An Experimental Performance Evaluation of Bluetooth Mesh Technology for Monitoring Applications

Eduardo De Leon*, Majid Nabi*
*Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
Email: e.a.de.leon.romero@student.tue.nl, m.nabi@tue.nl

***Abstract*—The introduction of Bluetooth Low Energy (BLE) in 2010 provided constrained devices with a wireless point-to-point communication standard. It facilitates the creation of pico-nets and reducing product development time and cost. It is until 2017 that the Bluetooth special interest group releases the Mesh Profile allowing a multi-hop interconnection through BLE's advertisements. Being a relatively new technology, this paper aims to experimentally evaluate its performance and investigate the limits of the technology in terms of data delivery capacity in monitoring applications. Several experiments are performed by deploying a number of BLE nodes in an office environment, making a multi-hop network. The performance of the network in terms of packet delivery to a base station is measured in each experiment. Moreover, experiments including mobile nodes are carried out under the multi-hop setup to test the behaviour of the protocol when some nodes move around. The experimental results show that the relay nodes impose critical limitations for message delivery in multi-hop networks, limiting the usage of the BLM technology for many monitoring applications.**

***Index Terms*—Bluetooth mesh, BLM, BLE, Multi-hop, Performance, Monitoring.**

I. Introduction

The Bluetooth Low-Energy (BLE) [1] technology was released in 2010 as a part of the v4.0 Core Specification [2] of the Bluetooth technology for short-range low-power wireless communication. It featured a highly standardized, robust and low-cost connectivity solution operating in the 2.4 GHz ISM band for power constrained devices. Nevertheless, BLE topology was limited to point-to-point and pico-nets (one master with a limited number of slaves) until the addition of the Bluetooth Mesh (BLM) profile [3] in 2017.

The BLM v1.0.1 specification [3] defines the mesh profile as a networking technology built on top of BLE. It allows up to 32,767 nodes and an 11-byte message payload for communication with a best-case throughput of 110 *bytes/sec*. Peer-to-peer interconnection is achieved with the aid of BLE's advertisements by operating under a controlled flooding technique for message relaying. The biggest advantage of this technique is its simplicity, as it does not require complex routing tables, thus reducing processing and memory requirements for devices. Also, it provides some level of robustness because of inherent redundancy of flooding.

Being a relatively new technology, not enough research has been carried out to test the performance of a BLM network under a multi-hop mobile scheme. The technology has been designed around BLE's key features and aims to provide an inter-operable, scalable, reliable and secure way of interconnecting many devices. Nevertheless, its performance may be limited by the technology's own design. Considering its flooding mechanism, and its reduced bandwidth of only the BLE's advertising channels, there are serious doubts whether it can be suitable for monitoring applications. For these reasons, it is valuable to implement a real-world representative BLM network and measure its performance to understand its potential and limitations.

This paper reports the results of several real-world experiments of BLM networks consisting of 33 nodes distributed in a 500 m^2 office area. A convergecast network is considered in which all nodes periodically generate data packets and propagate them towards a base station. Experiments including multiple mobile nodes are also performed. The results are analyzed to investigate the correlation between the provided packet delivery performance by the BLM technology and network characteristics such as data generation rate of the nodes, or nodes' distance to the base station. It is clear from the results that the BLM technology has serious limitations for large-scale monitoring applications in which the nodes have frequent data generation, while it can be a promising option for networks with low traffic load (e.g., event-driven applications).

The rest of the paper is organized as follows. Section II briefly introduces the BLM technology, discussing its node composition, features, and managed flooding mechanism. The related investigations and performance evaluation efforts are discussed in Section III. The methodology used to setup the network and calculate performance metrics under different data sampling rates is explained in Section IV. Section V presents and analyzes the results of the experiments. Section VI concludes.

II. BLE-Mesh Background

The Mesh profile is a *message-based* networking technology working under a *publish-subscribe* paradigm that sits on top of BLE. Some of the main characteristics of the BLM specification are as follows.

Nodes composition: a node's inner logical composition may consist of multiple parts called *elements* that implement standardized behaviors (or software-components) called *models*. Each of these models shares information through messages that trigger inner *states* of the model. These messages usually fall into either: GET, SET or STATUS categories.

Features: in addition to sending and receiving messages, a node may support the following soft-features:

- Relay: enables message forwarding on the nodes, thus allowing messages to transverse the entire network on a *multi-path* peer-to-peer interconnection.
- Proxy: enables communication between the mesh network and a non-mesh-supported BLE device through a BLE GATT.
- Friend: serves as a *broker* for a Low-Power Node (LPN) by buffering messages destined to the LPNs with which it has established a *friendship*.
- Low-Power: allows a node to stay asleep most of the time until it requires interaction with the network by establishing a *friendship* with a friend node.

Managed flooding: the BLE-Mesh profile specification defines a flooding technique for message relaying [4]. It is a broadcasting algorithm that does not require complex routing tables and provides a resilient network due to its possibly high redundancy with self-healing and multi-path capabilities. It is similar to a gossiping and controlled flooding algorithms which reduces packet forwarding in a probabilistic and deterministic manner respectively. In this case, relaying of messages is minimized by the following methods:

- Time-To-Live (TTL): restricts the number of *hops* a message can make through *relay* nodes before arriving to its destination.
- Message Caching: prevents unnecessary re-relaying of packets by keeping track of the previously processed messages, so they can be ignored or otherwise forwarded.

BLM supporting nodes communicate by using the GAP *advertising bearer* on the advertising channels 37, 38 and 39. This means that nodes switch between the advertising and scanning states of BLE's link layer as shown in Fig. 1. The scanner node defines the *scanInterval* as the interval by which it switches and selects a different channel to listen to. The *scanWindow* defines the time the scanner's radio is turned on to passively receive messages sent in the channel in which it is currently listening. To have a higher probability of messages reaching an observer during an advertising event, the advertiser node sends the same packet on each one of the three advertising channels in a sequential manner with an offset of at most 10 ms. Moreover, for consecutive broadcasts or message forwarding, a minimum advertising interval of *AdvInterval* $= 20$ ms is specified by Bluetooth 5.0+ controllers. An additional pseudo-random advertising delay (*advDelay*) from 0 to 10 ms is also generated for every *advertising event* to reduce the probability of collisions. It is also important to consider the maximum message rate of 100 Protocol Data Units (PDU) in any 10 $sec.$ window [4], and a maximum of 88 bits of useful payload in a mesh-message that give us a throughput of 110 $bytes/sec.$

III. Related Work

In this section, previous efforts for evaluation of the BLM technology are reviewed. These papers provide an insight on different setups, approaches and considerations for a mesh performed in an office environment. The main focus of these experimental evaluations is the performance of the network under different traffic loads.

[5] was carried out through simulation, no specifics are given regarding the virtual environment. The performance evaluation was based on a large-scale office scenario; 879 devices representing sensors and actuators laid through an area of 2,000 m^2. Relay nodes were located in open areas to improve coverage and network connection. However, redundant designs were also taken into account due to the uncertain propagation condition in an indoor space. After running simulations, the TTL and message repetition were optimized so that there would be less congestion on the network. It concludes that the best results were obtained when only 6 (1.5%) relays were deployed in a 1000 m^2 area.

[6] carries out a physical evaluation and analyzes the protocol's latency. The hardware nodes used for this experiment were EM35xx and EFR32 TM Mighty Gecko SoCs with a proprietary stack implementation. Nodes were deployed in a 2230 m^2 office area. Several tests are done using different payloads with both segmented and unsegmented messaging, network sizes up to 192 devices and up to 6 hop relaying in a real office where other radio technologies were present. One of the experiments was designed using small message size and optimized TTL values. The results show that the reliability of the network is greater than 99% under the following conditions: Unsegmented messages and low relay count; "When network size and number of hops increase, relay selection becomes critical for network performance" [6]. A bigger payload results in segmentation which directly affects latency and, for this reason, multi-casting should avoid using message segmentation.

[7] measures the performance of BLM in terms of latency through theoretical modeling and physical experimentation on a 1000 m^2 space. The experiment is carried out using 22 Nordic's nRF52832 SoCs under an acknowledge unsegmented messaging scheme. Measurements were done considering the Round Trip Time (RTT) against number of relay nodes and

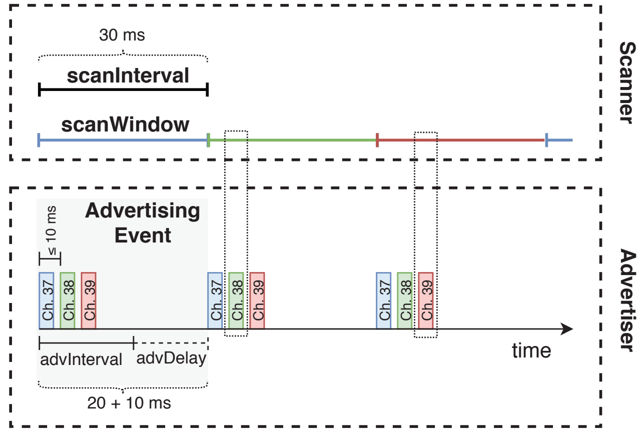

Fig. 1. BLM Scanning/Advertising GAP event - Zephyr configuration

number of hops. Conclusions state the positive impact of a proper backoff mechanism as well as network density in RTT [7]. They also point out the fact that *relay* features as the BLM backbone rely on non power-limited devices which makes the deployment of pure low-power devices unrealistic at this current state.

[8] developed an indoor localization experiment for hospital floors by utilizing BLM advertisements with the iBeacon packet structure containing Universally Unique Identifiers (UUID). Same as in this evaluation, Zephyr's OS stack implementation was used since it is an open source operating system which allows customization. Proximity, trilateration and fingerprinting localization algorithms were used in order to test their experiment. The limitation imposed by the specification is highlighted, which affects accuracy of the inferred position. This is because it is not possible to take the necessary amount of measurements required by the algorithms.

This paper performs an experimental evaluation for monitoring applications aiming to investigate the limits of BLM. Compared to [5], where a simulation is carried out, in this paper a real-world setup is deployed. Furthermore, unlike [6], which bases its analysis on latency measurements, this paper mainly focuses on packet delivery ratio and burst drops as performance metrics. Specifically, the focus is on the cases that multiple mobile nodes are present in coexistence with other sensor nodes and other wireless devices such as WiFi. As in [8], Zephyr's BLM stack is used and, by using mesh models, traffic generation is controlled. Additionally, considering [5], the amount of relays is minimized according to the physical limitations of the environment.

IV. Experimentation Methodology

For a monitoring application, requirements such as ubiquitous reliable communication and sampling rates may vary according to particular use-cases. Therefore, before it can be determined if BLM is suitable for such applications, experimentation is required to measure its performance and capacity to reliably deliver messages. By carrying out a real-world deployment, we encounter the same conditions that real applications may have. This allows us to have a solid ground when defining the limits of the technology which will ultimately determine the suitability for specific monitoring applications.

A. Experimental Setup

The evaluation is carried out by deploying a mesh consisting of 33 micro:bit [9] nodes programmed with Zephyr's BLM stack in a $500\ m^2$ office area. Periodic advertisements of 8 bytes data payload messages with a total packet size of 47 bytes is done by nodes in order to create traffic. Approximately, 500 messages containing the nodes' individual message counter are sent during each experiment. This counter is then used to measure a node's performance. In order to reduce interference between radios, transmission power has been fixed to $-4\ dBm$ and $0\ dBm$ for broadcasters and relay nodes respectively. The default advertisement and scanning settings for the Zephyr's stack implementation is shown in Fig. 1. Every message, regardless of the type of node sending it, is transmitted on three advertising events. Also note that the default *advDelay* is fixed to $10\ ms$ rather than a pseudo-random value, and that both *scanInterval* and *scanWindow* have the equal values resulting in a continuous channel scanning.

For the purpose of this evaluation, three different types of roles where devised and programmed into the nodes: infrastructure nodes solely consisting of the relay feature, generic sensors represented by advertising nodes, and a base station acting as the controller and data-logger of the whole network. Fig. 2 shows the office floor where the experiments take place. It shows the physical location of nodes as well as zones that different nodes are located in. Following is a brief description of the two major types of experiments carried out during the evaluation. Note that each experiment type is performed several times with different data generation rates (publish rates).

- **Multi-hop Static** experiments in which the nodes are fixed along the office space (Fig. 2). A total of 29 generic sensor nodes broadcasting at the same interval, 3 relay nodes to provide with message multi-hopping capability and a *base station* are installed.
- **Mobile** experiments in which three mobile nodes move together around the floor while the fixed nodes publish messages every 5 seconds. Depending on the nodes' distance to the base station, messages coming from these nodes may need to be forwarded by the relay nodes.

B. Performance Metrics

To analyze the performance of the mesh, the following metrics are focused as they give us an insight on the message delivery capacity from nodes and relays to the base station.

- End-to-end Packet Delivery Ratio (PDR): of a node is the number of successfully delivered data packets to the base station over all generated data packets by that source node in a time window. Eqn. 1 gives the PDR for source node s at time τ ($PDR^{\tau}(s)$) over a window of last H packet generation by node s.

$$PDR^{\tau}(s) = \frac{1}{H}\sum_{k=0}^{H} D^{\tau,k}(s) \tag{1}$$

  $D^{\tau,k}(s)$ is 1 if the k^{th} packet before time τ originated from node s has reached the base station. Note that, in this case, we are interested in the one-way communication from sensor nodes to the base station, as simple sensor nodes only broadcast measured data. It provides us the means to identify poor links caused by physical and/or technological limitations.
- *Burst Drops (BD)*: represents the number of packets losses in a row. It is an important metric as it directly affects the quality-of-service provided to the application. If this value is high for a specific node, we cannot rely on the network for a continuous monitoring application

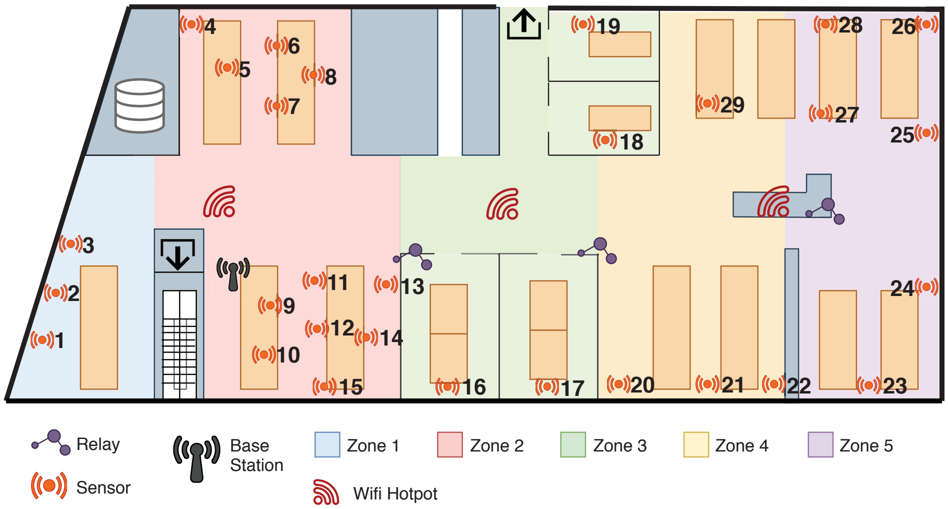

Fig. 2. Node placement in the experimental setup. This figure shows the blueprint of the office floor divided into color-coded zones.

since we cannot expect connectivity at all times. We are specifically interested to observe the maximum burst drop for different nodes in each experiment since this shows the worst case disconnection of the sensor nodes. A long period of not receiving updates/commands may lead to an application failure.

V. Results Analysis

In this section the results for different experiments are presented and discussed. For each of the experiments' types, details on PDR and BD for each node are shown in the corresponding graphs.

A. Multi-hop Static Experiments Results

The main objective of these experiments is to understand the network's behavior under a multi-hop communication and tuning it accordingly in preparation for the mobile experiments. The value of TTL for every node is set to 7 across all nodes regardless of their relative distance to the *base station*. Figures 3 and 4 present measured PDR and BD values for individual nodes for experiments with publish period of 5 sec. and 1 sec., respectively. For burst drops, the box plot of all burst packet drop instances happened during the experiments is shown to better observe their distributions. The bars and box-plots are colorized according to the average hop distance of the node (also numerically shown by the magenta numbers on top of the PDR bars).

For the first experiment, publish period for all nodes is fixed to 5 seconds. We can see at a first glance from Fig. 3 how, in general, the nodes located farther away in terms of the number of hops, already have a lower and fluctuating PDR compared to the ones located closer to the base station. It can be clearly seen how nodes with the highest PDR are close with a single hop distance. In accordance with BD results, the need to be relayed increases the probability of dropping a message. Farther a node from the base station is, the greater the burst drops may be due to congestion at the relay nodes. Nevertheless, whenever a burst drop occurs, on average, the next message reaches its destination. Some cases in which a physical obstacle may affect message reception are the case for nodes 1, 17 and 19 as we can see from their physical location in Fig. 2. On the other hand, nodes in Zone 5 with a clear line of sight (e.g., nodes 25 and 27) may be able to reach the base station through a single hop.

In the second experiment, nodes' publish period is set to 1 message per second and results are displayed in Fig. 4. The most evident difference in results compared to the previous experiment is the dramatic decrease in PDR values for the nodes that require more than one hop to reach the base station. The reason behind this decrease is the relays' limited capacity to process and transmit messages. If every node reaches at least a relay (assuming a best-case scenario), all relays would have to forward 29 messages per second due to the managed flooding mechanism. Furthermore, the results reveal how relay saturation is reflected on the average hop distance; basically any node with a single hop connection to the base station has an average of exactly one hop. A possible reason for this is that one of the three advertising packets coming from the nodes is more likely to be captured by the base station first, rather than a message being relayed.

The current stack implementation of BLM does not allow scanning preemption, thus limiting the capacity of scan-advertise for relays. For the current configuration, relays can manage at most 8 messages per second, which means that the nodes located at more than one hop distance will perceive severe burst drops. Such is the case for most nodes located outside of Zone 1 and 2 (refer to Fig. 2) beside nodes 18 and 25 which managed to reach the base station directly. On the other hand, nodes 17 and 27 have rather low PDR although they have low average distance. A possible reason for this may be the physical obstacles and RF environment being such that their messages either reach the base station within a single hop, or drop because of collisions.

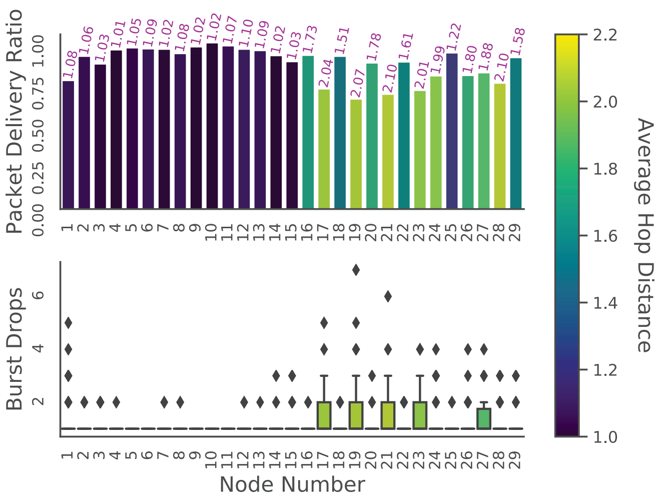

Fig. 3. Results for the multi-hop static experiment with 5 sec. publish period

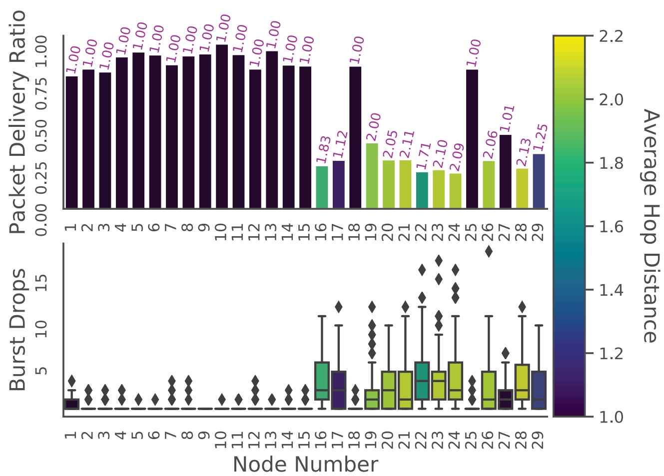

Fig. 4. Results for the multi-hop static experiment with 1 sec. publish period

B. Mobile Experiments Results

During this set of experiments, nodes 1, 2 and 3 (previously located in Zone 1 shown in Fig. 2) are carried together around the floor by a person (assuming a wireless body area network). The other nodes are kept static and set to a publish period of 5 seconds. Nodes located in Zone 2 are given a TTL of zero, whereas the rest are set to a TTL of 4 in order to reduce relay saturation. The mobile experiments are repeated with different message publish rates for mobile nodes (1 sec., 500 ms, 300 ms, and 200 ms).

Figure 5 shows the overall results of multiple experiments to provide a comparable view between them. Single-hop and multi-hop nodes' overall PDR are shown separately for each of the experiments. The individual PDR of each one of the mobile nodes is also shown as red dots in Fig. 5 for a clearer

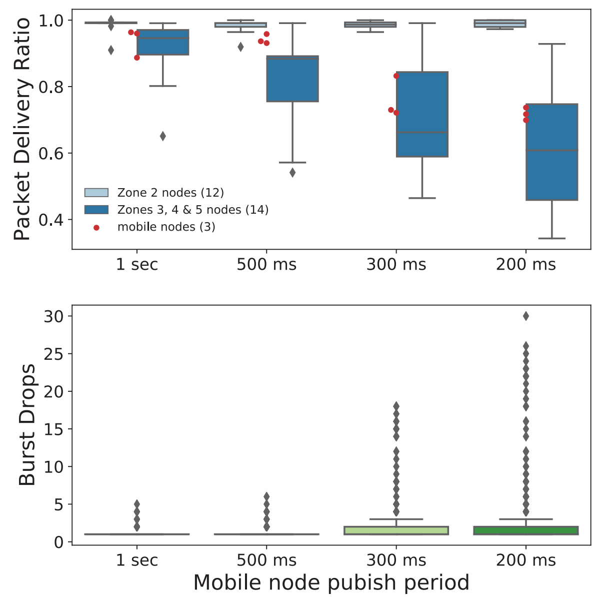

Fig. 5. Achieved Packet Delivery Ratio and Burst Drops in the Mobile experiments for different publish periods of the mobile nodes (static nodes have publish period of 5 sec.).

distinction. One of the major characteristics of these experiments is the PDR of the nodes in Zone 2 being greater than 95% throughout every experiment, basically being unaffected regardless of the mobile nodes' publish period. On the other hand, for experiments where the mobile publish period is set to 300 ms and 200 ms, the generated traffic reaches 12.8 and 17.8 messages per second, respectively. These messaging rates cause relay nodes to exceed their capacity of managing messages resulting in a considerable increase in burst drops for nodes located in Zones 3, 4 and 5.

Further detailed information about the mobile nodes' PDR and hop-distance over time while moving through the different zones is displayed in Fig. 6. From this figure, it is possible to correlate individual PDR values and the average hop distance computed from the gathered data over time. For 1 sec. and 500 ms experiments, it can be observed that moving farther away from the base station does not have a substantial impact on PDR, thanks to the proper functionality of the managed flooding mechanism. Also, it can be seen how the average hop distance increases when the mobile nodes reach Zone 3 and then continues until reaching Zone 5. However, during the 300 ms and 200 ms experiments, it can be clearly observed that Zone 2 is the only place where PDR for mobile nodes is still stable. When the nodes reach Zone 3, PDR starts decaying dramatically as the relay nodes are not able to manage the traffic load. A keen eye can see that during the 300 ms experiment, "Mobile 3" crashes at approximately 7^{th} minute from the start of the experiment. For this reason, we can see in Fig. 5 a red dot (Mobile node) having a significantly higher PDR compared to the other mobile nodes.

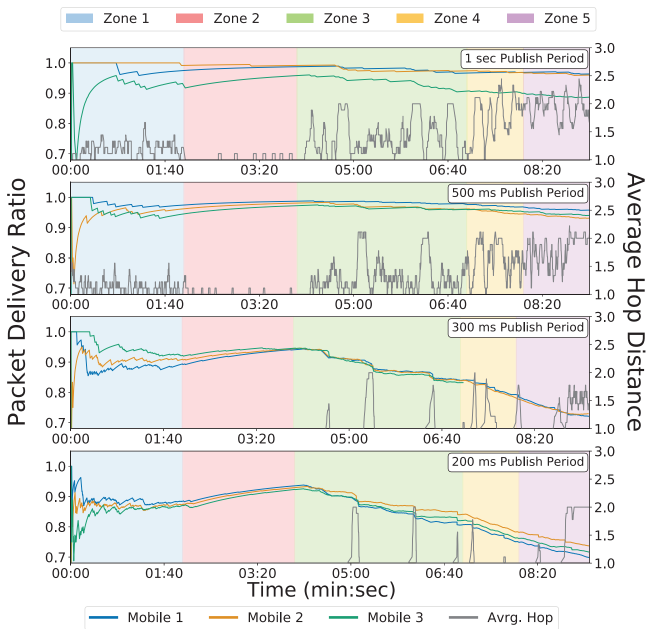

Fig. 6. Mobile Experiment Packet Delivery Ratio Over Time results

C. Discussion

The experiments in this paper are designed using a network which included static nodes as well as multiple mobile nodes in order to investigate a ubiquitous monitoring use-case. One of the main data delivery limitations found during this evaluation is that a node can send maximum 100 network PDUs in a 10 second window, each containing maximum 8 bytes of data payload. This gives us a best-case goodput of 80 bytes per seconds for each node. Furthermore, Zephyr's current BLM implementation in the particular configuration used in this work limits a node's publish period to at least 120 ms (or a maximum of 8.33 messages per second) due to the scanning and advertising non-preemptive task execution. For these reasons, special considerations for multi-hop communication should be taken to achieve a proper managed flooding functionality.

The results of the experiments shows that by utilizing BLE's advertising mechanism as in the current BLM standard, a robust wireless communication between low-rate devices is achieved for low message publish rates. Nevertheless, as of BLM Profile Specification v1.0.1, the technology relies mainly on less power-constrained devices (relays) to set its infrastructure due to the required continuous channel scanning. However, the technology allows low-power devices to join the network by utilizing the *friendship* mechanism. Considering the previously mentioned limitations of the BLM profile, the success of any implementation depends on the specific data rate and power requirements of the application.

VI. Conclusion

This paper evaluates the performance of the Bluetooth Mesh (BLM) technology, and investigates its limits for monitoring applications through real experiments. We encountered some of the current data rate limitations of the technology. The results clearly show that, to achieve a reliable packet delivery service, relay nodes should not get saturated. Otherwise, the managed flooding will not be effective. The experiments in this paper suggest that the BLM is better suited for very low-rate applications such as a sporadic or event driven monitoring. For monitoring applications with rather high periodic data generation rate by the nodes, the technology fails to provide an acceptable packet delivery performance. As a future work, the impact of the Time-to-Live (TTL) parameter of the flooding mechanism, and sub-networking feature on the packet delivery performance will be explored.

Acknowledgment

This work was supported by the SCOTT European project (www.scott-project.eu) that has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737422. The authors thank ALTEN Nederland B.V. for providing the means to carry out this evaluation.

References

[1] "Bluetooth Technology Website," accessed: 2019-10-17. [Online]. Available: https://www.bluetooth.com/bluetooth-technology/radio-versions/
[2] "Bluetooth Technology Website," accessed: 2019-07-04. [Online]. Available: https://www.bluetooth.com/about-us/our-history/
[3] "Bluetooth Technology Website," accessed: 2019-10-17. [Online]. Available: https://www.bluetooth.com/specifications/mesh-specifications/
[4] *Bluetooth Mesh Profile Specification*, vol. v1.0.1, mesh Working Group.
[5] P. Di Marco, P. Skillermark, A. Larmo, and P. Arvidson, "Bluetooth Mesh Networking," p. 13, accessed: 2017-07. [Online]. Available: https://www.ericsson.com/en/white-papers/bluetooth-mesh-networking
[6] Silabs, "AN1137: Bluetooth Mesh Network Performance," p. 18. [Online]. Available: https://www.silabs.com/documents/public/application-notes/an1137-bluetooth-mesh-network-performance.pdf
[7] "The Bluetooth Mesh Standard: An Overview and Experimental Evaluation," vol. 18, no. 8, p. 2409.
[8] M. Jrgens, D. Meis, D. Mllers, F. Nolte, E. Stork, G. Vossen, C. Werner, and H. Winkelmann, "Bluetooth Mesh Networks for Indoor Localization," in *2019 20th IEEE International Conference on Mobile Data Management (MDM)*, pp. 397–402.
[9] "Micro:bit Educational Foundation — micro:bit," accessed: 2019-09-09. [Online]. Available: https://microbit.org/