hyperspectral anomaly detection method on the basis of the GTVLRR model, which was the hybridization of graph regularization and TV regularization. This was developed with the intention of securing the spatial relationships and the geometrical structure of the hyperspectral images. They have conducted the experiments with the real and simulated hyperspectral data sets, and have shown the supremacy of the presented work over the extent anomaly detection methods.

In 2018, Qin *et al*. (2018) have developed a multi-target CNN with the intention of segmenting a variety of components in the train wheel system for detecting the faults in trains. Further, from the segmentation algorithm, the masks generated were assessed in an effective manner with the help of the semantic consistency matrix. Typically, this semantic consistency matrix had exhibited the spatial relationships existing in between the diverse train components with the quantitative measure.

In 2018, Sadreazami *et al*. (2018) have designed a novel distributed blind intrusion detection framework for anomaly detection in distributed sensor networks. They have realized the proposed model by means of constructing a graph signal on the basis of the placements as well as measurements of sensors. Further, on the basis of the hypothesis testing as well as log-likelihood ratio criterion, the intrusion detector was designed with Gaussian Markov random field distribution. They have validated the proposed intrusion detection scheme experimentally and have revealed its potential of detecting the measurement anomalies of the sensor over time.

In 2019, Farag *et al*. (2019) have developed a parallel outlier detection technique for detecting the outliers corresponding to the contextual anomalies in the sequential data. The authors have deployed the graph approach for detecting the contextual anomalies. The MST algorithm was deployed to find the MST of the graph. In addition, they have deployed a parallel threshold MST clustering technique to neglect out the inconsistent weights of edges from MST. The resultant of the presented work had exhibited accurate detection of the sequential data outliers even under the increasing count of processors.

The abnormal paired nodes are optimally selected using the improved Dragon Fly Algorithm (DA) based on the maximum mutual information. Since the weighting factors utilized in DA is based on fitness function, this algorithm is termed as Fitness Weighed Dragonfly Algorithm (FW-DA) proposed by Sravan Kumar Reddy et al (n.d.a). The effectiveness of this algorithm is substantiated by comparing it over the conventional models through various performance analyses. Another pioneer meta heuristic approach in detecting the anomaly detection is been proposed by Sravan Kumar Reddy et al. it aims to introduce a new idea of finding the abnormal or suspicious nodes, and this is done by modeling the graph structure using diverse nodes and links associated to the node in a particular distance through edges. Once the graph structure is framed, the ternary list formation is done, and the list of all nodes is based on the nearby list of neighborhood nodes. Subsequently, the logic is induced in this work to detect the abnormal nodes via optimizing the node pairs. For this, a new hybrid algorithm termed as Threshold-based LA Update (T-LAU) that hybridizes the concepts of Lion Algorithm (LA) and Firefly (FF), respectively. (Kumar, n.d.b)

In 2016, Seyedali Mirjalili et al propose Whale Optimization Algorithm (WOA) based on the behavioral inspiration of the whale namely mimicking the hunting behavior of humpback whales (Mirjalili & Lewisa, 2016). It is been a population based meta heuristic approach so it guides in identifying the prey through better exploration and exploitation. A proper balance is been found between exploration and exploitation that guides in searching the promising areas. The major cons observed in WOA are slow convergence and low accuracy.

In 2012 Seyed Habib et al. proposed BBO algorithm for testing differet classical problems related to FJSP (flexible job shop scheduling Problem) . In BBO the habitat is considered to be individual and it has habitat suitability index (HSI) which shows the degree of goodness. If HSI is low it shows a poor solution and vice versa. This BBO helps in solving this optimization problem through the presence of migration operators. This approach is been designed to portray the migration of the biological species based on the HSI value. The major cons of this approach it is found to have time complexity during the computation.

2.2. Extraction of the Liturature

The intelligent graph analysis methods are becoming more famous in the area of security. Table 1 describes the features and challenges of the semantic models discussed in the literature. In TG (Zhang et al., 2019), the highest degree of relevance is achieved even under the unlabeled regions. But, here the links in the nodes are not updated, while a new node arrives. In Graph-based semantic annotation (Vidal et al., 2014), the precision, as well as recall, is higher. This technique had no consideration on the context of the document while performing annotation. Further, the word embedding language model in (Albukhitan et al., 2019) is good in enhance annotation performance. But, here the time consumed for training is higher. Further, μPMUs (Cui et al., 2019) is good in detecting the anomalies in the temporal domain with a higher degree of accuracy. Yet, this technique is more prone to noise and hence fails in large-scale detection. The anomaly detection performance is significantly increased with GTVLRR (Cheng & Wang, 2020). Here, the hyperspectral anomaly was detected with high spectral correlations. Apart from these advantages, the computational cost and computational cost is higher. Moreover, multi-target CNN in (Qin, 2018) has a higher mean dice coefficient and hence can detect the possible defects in the system taken into consideration. The major drawbacks of this approach are its lower robustness to noise. The Distributed Graph-based Statistical Approach in (Sadreazami et al., 2018) provides the highest probability of detection for a given probability of false alarm. But, it is incapable of detecting sensor measurement anomalies over time. Moreover, MST in (Farag et al., 2019) has the potential of handling massive data even under the increase in the count of processors. This technique could be much better if the inconsistent weights of MST are neglected.

3. OVERALL ARCHITECTURE OF PROPOSED ANOMALY DETECTION IN THE SEMANTIC GRAPH

This section discusses about the proposed anomaly detection in the semantic graph and various steps involved in this methodology.

3.1. Steps Involved in Ternary-Oriented Anomaly Detection Mechanism

The steps involved in the abnormal node detection in the semantic graph are represented below:

Step 1: Initially, a semantic graph is constructed by considering each of the components as a node.
Step 2: In a similar graph, a ternary list is constructed to find the connectivity among the nodes.

Table 1. Features and Challenges of conventional semantic models

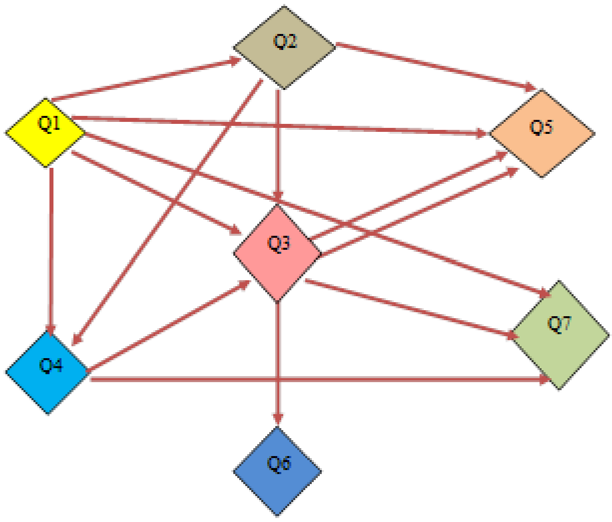
Author [Citation]	Methodology	Features	Challenges
Zhang *et al*.(2019)	TG	✓ Achieves more accurate and abundant annotation results. ✓ Achieve good performance on image region annotation	× High computational complexity
Vidal *et al*. (2014)	Graph-based semantic annotation	✓ Higher precision ✓ Higher recall	× Computationally expensive × Limit the exploration depth
Albukhitan *et al*. (2019)	word embedding language model	✓ Promising Precision and Recall ✓ Good improvement in multiple areas	× Requires performance improvement in each component × Training time higher
Cui et al. (2019)	μPMUs.	✓ Simple and tractable ✓ Able to accurately and robustly detect the anomalies in the temporal domain.	× Prone to noise × Low convergence
Cheng and Wang (2020)	GTVLRR	✓ Improve detection performance significantly. ✓ Enhance the discrimination between the intrinsic structure of the data and the noise.	× More computation time × Higher computational cost
Qin et al. (2018)	multi-target CNN	✓ Higher mean dice coefficient	× Total loss function is higher × Less robust to noise
Sadreazami et al. (2018)	Distributed Graph-based Statistical Approach	✓ Higher detection rate values	× Incapable of detecting sensor measurement anomalies over time
Farag et al. (2019)	MST	✓ Handle the massive data ✓ Computation time is decreased	× Process time need to be reduced further × Inconsistent weights of MST need to be removed
Al-Jadir et. al. (2018)	Memetic Algorithm Feature Selection (MAFS)	✓ To partition crime reports and criminal news	× Hybridization of algorithms × More computation time

Step 3: For ternary-oriented anomaly detection in semantic graphs, the nodes pairs are fed as input to the proposed optimization algorithm referred to BO-FBU, which is the extended version of BBO.
Step 4: Once, the abnormal nodes are detected, every pair on the gathered ternary list is updated using the proposed model.
Step 5: After updating, each pair is checked in terms of the objective function. If the objective function is satisfied, then the respective pair is chosen, if not move to the subsequent neighbour.

3.2. Semantic Graph Construction

The bibliography network encompassing seven nodes is exhibited in Figure 1. In the network, the nodes represent the biographer of papers, journals or organizations. The concerned work is published by means of combining a node with its adjacent node. This mechanism of merging two nodes is referred as co-authoring node or co-authoring pair. A node is said to be an abnormal node while two nodes combine with more or less than the limited count of nodes. As a novelty, the current research work pays attention in detecting these abnormal nodes (He et al., 2016) (Bhattacharya, Somayaji, Gadekallu et al, 2020).

Figure 1. The bibliography network encompassing seven nodes



Assumption: The nodes in the network are Q_i, where $i = 1,2,3,\ldots Q_B$. The notation Q_B is the count of nodes in the network. Here, seven nodes are taken into consideration, such that Q_B =7.

Illustration: Figure 1 exhibits the diagrammatic representation of semantic graph. Here, seven nodes (the bibliographers) are utilized and they are number from $Q1 - Q7$. The Node $Q1$ is connected to its adjacent nodes $Q2$, $Q3$, $Q4$, $Q5$ and $Q7$. Further, $Q2$ is connected with its neighbours $Q3$, $Q4$ and $Q5$. The adjacent nodes of $Q3$ is $Q5$, $Q6$ and $Q7$, and $Q5$ as well as $Q7$ are the neighbours of $Q4$. Moreover, $Q5$, $Q6$ and $Q7$ are free (not connected).

Objective Function: The major objective of this research work is to detect the abnormal node in the semantic graph, thereby reducing B. The objective function $\left(Ob\right)$ of the current research work is expressed mathematical in Equation (1), in which B is computed using Equation (2).

$$Ob = \min(B) \tag{1}$$

$$B = \frac{1}{P\left(Q_i, \left|Q_j^{(i)}\right.\right) * P\left(\left|Q_i, Q_j^{(i)}\right.\right)} + \mu I \tag{2}$$

The mathematical formula for mutual information $\left(\mu I\right)$ is shown in Equation (3)

$$\mu I = P\left(Q_i, Q_j\right) \log \left| \frac{P\left(Q_i, Q_j\right)}{P\left(Q_i\right) P\left(Q_j\right)} \right| \tag{3}$$

Here,

$Q_i \rightarrow i^{th}$ Node

$Q_j^{(i)} \rightarrow j^{th}$ Neighbour of i^{th} node

$\left(Q_i, \left|Q_j^{(i)}\right.\right) \rightarrow Q_i$ Co-authors other than the $Q_j^{(i)}$ co-authors

$\left|Q_i\right., Q_j^{(i)} \rightarrow Q_j^{(i)}$ Co-authors other than the Q_i co-authors

3.3. Ternary List Formation

In order to achieve the objective of abnormal node detection, a new contribution is deployed in the current research work, the ternary list formation. The connectivity established among the nodes is detected by the ternary list formation mechanism. The ternary list is typically formed by means of determining the neighbours of each node. The intersection between i^{th} node and j^{th} node is symbolized as $T_i\left(j\right)$ and it is determined using Equation (4).

$$T_i\left(j\right) = A_i \cap A_j^{(i)} \tag{4}$$

Here,

$A_i \rightarrow$ Adjacent list of i^{th} node

$A_j^{(i)} \rightarrow$ Adjacent list of j^{th} neighbour of i^{th} node

Once, the intersection $T_i\left(j\right)$ is computed, the ternary list $TL_{ij}\left(k\right)$ is determined using Equation (5).

$$TL_{ij}\left(k\right) = \left[Q_i \;\; Q_{A_j^{(i)}} \quad TL_i\left(j\right)\left(k\right)\right] \tag{5}$$

Lemma 1: As per Equation (6), the count of rows of ternary list $TL_{ij}\left(k\right)$ needs to be equivalent to the count of intersection $T_i\left(j\right)$.

The ternary nodes formed by combining the nodes are denoted as Δ_i and it is expressed mathematically in Equation (6)

$$\Delta_i = \sum_{j=1}^{A_i} \left|TL_{i,j}\right| \tag{6}$$

The neighbouring list of node 1 $Q1$ is $Q2$, $Q3$, $Q4$, $Q5$ and $Q7$. In addition, $Q3$, $Q4$ and $Q5$ are the neighbours of $Q2$. The neighbouring list between $Q1$ and $Q2$ is $QL_1\left(2\right)$, and it is formed by taking their intersections between them $\left[i.e. \left\{Q2, Q3, Q4, Q5,\ Q7\right\} \cap \left\{Q3, Q4, Q5\right\}\right]$. Thus, the ternary list formed is denoted as $T_1\left(2\right) = \left\{Q3, Q4, Q5\right\}$. The formulation of the ternary list is as per Equation (5). The formulated ternary node $TL_1\left(2\right)$ is shown in Equation (7).

$$TL_1\left(2\right) = \begin{bmatrix} Q1, Q2, Q3 \\ Q1, Q2, Q4 \\ Q1, Q2, Q5 \end{bmatrix} \tag{7}$$

In addition, the interaction between $Q1$ and $Q3$ for ternary list formation is $TL_1\left(3\right)$. It is formed by taking the intersections $\left[i.e \left\{Q2, Q3, Q4, Q5,\ Q7\right\} \cap \left\{Q5, Q6, Q7\right\}\right]$. Thus, the ternary list between $Q1$ and $Q3$ is $T_1\left(3\right) = \left\{Q5, Q7\right\}$. The ternary list formed between $N1$ and $N3$ is shown in Equation (8).

$$TL_1\left(3\right) = \begin{bmatrix} Q1, Q3, Q5 \\ Q1, Q3, Q7 \end{bmatrix} \tag{8}$$

Similar to this, the ternary list between $Q1$ and $Q4$ is $T_1\left(4\right) = \left\{Q5, Q7\right\}$. This ternary list $TL_1\left(4\right)$ can be expressed as per Equation (9)

$$TL_1\left(4\right) = \begin{bmatrix} Q1, Q4, Q5 \\ Q1, Q4, Q7 \end{bmatrix} \tag{9}$$

Since, there exist no intersections between nodes $Q1$ and $Q5$, the ternary list $TL_1\left(5\right)$ is denoted as per Equation (10).

$$TL_1\left(5\right) = \Phi \tag{10}$$

Alike the above condition (i.e. between $Q1$ and $Q5$), there exists no interconnection between $Q1$ and $Q7$. It is expressed as per Equation (11)

$$TL_1\left(7\right) = \Phi \tag{11}$$

Thus, Equation (12) shows the model formed by the combination of $Q1$ nodes.

$$\Delta_1 = \begin{bmatrix} Q1, Q2, Q3 \\ Q1, Q2, Q4 \\ Q1, Q2, Q5 \\ Q1, Q3, Q5 \\ Q1, Q3, Q7 \end{bmatrix} \tag{12}$$

The ternary list determination of the entire nodes that is on the basis of the adjacent nodes is defined as in Equation (13) and Equation (14), respectively.

$$\Delta_2 = \begin{bmatrix} Q2\,Q3, Q5 \\ Q2\,Q4, Q5 \end{bmatrix} \tag{13}$$

$$\Delta_3 = \Phi \tag{14}$$

4. ABNORMAL NODE DETECTION USING PROPOSED OPTIMIZATION CONCEPT

This section discusses about detecting the abnormal behaviour of nodes with the help of the proposed optimising methodology

4.1. Solution Encoding

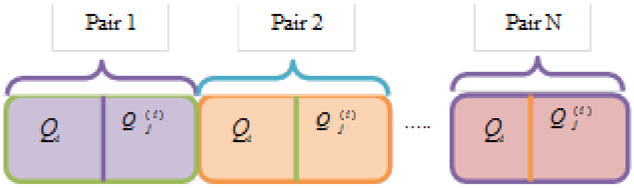
The abnormal nodes or abnormal co-authoring detection in the semantic graphs being the objective of current research is achieved with the aid of the proposed optimization concept. The input (solution) to the proposed optimization algorithm is the variables Q_i and $Q_j^{(i)}$. Here, the length of the given solutions is user defined (any number of pairs).The probability of variables Q_i and $Q_j^{(i)}$ is determined using Equation (15), Equation (16) and Equation (17), respectively.

$$P\left(Q_i\right) = \frac{T_{Q_i}}{T_c} \tag{15}$$

$$P\left(Q_j^{(i)}\right) = \frac{T_{Q_j^{(i)}}}{T_c} \tag{16}$$

$$P\left(Q_i, Q_j^{(i)}\right) = \frac{T_{Q_{(i,j)}}}{T_c} \tag{17}$$

Figure 2. The solution given to the algorithm



Here, T_{Q_i} →count of occurrence in the ternary list

T_c → Cardinality of ternary list

$T_{Q_{(i,j)}}$ → Count of joint appearance of Q_i and $Q_j^{(i)}$ in ternary list.

The solution given to the algorithm is shown in Figure 2, where Q denotes the number of pairs.

4.2. Proposed BO-FBU

BBO is one among the recent evolutionary algorithms that is evolved from the biogeography (or zoogeography) studies related to the biological organisms. The immigration, emigration, and mutation of different species are the main inspiration of the BBO algorithm. Since, the convergence of the standard BBO is lower, it is essential to have improvement over it. The optimization algorithms have undergone various improvements in terms of many factors. One among them is by introducing adaptive operators or adaptive functions (Rajakumar, 2013a) (Rajakumar, 2013b) (Swamy et al., 2013) (George & Rajakumar, 2013) (Rajakumar & George, 2012).The abnormal nodes or abnormal co-authoring detection in the semantic graphs is accomplished with the proposed BO-FBU model (He et al., 2016) (Bhattacharya, Somayaji, Gadekallu et al, 2020).

Mathematical model of proposed BBO: Initially, the population of the habitats is generated in a random manner. Every habitat exists in the v - dimensional search space, in which v is a count of variables required for the optimization process. The fitness of all the search agents (overall population) are evaluated and sorted in the increasing order. Then, the first best half fitness of search agents are updated on the basis of the random position using a new variable $update$. The $update$ variable encloses two random position referred as 1 and 2.

(i.e. $update = rand.Position(2,1)$).

Case 1 of $update == 1 \rightarrow$ update the immigration, emigration and mutation rate of the search agent using Equation (18), Equation (19) and Equation (20), respectively. The immigration, emigration and mutation rate represented as $\beta_{\hat{k}}$, $\chi_{\hat{k}}$ and mut_v, respectively in order to exhibit both the exploratory and exploitative behavior of each and every individual species. The mathematical formula for $\beta_{\hat{k}}$ and $\chi_{\hat{k}}$ is expressed in Equation (18) and Equation (19), respectively. The maximum count of habitat, the current count of habitat, maximum immigration rate and the maximum emigration rate is depicted using the term $\hat{N}$, v In and E, respectively.

$$\beta_{\hat{k}} = In\frac{1-v}{\hat{N}} \tag{18}$$

$$\chi_{\hat{k}} = \frac{E * v}{\hat{N}} \tag{19}$$

The mathematical formula for mutation rate mut_v of a habitat is shown in Equation (20), in which the maximum count of the mutation is represented as $\hat{M}$ and the mutation probability in v^{th} habitat is indicated using the term $\hat{p}_v$. The mathematical formula for $\hat{p}_{\max}$ is expressed in Equation (21).

$$mut_v = \hat{M}\left(1 - \frac{\hat{p}_v}{\hat{p}_{\max}}\right) \tag{20}$$

$$\hat{p}_{\max} = \arg\max\left(\hat{p}_v\right),\ v = 1,2,\ldots,\hat{N} \tag{21}$$

Case 2 If $update == 2 \rightarrow$ update the position of the search gent using Equation (22).

$$Pos = \left(Pos * rand\right) + Pos_{best} \tag{22}$$

Here, $Pos_{best} \rightarrow$ best position of the search agent

Further, the next half of the best solutions is updated using Equation (23).

$$Pos = X_{\min} + \left(\left(X_{\max} - X_{\min}\right) * rand\,n(1, size\left(X_{\min}\right)\right) \tag{23}$$

Here, $X_{\min}$ and $X_{\max}$ are the minimum and the maximum solution bounds, respectively. The pseudo-code of the proposed model is shown in Algorithm 1. The flow chart of the proposed work is exhibited in Figure 3.

4.3. Updating Procedure

Each pair of the variables Q_i and $Q_j^{(i)}$ in the collected ternary list is updated using the proposed model. Further, while updating the node Q_i, the count of adjacent nodes of Q_i from the ternary list is verified. Then, it is updated with every twosome of its adjacent list.

For illustration: If $Q1$ node is updated, then its neighbours, viz. $Q2$, $Q3$, $Q4$, $Q5$ and $Q7$ are updated. Once the nodes are updated, every pair is verified in terms of minimized objective function (shown in Equation (1)). If the objective function is satisfied, then the concern pair is selected otherwise, leave out to the subsequently neighbour.

5. RESULTS AND DISCUSSION

This section discusses about the simulation parameters that has been considered for simulating the proposed and existing methods and the obtained results has been portrayed.

Algorithm 1. Pseudo code of Proposed model

Initialize population of the search agent Pop		
For $i = 1 : Pop$		
If 1 $Pop <= Pop / 2$		
	$update = rand.Position(2,1)$	
	If2 $update == 1$	
		Update the immigration, emigration and mutation rate of the search agent using Equation (18), Equation (19) and Equation (20), respectively.
	else	
		Update the position of the search gent using Equation (21).
	End if2	
Else		
	Update the position of the search agent using Equation (22).	
End if1		
End for		

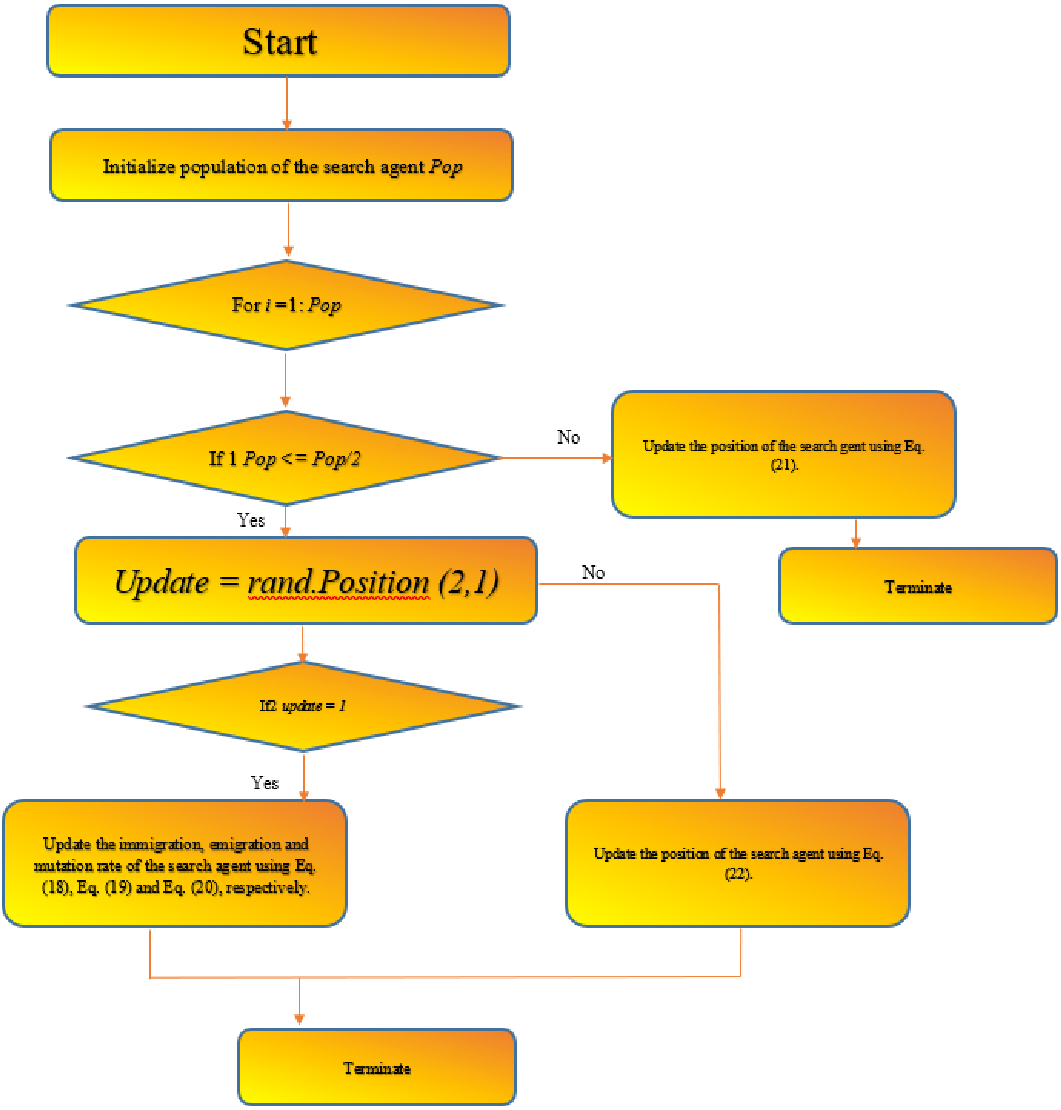
5.1. Simulation Procedure

The proposed ternary-oriented anomaly detection in semantic graphs with BO-FBU was implemented in **MATLAB** and the resultants acquired are noted. The evaluation is undergone with the datasets collected from https://snap.stanford.edu/data/ca-HepTh.html [Access Date: 2020-02-29]. With this collected datasets, the evaluation of the presented work (BO-FBU) was made, by comparing it over the extant model like WOA (Mirjalili & Lewisa, 2016), BBO (Seyed Habib, 2012), DA (Mirjalili, 2016), FW-DA (Kumar, n.d.a) and T-LAU (Kumar, n.d.b) for five pairs namely, 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs of nodes. Here, the evalution is made in tersm of statistical measures and convergence analysis.

5.2. Convergence Analysis

The convergence of the presented work is compared over the existing works in order to exhibit the achievement of the objective function. The convergence analysis of 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs of nodes are evaluated by varying the in Figure 4. In Figure 4(a), the cost function of the presented work for 5 pairs is the lowest one, while comapred to the other measures. The cost fucntion here is below the range of 1000 in the overall iteartions and it is 83.3%, 85.7%, 80%, 50% and 44.4% better than the tarditional models like WOA, BBO, DA, FW-DA and T-LAU, respectively at 100th iteration. Moreover, from Figure 4(b), the cost function of the presented work is initially higher for 7 pairs, but still lower than the traditional approaches. This cost function tends to diminish over the increasing course of iterations and finally at 100th iteration, the presented work achieves the lowest value and it is 94.2%, 96%, 93.3%, 90% and 86.65 better than WOA, BBO, DA, FW-DA and T-LAU, respectively. On observing Figure 4(c), the convergence of 10 pairs is lower all throughout the iterations and hence archives the objective function. The convergence of 12 pairs of nodes is shown in Figure 4(d) and here the cost function of the presented work is initially higher between the ranges 0 to 60 iterations and then it tends to lessen and finally reaches the lowest value than the other models.

Figure 3. The flow chart of the proposed work

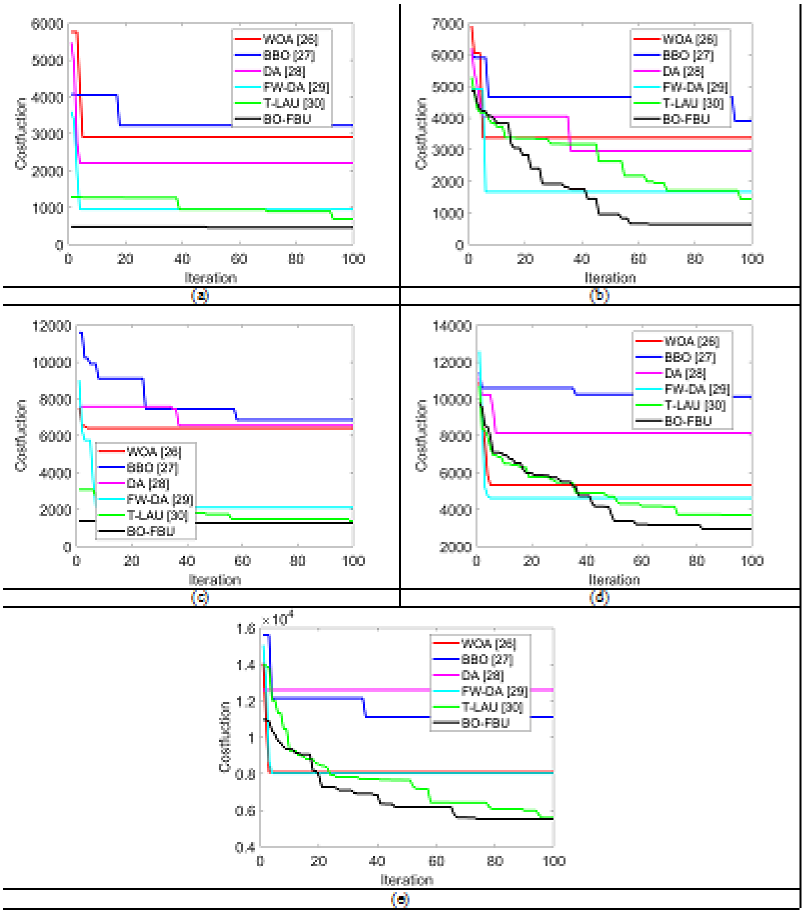


The convergence of 15 pairs of nodes is revealed in Figure 4(e), in which the lowest cost function is revelled at increasing iterations. Since the objective is being the minimization function, the cost function of the presented model is low in all the five cases (viz. 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs of nodes) and hence the presented work is said to be flexible for abnormal node detection.

5.3. Analysis on Performance Metrics

Typically, the anomaly detection in the semantic graph is based on certain quality measures, which are expressed mathematically in Equation (24) - Equation (29), respectively. The contribution of 1st and 2nd author is represented as $Con-1$ and $Con-2$, and their mathematical expression is shown in Equation (24) and Equation (25), respectively. The papers with author 1 and 2 are denoted as $Pap-1$ and $Pap-2$, respectively. In Equation (24), the co-authored by author 1, 2 is denoted as $Co-1,2$. Further, TP denotes the total papers and number of (1,2) pairs is denoted as $NP(1,2)$.

Figure 4. The convergence analysis of 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs of nodes are evaluated



$$Con-1=\frac{Pap-1}{TP} \tag{24}$$

$$Con-2=\frac{Pap-2}{TP} \tag{25}$$

$$Metric1=\frac{Co-1,2}{TP} \tag{26}$$

$$Metric2=\frac{NP\left(1,2\right)}{Number\ \ of\ \ pairs} \tag{27}$$

The frequency of Co-author pair is denoted as f . In addition, F_{Co-1} , F_{Co-No} and $F_{proposed\ pair}$ are the frequency of the Co-author 1, Co-author No and proposed pair, respectively.

$$f_1 = \frac{F_{Co-1}}{F_{proposed\ pair}}$$

.

.

.

$$f_{NO} = \frac{F_{Co-No}}{F_{proposed\ pair}} \tag{28}$$

$$Metric3 = \mu f_{Q0},\ and\ SD \tag{29}$$

The metric based evaluation of the presented and extant works for the considered 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs is exhibited in Table 2. On observing 5 pairs, the metric 3 (mean) of the presented work is the lowest (0.050533), which is 97%, 67.17%, 97.8%, 97.7% and 43.3% better than WOA, BBO, DA, FW-DA and T-LAU, respectively. In case of 10 pair of nodes, the metric 3 (mean) of the presented work is 0.069086, while the mean of WOA=1.2648, BBO= 0.25054, DA= 2.6945, FW-DA= 1.1548 and T-LAU= 0.075878.From this, the lowest metric 3 (mean) is observed by the presented work and hence proves to be significant in abnormal node detection. Alike this, for 15 pair of nodes, the metric 3 (mean) of presented work is 0.17288 (lowest) and it is 84.7%, 40.3%, 92.9%, 83.8% and 13.3% better than the extant models like WOA, BBO, DA, FW-DA and T-LAU, respectively.

5.4. Statistical Evaluation

As the meta-heuristic models are stochastic in nature, the algorithms are executed for 5 times to prove the effectiveness in the defined application by determining the best, worst, mean, median and standard deviation, respectively. The statistical analysis of 5 pairs, 7 pairs, 10 pairs, 12 pairs, and 15 pairs of nodes are tabulated in Table 3. The mean of the presented work for 7 pair of nodes is 1520.5, which is the lowest value, while compared to the existing models. The overall evaluation shows that the presented work is 45.9%, 66.6%, 64%, 25.9% and 2.1% better than the existing models like WOA, BBO, DA, FW-DA and T-LAU, respectively. Further, for 10 pair of nodes, the mean of the presented works is 1341.3, which is 69.1%, 81.05%, 78.2%, 58% and 27.75 better than the conventional models like WOA, BBO, DA, FW-DA and T-LAU, respectively. Thus, the improved performance of the presented work is confirmed from the attained outcomes.

6. CONCLUSION AND FUTURE WORK

This paper had implemented a graph structure for generating the semantic profile of each node by numerous kinds of nodes and links that associated to node in a specific distance via edges. After the framing of the graph structure, the ternary list was formed on the basis of its adjacent nodes. The abnormalities in the nodes were detected by introducing a new optimization concept referred as BO-FBU, which was the extended version of standard BBO. The abnormal behavior in the network

Table 2. Performance Analysis of Presented work over the Conventional model for 5 pairs, 7 pairs, 10 pairs, 12 pairs and 17 pairs of nodes

5 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Metrics1	0.000729	0.007897	0.001458	0.002248	0.018751	0.02361
Metrics2	0.000139	0.001501	0.000277	0.000427	0.003564	0.004487
Metric3-Mean	1.7047	0.15397	2.3443	2.1996	0.089155	0.050533
Metric3-std	2.6709	0.24124	3.673	3.4463	0.13969	0.079175
7 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Metrics1	0.000824	0.005641	0.000304	0.001128	0.01241	0.025398
Metrics2	0.000157	0.001072	5.77E-05	0.000214	0.002359	0.004827
Metric3-Mean	1.2569	0.23567	3.2994	1.0998	0.1262	0.054417
Metric3-std	1.9693	0.36925	5.1695	1.7232	0.19773	0.08526
10 pairs of nodes						
	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Metrics1	0.00082	0.006348	0.000516	0.001063	0.021677	0.022102
Metrics2	0.000156	0.001206	9.81E-05	0.000202	0.00412	0.0042
Metric3-Mean	1.2648	0.25054	2.6945	1.1548	0.075878	0.069086
Metric3-std	1.9816	0.39254	4.2217	1.8093	0.11889	0.10824
12 pairs of nodes						
	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Metrics1	0.000861	0.006556	0.000557	0.002227	0.006362	0.006792
Metrics2	0.000164	0.001246	0.000106	0.000423	0.001209	0.001291
Metric3-Mean	1.987	0.29869	2.4379	1.2831	0.19085	0.19597
Metric3-std	3.1133	0.46799	3.8197	2.0103	0.29903	0.30705
17 pairs of nodes						
	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Metrics1	0.000891	0.006014	0.001033	0.001681	0.00837	0.008869
Metrics2	0.000169	0.001143	0.000196	0.000319	0.001591	0.001686
Metric3-Mean	1.1365	0.28991	2.4635	1.0705	0.15246	0.17288
Metric3-std	1.7806	0.45423	3.8599	1.6772	0.23887	0.27088

was identified y the similarities among the derived rule features. Further, the performance of the proposed model was compared over the other classical models in terms of certain performance measures like convergence as well. the cost function of the presented work for 5 pairs is the lowest one, while compared to the other measures. The cost function here is below the range of 1000 in the

Table 3. Statistical Analysis of Presented work Conventional models for Pair 5, PAir 7, Pair 10, Pair 10, Pair 12 and pair 15 nodes

5 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Best	113.26	2856.4	624.8	354.28	256.4	281.46
Worst	2915.1	3225.5	2216	3353.1	842.01	1099.8
Mean	1112.9	3016.1	1593.1	1112.7	503.86	508.55
Median	908.75	2977.9	1879.5	504.5	448.16	371.04
Standard Deviation	1078.2	136.73	705.37	1274.7	256.27	336.61
7 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Best	1649.5	3913.6	2965.9	1611.9	1243.2	639.06
Worst	4466.3	5258.5	5528.1	3131.8	1890.4	2276.5
Mean	2810.7	4564.5	4226.4	2054.4	1554.5	1520.5
Median	2331.5	4433.4	4624.3	1665.6	1529.6	1671.6
Standard Deviation	1115.6	547.48	1105.4	653.3	241.68	633.17
10 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Best	1766.8	6460.2	3338.5	1589.8	1354.3	981.01
Worst	8587.3	7839.5	10148	7949.2	2605.4	1592.7
Mean	4347.5	7081.5	6162.4	3199.6	1856.2	1341.3
Median	2911.8	6902.7	5522.9	2109.6	1918.8	1343.7
Standard Deviation	3009.8	526.14	2513.7	2671.4	521	240.38
12 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Best	3171.3	8769.2	6490.3	3178.4	3328.5	2949.9
Worst	6908.3	10365	8138.6	4932.2	5196	4648.4
Mean	4983.1	9747.5	7510	4387.4	3905.2	3759.4
Median	5339.7	10112	7692.7	4624.3	3705.2	3541.2
Standard Deviation	1412.3	664.75	629.66	701.99	744.92	683.51
15 pairs of nodes						
Metrics	WOA(Mirjalili & Lewisa, 2016)	BBO(Seyed Habib, 2012)	DA(Mirjalili, 2016)	FW-DA(Kumar, n.d.a)	T-LAU(Kumar, n.d.b)	BO-FBU
Best	8090.7	10036	12598	8023.4	5589.9	4268.8
Worst	17611	12147	20746	18980	6962.7	5532.6
Mean	12446	11348	16629	13444	6019.1	5045.1
Median	11513	11301	15033	14178	5801.2	5028.1
Standard Deviation	4315.1	874.05	3725.8	4478.8	550.55	524.59

overall iterations and it is 83.3%, 85.7%, 80%, 50% and 44.4% better than the traditional models like WOA, BBO, DA, FW-DA and T-LAU, respectively at 100th iteration. As a part of the future work the proposed BO-FBU optimisation algorithm can be deployed in detecting the anomaly behaviour of vehicle users under different scenarios like urban and highway as it involves human lives at more care must be taken in identifying the anomaly behaviours of the vehicle users.

REFERENCES

Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017, November 1). Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, *262*, 134–147. doi:10.1016/j.neucom.2017.04.070

Al-Jadir, I. (2018). Enhancing digital forensic analysis using memetic algorithm feature selection method for document clustering. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. doi:10.1109/SMC.2018.00621

Alazab, M., Huda, S., Abawajy, J., Islam, R., Yearwood, J., Venkatraman, S., & Broadhurst, R. (2014). A hybrid wrapper-filter approach for malware detection.*Journal of Networks,9*(11), 2878-2891.

Albukhitan, S., Alnazer, A., & Helmy, T. (2019). Semantic Web Annotation using Deep Learning with Arabic Morphology. *Procedia Computer Science*, *151*, 385–392. doi:10.1016/j.procs.2019.04.053

Amin, R., Islam, S. H., Biswas, G. P., Khan, M. K., & Kumar, N. (2015). An efficient and practical smart card based anonymity preserving user authentication scheme for TMIS using elliptic curve cryptography. *Journal of Medical Systems*, *39*(11), 180. doi:10.1007/s10916-015-0351-y PMID:26433889

Antonello, M., Chiesurin, S., & Ghidoni, S. (2020, April). Enhancing semantic segmentation with detection priors and iterated graph cuts for robotics. *Engineering Applications of Artificial Intelligence*, *90*, 103467. doi:10.1016/j.engappai.2019.103467

Assi, A., Mcheick, H., Karawash, A., & Dhifli, W. (2019, December). Context-aware instance matching through graph embedding in lexical semantic space. *Knowledge-Based Systems*, *186*, 104925. doi:10.1016/j.knosys.2019.104925

Azab, A., Layton, R., Alazab, M., & Oliver, J. (2014, November). Mining malware to detect variants. In*2014 Fifth Cybercrime and Trustworthy Computing Conference*(pp. 44-53). IEEE. doi:10.1109/CTC.2014.11

Bali, R. S., & Kumar, N. (2016). Secure clustering for efficient data dissemination in vehicular cyber–physical systems. *Future Generation Computer Systems*, *56*, 476–492. doi:10.1016/j.future.2015.09.004

Bhattacharya, S., Kaluri, R., Singh, S., Alazab, M., & Tariq, U. (2020). A Novel PCA-Firefly based XGBoost classification model for Intrusion Detection in Networks using GPU. *Electronics (Basel)*, *9*(2), 219. doi:10.3390/electronics9020219

Bhattacharya, S., Somayaji, S. R. K., Gadekallu, T. R., Alazab, M., & Maddikunta, P. K. R. (2020). A review on deep learning for future smart cities.*Internet Technology Letters*, e187.

Bounhas, Soudani, & Slimani. (2019). Building a morpho-semantic knowledge graph for Arabic information retrieval. *Information Processing & Management*.

Ch, R., Gadekallu, T. R., Abidi, M. H., & Al-Ahmari, A. (2020). Computational System to Classify Cyber Crime Offenses Using Machine Learning. *Sustainability*, *12*(10), 4087. doi:10.3390/su12104087

Challa, S., Das, A. K., Odelu, V., Kumar, N., Kumari, S., Khan, M. K., & Vasilakos, A. V. (2018). An efficient ECC-based provably secure three-factor user authentication and key agreement protocol for wireless healthcare sensor networks. *Computers & Electrical Engineering*, *69*, 534–554. doi:10.1016/j.compeleceng.2017.08.003

Chen, J., Chen, Y., Zhang, X., Du, X., & Wen, J.-R. (2018, October). Entity set expansion with semantic features of knowledge graphs. *Journal of Web Semantics*, *52-53*, 33–44. doi:10.1016/j.websem.2018.09.001

Cheng, T., & Wang, B. (2020, January). Graph and Total Variation Regularized Low-Rank Representation for Hyperspectral Anomaly Detection. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(1), 391–406. doi:10.1109/TGRS.2019.2936609

Cui, M., Wang, J., Florita, A. R., & Zhang, Y. (2019, September). Generalized Graph Laplacian Based Anomaly Detection for Spatiotemporal MicroPMU Data. *IEEE Transactions on Power Systems*, *34*(5), 3960–3963. doi:10.1109/TPWRS.2019.2917586

Etaiwi, W., & Awajan, A. (2020, May). Graph-based Arabic text semantic representation. *Information Processing & Management*, *57*(3), 102183. doi:10.1016/j.ipm.2019.102183

Farag, Abdelkader, & Salem. (2019). Parallel graph-based anomaly detection technique for sequential data. *Journal of King Saud University - Computer and Information Sciences in Communication*.

George & Rajakumar. (2013). APOGA: An Adaptive Population Pool Size based Genetic Algorithm. *AASRI Procedia - 2013 AASRI Conference on Intelligent Systems and Control (ISC 2013), 4*, 288-296. doi:10.1016/j.aasri.2013.10.043

Guesmi, S., Trabelsi, C., & Latiri, C. (2016). CoMRing: A Framework for Community Detection Based on Multi-relational Querying Exploration. *Procedia Computer Science*, *96*, 627–636. doi:10.1016/j.procs.2016.08.244

Gupta, B. B., & Sheng, Q. Z. (Eds.). (2019). *Machine learning for computer and cyber security: principle, algorithms, and practices*. CRC PRESS. doi:10.1201/9780429504044

He, D., Kumar, N., & Lee, J. H. (2016). Privacy-preserving data aggregation scheme against internal attackers in smart grids. *Wireless Networks*, *22*(2), 491–502. doi:10.1007/s11276-015-0983-3

He, D., Ma, M., Zeadally, S., Kumar, N., & Liang, K. (2017). Certificateless public key authenticated encryption with keyword search for industrial internet of things. *IEEE Transactions on Industrial Informatics*, *14*(8), 3618–3627. doi:10.1109/TII.2017.2771382

Iwendi, C., Jalil, Z., Javed, A. R., Reddy G, T., Kaluri, R., Srivastava, G., & Jo, O. (2020). KeySplitWatermark: Zero Watermarking Algorithm for Software Protection Against Cyber-Attacks. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 72650–72660. doi:10.1109/ACCESS.2020.2988160

Javed, A. R., Beg, M. O., & Asim, M. (2020). *AlphaLogger: detecting motion-based side-channel attack using smartphone keystrokes*. J Ambient Intell Human Comput. doi:10.1007/s12652-020-01770-0

Javed, A. R., Usman, M., Rehman, S. U., Khan, M. U., & Haghighi, M. S. (n.d.). Anomaly Detection in Automated Vehicles Using Multistage Attention-Based Convolutional Neural Network. IEEE Transactions on Intelligent Transportation Systems. doi:10.1109/TITS.2020.3025875

Kumar. (n.d.a). *Ternary-based Feature Level Extraction for Anomaly Detection in Semantic Graphs: An Optimal Feature Selection Basis*. Academic Press.

Kumar. (n.d.b). *Lion plus Firefly Algorithm for Ternary based Anomaly Detection in Semantic Graphs.* Academic Press.

Lampropoulos, Keramopoulos, & Diamantaras. (2020). Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: A review. *Visual Informatics*.

Lhioui, C., Zouaghi, A., & Zrigui, M. (2017). A Rule-based Semantic Frame Annotation of Arabic Speech Turns for Automatic Dialogue Analysis. *Procedia Computer Science*, *117*, 46–54. doi:10.1016/j.procs.2017.10.093

Lin, S., & Chalupsky, H. (2008, August). Discovering and Explaining Abnormal Nodes in Semantic Graphs. *IEEE Transactions on Knowledge and Data Engineering*, *20*(8), 1039–1052. doi:10.1109/TKDE.2007.190691

Lugowski, A., Kamil, S., Buluç, A., Williams, S., & Gilbert, J. R. (2015, May). Parallel processing of filtered queries in attributed semantic graphs. *Journal of Parallel and Distributed Computing*, *79–80*, 115–131. doi:10.1016/j.jpdc.2014.08.010

Mirjalili, S. (2016, May). Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Computing & Applications*, *27*(4), 1053–1073. doi:10.1007/s00521-015-1920-1

Mirjalili, S., & Lewisa, A. (2016, May). The Whale Optimization Algorithm. *Advances in Engineering Software*, *95*, 51–67. doi:10.1016/j.advengsoft.2016.01.008

Mittal, M., Iwendi, C., Khan, S., & Rehman Javed, A. (2020). Analysis of security and energy efficiency for shortest route discovery in low-energy adaptive clustering hierarchy protocol using Levenberg-Marquardt neural network and gated recurrent unit for intrusion detection system. *Transactions on Emerging Telecommunications Technologies*, 3997. doi:10.1002/ett.3997

Numan, M., Subhan, F., Khan, W. Z., Hakak, S., Haider, S., Reddy, G. T., Jolfaei, A., & Alazab, M. (2020). A systematic review on clone node detection in static wireless sensor networks. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 65450–65461. doi:10.1109/ACCESS.2020.2983091

Qin, Z. (2018). Knowledge-Graph Based Multi-Target Deep-Learning Models for Train Anomaly Detection. *2018 International Conference on Intelligent Rail Transportation (ICIRT),* 1-5. doi:10.1109/ICIRT.2018.8641687

Rajakumar & George. (2012). A New Adaptive Mutation Technique for Genetic Algorithm. *Proceedings of IEEE International Conference on Computational Intelligence and Computing Research (ICCIC),* 1-7. 10.1109/ICCIC.2012.6510293

Rajakumar, B. R. (2013a). Impact of Static and Adaptive Mutation Techniques on Genetic Algorithm. *International Journal of Hybrid Intelligent Systems*, *10*(1), 11–22. doi:10.3233/HIS-120161

Rajakumar, B. R. (2013b). Static and Adaptive Mutation Techniques for Genetic algorithm: A Systematic Comparative Analysis. *International Journal on Computer Science and Engineering*, *8*(2), 180–193. doi:10.1504/IJCSE.2013.053087

Rehman Javed, A., Jalil, Z., Atif Moqurrab, S., Abbas, S., & Liu, X. (2020). Ensemble adaboost classifier for accurate and fast detection of botnet attacks in connected vehicles. *Transactions on Emerging Telecommunications Technologies*, 4088.

RM, , S. PMaddikunta, , P. K. RParimala, , MKoppu, , SReddy, , TChowdhary, , C. LAlazab, , M. (2020). An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. *Computer Communications*.

Sadreazami, H., Mohammadi, A., Asif, A., & Plataniotis, K. N. (2018, March). Distributed-Graph-Based Statistical Approach for Intrusion Detection in Cyber-Physical Systems. *IEEE Transactions on Signal and Information Processing Over Networks*, *4*(1), 137–147. doi:10.1109/TSIPN.2017.2749976

Seyed Habib, A. (2012, February). A new biogeography-based optimization (BBO) algorithm for the flexible job shop scheduling problem. *International Journal of Advanced Manufacturing Technology*, *58*(9–12), 1115–1129.

Sun, Huo, Huan, & Vitter. (2020). Feature reduction based on semantic similarity for graph classification. *Neurocomputing in Communication, 397*, 114-126.

Swamy, S. M., Rajakumar, B. R., & Valarmathi, I. R. (2013). Design of Hybrid Wind and Photovoltaic Power System using Opposition-based Genetic Algorithm with Cauchy Mutation. *IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013).* doi:10.1049/ic.2013.0361

Vela, A. P., Ruiz, M., & Velasco, L. (2017, July). Distributing data analytics for efficient multiple traffic anomalies detection. *Computer Communications*, *107*, 1–12, 15. doi:10.1016/j.comcom.2017.03.008

Vidal, J. C., Lama, M., Otero-García, E., & Bugarín, A. (2014, January). Graph-based semantic annotation for enriching educational content with linked data. *Knowledge-Based Systems*, *55*, 29–42. doi:10.1016/j.knosys.2013.10.007

Vlietstra, W. J., Zielman, R., van Dongen, R. M., Schultes, E. A., & Kors, J. A. (2017, July). Automated extraction of potential migraine biomarkers using a semantic graph. *Journal of Biomedical Informatics*, *71*, 178–189. doi:10.1016/j.jbi.2017.05.018 PMID:28579531

Wang, S., Li, X., Ye, Y., Huang, X., & Li, Y. (2018, January 24). Multi-attribute and relational learning via hypergraph regularized generative model. *Neurocomputing*, *274*, 115–124. doi:10.1016/j.neucom.2016.06.003

Yao, Y., Wang, Z., Gan, C., Kang, Q., & Zhang, L. (2016, October 5). Multi-source alert data understanding for security semantic discovery based on rough set theory. *Neurocomputing*, *208*, 39–45. doi:10.1016/j.neucom.2015.12.127

Zhang, J., Tao, T., Mu, Y., Sun, H., & Wang, Z. (2019, May). Web image annotation based on Tri-relational Graph and semantic context analysis. *Engineering Applications of Artificial Intelligence*, *81*, 313–322. doi:10.1016/j.engappai.2019.02.018

Zhao, Q., Kang, Y., Li, J., & Wang, D. (2018, October). Exploiting the semantic graph for the representation and retrieval of medical documents. *Computers in Biology and Medicine*, *101*, 39–50. doi:10.1016/j.compbiomed.2018.08.009 PMID:30099238

APPENDIX

Table 4. Nomenclature

Abbreviation	Description
TG	Tri-Relational Graph
GGL	Generalized Graph Laplacian
PMU	Phasor Measurement Unit
μPMU	Distribution-Level Phasor Measurement Unit
GTVLRR	Graph And TV Regularized LRR
TV	Total Variation
MST	Minimum Spanning Tree
FW-DA	Fitness Weighed - Dragonfly Algorithm
T-LAU	Threshold Based LA Update
BBO	Biogeography Optimization
WOA	Whale Optimization Algorithm
BO-FBU	Biogeography Optimization With Fitness Sorted Update
MRN	Multi-Relational Networks
SNA	Social Network Analysis

M. Sravan Kumar Reddy is pursuing hisPh.D.form SITE, Vellore Institute of Technology, Vellore, India. He Received M.Tech in SE from JNTU, Hyderabad, India, in 2010 and B.Tech Degree in CSIT from JNTU, Anantapur, India. His research areas are Data Mining and Big Data Analytics. He has published 8 reputed Journal Papers and 2 papers presented in reputed international conference. He is having an experience of 11 years in teaching and 3 years in Research. He is the Life member of ISTE and presently working as Assistant Professor in CSE Dept at R.G.M.C.E.T, Nandyal, India.

Dharmendra Singh Rajput (PhD) is working as Associate Professor in Department of Software and Systems Engineering, SITE, VIT, Vellore, India. He Completed PhD (Jan 2015) from NIT, Bhopal, India. His research areas are Data Mining and Big Data Predictive Analytics. He has published 15+ reputed Journal Papers and 17 papers presented in reputed international conference. He has received various awards from Indian Government like DST-SERB, CSIR Travel Grant and MPCST Young Scientist Fellowship. He visited various countries UK, France, Singapore, UAE, China, and Malaysia for academic purpose.