Journal of Advances in Science and Technology
Vol. 19, Issue No. 1, March-2022, ISSN 2230-9659

Security and efficiency in wireless sensor Network Data Aggregator Design

Phakade Shirish Vitthalrao[1]*, Dr. C. Ram Singla[2], Dr. Omprakash Rajankar[3]

[1] PhD Student

[2] Sunrise University, Alwar

[3] NBN,S.C.O.E., Pune

Abstract - Most wireless sensor networks are composed of a large number of low-cost sensor nodes with very limited sensing, computing and communication capabilities. Since sensor nodes often have limited resources, it is necessary to maximize both the average sensor lifetime and total bandwidth usage by minimizing the amount of data sent. Information overload in a network can be reduced by data aggregation, which involves compiling and synthesizing input from multiple sensors. Sensor nodes in wireless sensor networks are vulnerable to node compromise attacks, and security concerns such as data confidentiality and integrity are of paramount importance due to the sensitive nature of the information being transmitted. Thus, security should be a primary consideration in the development of protocols for use in wireless sensor networks, such as those used to collect data.In a wireless sensor network, data aggregation is an essential approach. Because by pooling together data, we can cut down on wasteful duplication and thereby save energy. In wireless sensor networks, extending the network's lifetime is the most difficult challenge, however this may be achieved by data aggregation.

Keywords - Wireless Sensor Network, Security, Data Aggregation, Communication.

----------X----------

1. INTRODUCTION

Wireless sensor networks, or WSNs for short, are one of the top ten emerging technologies that might fundamentally alter the path of human history. Wireless sensor networks (or WSNs) are a special type of network made up of several extremely tiny sensor nodes. These sensor nodes may only be allowed a small fraction of the network's total bandwidth, and the network's architecture may need a large number of hops before reaching any one node. The devices may gather information from a variety of sources on a variety of aspects, such as heat, temperature, sound, vibration, pressure, motion, and pollution. WSNs are unique among networking technologies in a number of ways, including their low power consumption, resistance to extreme weather, capacity to manage node failures, mobility of nodes, dynamic network topology, communication failures, node heterogeneity, large-scale deployment, and unattended operation. Another feature that sets WSNs apart is their adaptability to different types of nodes joining the network. To improve object tracking and battlefield surveillance were the initial motivations for research and development in the field of WSNs. [1]

WSN was first implemented for a variety of purposes; surveillance, facility tracking, and ecological monitoring were among the earliest uses of the technology. As time goes on, the role of humans in sensing, data collecting, and computing will only grow in importance. Users may collect data from their immediate surroundings, do analysis on that data, and then share their findings with other users to get a more in-depth understanding of the environment, whether it social or physical. Anyone, no matter how much or how little experience they have, should be able to complete this assignment properly. It is possible that WSN may eventually prove to be a more efficient replacement for specialised infrastructure and niche networks. When WSNs are used in applications, people are responsible for sensing, collecting data, and analysing data. These programmes herald a new era in which omnipresent computing and networking are not the exception but the rule. [2]

1.1 Communication Systems

As was said before, the notion of communication centres on the evidence of information flow between at least two devices. Generally speaking, data communication refers to the exchange of digital information between at least two different types of transportation hardware. [3]

As early as the mid-1960s, it was proposed that computers couldn't exchange data with one another

unless they were in the same room. This guidance was provided to facilitate the smooth transfer of data from one system to another. A framework has been established by the research community that, while not always feasible, enables computers to make better use of their resources and increase the rate at which data can be sent between them, even if they are not in the same physical place. This was completed despite the fact that it isn't always feasible to do so. This type of system is commonly referred to as a data communication network.

At the very least, two devices must be linked together to form a communication network. As a result, we can now pool our resources, have conversations on computers, and send and receive data across great distances. When considering the many ways in which connections may be made, a communication network can be broken down into two categories. Here are the classes in question: Wire- and wireless-based solutions are both part of this bundle. [4]

1.2 Advantages ofWireless Networks

As seen in Figure1, there are several advantages to using wireless networks. There are several advantages to using a wireless network, but among the most prominent are the low barrier to entry for installation, the high degree of mobility, the wide availability, the low cost, and the scalability. [5]

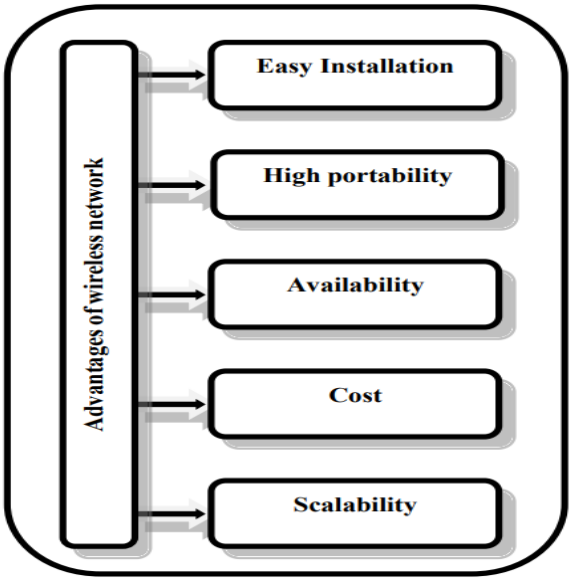

Figure 1: Advantages of wireless network

1.3 Types of Wireless Networks

Considering the increasing popularity of wireless network deployment, its potential uses may be identified in several contexts. In Figure 2 we can see the differences between the four different wireless networks.

- Cellular networks are the backbone of PSTN (Public Switched Telephone Network), which coordinates all mobile phone calls. This may be discerned from data given by cellular systems (PSTN). Each "cell" in this type of network is really made up of a huge number of tiny "sub-cells," and the "base station" provides the cells with their electrical power. Because digital communication wasn't yet available, the first generation of cellular networks relied on analogue communication to function. Networks of the second generation encrypt digital speech as one way they contribute to the greater security of such networks. One of the ways in which these networks have advanced is in this respect. The third generation of mobile phones utilise a network connection to the Internet to speed up the transfer of packets via the use of packet switching and the transmission of sounds through the use of circuit switching. This was done to satisfy the requirement for a faster data transmission rate.
- A WSN is made up of independent nodes that work together to provide a holistic perspective and facilitate interaction between people, machines, and their environments. The most prevalent use of WSNs is in the study of natural phenomena and processes.
- Each node in a WSN can communicate with the other nodes in the network. WSN is organised in such a way as to monitor temperature, humidity, and other conditions, as well as classify and recognise targets. Classifying and identifying targets on a worldwide scale is what WSN is used for.

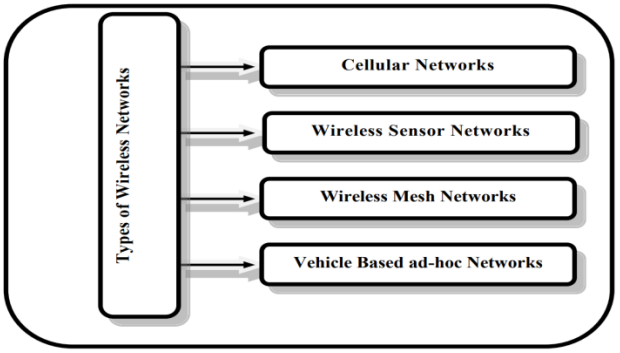

Figure 2: Explanation of the Various Types of Wireless Sensor Networks

- t has reasonable prices and may be utilised for a wide variety of purposes, including military surveillance and target tracking, healthcare and environmental monitoring, fire prevention, and traffic control. Surveillance in the battlefield, monitoring of targets, the environment, healthcare, and the identification of wildfires are all examples of this. A WSN's overall power is constrained since its sensor nodes aren't powerful enough to

perform computations without additional power.

- SNs may be set up anywhere, whether on land, underwater, or in the air. To operate effectively, a sensor network must first surmount a number of obstacles and constraints that are unique to the setting in which it was designed. These difficulties and constraints are unique to the setting in which the network was first created. These challenges and restrictions are specific to the setting in which the network was originally created. WSNs may be broken down into four main categories: mobile, terrestrial, subsurface, and submerged. There is a vast range of WSN flavours available. In addition to the ubiquitous WSN, variants such as the multimodal WSN and the underwater WSN exist.[6]

1.4 Classification of WSNs

Several types of WSNs are shown in Figure3. All of these classes are represented in the diagram. [7]

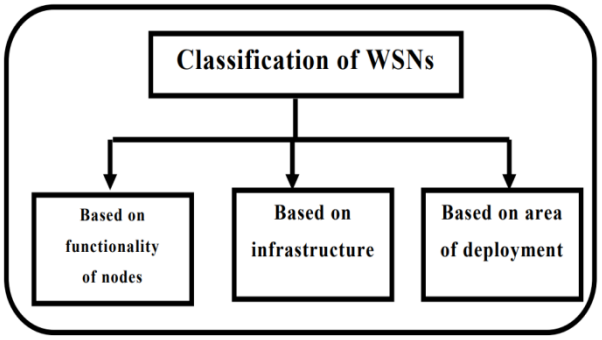

Figure 3: Categories of WSNs based on different functions.

- Basedonfunctionalityofnodes
- Basedoninfrastructure
- Basedonareaofdeployment

The way in which individual sensor nodes carry out their tasks determines whether a wireless sensor network, or WSN, is considered harmonised or heterogeneous. In a standardised WSN, all of the sensor nodes have the same amount of memory, transmission capacity, and battery life. It's possible for a WSN to consist of two or more different kinds of nodes, each of which may be distinguished from the others by its own set of features, such as its memory capacity, its power output, or some other set of abilities. While a diversified WSN may be less expensive in the long run, a harmonised WSN will likely use more energy efficiently because of its focus on reducing variation across nodes. [8]

Furthermore, depending on the underlying network architecture, WSNs can be classified as either structured or unstructured networks. In contrast, unstructured WSNs often consist of a huge number of nodes, each of which is placed in the field on the fly, without any sort of preparation. A structured WSN is employed to set up nodes according to a preset plan. In contrast to unstructured WSNs, which could still have some uncovered areas, structured WSNs provide excellent coverage even with a small number of nodes. Even the most well-connected unstructured WSNs may have blind spots. [9]

1.5 Wireless Sensor Networks: Design Issues

Multiple aspects of traditional network design, including the network's architecture, deployment area, performance, and so on, receive extensive attention. These factors aren't as heavily weighted in the design process of other kinds of networks. Design requirements should be revaluated to accommodate the development of WSNs. Figure4 depicts some of the design issues that have arisen. There needs to be more research and discourse on these topics.[10]

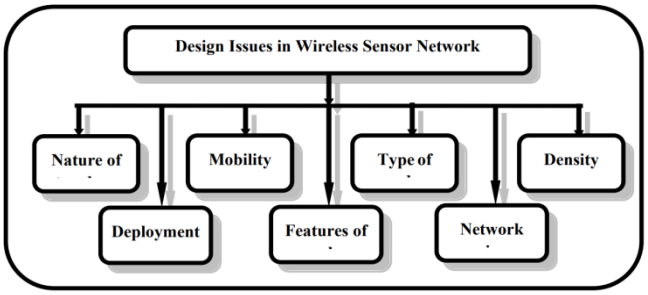

Figure 4: Some of the design challenges faced by WSNs.

Deployment: WSNs may be used in many different types of businesses and marketplaces. The nodes used by these apps may be randomly placed, as is the case with certain others, or they may follow a present architecture. Nodes are distributed at random in military applications by flying them to different locations.

Topology's fundamental building components There are a variety of uses for nodes that remain in the same location for the duration of the network. There will be no further opportunities to join when the deployment is finished. There is constant evolution as nodes are added and removed from various initiatives.[11]

Mobility: Once a terrestrial network is established, its nodes do not move from their original locations. When compared to wireless networks, where nodes are free to roam, this is a key difference. One's freedom of movement is hampered by this. There are no restrictions on the location that a node in a WSN can relocate to. The WSN approach benefits from mobility; hence, it is a crucial part of the plan. Characteristics of the node are as follows: Nodes used in the deployment of WSNs should be able to

run independently for long stretches of time, cost a reasonable amount, and have a sufficient amount of processing power compared to other nodes in the same class. The requirements for a robust node are incompatible with one another because of their large size, expensive cost, and high degree of energy consumption.[12]

Nodes that are categorised as type: By employing sensor nodes ranging in size from a few millimetres to a block, WSN is able to carry out the numerous environmental monitoring applications it supports.[13]

Network size: The protocols and algorithms that control a sensor network must provide scalability as the number of nodes in the network increases from tens to hundreds to thousands.

The density of a network depends on two factors: the total number of nodes and the total area of the region covered by the network. High-density networks are undeniably more expensive and produce less dependable outcomes than low-density networks.[14]

2. METHODOLOGY

A sensor network is a set of interconnected sensors that performs a certain function. These sensors, however autonomous and equipped with their own data, share a common network. The aggregation tree, which includes all of the nodes that contributed to the aggregate, finishes the job and then proceeds to the BS. Common types of nodes in a sensor network include "leaf" nodes, "intermediate" nodes (sometimes called "aggregators"), and "base station" (BS) nodes (also called "sinks"). As the most fundamental type of node, leaves are the starting point for all other types. By taking use of the operation slices supplied by the leaf node, it is feasible to aggregate data while still maintaining users' privacy. It then performs data sensing and data aggregation, as mentioned above, before passing the resulting data on to its parent. The central node is in charge of collecting data, performing MAC, and then sending it on to a higher-level aggregator or a sink. The sink's job is to take the network's aggregated result and turn it into field-specific data. Not only does this guarantee the integrity of the network as a whole, but it also guarantees the integrity of any individual findings that are derived from that network.

2.1 Key Distribution in EPSDA

The implementation of asymmetric cryptographic methods requires a larger time and financial investment than symmetric cryptographic procedures. The EPSDA's method of key distribution is outlined below; it is secure and energy efficient. Cluster topology frequently employs ESPDA key distribution, therefore there must be some link between the two. Every node is given a unique identity (Idi), a node-specific key (Ki), and a shared secret key during production (K). The sink is given with a set of ID, secret Key (Idi-Ki) pairs for each node in the network in addition to a shared secret key (K), a session key (Ks), and a common secret key (K). The sink gets these things before it's put into the network. These products are offered for sale in pairs consisting of two of the same item.

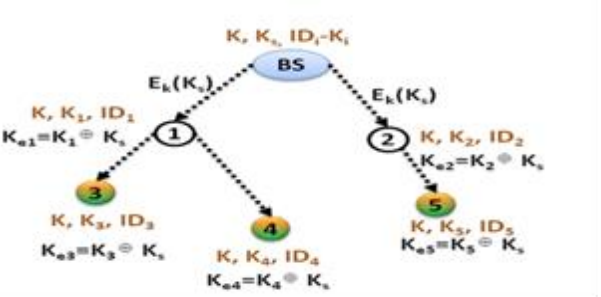

Figure 5: Key Distribution

When the BS receives the aggregated encrypted result, it first determines the secret Key (Ki) corresponding to the node ID's, and then generates the decryption key (Kei) by XORing it with the session key transmitted by the sink to the network. This eliminates the need for asymmetric cryptography for encryption and decryption. The reason for this is that when the BS receives the combined encrypted result, it first determines the secret Key (Ki Data is encrypted using a unique encryption key for each session in order to safeguard its confidentiality, ensure that it is kept as up to date as is practically possible, and so on.

2.2 Energy Efficient Privacy Preserving Secure Data Aggregation

In this part, we go out the specifics of our proposed approach to EPSDA, or Energy Efficient, Privacy Protecting, Secure Data Aggregation. Security Data Aggregation that Saves Power and Protects Individual Privacy is abbreviated as EPSDA. Aggregation There are a total of five phases to the process: the first, in which the tree is created; the second, in which it is sliced; the third, in which it is mixed; and the fourth, in which it is aggregated.

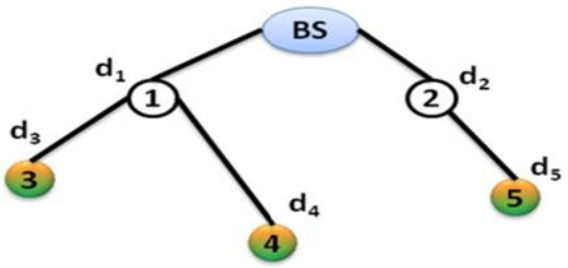

Figure 6: Aggregation Tree

Encryption $C_{IJ} = (\sum_{i,j=1}^{N}(d_{ij} + K_{ei}))modM$

If communication between nodes I and j is cut off, the value of the data slice (denoted by dij) that

moves from node I to node j will be 0. The initials Kei stand for the node I encryption key.

Aggregation $d_{IA} = (\sum_{I,J=1}^{N} c_{ij}) \text{mod} M$

Decryption $f_R = (\sum_{I,J=1}^{N} c_{ij} - \sum_{I,J=1}^{N} K_{ei}) \text{mod} M$

$$d_I = \sum_{I,J=1}^{N} d_{ij}$$

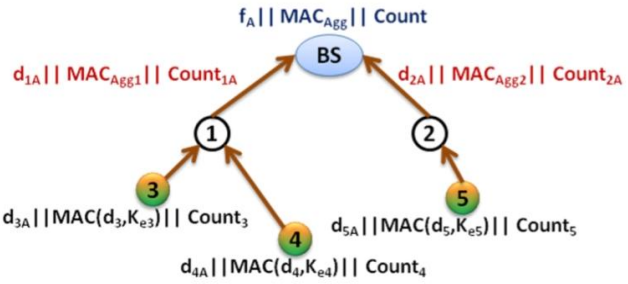

Figure 7: Aggregation

Each intermediate node generates a MAC for the data it has gotten by first using its own private encryption key to establish a MAC for the data, and then use privacy homomorphism to combine that MAC with the MACs it has collected from its offspring in order to complete the process. In order to accomplish the process of MAC aggregation, the CMT approach is utilised.

$$MAC_{Ag} = \left(\sum_{i=1}^{N} MAC(d_{ij}, K_{ei}) \right) \text{mod} M$$

3. RESULTS

NS2 can provide you with an overview of both the EEHA and EPSDA systems if that is something you are interested in. In the testing area, which measured 1000 metres on one side and 500 metres on the other, each of the 15 sensor nodes that make up the WSN were put through their paces. After it has been determined which of the surviving nodes will serve as the sink, the nodes that are still standing will work together to build a tree with the sink Node serving as the tree's root. The simulation was set up using the specifications for 0.395 watts of idle power, 0.395.0 watts of receiving power, and 0.660 watts of transmission power. A total of 100 J of energy will be distributed to each node. In order to acquire a better knowledge of both the power consumption and the level of security that EPSDA possesses, we conducted an analysis using the EEHA methodology.

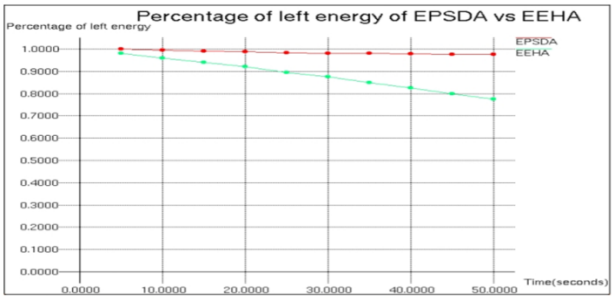

Figure 8: Comparison between the amounts of Energy Consumed by EPSDA, EEHA

Energy: depleting all reserves of usable power Figure8 provides a visual depiction of the percentage of a node's total energy that may still be used for each of the possible methods. Since the EEHA system requires far more communication and processing overhead than the EPSDA system, we reasoned that it must use a greater total amount of energy. Total communication overhead is reduced since the volume of communications is reduced and the time it takes for communications to occur is shortened due to end-to-end encryption. In less intensive decryption operations to run, the EPSDA is more efficient with its power use.

Security: As end-to-end secrecy, data freshness, data privacy, and message authentication are all supported by the EPSDA system, it is ideally suited for usage in applications that place a premium on securing sensitive information. Because of this, it may be used confidently in any setting where confidentiality of information is a priority. Due to the high communication and processing costs associated with the EEHA approach, it was best reserved for less privacy-sensitive applications. The EEHA method is too expensive and hence unsuitable for widespread usage as a privacy safeguard.

Data Aggregation

We take into consideration the scenario in which characteristics, parents, energy status, and gradients are conveyed once per fifty seconds in order to conform to the requirements of HDA. We make use of criteria such as the power dissipation during idle time (35 mW), the power dissipation when receiving (395 mW), and the power dissipation when sending (660 mW), all of which are in accordance with DD. At this rate of sampling, a single sample was collected once every single second. In this study, we investigate the correlations between the total amount of energy spent and the number of source nodes, cardinality of the network, and size of the network. Additionally, we look at how these three factors are related to one another.

Due to this, a larger WSN will consume more power than a smaller one. Our DP technology consistently yields superior results compared to HDA and DD methods. The DP method does this by minimising

the production of network messages by avoiding the generation of unnecessary traffic during data transfer to the sink node. That's important because it allows us to reach our ultimate aim, which we'll talk about in more detail below. It has been shown that when network capacity increases, the performance gap between DP and DD and DP and HDA rises. And this held true for both of those aforementioned differences in performance. This shows that our system's data aggregation capabilities are superior than those of HDA, DD in terms of scalability.

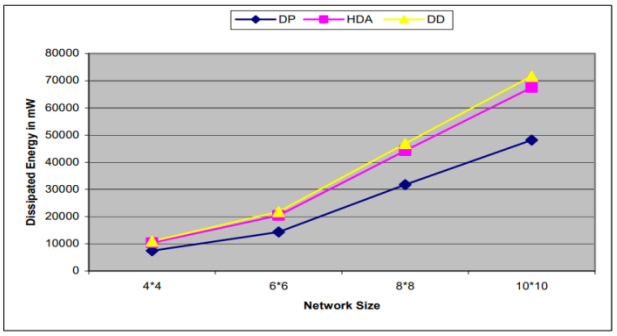

Figure 9: The quantity of energy used by networks of varying sizes

The density of the source nodes was increased from 10 to 50 for a 1010 WSN of constant size. As seen in Figure10, the total number of source nodes in the network has increased from 10 to 50. This means a greater overall quantity of energy will be required for data transmission to the sink node. In this respect, there is no difference in how DP, HDA, or DD systems function. Having more nodes to send from resulted in more messages being generated by the network. Due to the increased energy requirements for sending these messages, a higher number of source nodes was required. However, the rate of increase in the overall quantity of energy wasted by the DP system slows down as the number of source nodes increases. In contrast, the rate of increase in the amount of dissipated energy develops more rapidly in HDA, DD systems. This feature is not present in any of these systems. In particular, networks with a larger number of source nodes benefitted from this adjustment, as their performance using the DP technique improved. This substantiates the usefulness of our DP approach for its intended purpose of data aggregation in WSNs.

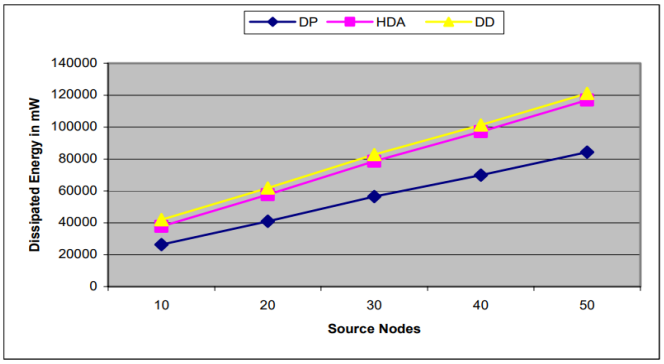

Figure 10: Quantity of energy that was used by a variety of source nodes.

The network may be described by its cardinality, which is: The 1010 WSN was the standard, with the number of source nodes equaling around 15% of the total number of sensor nodes. The network's cardinality has increased from 3 to 5, as seen in Figure11. The term "cardinality" was coined to define the average number of parent nodes linked to each sensor node in a WSN and their offspring. The average cardinality was calculated to be this value. This number was calculated when building the hierarchical structure of the network with multiple parents and multiple offspring. Figure11 shows that our DP approach outperforms HDA, DD systems, despite the fact that all three techniques reduce their energy loss as the network's cardinality grows. This was shown by the fact that, despite the fact that the total amount of energy spent by all three systems lowers with increasing network cardinality, this was still the case. This happened despite the fact that all three strategies share a common property, which is as follows: The reason for this is because as the number of nodes in a network increases in size (its cardinality), the quantity of coverage provided by the sensor nodes also increases. Consequently, the network generates fewer messages each time data is transferred to the sink. This resulted from the circumstance node directly.

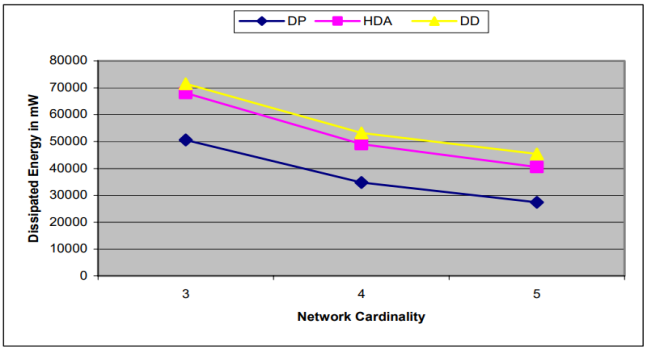

Figure 11: The amount of energy used by different network cardinality configurations

When it comes to the aggregation of data in WSNs, the DP scheme is more energy efficient than the HDA scheme and the DD scheme, as demonstrated by the findings of the analytical performance assessments that were previously reported. To achieve this objective, the DP plan was proposed.

4. CONCLUSION

It is expected that civilian applications of WSNs would usher in a revolutionary social and technological shift that will set a new paradigm for ubiquitous computing. With the advent of so many connected, embedded devices, it is now feasible to gather detailed information about a sizable population. Additionally, this may happen in the not-too-distant future. Extensive study has been

conducted in recent years on various data processing techniques, including data aggregation and data mining. Unfortunately, much more research and development is required before WSNs can realise their full potential in practical applications. Data privacy and integrity issues are two of the most significant obstacles hindering the widespread use of WSNs in civilian applications, and may lead users to conceal crucial data during data collecting. Data integrity and confidentiality concerns might be to blame. Worries about how their data would be used in the future may have influenced this decision.

REFERENCES

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Commun. Mag. 40 (8) (2002) 102– 114.
2. J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Comput. Networks 52 (12) (2008) 2292–2330.
3. K. Akkaya, M. Demirbas, R.S. Aygun, The Impact of Data Aggregation on the Performance of Wireless Sensor Networks, Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008) 171–193.
4. J. Newsome, E. Shi, D. Song, A. Perrig, The Sybil attack in sensor networks: analysis and defenses, in: Proceedings of the Third IEEE/ ACM Information Processing in Sensor Networks (IPSN'04), 2004, pp. 259–268.
5. Perrig, R. Szewczyk, D. Tygar, V. Wen, D. Culler, SPINS: security protocols for sensor networks, Wireless Networks J. (WINE) 2 (5) (2002) 521–534.
6. E. Fasolo, M. Rossi, J. Widmer, M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, IEEE Wireless Commun. 14 (2) (2007) 70–87.
7. R. Rajagopalan, P.K. Varshney, Data aggregation techniques in sensor networks: a survey, IEEE Commun. Surveys Tutorials 8 (4) (2006).
8. S. Madden et al., TAG: A Tiny AGgregation Service for Ad Hoc Sensor Networks, OSDI, Boston, MA, 2002.
9. B. Zhou et al., A Hierarchical Scheme for Data Aggregation in Sensor Network, IEEE ICON 04, Singapore, 2004.
10. M. Lee, V.W.S. Wong, An Energy-Aware Spanning Tree Algorithm for Data AggrAegation in Wireless Sensor Networks, IEEE PacRrim, Victoria, BC, Canada, 2005.
11. S. Lindsey, C. Raghavendra, K.M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, IEEE Trans. Parallel Distrib. Sys. 13 (9) (2002) 924–935.
12. G. Di Bacco, T. Melodia, F. Cuomo, A MAC Protocol for DelayBounded Applications in Wireless Sensor Networks, Med-Hoc-Net, Bodrum, Turkey, 2004.
13. W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Trans. Wireless Commun. 1 (4) (2002) 660– 670.
14. Y. Yao, J. Gehrke, The Cougar approach to in-network query processing in sensor networks, ACM SIGMOD Rec. 31 (3) (2002) 9– 18.

Corresponding Author

Phakade Shirish Vitthalrao*

PhD Student