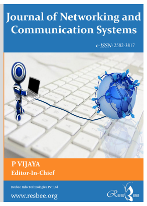
Vol.5 No.2 April 2022

Journal of Networking and Communication Systems

Received 17 December, Revised 26 February, Accepted 28 March



Adaptive Whale Algorithm: Big data Classification in IoT Networks

Aravinth J S
Department of Cyber Security, Dublin Business School,
13/14 Aungier St, Dublin 2, D02 WC04, Ireland

Abstract: In the Internet of Things (IoT)-enabled networks, because of the open nature of wireless data transmission, data security and routing faces a significant issue. Nevertheless, characteristic features, such as heterogeneity, constrained resources, scalability requirement, and uncontrolled environment, construct the security problems highly difficult. Therefore, to augment energy effectiveness as well as nodes lifetime, an efficient, as well as secure routing protocol called Enhanced Energy Harvesting Trust Aware Routing Algorithm (Enhanced-EHTARA), is presented. Nevertheless, developed EHTARA is modeled by using Link Lifetime (LLT) with conventional EHTARA. Using adopted enhanced-EHTARA, the optimal secure routing path is effectually selected by cost metric that takes into consideration the metrics, namely energy, LLT, delay, as well as trust. Nevertheless, by the MapReduce framework, the process of big data classification is done at BS. As a result, the Adaptive Whale Algorithm trains big data classification in progress by the stacked autoencoder. Finally, the adopted enhanced-EHTARA reveals superior performance which attains maximum energy.

Keywords: Base station, IoT, Energy, Routing Algorithm, Security.

Nomenclature

Abbreviations	Descriptions
SEAT-DSR	Spatial and Energy-Aware Trusted Dynamic Distance Source Routing
IoT	Internet of Things
BS	Base Station
QoS	Quality of Service
ROLL	Routing Over Low power and Lossy networks
RFID	Radio-Frequency Identification
Enhanced-EHTARA	Enhanced Energy Harvesting Trust Aware Routing Algorithm
ATSR	Ambient Trust Sensor Routing
LQSR	Link Quality Source Routing
IETF	Internet Engineering Task Force
TLSRP	Trust dependent Link State Routing Protocol
LLNs	Low-power and Lossy Networks
MRTS	Metric-based RPL Trustworthiness Scheme
RPL	Routing Protocol for Lowpower and Lossy Networks
CHS	Cluster Head Selection
TARF	Trust-Aware Routing Framework
LEACH	Low Energy Adaptive Clustering Hierarchy
SN	Sensor Nodes
WSN	Wireless Sensor Networks
Taylor C-SSA	Taylor Cat Salp Swarm Algorithm
SecTrust	Secure Trust
EOSR	Energy Optimized Secure Routing

https://doi.org/10.46253/jnacs.v5i2.a4

1. Introduction

IoT is an innovative communication model which affects the daily lives in many domains, namely home, healthcare, building automation, urban, automobile, and industrial applications [15]. IoT-based networks are possibly created of LLNs that are collected from several heterogeneous wireless technologies (objects), namely sensors, RFID tags, actuators, and so on. In these technologies, communication, as well as computing systems, is effortlessly embedded. IoT objects are characterized both by their strong resource constraints and by their lossy communication connections [17]. Certainly, these objects have restricted memory, processing power, and energy providers, as well as a high loss rate, an imperfect frame size, a low throughput, and diminutive communication ranges [20]. Such restrictions arise in various confronts for the industry and academic research community, such as scalability, security, and routing [1].

In WSN, The routing models are significant as they present minimum energy utilization, QoS, latency, as well as data throughput [2]. To address the issues caused when routing data packets, several protocols are formulated because WSN is application-oriented. In WSN, the conventional protocols resolve the energy problems which range from physical to application layer. For routing the data, various protocols were formulated which start from the primary node to the target node including LQSR, and are taken into for routing data packets are called conventional routing in WSN. A multipath routing protocol called TARF is modeled by calculating the dependability of neighboring nodes in WSN [13]. The protocol avoids unreliable nodes as well as carries out the routing based on trust evaluation and energy efficiency. By exploiting the ATSR, the distributed approach was modeled to compute the reliability of the node. Here, to monitor the neighbor's activities, every node is used based on the particular criterion of trust and calculates the direct trust value by exploiting the neighboring nodes. The TLSRP approach was modeled for multi-hop routing based on direct and indirect trust. The multi-hop routing protocols are on basis of two diverse classifications, such as data-centric routing and Location routing protocols. The data-centric routing exploits the base station to transmit queries to exact areas. Similarly, hierarchical routing is exploited to maintain the energy utilization between the SNs by exploiting multi-hop communication to reduce transmitted messages to BS. For sensor networks, diverse routing protocols are modeled that require the position information of nodes for further processing [14]. Moreover, position information is exploited to compute the distance between two nodes as well as calculate utilized energy [3]. Over the past decade, for LLNs, various routing solutions were suggested. At last, the IETF ROLL working group modeled and standardized the RPL. For the IoT [18], the main problem is the routing's security which researchers have taken into consideration as a dangerous requirement [16]. Although the RPL specification states cryptography-based models to assure control messages integrity and confidentiality over outsider attackers, RPL is still susceptible to diverse known and novel internal threats, which were broadly studied in the state of the art [5].

The major aim of this study is to work on an energy harvesting and secure routing protocol called enhanced-EHTARA, which is adopted to discover the optimal routing paths to transfer the data packets to BS. Nevertheless, by utilizing the cost metric function, the routing path is optimally started which takes into consideration the metrics such as delay, LLT, energy, and trust. In addition, big data classification is done by stacked autoencoder by encoding and decoding model so that it minimizes the computational time and complexity more precisely.

2. Literature Survey

In 2019, V. Mythili et al [1] modeled a secure routing model named the SEAT-DSR approach to in improving the lifespan of the network for WSN. Moreover, the level of energy, spatial information, and quality of data efficiency was equalized using the QoS on the basis of the energy-aware routing methods. Besides this technique, a standard clustering approach was integrated to group the WSN on the basis of the spatial information, trust score, energy level as well as distance amid nodes. In addition, a novel hierarchical trust model was developed in this technique that uses multi-attributes of a lot of WSNs consistent with the data communication speed, energy consumption, data size, and suggestion. In 2019, A. Vinitha et al [2], recognized the energy problem and develops an energy-efficient multi-hop routing in WSN referred Taylor C-SSA. It was the integration of the C-SSA with Taylor series. This technique encountered two stages to accomplish the multi-hop routing consisting of the CHS, and data transmission. Initially, using LEACH protocol, energy-competent CHS was performed for effective data transmission, SNs transmit data over CH that transfers data to BS through the elected optimal hop. By exploiting Taylor C-SSA, the optimal hop selection was done. Additionally, security-aware multi-hop routing was carried out using a developing trust model that involves the integrity factor, direct trust, indirect trust, as well as data forwarding rate. In 2020, Nabil Djedjig et al [3], worked on the resource-constrained nature of IoT objects that create the RPL susceptible to diverse attacks. To control messages although RPL specification presents encryption security, RPL was still vulnerable to internal attackers and self-centered behaviors. To cope with the lack of a robust security model in RPL, a novel MRTS model was designed which develops trust assessment for secure routing topology model. In 2018, David Airehrour et al [4], worked on the RPL, in effect routing protocol for IoT presents small protection over several types of routing attacks. An attacker can use the routing system of RPL to open devastating and critical attacks over an IoT network. Well-liked amongst these IoT attacks were Sybil as well as Rank attacks. A time-based trust-aware RPL routing protocol SecTrust-RPL was developed and examined to secure IoT networks from routing attacks. To present protection over Sybil as well as Rank attacks SecTrust trust system was entrenched into the RPL routing protocol. In 2018, Tao Yang et al [5], developed an EOSR on the basis of the distributed trust estimate recognize to recognize and segregate Malicious nodes. A multi-factor routing scheme was developed for the EOSR routing protocol and the node's trust level, residual energy, and path length. This scheme not only assures that data was transferred during the trusted nodes but also balances energy utilization between trusted nodes.

3. Description of IoT Model

IoT networks consist of various objects namely smart devices which are connected with each other to transmit data packets via a network [19]. In the IoT model, many network devices are exploited as resource-constrained which carries out the data processing model as well as communication procedure to transmit data amid two entities. Moreover, various nodes are experimented within the network to transfer data packets to BS by the optimal path. Nevertheless, the source node transfers data packets to BS by examining the path with maximum energy nodes so that node needs a maximum number of energy to ahead packets with other nodes. The nodes can be represented as routers or sensors in IoT networks. The SN produces data packets as well as forwards them to BS. Nevertheless, to forward packets to BS, the router helps SN. In mobility scenarios, routing is necessary to transmit that data to BS.

Consider $B = \{p_1, p_2, \ldots p_c, \ldots p_r\}$ as network model and indicates the set of nodes possessing r count of IoT nodes so that $1 \le c \le \ell$, correspondingly. Nevertheless, IoT nodes transfer data to BS A which is positioned at $(0.5T_i, 0.5Y_i)$. For that reason, connection exploited to link network nodes is indicated as k and sender, as well as receiver node of IoT network, is represented as P as well as Q. By exploiting the routing model, the optimal path employed to route data from sender to receiver is turned on. In addition, the routing process is attained by exploiting factors, namely the process of energy harvesting, trust model, energy approach, as well as LLT approach.

3.1 Energy Model

To calculate the energy needed the energy approach [6] is exploited by a node to broadcast data packets. It is considered significant to state energy utilized by nodes to model energy-competent routing protocol. Nevertheless, the node needs a significant number of energy to transmit, receive and forward the data packets in the chosen optimal path. Though, energy utilized by a node c on link $k(c,s) \in D$ to procedure the data packet is indicated as,

$$D_S^c = D_b^c + D_e^c + D_u^c + D_v^c \tag{1}$$

wherein, D_b^c signifies energy utilized by the node c at pay attention period, D_e^c simplifies energy utilized at receiving period, D_e^c represents energy utilized at the transferring period, and D_v^c simplifies energy utilized at the sleeping period.

$$D_S^c = \left(x_b^c J_b + (J_e + J_u)\frac{\tau}{\delta} + x_v^c J_v \right)\lambda \tag{2}$$

wherein, x_v^c simplifies the duration by the node at sleeping mode, J_e, J_b, J_u, as well as J_v simplifies current drawn throughout transmission, listening, receiving, and sleeping mode, “and x_b^c simplifies time taken by the node at listing mode. Moreover, λ simplifies nodes' battery voltage” τ simplifies packet length which is indicated in bits, as well as δ simplifies the data rate that is indicated in Kbps.

Nevertheless, x_v^c and x_b^c is denoted by exploiting the following formulation as,

$$x_v^c = l - w \tag{3}$$

$$x_b^c = l - \left(x_e^c + x_u^c + x_v^c\right) \tag{4}$$

wherein, l denotes beacon interval, w denotes superframe time period, x_u^c denotes the time exploited to receive data packets, and x_e^c denotes the time required to transmit packets, correspondingly.

Assume $D_u^c = 0$ or $x_u^c = 0$ if c indicates source node, as well as $D_e^c = 0$ or $x_e^c = 0$ if c indicates receiver node, subsequently eq. (2) is reformulated as,

$$D_S^c = \begin{cases} \left(\left(x_b^c J_b + J_e \dfrac{\tau}{\delta} + x_v^c J_v\right)\lambda \right. & ; \text{if } c \text{ is sender node} \\ \left(\left(x_b^c J_b + J_u \dfrac{\tau}{\delta} + x_v^c J_v\right)\lambda \right. & ; \text{if } c \text{ is receiver node} \end{cases} \tag{5}$$

At last, the remaining node energy c is indicated as,

$$D_{resdl}^c = D_h^e - D_S^c + D_m^c \tag{6}$$

wherein, D_m^c signifies harvested energy, D_h^c signifies initial energy, D_S^c implies utilized energy, and D_{resdl}^c signifies remaining energy.

3.2 Energy Harvesting Process

By exploiting node battery levels the process of Energy harvesting [6] is indicated. Moreover, the battery of the node is divided into 2 areas and 3 levels, such as level-1, 2, and 3. Nevertheless, level-1 indicates utmost energy level; level-2 indicates reasonable energy level, as well as level-3, implies minimum energy level. The minimum energy level is exploited to guarantee that node can function in 4 operating modes, such as idle listening, sleeping, as well as transmitting, and receiving mode. Moreover, three diverse steps are examined in the energy harvesting procedure that is described as below:

Step: 1: $level\_2 < D_{resdl}^c \le level\_1$

The nodes remaining energy has to be maximum to preserve their operations. Therefore, it is not needed to identify the energy utilization procedure in this scenario.

Step: 2: $level\_3 < D_{resdl}^c \le level\_2$

In this scenario, the sleeping mode is used to facilitate nodes to harvest extra energy. Moreover, the node minimizes energy utilization as well as from ambient energy sources it harvests a smaller amount of energy. Therefore, a node can augment the lifespan of the network. Nevertheless, at sleeping mode x_v^c , the energy harvested using node c is indicated as,

$$D_m^c = x_m^c L_m^c = x_v^c L_m^c \tag{7}$$

wherein, x_m^c indicates the harvesting time and L_m^c indicates the energy harvesting rate.

Step: 3: $0 < D_{resdl}^c \le level\_3$

In this scenario, the node does not possess sufficient energy to preserve normal operations. Consequently, nodes' ramp transceiver mode in nodes can minimize utilized energy as well as increase node lifetime. Moreover, to improve the node back-off period, the energy back-off procedure is exploited. Hence, to increase the lifespan of a network, the process of energy utilization exploits the energy back-off period. To carry out the energy back-off procedure, the time required is indicated as,

$$x_{energyboff}^c = \frac{D_{S,cur}^c}{L_{m,pre}^c} \tag{8}$$

wherein, $D_{S,cur}^c$ represents u node energy c at present beacon interval as well as $L_{m,pre}^c$ represents the rate of energy harvesting at preceding beacon interval.

Nevertheless, node c harvesting time, and number of harvested energy at c is indicated as,

$$x_m^c = x_{back\ off}^c + x_v^c \tag{9}$$

$$D_m^c = x_m^c L_m^c = \left(x_{back\ off}^c + x_v^c\right) L_m^c \tag{10}$$

From Eq. (7) as well as Eq. (10), energy harvested is indicated as,

$$D_m^c = \begin{cases} x_v^c L_m^c & \text{if level} - 3 \le D_{resdl}^c \le \text{level} - 2 \\ \left(x_{back\,off}^c + x_v^c\right) L_m^c & \text{if } 0 < D_{resdl}^c < \text{level} - 3 \end{cases} \tag{11}$$

3.3 Link Lifetime Approach

Because of the dynamic topology structure, a node must calculate the lifespan in the IoT network [7]. Let the nodes as c well as s which lies within the transmission of the data range. Nevertheless, node LLT is calculated as,

$$H_{cs} = \frac{-\left(ly + KZ + \sqrt{\left(l^2 + K^2\right)d^2 - \left(lZ - Ky\right)^2}\right)}{\left(l^2 + K^2\right)} \tag{12}$$

wherein, (O_c, T_c) implies node coordinate c, M_c implies node mobility speed c, (O_s, T_s) implies the node coordinate s, θ_s implies the motion direction of node s, M_s implies s node mobility speed, θ_c implies the motion direction of node c, as well as d implies transmission range.

3.4 Trust Model

Trust factors [8] [9] present security for the adopted routing model to set up trust-based secure routing. The trust is calculated to measure the dependability of nodes and it is computed as,

$$X_{c,s}(f) = XD_{c,s}(f) + XI_{c,s}(f) + XE_{c,s}(f) + XA_{c,s}(f) \tag{13}$$

wherein, $XD_{c,s}(f)$ indicates node direct trust c on the node s at a time f, $XE_{c,s}(f)$ indicates node integrity factor, $XI_{c,s}(f)$ indicates node indirect trust, as well as $XA_{c,s}(f)$ indicates node availability factor, correspondingly.

4. Proposed Model for Big Data Classification

Exploiting trustworthy nodes by means of energy harvesting procedures maximizes energy effectiveness as well as network lifetime to set up secure routing. In this paper, an enhanced-EHTARA is developed to transmit data steadily to BS by exploiting the energy harvesting procedure. Nevertheless, by the MapReduce framework at BS "big data classification" is performed. At first, the IoT nodes are distributed to find a secure route to transfer data to BS in the IoT network. By routing algorithm it is extremely essential to identify the optimal path to transmit data safely. Moreover, the routing process is performed by exploiting the enhanced-EHTARA. The developed enhanced EHTARA identify the optimal path by exploiting cost metric that is regarded as factors, such as trust, LLT, energy, as well as delay. The data packets are transferred to BS with the secure route, whereas "big data classification is attained by exploiting MapReduce framework" with an optimization algorithm stacked autoencoder [20]. Nevertheless, by optimization approach, the feature selection procedure is performed at the mapper stage, and at the reducer stage, "big data classification is developed by exploiting a stacked autoencoder" [20]. Fig. 1 demonstrates the architectural model of the secure routing approach.

4.1 Adopted Modified EHTARA

The major aim of adopted model is to identify an optimal path for routing data packets to BS. Nevertheless, nodes with the least cost as well as utmost energy level are taken into consideration to ascertain the secure path to transmit the data to other nodes. The adopted enhanced EHTARA identifies the optimal path by exploiting cost metric that may exploit the metrics such as, LLT energy, delay, as well as trust.

Cost metric: The adopted technique exploits two diverse measures, such as node cost Y_c and link cost $Y_{c,s}$ with energy harvesting procedure. For link the cost metric is calculated by exploiting the link distance as well as the energy harvesting process as stated below:

$$Y_{c,s} = \begin{cases} \frac{U_{c,s}}{4}\left(\frac{D_S^c}{D_{resdl}^s} + X_{c,s} + E_{c,s} + H_{cs}\right) & \text{for case} - 2 \\ \frac{U_{c,s}}{4}\left(\frac{D_S^c}{D_{resdl}^s} + \log_2\left(\frac{x_{back\,off}^s}{\chi}\right) + X_{c,s} + E_{c,s} + H_{cs}\right) & \text{for case} - 3 \end{cases} \tag{14}$$

wherein, D_{resdl}^s implies remaining energy of node s, $U_{c,s}$ implies geographical distance amid the node c as well as s, D_S^c implies energy utilized by node c, χ implies unit back-off period, and H_{cs} implies LLT. The secure route is efficiently chosen to transfer data packets to BS by the aforesaid cost metric. Moreover, $E_{c,s}$ implies end-to-end delay as well as it is stated as below,

$$E_{c,s} = \frac{I}{p_r} \tag{15}$$

wherein, p_r implies the total count of nodes in the network, and I implies total count of nodes chosen in route to attain routing process. Nevertheless, the node cost Y_c permits the adopted enhanced-EHTARA to choose the node with utmost remaining energy to attain transmission of data by preserving link availability to secure routing. Therefore, node cost metric Y_c is calculated as,

$$Y_c = \frac{1}{D_{resdl}^s} \tag{16}$$

wherein, D_{resdl}^s implies remaining node energy c. The adopted enhanced EHTARA includes 2 diverse stages, such as forwarder selection and link selection to attain effectual routing.

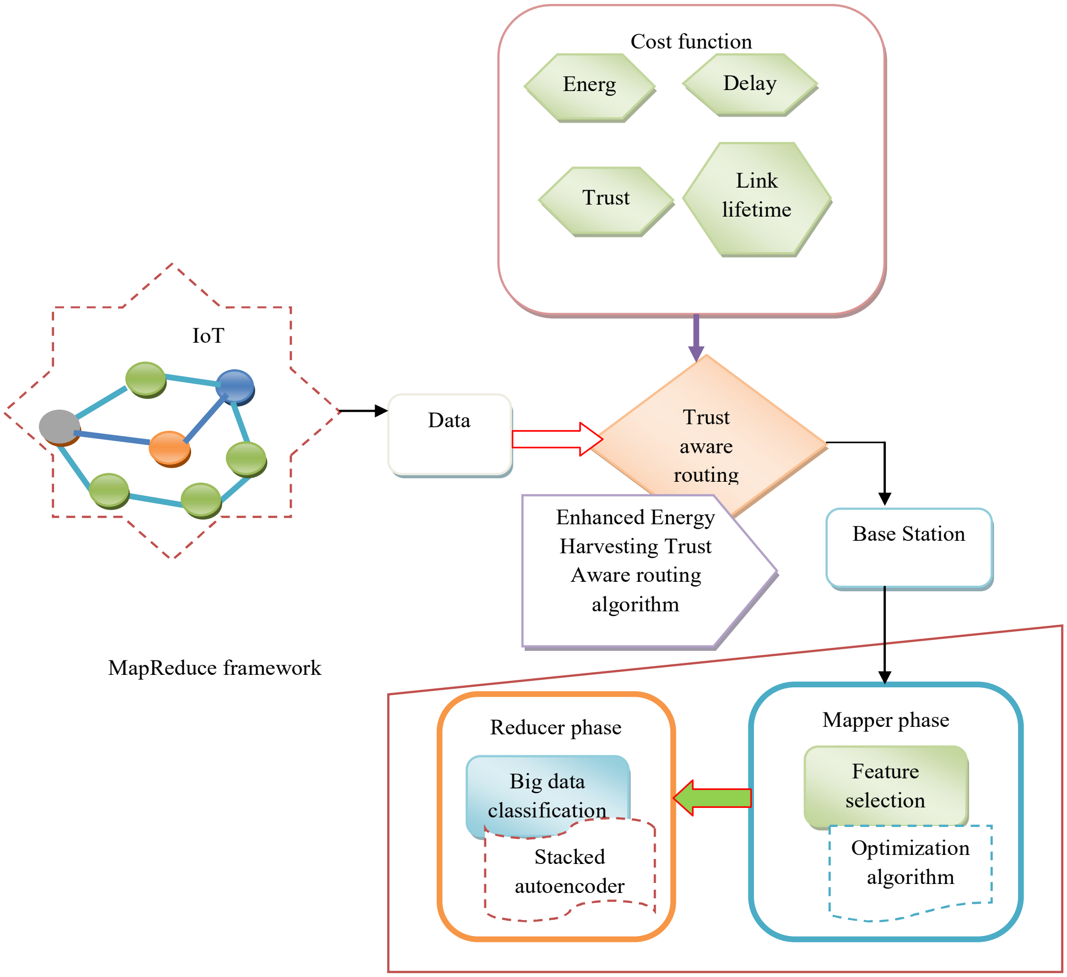


***Fig. 1** Architectural diagram of secure routing model by the proposed method*

5. MapReduce Framework

Let the data "which is being routed from a sender node c to BS as A that is input data exploited to carry out the big data classification". Nevertheless, by exploiting the MapReduce framework, the BS carries out "big data classification". Big data classification is important to permit simple as well as effectual data analysis from several data sources. To carry out the big data classification, numerous techniques are used; however, the conventional techniques undergo complexity problems because of extra data and missing attributes. This paper is mainly concentrated on carrying out "big data classification" by exploiting the MapReduce framework to minimize computational time as well as to manage the data more efficiently. Moreover, computational complexities are minimized by choosing optimal features by exploiting an optimization algorithm that efficiently minimizes the feature dimensionality. The MapReduce framework comprises 2 stages, as the reducer phase and mapper phase. By exploiting the optimization technique, the feature selection procedure is performed at the mapper phase, as well as big data classification process is done by a stacked autoencoder at the reducer phase.

5.1 Adaptive Whale Algorithm for Feature Selection

By exploiting the input data, the feature selection procedure is carried out at the mapper phase which is being transmitted to BS. The mapper stage uses input data A as well as carries out a feature selection model by exploiting the optimization approach. From the input data by choosing unique and necessary features, data classification performance can be maximized. Nevertheless, features chosen by exploiting the optimization approach are indicated as J with $[\mathrm{U} \times \mathrm{V}]$ dimension that is transmitted as input to the classification module.

5.2 Stacked Autoencoder for Big Data Classification

From the input data, the optimal feature J is chosen; the subsequent step is to carry out the big data classification by exploiting the stacked autoencoder. In the reducer phase, the big data classification process is done by exploiting the stacked autoencoder [10] [11].

5.2.1 Adaptive Whale Approach to Training Process

By exploiting the Adaptive Whale approach "the training process of the stacked autoencoder" is performed that is adopted by combining the adaptive idea with the Whale algorithm [12]. The weight update procedure can be effectually performed to attain superior classification outcomes by combining the features of adaptive ideas.

Presently, the WOA approach research is mostly from 3 stages: the performance of the approach enhancement, the integration of Whale Optimization Algorithm (WOA) approach and other intelligent approaches, as well as solution of practical issues by the WOA approach. In this paper, an adaptive scheme for the WOA is introduced. The WOA primarily comprises three phases such as bubble attack, encircle predation, and prey search.

In a random manner, the whales can search for food. In reality, individual whales search arbitrarily based on each other's location, and the formulation is stated as below:

$$\mathrm{Y}(\mathrm{t}+1)=\mathrm{Y}_{\mathrm{rand}}(\mathrm{t})-\mathrm{A}\,|\,\mathrm{C}\times \mathrm{Y}_{\mathrm{rand}}(\mathrm{t})-\mathrm{Y}(\mathrm{t})\,| \tag{17}$$

wherein, $\mathrm{Y}_{\mathrm{rand}}(\mathrm{t})$ indicates the arbitrarily chosen individual whales' location in the current population.

$$\mathrm{a}=\mathrm{a}_1+\mathrm{a}_2\times\left(\cos\left(\frac{2\pi \mathrm{t}}{\mathrm{t}_{\max}}\right)+\frac{\mathrm{t}}{\mathrm{t}_{\max}}\times\frac{1}{\mathrm{f}_{\mathrm{obj}}^{\max}\left(\mathrm{y}_{\mathrm{i}}^{\mathrm{t}}\right)-\mathrm{f}_{\mathrm{obj}}^{\min}\left(\mathrm{y}_{\mathrm{i}}^{\mathrm{t}}\right)+\zeta}\right) \tag{18}$$

Wherein, $\mathrm{t}_{\max}$ signifies a maximum number of iterations, [2] t signifies the current count of iterations, ζ signifies an arbitrary number among [1], $\mathrm{f}_{\mathrm{obj}}^{\min}$ and $\mathrm{f}_{\mathrm{obj}}^{\max}$ indicates the minimum and maximum values of the fitness value.

6. Experimentation Procedure

In this section, an analysis of the adopted model was discussed regarding energy. For the experimentation, the "heart disease dataset comprises 76 attributes, but only 14 attributes were extensively exploited". Here, the adopted model was evaluated with traditional techniques like LASer, EHARA, centralized routing, and EHTARA.

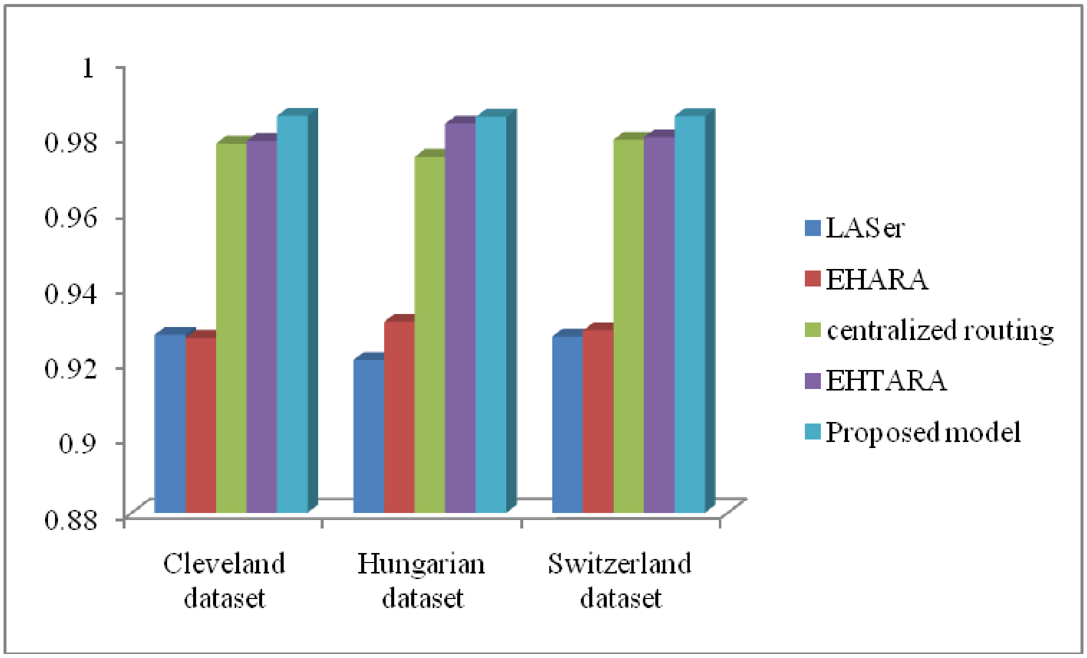


Fig.2. *Analysis of developed and conventional techniques concerning energy for 3 datasets*

Fig. 2 indicates the performance analysis of the adopted technique over traditional techniques concerning energy. While the experimentation time is represented as 5sec, the energy values of conventional models are better than the proposed model for three data sets. Therefore, it is obviously exhibited that the proposed method attained maximum energy for 3 data sets. For Cleveland dataset, the proposed model is 22% superior to LASer, 21% superior to EHARA, 15% superior to centralized routing, and, 11% superior to EHTARA. For the Hungarian dataset, the proposed model is 31% superior to LASer, 29% superior to EHARA, 26% superior to centralized routing, and, 21% superior to EHTARA. For the Switzerland dataset, the proposed model is 22% superior to LASer, 21% superior to EHARA, 28% superior to centralized routing and, 11% superior to EHTARA.

7. Conclusion

An effectual and secure routing protocol was proposed in this paper by exploiting the adopted enhanced EHTARA technique on the basis of the cost metric function. At first, to discover a secure route as well as to transfer data to BS, IoT nodes were distributed in the IoT network. It was very effective to identify the optimal path by the routing model to transfer the data safely. The optimal path selection was performed by exploiting the adopted enhanced-EHTARA that ascertains a secure path on basis of cost metric function. The cost function was taken into consideration the metrics, such as energy, trust, LLT, as well as delay to start the secure routing. By exploiting trust metrics, namely direct trust, integrity factor, indirect trust, and availability factor, the trust model was calculated. "If the data was arrived at the BS, subsequently by MapReduce framework, the process of big data classification was done. For that reason, at the mapper phase, the feature selection process was performed and big data classification was attained at the reducer phase" exploiting the stacked autoencoder that was trained using the Adaptive whale algorithm. In addition, the optimization approach was exploited for the feature selection process. Moreover, the adopted model showed superior performance by attaining maximum energy.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

Reference

[1] V. MythiliA. SureshR. Dhanasekaran, "SEAT-DSR: Spatial and energy aware trusted dynamic distance source routing algorithm for secure data communications in wireless sensor networks", Cognitive Systems, 15 May 2019.

[2] A. VinithaM. S. S. RukminiDhirajsunehra, "Secure and energy aware multi-hop routing protocol in WSN using Taylor-based hybrid optimization algorithm", Journal of King Saud University - Computer and Information Sciences Available online, 21 November 2019.

[3] Nabil DjedjigDjamel TandjaouiImed Romdhani, "Trust-aware and cooperative routing protocol for IoT ", Journal of Information Security and Applications, 28 February 2020.

[4] David AirehrourJairo A. GutierrezSayan Kumar Ray, "SecTrust-RPL: A secure trust-aware RPL routing protocol for Internet of Things", Future Generation Computer Systems, 26 March 2018.

[5] Tao YangXu XiangyangPan Leina, "A secure routing of wireless sensor networks based on trust evaluation model", Procedia Computer Science11 May 2018.

[6] Nguyen, T.D., Khan, J.Y. and Ngo, D.T., “A distributed energy-harvesting-aware routing algorithm for heterogeneous IoT networks”, IEEE Transactions on Green Communications and Networking, vol. 2, no. 4, pp. 1115-1127, 2018.

[7] Chen, Z., He, M., Liang, W. and Chen, K., "Trust-aware and low energy consumption security topology protocol of wireless sensor network," Journal of Sensors, 2015.

[8] Balachandra, M., Prema, K.V. and Makkithaya, K., “Multiconstrained and multipath QoS aware routing protocol for MANETs”, Wireless networks, vol. 20, no. 8, pp. 2395-2408, 2014.

[9] Zhu, J., “Wireless Sensor Network Technology Based on Security Trust Evaluation Model”, International Journal of Online and Biomedical Engineering (iJOE), vol. 14, no. 04, pp. 211-226, 2018.

[10] Liu, G., Bao, H. and Han, B., “A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis”, Mathematical Problems in Engineering, 2018.

[11] S.Md. Mujeeb ; R. Praveen Sam ; K. Madhavi, " Adaptive hybrid optimization enabled stack autoencoder-based MapReduce framework for big data classification", In Proceedings of the International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020.

[12] X. Chen, "Research on New Adaptive Whale Algorithm," in IEEE Access, vol. 8, pp. 90165-90201, 2020.

[13] G.Gokulkumari,"An overview of Big Data Management and its Applications", Journal of Networking and Communication Systems, vol. 3, no.3, July 2020.

[14] Dr.Sivaram Rajeyyagari, "Automatic Speaker Diarization using Deep LSTM in Audio Lecturing of e-Khool Platform",vol. 3, no.4, October 2020.

[15] P. Khatiwada, H. Bhusal, A. Chatterjee and M. W. Gerdes, "A Proposed Access Control-Based Privacy Preservation Model to Share Healthcare Data in Cloud", 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 40-47,2020.

[16] P. Khatiwada, A. Chatterjee and M. Subedi, "Automated Human Activity Recognition by Colliding Bodies Optimization (CBO) -based Optimal Feature Selection with RNN," IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), pp. 1219-1228, 2021.

[17] Shende, D. K., Angal, Y. S., & S.C. Patil, "An Iterative CrowWhale-Based Optimization Model for Energy-Aware Multicast Routing in IoT", IJISP vol.16, no.1, pp.1-24, 2022.

[18] Dipali K. Shende and Sonavane S.S and Yogesh S. Angal, "A Comprehensive Survey of the Routing Schemes for IoT applications",Scalable Computing Practice and Experience", vol.21, pp.203-216, 2020.

[19] Gadekar, Prakash & Verma, Avnish & Dhotre, Virendrakumar, “Multicast Routing Protocols for Internet of Things (IoT) Applications”, 2020.

[20] D Sahija, "Impact of IoT integration with Mixed Reality on Manufacturing Operations", International Journal of Research Culture Society, vol. 24, pp. 78-85, 2021.