International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 8 (2020), pp. 1992-1998
 https://dx.doi.org/10.37624/IJERT/13.8.2020.1992-1998

Using Unstructured Search Algorithms for Data Collection in IoT-Based WSN

Nidal A. Al-Dmour
Faculty of Engineering and Informaion Technology, Ajman University,
Ajman, PO (346), UAE

Corresponding Email: (n.aldmour@ajman.ac.ae)

Abstract: IoT is used to connect things and these things "devices" could be sensors, actuators, house appliances, smartphones, or any electronic devices [3]. These devices are provided with sensing equipment's and generate various readings and measurements. Keeping the number of the sensors low and at the same time implementing efficient data collection is very essential. Many different search protocols have been developed for finding resources across the network to satisfy user's requests. This paper presents the results of using unstructured search protocols for data collection generated from the sensors in IoT-Based WSN. Experiments were conducted on a WSN with 10000 sensors and assuming certain percentage of the sensors generate unexpected important readings. Four different search algorithms are used: BFS, IntBFS, DRLP and Random Walks. A comparison is made between these algorithms and experiments show that IntBFS was the best approach for data collection in IoT-Based WSN.
Keywords: IoT, WSN, Cloud Computing, and Search Algorithms.

I. INTRODUCTION

Internet is defined as a Network of Networks [1]. Kevin Ashton in 1999 has introduced the term Internet of Things (IoT) which refers to uniquely identifiable objects and their virtual representations in an "internet-like" structure [2]. IoT is a system of interrelated devices with computing capabilities and can communicate and share information amongst them. IoT can also be defined as a connection of physical things and interact with each other using the Internet [3, 4].

Devices in IoT could be: sensors, actuators, house appliances, smartphones, or any electronic devices. In year 2003, the number of things connected to the Internet exceeded the number of the people [5]. In year 2020, the number of things connected to the Internet has reached 50 Billions and 4.4 zettabyte of data has been created [4

The Internet of Things (IoT) market is experiencing rapid growth, due to the adoption ioT devices in various types industries. Devices in IoT are seamlessly communicate with each other. Figure 1 shows the size of the IoT market in the world from year 2016 to 2020 [5].

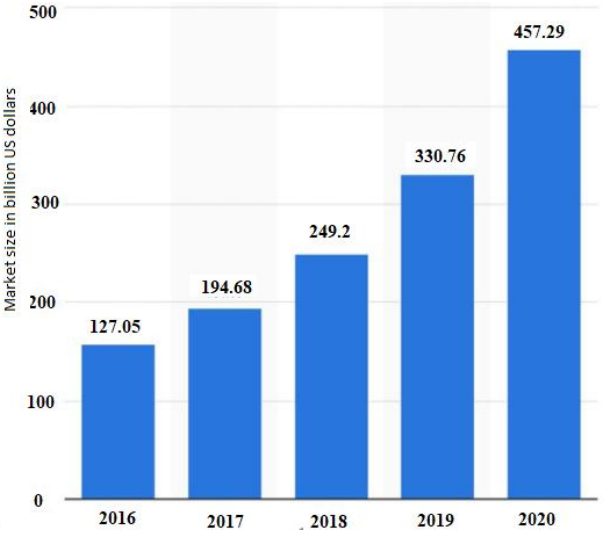


Figure1: IoT market [5].

The characteristics of the IoT as follows [8, 9]:

a. Interconnectivity: The devices in IoT are connected to each other to provide us with the required information at anytime and anyplace. They are interconnected with communication infrastructure using various types of sensors.
b. b. Heterogeneity: The devices in the IoT are heterogeneous and from different vendors as well as different platforms. They can interact with each other using various protocols such as IEE 802.11, IEEE 802.3, IEEE 802.15.4, or BLUETOOTH.
c. Dynamic changes: The devices in IoT join and leave at any time without notice. Additionally, the number of devices changes dynamically in IoT, devices goes into sleeping mode and wake up at irregular interval of times.
d. Scalability: IoT characterized of having large number of devices, users, and applications that could be spread over large area.

IoT is characterized of having four models as shown in figure 1 [10] which are:

a. Integrated Application.
b. Information Processing.
c. Network Construction.
d. Sensing and Identification.

International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 8 (2020), pp. 1992-1998
 https://dx.doi.org/10.37624/IJERT/13.8.2020.1992-1998

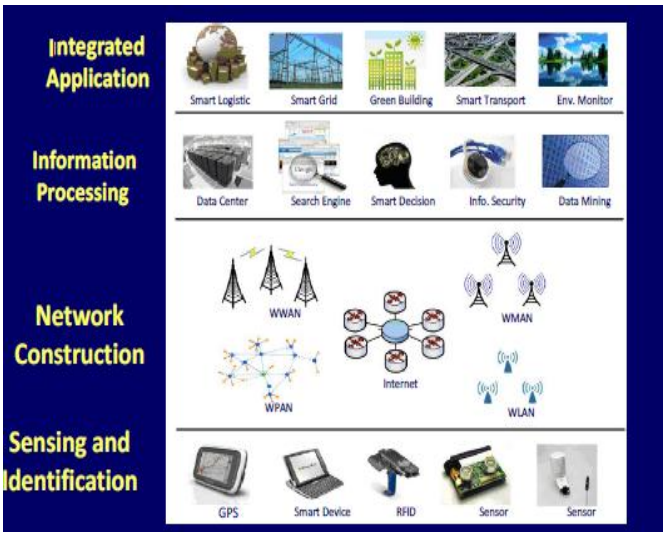


Figure 2: Layers Model of IoT [10]

Additionally, there are several types for IoT such as: Internet of Underwater Things, Internet of Underground Things, Internet of Battlefield Things, Internet of Space Things, Internet of Nanothings, and Internet of NanoBio Things [7] [8].

Cloud computing has helped the IoT to increases its deployment and ease its management. Information Technologies (IT) companies such as Microsoft (Azure IoT) [12], Amazon (AWS IoT) [13], and Cisco IoT Control Center [14] have emerged in the IoT market in the last few years.

There are several challenges and obstacles which affect the deployment of IoT. These are the main challenges:

a. Scalability: It refers to "the capability of a system, network, or process to handle a growing amount of work, or its potential to be enlarged in order to accommodate that growth" [22].
b. Security and Privacy: Data generated by sensors and transmitted for storing and processing are vulnerable for many types of attacks which endanger the Confidentiality, Integrity and Availability. Security is considered the major concern for the deployment of IoT. IoT has the potential to provide large amount of personal information which may be misused by criminals or cause disclosure to personal information.
c. Big Data Processing and Storing: Hundreds of billions of sensors in IoT are generating vast amount of data, this data need to be managed and analyzed. IoT has increased the demand for Data Analytics [7].
d. Interoperability [14]: There are many platforms in IoT provided by different vendors and these platforms are independent from each other's and has its own standards and protocols. It is a challenge to make these different platforms to work together and drives the demand for standardization to unify these platforms
e. Locating Services in IoT: Majority of the devices in IoT are sensors and characterized of having of limited battery life (power), memory and processing capabilities. It is very important to find the sensor with the required data in mimum amount of time.

Wireless Sensor Network (WSN) has been basis for any IoT system. WSN are used for many applications, however, the implementations generally differ under various aspects [29]. Therefore, WSN solutions are often application-specific prototype [29]. Additionally, it is difficult to define WSN requirement for hardware and software because sensors are continuously adapt to very diverse application [30]. WSN is composed of large number of sensors and their job is to gather data and forward it to the Internet using various routing protocols in an IoT system. WSN is having the following characteristics [15]:

a. Number of sensor nodes could be hundreds or thousands and more than the number of nodes in Ad Hoc networks.
b. Sensor nodes are vulnerable to failure due to the limited battery life. Therefore, WSN topology is highly dynamic.
c. Sensor nodes have limited capabilities such as processing speed, memory size and limited battery life.

Sensor nodes in WSN are responsible for collecting the data and a management node is responsible for data collections. There are two scenarios for data collection: the first scenario, the sensor nodes forward the collected data along other sensor nodes by hopping in regular intervals. The second scenario, management nodes send queries to the sensor nodes and request its data "readings". These queries forwarded across the nodes in WSN using various search mechanisms.

The search for the required data is one of the major problems of computer networks. This paper presents the usage of unstructured search algorithms to collect the readings from sensors nodes in IoT-Based WSN environment. A simulated environment was created with ten thousands sensors and four search algorithms were used to search for sensors with specific readings. Experiments were conducted to compare between the efficiency of these algorithms with regard to the data collection in IoT-Based WSN. Section II presents related work for searching the sensory data in IoT-Based WSN. Section III presents the simulation results and discussions. Finally, the conclusion was presented in section IV.

II. Related Work

The authors in [13], describes an application of semantics to sensors to get the functional and non-functional properties of sensors, as well as the composition of sensors. Semantic sensor composition allows data flow between sensors to be specified based on semantic constraints of sensors and user goals [13].

Sense2Web [14] supports flexible and interoperable descriptions, implementing the association of various sensor data. Different sensor data ontologies are connected to domain data and resources on the semantic Web and the Web of data. Sense2Web interconnects data generated by sensors in IoT-based Wireless Sensor Networks (WSNs) by defining relationships between ontologies and schemas.

Dyser [15] is a real-time search engine for the Web of Things assumes that each sensor includes a metadata description within its HTML page, making it accessible for crawling. This search engine employs a probabilistic model to determine which sensor is most likely to deliver relevant data when responding to a user's query.

International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 8 (2020), pp. 1992-1998
 https://dx.doi.org/10.37624/IJERT/13.8.2020.1992-1998

Snoogle is another IoT-Based WSN approach for indexing and ranking entities [16]. It is based on using nodes addresses (mainly IP address) and building inverted indexes for all connected entities

IETF standardization presented Constrained Application Protocol (CoAP) [21]. CoAP is characterized of having small number of messages to minimize message fragmentation. The communication in CoAP is based on the client/server paradigm by using HTTP functionalities. Sensors register their resources by using push model whereas sensors in IoT-Based WSN act as a web servers sending their resources to the RD

III. Simulation Results and Discussions

The unstructured search algorithms have an effective role in the world of the Internet of Things. There are millions of devices with a large amount of data generated by these devices leads to great challenges in Internet of Things environments. Unstructured search algorithms are capable to find the information effectively without the need for a pre-defined structure. Many applications in IoT systems are going to benefit from using unstructured search algorithms.

In IoT, unstructured search algorithms can be used to process data generated from various types of devices, such as smart home devices, industrial sensors, and wearable technology. Unstructured search algorithms can be useful in traffic flow analysis and monitoring congestion patterns. However, there are various challenges that are encountered when using unstructured search algorithms as the number of devices increases, the volume of data generated can become overwhelming, thus, ustructured search algotithms might not be capable to find resources in IoT systems.

There are many types of unstructured search algorithms in literature, and we selected four different search algorithms for data collection in IoT-Based WSN which are: Breadth First Search (BFS) [17], Intelligent BFS, Random Walks [17], and Distributed Resource Location Protocol (DRLP) [17].

Wireless Sensor Network (WSN) [18] is the basis for IoT environment and can be represented by a graph formed by a set of nodes, in which each node in this network within the coverage area of certain number of sensory nodes. Nodes in a WSN are deployed in an ad hoc manner for majority IoT-Based WSN applications. Messages will be forwarded between sensory nodes through multi-hop connections. We assume that there are *m* sensors has unexpected important readings or data denoted by "Rd".

We present results for four searching algorithms: BFS, Intelligent BFS, Random Walks, and DRLP. We experimented with 10000 sensors, and 10 management nodes in a simulated IoT-Based WSN. Management nodes are responsible for storing and forwarding the data to the outside world through the Internet and are considered as a gateway to the WSN. The application in IoT can be categorized into three domains: Personal, Enterprise and Utilities. In our simulation environment, we assume the sensors are deployed in a personal environment and each sensor has detected or produced unexpected Rd.

WSN is very dynamic network and nodes join and leave at any time as well as some nodes become inactive and go into sleep mode to save battery life. Therefore, we keep only 75% of sensory nodes active while the simulations are running.

Each node in our simulation is configured to use IPv4 and ICMPv4. Nodes are able to transmit echo messages to all its neighbours in the topology. Any node in the network receives an echo request will reply back the same message. Each sensory nodes has 5 to 6 other nodes within its coverage area. The Time to Live (TTL) parameter was set to ten for all algorithms in our simulations.

The communication of nodes is based on protocols IEEE 802.15.4. The payload length of the echo message delivered between the sensory wireless nodes was 10 bytes. The simulation was repeated several times for different values of frequency of the ping messages. Table 1 shows parameters of each sensory nodes.

Table 1: Parameters for each sensory node.

Wireless Protocol	802.15.4
Bitrate	300 Kbps
Transmitting Power	60 mW
Receiving Power	69 mW
Sleep Power	0.06 mW

We assumed all the sensors have generated different readings, however, there are number of sensors generated unexpected important results *Rd*. The searching algorithms are going to be used to find these unexpected results. We experimented with two scenarios: For the first scenario, two hundreds sensors generated unexpected readings. For the second scenario, 400 sensors generated unexpected readings. Two metrics were measured:

- The number of sensors with important readings found after sending 2000, 4000, and 6000 queries by the management nodes.
- The number of sensory nodes the queries passes through during the search for *Rd*.

Figures 3 and 5 show that the IntBFS was the best approach for finding the sensors which have unique readings in our simulation networks. The queries are forwarded in WSN to other sensory nodes which have answered most of the queries. Once a new query arrives, the sensor node forward it to a set number of its neighbours that have returned the most results for these queries.

Both DRLP and BFS exhibit interesting characteristics in data collection in IoT-Based WSN. Both have achieved approximately the same results for data collections although BFS is sending more queries and this is due to the fact storing the resource location at the intermediate sensory nodes once a node with required data has been found. Random Walks was the lowest algorithm in data collections in our simulated IoT-Based WSN.

Figures 4 and 6 show the number of sensory nodes visited during the search for *Rd*. BFS has visited the highest number of sensory nodes as some of the nodes were visited more than once. Both IntBFS and DRLP have achieved the same number of visited sensory nodes.

Random Walks was the lowest with regard to the number of visited sensory nodes.

Conclusions

Internet of Things (IoT) Based Wireless Sensor Network is characterized of having limited power, memory, and processing capabilities. These sensors are deployed to measure a certain physical environment. Unstructured search algorithms were presented in the literature to be used for data collection as well as resource discovery in computer networks. Search algorithms were tested for gathering the results generated from the sensors in IoT-Based WSN in this paper. The focus was on the search for unexpected important data *Rd* generated from the sensors in IoT-Based WSN. We experimented with four different search algorithms: BFS, IntBFS, DRLP, and Random Walks. A simulated environment of IoT-Based WSN was created to conduct our experiments. A comparison was made between these algorithms to find the efficient algorithm capable to find the maximum number of sensory nodes with important readings. IntBFS was found to be the best approach for data collection in IoT-Based WSN in comparison to other search algorithms.

References

[1] Paul Tulenko. INTERNET IS A NETWORK OF NETWORKS, https://www.deseret.com/1995/5/9/19174405/internet-is-a-network-of-networks. Desert News, 1995. Last visited 9th of August 2020.

[2] Ashton, K. That "Internet of Things" Thing: In the Real World Things Matter More than Ideas. RFID Journal, 2009.

[3] Atzori L, Iera A, and Morabito G. The Internet of Things: A survey. Computer Networks 54(15):2787–2805, 2010.

[4] Gubbi J, Buyya R, Marusic S, and Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computing System. 29(7):1645–1660, 2013.

[5] Noor A. Hussein, Mohamed Ibrahim Shujaa. DNA computing based stream cipher for internet of things using MQTT protocol. International Journal of Electrical and Computer Engineering (IJECE) Vol. 10, No. 1, February 2020.

[6] Comprehensive Guide to IoT Statistics You Need to Know in 2020, https://www.vxchnge.com/blog/iot-statistics. Last visited 9th of August 2020.

[7] Statista, https://www.statista.com/statistics/764051/iot-market-sizeworldwide. Last visited 9th of August 2020.

[8] Dr. Ovidiu Vermesan SINTEF, Norway, Dr. Peter FriessEU, Belgium. Internet of Things–From Research and Innovation to Market Deployment. river publishers' series in communications, 2014.

[9] V. R. Konduru and M. R. Bharamagoudra, Challenges and solutions of interoperability on IoT: How far have we come in resolving the IoT interoperability issues. International Conference on Smart Technologies for Smart Nation (SmartTechCon), Bangalore, 2017.

[10] Dr. I.F. Akyildiz gives a keynote speech entitled, Internet of Things: Trends, Directions and Opportunitie. Annual IoT Forum, Tunisia, April 28, 2017.

[11] Keyur K Patel, Sunil M Patel, P G Scholar, and Carlos Salazar. Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges. International Journal of Engineering Science and Computing IJESC, ISSN 2321 3361 May, 2016.

[12] Microsoft Azure, https://azure.microsoft.com/en-us/overview/iot/. Last visited 9th of August 2020.

[13] Amazon AWS, https://aws.amazon.com/iot/. Last visited 9th of August 2020.

[14] CISCO–IOT, https://www.cisco.com/c/en/us/solutions/internet-of-things/iot-control-center.html. Last visited 9th of August 2020.

[15] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E, Cayirci. A Survey on Sensor Networks". IEEE Communications Magazine, August 2002.

[16] Tran, K.N., Compton, M., Wu, J., Gore, R. Short paper: Semantic sensor composition. In: Proceedings of the 3rd International Conference on Semantic Sensor Networks-Volume 668. CEUR-WS. org; 2010.

[17] Barnaghi, P., and Presser, M. Publishing linked sensor data. In: Proceedings of the 3rd International Conference on Semantic Sensor Networks Volume 668, CEUR-WS. org; 2010.

[18] D. Russomanno, C. Kothari, and O. Thomas. Sensor ontologies: from shallow to deep models. System Theory, 2005.

[19] ORGC Project, https://www.ogc.org/node/698. Last visited 9th of August 2020.

[20] B. Ostermaier, K. Römer, F. Mattern, M. Fahrmair and W. Kellerer. A real-time search engine for the Web of Things. Internet of Things (IOT), Tokyo, 2010.

[21] Haodong Wang, Chiu C Tan, and Qun Li. Snoogle: A search engine for pervasive environments. IEEE Transactions on Parallel and Distributed Systems, 21(8):1188–1202, 2010.

[22] Maciej Kurant, Athina Markopoulou, and Patrick Thiran, on the bias of BFS (Breadth First Search). Teletraffic Congress (ITC), 2010.

[23] V. Kalogeraki, D. Gunopulos, and D. Zeinalipour-Yazti. A Local Search Mechanism for Peer-to-Peer Networks. Proceedings of the Eleventh International Conference on Information and Knowledge Management (CIKM), McLean, Virginia, USA, ACM Press, Pages: 300-307, 2002.

[24] C. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker. Search and Replication in Unstructured Peer-to-Peer Networks. Proceedings of the Sixteenth Annual ACM International Conference on Supercomputing (ICS), New York, USA, ACM Press, Pages: 84-95, June 2002.

[25] D. Menascé and G. Mason. Probabilistic Scalable P2P Resource Location Services. ACM SIGMETRICS Performance Evaluation Review archive, ACM Press, Volume: 30, Issue: 2, Pages: 48-58, September 2002.

[26] J. Yick, B. M., and D. G. Wireless sensor network survey. Computer networks, vol. 52, no. 12, pp. 2292–2330, 2008.

[27] M. Z. Khan, B. Askwith, F. Bouhafs, and M. Asim, Limitations of simulation tools for largescale wireless sensor networks. IEEE Workshops of International Conf. on Advanced Information Networking and Applications, pp. 820–825, 2011.

[28] P. Chhimwal, D. S. Rai, and D. Rawat, Comparison between different wireless sen- sor simulation tools. IOSR Journal of Electronics and Communication Engineering, vol. 5, no. 2, pp. 54– 60, 2013.

[29] Lazarescu M.T. Wireless Sensor Networks for the Internet of Things: Barriers and Synergies. In: Keramidas G., Voros N., Hübner M. (eds) Components and Services for IoT Platforms. Springer, Cham. https://doi.org/10.1007/978-3-319-42304-3_9, 2017.

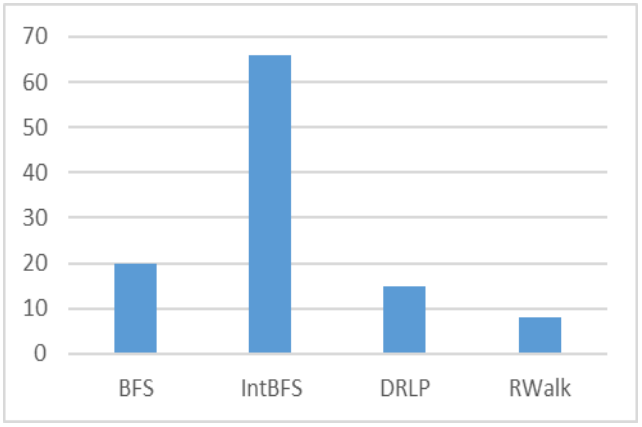
[30] K. Romer and F. Mattern. The design space of wireless sensor networks. IEEE Wireles Comm., 11(6):54–61, December 2004.

[31] Z. Shelby, K. Hartke, and C. Bormann. (August 2020). Constrained Application Protocol (CoAP). RFC 7252 Internet Engineering Task Force [Online]. Available: http://tools.ietf.org/html/rfc7252.

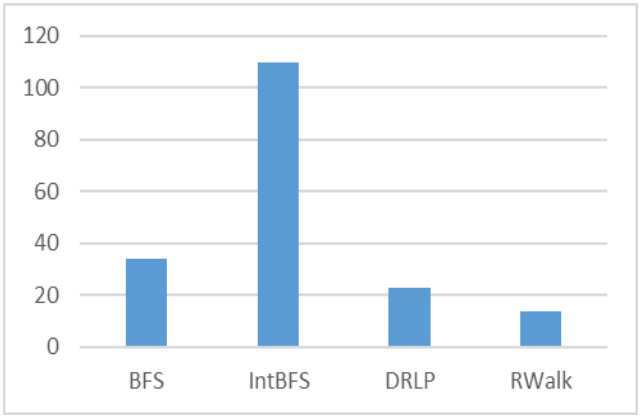
[32] Vandana C.P and Ajeet A. Chikkamannur. Study of Resource Discovery trends in Internet of Things (IoT). Int. J. Advanced Networking and Applications Volume: 08 Issue: 03 Pages: 3084-3089 (2016) ISSN: 0975-0290.

[33] Hugo Barrigas, Daniel Barrigas, Melyssa Barata, Jorge Bernardino, and Pedro Furtado. Scalability of Facebook Architecture. Advances in Intelligent Systems and Computing book series, Springer International Publishing Switzerland 2015.

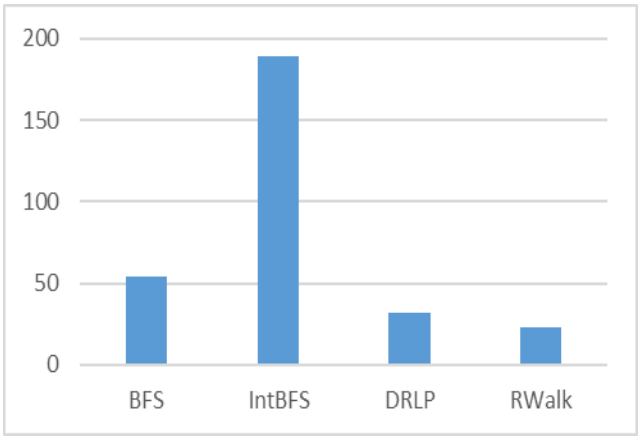
International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 8 (2020), pp. 1992-1998
 https://dx.doi.org/10.37624/IJERT/13.8.2020.1992-1998



(a)

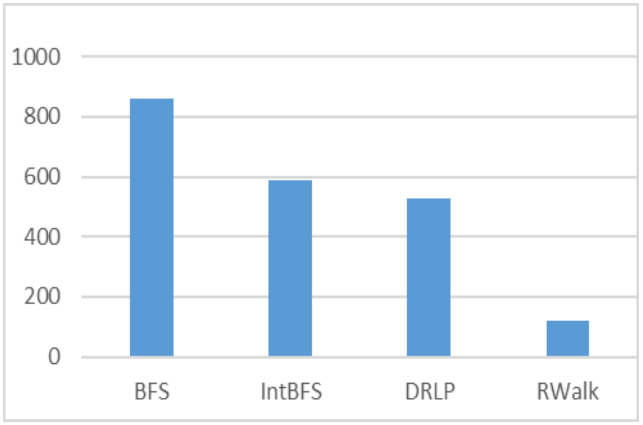


(b)

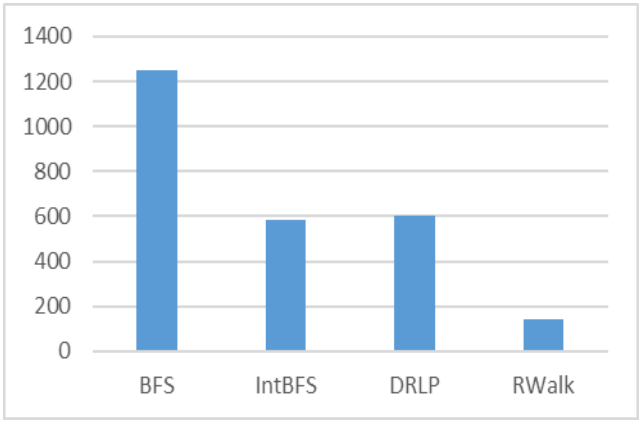


(c)

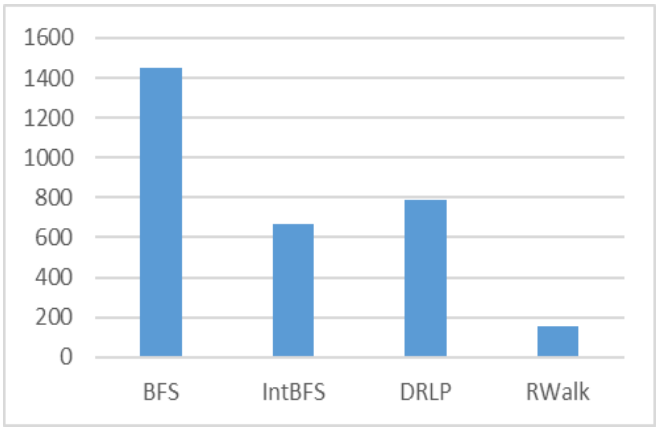
Fig 3. 200 sensors have generated unexpected data *Rd*. The number of sensors with important readings found after sending (a) 2000, (b) 4000, and (c) 6000 queries by the management nodes.



(a)



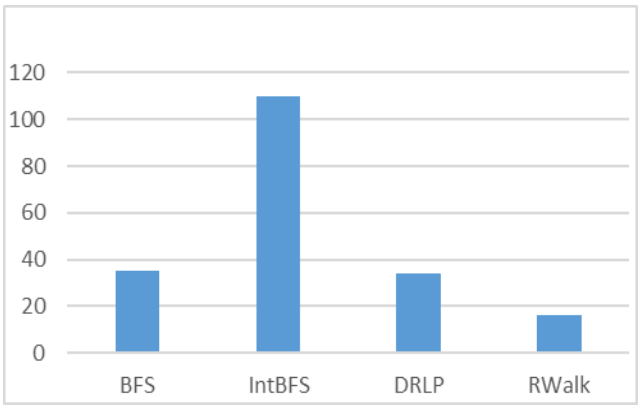
(b)



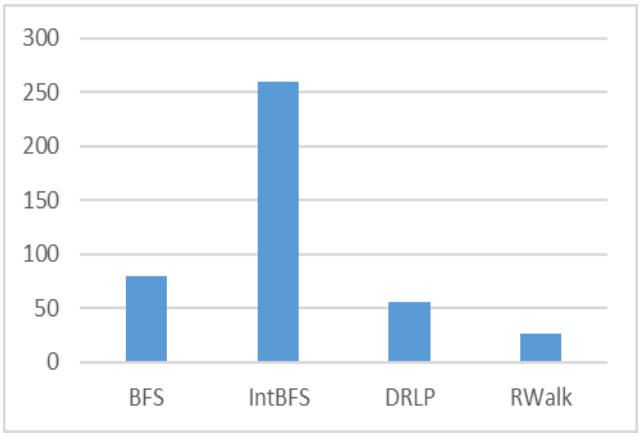
(c)

Figure 4: The number of sensory nodes the queries passes through during the search for sensory nodes with unexpected data Rd after sending (a) 2000, (b) 4000, and (c) 6000 queries by the management nodes.

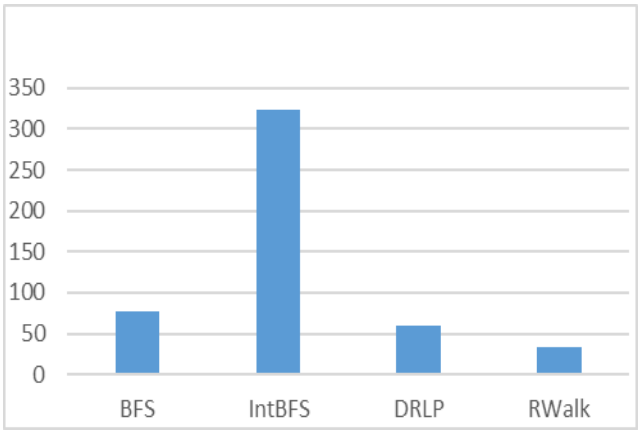
International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 8 (2020), pp. 1992-1998
 https://dx.doi.org/10.37624/IJERT/13.8.2020.1992-1998



(a)

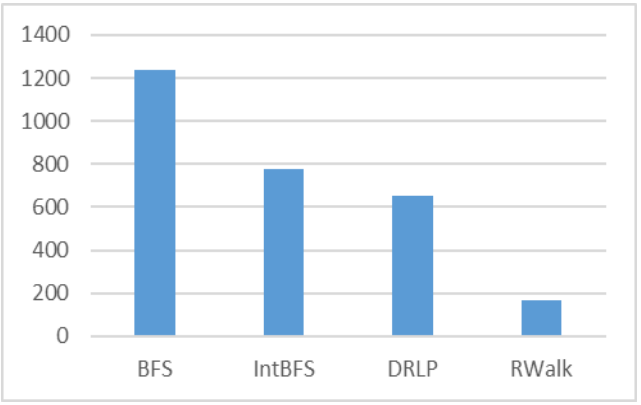


(b)

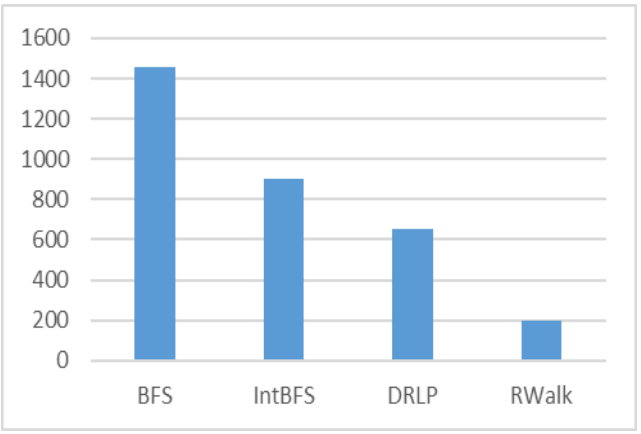


(c)

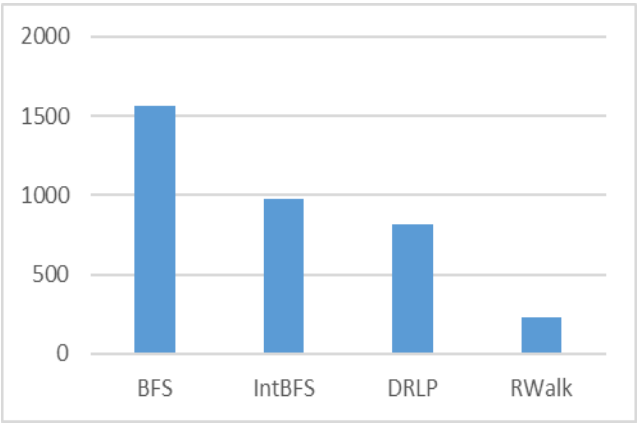
Figure 5: 400 sensors have generated unexpected data *Rd*. The number of sensors with important readings found after sending (a) 2000, (b) 4000, and (c) 6000 queries by the management nodes.



(a)



(b)



(c)

Figure 6: 400 sensors have generated important readings. The number of sensory nodes the queries passes through during the search for sensory nodes with unexpected data *Rd* after sending (a) 2000, (b) 4000, and (c) 6000 queries by the management nodes.