1

Edge AI LoRa Mesh Technologies

Ovidiu Vermesan[1], Kai vorm Walde[2], Roy Bahr[1], Cordula Conrady[2], Janis Judvaitis[3], Gatis Gaigals[3], Tore Karlsen[4], Marcello Coppola[5], and Hans-Erik Sand[6]

[1]SINTEF AS, Norway
[2]IMST GmbH, Germany
[3]Institute of Electronics and Computer Science, Latvia
[4]ProLux AS, Norway
[5]STMicroelectronics, France
[6]NXTECH AS, Norway

Abstract

Intelligent connectivity at the edge combines wireless communication, edge artificial intelligence (AI), edge computing and internet of things (IoT) technologies to perform machine learning (ML) and deep learning (DL) on connected edge devices. Low latency, ultra-low-energy intelligent IoT devices with on-board computing, and a distributed architecture and analytics are essential to drive intelligent connectivity.

Intelligent wireless mesh technologies exploit multiple interconnected devices, or nodes, to create a distributed network integrated with edge AI analytics using ML and DL algorithms. In an intelligent wireless mesh network (WMN), each node has embedded intelligence and can communicate directly with its neighbouring nodes and transfer data efficiently to other nodes. Compared with traditional point-to-point wireless networks, the intelligent wireless mesh approach offers several advantages, including increased coverage, redundancy, scalability and resilience.

The convergence of multiple technologies (connectivity, edge AI, IoT, distributed architectures and federated learning) delivers intelligent edge

mesh communication systems that perform efficient connectivity by optimising data rates, coverage, energy, and interference.

This article overviews the latest advancements in edge AI long-range mesh technologies and applications, highlights state-of-the-art mesh communication requirements and implementations and identifies future research challenges and directions.

Keywords: mesh communication technologies, edge artificial intelligence, LoRaWAN, LoRa mesh.

1.1 Introduction

Star, tree and mesh networks are examples of topologies used in communication networks. Each is suitable for different application scenarios. An illustration of the different network architectures is shown in Figure 1.1.

Star networks are simple to set up and manage because they have centralised control points. However, this makes them more susceptible to single-point failures. Mesh networks offer high redundancy and self-healing (e.g., recovery from a link failure), making them more reliable and fault tolerant at the cost of increased complexity.

Wireless mesh technologies play an essential role in creating robust and flexible wireless networks that address modern connectivity challenges.

In a star topology, all nodes are directly connected to a single central root node, often referred to as a hub. Direct peer-to-peer communication is not supported; all nodes must communicate through this central hub.

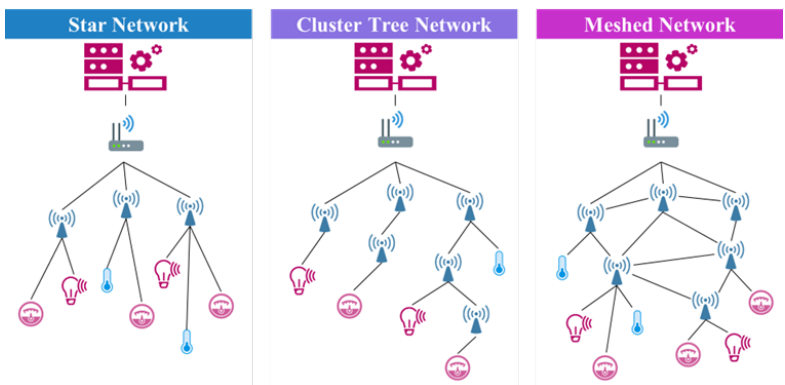

Figure 1.1 Network Topologies

Networks with cluster-tree topologies are divided into so-called clusters. Each cluster consists of a group of nodes connected to a local central node referred to as the cluster head. The cluster head coordinates the communication within its cluster. The tree terminology refers to the cluster heads which are organised in such a hierarchical structure. Communication from node to node may involve routing via multiple cluster heads.

Wireless mesh technologies use multiple interconnected nodes which can communicate directly with their neighbours nodes. This approach offers several advantages, including increased coverage, redundancy, resilience, and scalability.

Mesh communication technologies use distributed networking methods that typically create a decentralised and self-configuring network. Each node can also act as a repeater to extend network coverage and improve resilience.

The convergence of edge computing, edge AI, federated learning and IoT can create multi-dimensional architectures consisting of a wide range of heterogeneous entities with different sensing/actuating, connectivity, processing, and intelligence capabilities connected with applications in a dynamic mesh network linked by platforms and distributed services located at the edge level. Some of the technologies contributing to enhancing the capabilities of intelligent mesh connectivity include:

Edge AI - The deployment of AI algorithms and data processing capabilities directly on edge devices, rather than relying on centralised cloud servers, brings the following benefits:

- Real-time decision making – By processing data locally, AI models can make fast decisions without the latency of sending data to remote servers, enabling rapid responses to critical events.
- Data privacy and security – Edge AI reduces the need to transmit sensitive data to the cloud, increasing privacy and decreasing the consequence of data breaches.
- Bandwidth efficiency – Edge AI can filter and prioritise data before transmission, reducing bandwidth demands.

Federated Learning (also called collaborative learning) is a machine learning method in which edge devices collaboratively contribute to a global model while keeping their data locally. Federated learning can play a significant role in enhancing the combination of edge AI, IoT and communication technologies.

Privacy preservation: Federated learning avoids the transmission of raw data to a central server. This ensures that sensitive data remain on the edge devices, addressing privacy concerns and complying with data protection regulations.

Bandwidth efficiency: By training models locally on edge devices, federated learning reduces the need to send large amounts of data to the cloud for model training. This optimises bandwidth usage, making it more efficient for IoT devices with limited communication capabilities, such as long-range (LoRa)-based communication devices. LoRa is a wireless spread spectrum modulation technique derived from the chirp spread spectrum (CSS), which enables long-range communication between devices with low power consumption. The technology was initially developed by a company called Cycleo SAS and later acquired by Semtech Corporation [1], a semiconductor company specialising in analogue and mixed signal circuits.

Improved model performance: Federated learning allows IoT devices to continuously improve their local models. This can result in better model performance and adaptability over time, as each device benefits from the collective intelligence of the entire network.

Decentralised intelligence: Federated learning distributes intelligence across edge devices, promoting decentralised data processing and decision-making. This leads to increased resilience in the overall system.

Collaboration and knowledge sharing: By collaborating on model training, edge devices share knowledge and insight. This collaborative approach fosters rich and diverse learning experiences.

Reducing infrastructure costs: Federated learning reduces the need for large-scale cloud infrastructure for centralised model training. This results in cost savings in respect of data transmission and cloud computing resources.

Versatility and scalability: Federated learning can be adapted to many different edge devices and network architectures. It can scale efficiently making it suitable for IoT networks with diverse deployments and configurations.

Federated learning complements the combination of edge AI, IoT and LoRa by enhancing privacy, efficiency, model performance and collaboration. It empowers IoT networks with intelligent decision-making capabilities while respecting data privacy and promoting decentralised data processing.

Internet of Things (IoT) is related to the network of interconnected devices and sensors that collect, exchange, and analyse data. By integrating IoT with edge AI and LoRa technology, it becomes a powerful enabler across various domains:

- Remote monitoring and control – IoT sensors can collect data from different environments, enabling remote monitoring and control of processes, infrastructure, and assets.
- Predictive maintenance – IoT data, when combined with edge AI analytics, allows the prediction of equipment failures, optimisation of maintenance schedules and reductions in downtime.
- Energy management – IoT deployments combined with edge AI enable efficient energy management, waste reduction and improved urban services in smart city applications.

The combination of edge AI and mesh communication has several benefits, especially when infrastructure is impracticable or unavailable. Mesh networks enable flexible, reliable, and scalable networks. They are increasingly used in industrial IoT, energy, smart homes, agri-food and beverage, disaster recovery operations and smart city applications.

This chapter is organised into the following sections. Section 1 introduces the research area and the state of play of technology development. Sections 2 and 3 provide an overview of the state of the art of existing wireless mesh technologies and their primary functions, operating characteristics and actual advantages and disadvantages. Section 4 describes the LoRa wireless modulation technique and the long-range wide area network (LoRaWAN) technology, the main architectures, the architectural building blocks, and their characteristics. Section 5 covers enabling technologies (e.g., edge AI, edge computing, internet of intelligent things, artificial intelligence of things) and integration with LoRa mesh to enhance and optimise communication performance and mesh-based systems' collaborative and cooperative capabilities. Section 6 presents potential applications for LoRa mesh connectivity, edge AI and IoT systems and emphasises the requirements for intelligent communication and convergence with other technologies. Section 7 outlines the conceptual edge AI LoRa mesh device architecture. Section 8 analyses the state of play and future research directions and highlights several challenging open issues for intelligent edge LoRa meshes. Finally, Section 9 summarises the main points for discussion.

1.2 Overview of the State-of-the-Art Wireless Mesh Technologies

Meshes are networks that create a decentralised and robust structure where each node can communicate directly with neighbouring nodes.

Nodes are interconnected and, depending on the network topology, there can be multiple connection pathways for each node. Connections between nodes may be dynamically updated and optimised through a built-in mesh routing table. As nodes enter and exit the network, the mesh topology enables the nodes to reconfigure routing paths based on the new network configuration.

Mesh topology and ad-hoc routing assures stability in the face of changing communication conditions or node failure.

Mesh networks use a distributed approach, where each node can act as a repeater to extend network coverage and improve resilience. The critical characteristics of mesh communication technologies include:

- **Decentralisation** – mesh networks are not dependent on a single central point of control. Each node can communicate with its neighbour, allowing messages to bounce from one node to another until they reach their destination.
- **Self-configuration** – mesh networks are capable of self-organisation. When nodes are added or removed the network can dynamically reconfigure itself to accommodate these changes.
- **Redundancy and reliability** – due to their decentralised nature and self configuration capability, mesh network topologies are more resilient to node failure or network disruption.
- **Extended coverage** – mesh networks can cover an extended area by using multiple nodes as relays. This provides an advantage in cases when establishing a traditional infrastructure might be challenging or costly.
- **Ad-Hoc networking** – mesh communication technologies enable ad-hoc networking, where devices can spontaneously create a network without relying on pre-existing infrastructure.
- **Geographical scalability** – mesh networks can quickly expand their coverage by adding more nodes which do not need to be in direct communication.

1.2.1 Mesh components and roles

Wireless mesh networks usually consist of routers, nodes, and coordinators as described below:

- **Routers** – these devices form the backbone of a wireless mesh network. They are typically more powerful than simple nodes with enhanced processing capabilities and are responsible for routing data within the

whole network. Mesh routers communicate with other routers and nodes in the network to forward data packets along the most efficient path to reach their intended destination.

- **Nodes** – these are individual devices connected to the mesh network. They can be computers, smartphones, sensors, IoT devices, or any other device capable of wireless communication. Mesh nodes are typically senders, receivers, or relay points. Unlike traditional networks, mesh nodes in a wireless mesh network can communicate directly with each other, creating multiple data transmission paths. This decentralised communication architecture enhances the network's reliability and overall performance.
- **Coordinators** – mesh coordinators are nodes with specialised roles in some wireless mesh network protocols. They act as central control points for the entire mesh network. A coordinator is responsible for managing and organising the network, assigning roles to other nodes (such as routers or end devices), and controlling aspects of the network's operation. They handle tasks like channel allocation, network formation, and security management. In some mesh network implementations, coordinators have a critical role in preserving the network's stability and performance. On one hand, central coordinators can offer efficient control and coordination; on the other hand, they can also become a single point of failure, potentially disrupting the entire network and compromising one of the key advantages of mesh topologies.
- **Decentralised functionality** – this approach eliminates the central mesh coordinator. Instead, the process of decision-making and control is distributed across multiple nodes. Nodes may possess a degree of autonomy, enabling them to make local decisions based on independent observations and interactions with neighbouring nodes. Local decisions collectively contribute to the overall behaviour of the network.

1.2.2 Wireless routing concepts

One of the key elements for wireless mesh communication, routing protocols are designed to enable communication and data exchange between devices in a wireless network. These protocols establish routes for data transmission and determine the best paths for information to flow from a source to a destination. The functions of a wireless routing protocol vary depending on the specific protocol used and the type of wireless network. We present a

general overview of the common functionalities of these wireless routing protocols:

- **Neighbour discovery** – in wireless networks, devices must discover neighbours to establish direct communication links.
- **Route discovery** – when a device wishes to send data to another device, a route discovery process is initiated. During the process, the device searches direct links or for potential intermediate devices (routers) that can relay the data towards the destination. This process can involve broadcasting or multicasting route request packets to nearby devices to find potential routes.
- **Route maintenance** – once a route is established, the routing protocol is responsible for maintaining the health and stability of it. This includes monitoring the status of the intermediate devices along the path and detecting any changes, such as link failures or device mobility. If a route becomes unavailable, the routing protocol triggers a route repair process to find an alternative path.
- **Routing metrics** – wireless routing protocols use various metrics to determine the quality and efficiency of potential routes. Metrics include signal strength, link quality, distance, and available bandwidth. The routing protocol uses these metrics to select the preferred routes based on network conditions and requirements. The current battery state of a node may also be a metric to implement a kind of energy-balancing policy.
- **Data forwarding** – once a route is established, the data packets are forwarded from one router to the next until they reach their destination. Each router in the path makes a forwarding decision based on the routing table and the packet's destination address.
- **Adaptation to network changes** – wireless routing protocols are constructed to adapt to changes in the network topology, such as device mobility, link quality fluctuations, or node failures. They continuously monitor the network and adjust the routing paths to ensure reliable and efficient data transmission.

1.3 Routing protocols

Some standard wireless routing protocols, include Optimised Link State Routing (OLSR) [29][30][31][33], Ad hoc On-Demand Distance Vector

(AODV) [34][35], Dynamic Source Routing (DSR) [36][37] and Routing Protocol for Low-Power and Lossy Networks (RPL) [38][39][40]. Each protocol has specific features, advantages, with use cases tailored for different wireless networks and applications. There follow some details about the algorithms and their pros and cons.

1.3.1 Ad hoc on-demand distance vector (AODV)

AODV is a demand-driven reactive wireless routing protocol that establishes routes only when needed. When a source node requests to send data to a destination node, it initiates a route discovery process to find the most efficient path. The protocol uses sequence numbers to ensure loop-free routes and maintains a routing table to store information about discovered routes.

Pros:

- **Reduced overhead** – AODV minimises control message overhead by initiating route discovery only when necessary. This helps conserve network resources and reduces unnecessary traffic.
- **Loop-free routes** – using sequence numbers ensures that routes are loop-free, improving route stability and reliability.
- **Proactive link failure detection** – AODV employs proactive link failure detection to quickly identify failed links and initiate route repair, ensuring data continues to flow via alternative paths.
- **Scalability** – AODV performs well in moderately sized networks and maintains route information for frequently used paths, reducing route discovery latency.

Cons:

- **High latency for new routes** – AODV's on-demand route discovery process can introduce delays in finding a new route, especially in large networks or sparse topologies.
- **Route rediscovery** – several cases (link changes, node mobility, malicious nodes, battery depletion, network congestion or topology changes) lead to frequent route rediscovery, increasing control message overhead.
- **Suboptimal routes** – sometimes, AODV may not find the shortest path in specific network scenarios, leading to less efficient data transmission.

AODV balances control message overhead and route discovery latency, making it suitable for dynamic networks with changing topologies. However, its performance may vary depending on network size, mobility patterns, and the frequency of route changes.

1.3.2 Optimized link state routing (OLSR)

OLSR is a proactive routing protocol that uses a hybrid approach, combining both proactive and reactive mechanisms. It optimises link-state information exchange to minimise overhead while ensuring efficient route computation and maintenance. OLSR uses Multi-Point Relays (MPRs) to reduce control message flooding and speed up route discovery.

Pros:

- **Reduced control message overhead** – OLSR uses MPRs to limit the number of nodes participating in control message dissemination. This decreases control overhead and improves scalability, making it suitable for large networks.
- **Proactive and reactive hybrid approach** – OLSR combines proactive link-state information with reactive route discovery. It provides real-time responsiveness while minimising the amount of control traffic generated.
- **Loop-free routes** – OLSR guarantees loop-free routes and enhances route stability and reliability.
- **Fast route recovery** – MPRs and proactive topology updates enable quick route recovery and repair in case of link failures.
- **Better convergence** – OLSR converges quickly and efficiently, enabling devices to find optimised routes with lower latency.

Cons:

- **Memory and computation requirements** – OLSR requires storing and managing additional topology information due to MPRs. This imposes overhead which might be critical on devices with limited resources.
- **Increased initial setup overhead** – the initial setup phase in OLSR involves the exchange of control messages to determine MPRs which leads to higher overhead during network initialisation.
- **Relatively complex implementation** – compared to other protocols, the implementation and management of OLSR can be more complex due to its hybrid nature and the need to optimise MPR selection.

OLSR balances proactive and reactive mechanisms, making it suitable for dynamic networks with varying traffic patterns and topology changes. Its efficiency in controlling message overhead and quick route convergence makes it a viable choice for both small and large-scale wireless networks.

1.3.3 Dynamic source routing (DSR)

DSR is an on-demand routing protocol that establishes routes between nodes only when needed. When a source node requests to send data to a destination node, it initiates a route discovery process to find a path. The route discovery process is based on source routing, which includes the complete route in the data packet. Intermediate nodes use this route information to forward the packet to the next hop until it reaches the destination.
Pros:

- **Reduced overhead** – DSR minimises control message overhead since route discovery is initiated only when needed, conserving network resources and reducing unnecessary traffic.
- **Loop-free routes** – DSR ensures loop-free routes through sequence numbers and route caching, enhancing route stability and reliability.
- **Efficient source routing** – including the complete route in the data packet enables efficient source routing, eliminating the need for intermediate nodes to maintain routing tables.
- **Route repair** – DSR supports quick route repair in case of link failure, as the source node can initiate a new route discovery process to find an alternative path.

Cons:

- **Route discovery latency** – the route discovery process in DSR can introduce delays, especially in large networks or sparse topologies, as it requires time to find a route to a new destination.
- **Increased packet overhead** – including the complete route in the data packet leads to larger packet sizes, especially for long routes, resulting in increased packet overhead.
- **Route maintenance overhead** – frequent mobility or link changes can lead to higher route maintenance traffic, as DSR requires regular route updates to adapt to topology changes.
- **Source routing overhead** – While source routing eliminates the need for routing tables in intermediate nodes, it increases the size of data packets, which can be a concern for resource-constrained devices.

DSR offers a simple and efficient approach to routing in Mobile Ad-hoc Networks (MANETs), particularly for networks with moderate mobility and communication demands. Its reactive nature allows it to adapt to changing network conditions, while the use of source routing eliminates the need for routing tables in intermediate nodes. The trade-offs include potential

overhead from route discovery and maintenance, which should be considered when selecting DSR as the routing protocol for specific MANET deployments.

1.3.4 Routing protocol for low-power and lossy networks (RPL)

RPL is a specialised routing protocol for low-power and lossy networks (LLNs) as commonly been in IoT and wireless sensor networks. RPL is a proactive routing protocol that forms a directed acyclic graph (DAG) to route data in LLNs efficiently. It organises devices into a tree-like structure, with a root node at the top. It optimises routes using objective functions based on specific metrics, such as energy efficiency or latency. RPL is tailored for devices with limited resources, making it well suited for battery-powered IoT devices that require reliable and energy-efficient communication.
Pros:

- **Energy efficiency** – RPL is designed to minimise energy consumption in resource-constrained devices. It optimises routes to ensure that energy is conserved during data transmission, thus prolonging the battery life of IoT devices and the entire IoT system.
- **Adaptability to LLNs** – RPL's tree-like DAG structure is well-suited for LLNs, where devices may have limited processing power and intermittent connectivity.
- **Objective function flexibility** – RPL allows network designers to choose different objective functions based on their specific requirements, such as energy efficiency, latency, or reliability.
- **Self-configuring and self-healing** – RPL networks can self-configure and adapt to changes in network topology, including the addition or removal of devices. It also supports self-healing, where the network finds alternative routes if link failures occur.

Cons:

- **Complex configuration** – configuring RPL for specific use cases can be complicated due to the various parameters and objective functions that must be considered. Proper tuning and optimisation may require expertise and considerable time.
- **Scalability for large networks** – while RPL performs well in small to medium-sized LLNs, it may face challenges in large networks, where the tree-like structure can lead to increased control traffic and reduced scalability.

- **Overhead in highly mobile networks** – in highly mobile LLNs frequent changes in the network topology may result in increased control message overhead as the network adapts to mobility.

Overall, RPL's focus on energy efficiency and adaptability to low-power and lossy networks makes it a strong choice for IoT and wireless sensor networks. It effectively addresses the unique challenges posed by resource-constrained devices, allowing them to form reliable and efficient communication links while optimising energy consumption. However, careful configuration and consideration of scalability in large networks are essential to ensure the protocol's effectiveness for specific deployment scenarios.

1.3.5 Wireless mesh protocols

Mesh communication technologies offer flexible, reliable, and scalable networking solutions, and several protocols include mesh topologies. A short overview of mesh protocols such as B.A.T.M.A.N., Bluetooth Mesh, OpenThread, Thread, ZigBee, Wi-Fi, Wi-SUN, WirelessHART, Z-WAVE and 6LoWPAN is presented before focusing on the LoRa mesh protocol and applications.

1.3.5.1 B.A.T.M.A.N

The protocol Better Approach To Mobile Ad-hoc Networking (B.A.T.M.A.N.) [11] is a multi-hop ad-hoc mesh network routing protocol where each node transmits broadcast or originator messages (OGMs) to notify neighbouring nodes about its presence. These neighbours re-broadcast the OGM.s based on specific rules to inform their neighbours about the presence of the original initiator. The network is steeped with OGM.s that are small, with a typical raw packet size of 52 bytes, including IP and UDP overhead. OGMs contain at least the originator's address, the address of the transmitting packet's node, a Time to Live (TTL) and a sequence number.

The approach of the B.A.T.M.A.N. algorithm is to divide the knowledge about the best end-to-end paths between nodes in the mesh to all participating nodes.

B.A.T.M.A.N. uses a proactive routing approach, which means it continuously maintains up-to-date routing information without waiting for a specific request to transmit data. Instead of relying on global routing tables, each node perceives and retains only the information about the best next hop towards all other nodes. Thereby the condition for overall network knowledge about local topology changes is unnecessary. Since wireless mesh networks are subject

to frequent changes, B.A.T.M.A.N. is designed to be adaptive and capable of quickly reconfiguring routes when nodes join, leave, or move within the network.

The protocol also supports load balancing by distributing traffic across multiple paths to prevent congestion and optimise the overall network performance.

1.3.5.2 Bluetooth Low Energy

Bluetooth Low Energy (BLE) [13] is optimised for low power consumption to address small-scale consumer IoT applications. BLE is integrated into several IoT devices, and data is conveniently communicated to and visualised on smartphones. The Bluetooth Mesh specification aims to enable a scalable deployment of BLE devices.

BLE provides versatile indoor localisation features, and IoT beacon networks are used for different IoT service applications. BLE is incompatible/non-interoperable with Bluetooth, and a dual-mode device is required to achieve interoperability.

BLE uses multiple techniques to ensure low power consumption implementing the data protocol to create low-duty-cycle transmissions, combined with very low-power sleep modes.

Bluetooth Low Energy Mesh [12] protocol is a networking technology built on the BLE standard. It enables large-scale, reliable, secure communication between many devices, forming a mesh network. This mesh network allows devices to communicate with each other and extend the range of the network.

A device can have one or more logical elements in the Bluetooth Mesh network. Each element represents a specific functionality or component of the device, and each element is assigned a unique address within the network.

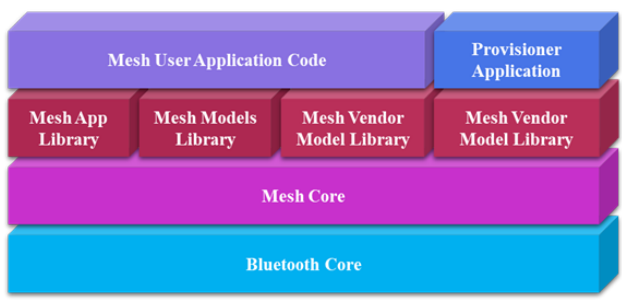

Figure 1.2 BLE Mesh Layered Architecture

Bluetooth Mesh devices use models to define their behaviour and capabilities. Models represent how a device handles messages, what types of messages it supports, and how it behaves in the mesh.

Provisioning is the process of securely adding a new device to the mesh network. Encryption keys and other necessary information are exchanged between the new device and the network during this process.

1.3.5.3 OpenThread and Thread

Thread [14] is a mesh networking low-power wireless protocol based on Internet Protocol version 6 (IPv6), designed to address the interoperability, security, power, and architecture challenges of the IoT. Thread utilises 6LoWPAN that employs the IEEE 802.15.4 wireless protocol with mesh communication. Thread is IP-addressable, with cloud access and advanced encryption standard (AES).

Thread uses a mesh network topology in the 2.4 GHz frequency spectrum, providing data rates of 250 kbps with a coverage range of 30 m. Security uses a 128-bit AES encryption system and the encryption cannot be disabled.

Thread utilises a network-wide key for inscription that is applied at the Media Access Layer (MAC). The key is employed as specified in IEEE 802.15.4. Attacks on Thread network originating over-the-air from outside the network are protected by IEEE 802.15.4 security mechanisms. The Thread network's nodes exchange frame counters with their neighbours via a Mesh Link Establishment (MLE) handshake. The protection against replay attacks is done via frame counters. Thread lets the application use various internet security protocols for end-to-end communication and can connect up to 250 devices.

OpenThread, released by Google, is an open-source implementation of Thread that implements all Thread networking layers (IPv6, 6LoWPAN, IEEE 802.15.4 with MAC security, Mesh Link Establishment, Mesh Routing), device roles, and Border Router support.

1.3.5.4 ZigBee

ZigBee [15] is a short-range, low-power, wireless standard deployed in a mesh topology to extend coverage by relaying IoT sensor data over multiple sensor nodes.

The Zigbee standard works on the IEEE 802.15.4 physical radio specification and runs in unlicensed bands such as 2.4 GHz, 915 and 868 MHz.

Zigbee 3.0 sustains wireless networks' increasing scale and complexity and deals with extensive local networks of over 250 nodes. The data rates

provided are 250 kbps (2.4 GHz), 40kbps (915 MHz) and 20kbps (868 MHz). Zigbee also handles the dynamic behaviour of the networks (with nodes disappearing, appearing, and re-appearing in the network topology) and permits orphaned nodes, resulting from the loss of a parent to rejoin the Zigbee network through another parent.

The self-healing structure of state-of-the-art Zigbee Mesh networks permits nodes to drop out of the network without disrupting internal routing. Zigbee supports over-the-air (OTA) upgrades during device operation and provides enhanced network security by employing a coordinator/trust centre, which creates the network and oversees the allocation of network and link security keys to joining nodes or distributed security where there is no coordinator/trust centre. The Zigbee router node can provide the network key to joining nodes.

1.3.5.5 Wi-Fi

Wi-Fi (IEEE/ISO/IEC 8802-11-2022) is a standard defining the characteristics of a wireless local area network (WLAN). The name Wi-Fi (short for "Wireless Fidelity") relates to the name provided by the Wi-Fi Alliance, formerly WECA (Wireless Ethernet Compatibility Alliance). This group assures compatibility between hardware devices that use the 802.11 standards. Wi-Fi networks must comply with the 802.11a-x specifications.

Wi-Fi mesh [16] protocol IEEE 802.11s creates a mesh network that extends Wi-Fi coverage over a larger area and enhances overall network performance and reliability. Traditional Wi-Fi networks are based on a single wireless access point (router) communicating directly with Wi-Fi-enabled devices. They may suffer from limited range and dead zones in larger spaces.

A Wi-Fi mesh network consists of multiple interconnected access points that work together to create a seamless and continuous network. These access points, often referred to as "nodes" or "mesh nodes", communicate with each other wirelessly, forming a self-healing network that can automatically reroute data packets to find the most efficient path to reach the destination device.

The system architecture for WLAN mesh network technology is described in IEEE 802.11 functional requirements and scope [17] and illustrated in Figure 1.3.

The functional blocks of the architecture include the following:

- The Mesh Topology Learning, Routing, and Forwarding block includes a function for discovering neighbouring nodes, a function for obtaining radio metrics, which deliver information on the quality of wireless links,

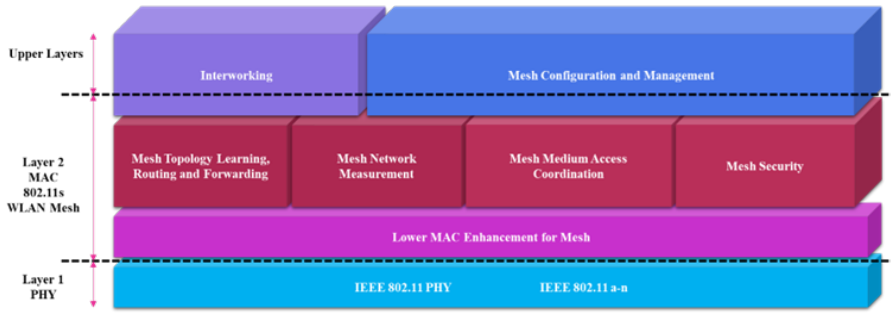

Figure 1.3 Wi-Fi Mesh Layered Architecture

a routing protocol for determining routes to transfer packets to their destinations using MAC addresses as identifiers, and a packet forwarding function. The routing protocol must use radio metrics and multiple frequency channels according to radio conditions to efficiently use radio resources.

- The Mesh Network Measurement block includes functions for calculating radio metrics used by the routing protocol and measuring radio conditions within the WLAN mesh network for frequency channel selection.
- The Mesh Medium Access Coordination block contains functions for preventing degraded performance due to hidden and exposed terminals, procedures for performing priority control, congestion control, and admission control, and a function for achieving spatial frequency reuse.
- The Mesh Security block comprises security functions (e.g., WLAN security schemes defined by the IEEE 802.11 standard) for protecting data frames carried on the WLAN mesh network and management frames used by control functions such as routing protocol.
- The Interworking block implements the function that supports WLAN mesh network to conform to IEEE 802 network architecture and connect to other networks by implementing a transparent bridge function enforced in the mesh portal situated at the network boundary. Each WLAN mesh network must operate as a broadcast network to deliver forwarded packets to all terminals connected to the LANs.
- The Mesh Configuration and Management block comprises a WLAN interface for the automatic setting of each mesh point's RF parameters (transmit power, frequency channel selection, etc.) and quality of service (QoS) policy management.

Wi-Fi mesh protocol is designed to address the limitations of traditional Wi-Fi networks, making them ideal for large homes, offices, or public spaces where extended coverage and high-performance connectivity are required.

1.3.5.6 Wi-SUN

Wi-SUN [18] stands for Wireless Smart Ubiquitous Network and is a mesh network protocol developed by Wi-SUN Alliance. Wi-SUN is one of the most popular IPv6 sub-GHz mesh technologies for smart utility and smart city applications. The target networks are named Field Area Networks (FANs), and they deliver a communications infrastructure for large-scale outdoor networks, usually outdoor IoT devices. FANs let industrial devices such as smart meters and streetlights interconnect onto one common network.

Wi-SUN is based on the IEEE 802.15.4g standard for the physical layer (PHY) and the IEEE 802.15.4e standard for the medium access control layer (MAC). It supports multiple data rates and frequency bands to meet regulatory requirements worldwide.

Wi-SUN makes interoperable, multi-service, secure wireless mesh networks available to service providers, utilities, municipalities/local governments, and other businesses. Wi-SUN can be used in various line-powered and battery-powered applications for large-scale outdoor IoT wireless communication networks. With the help of Wi-SUN, developers can add new features to existing infrastructure platforms by extending open standard internet protocols (IP) and APIs. With its long-range capabilities, high data throughput, and support for IPv6, Wi-SUN is designed to scale and makes wireless infrastructure easier for commercial applications and the development of smart cities.

1.3.5.7 WirelessHART

WirelessHART [19][20] is a process automation application wireless communications protocol that provides wireless capabilities to extend Highway Addressable Remote Transducer (HART) by keeping compatibility with existing HART commands, tools, and devices.

The architecture of the WirelessHART protocol stack according to the OSI 7-layer communication model is illustrated in Figure 1.4.

The WirelessHART protocol stack addresses five layers: physical layer, data link/MAC layer, network layer, transport layer and application layer. A central network manager is added for arbitrating the communication schedule and manage the routing.

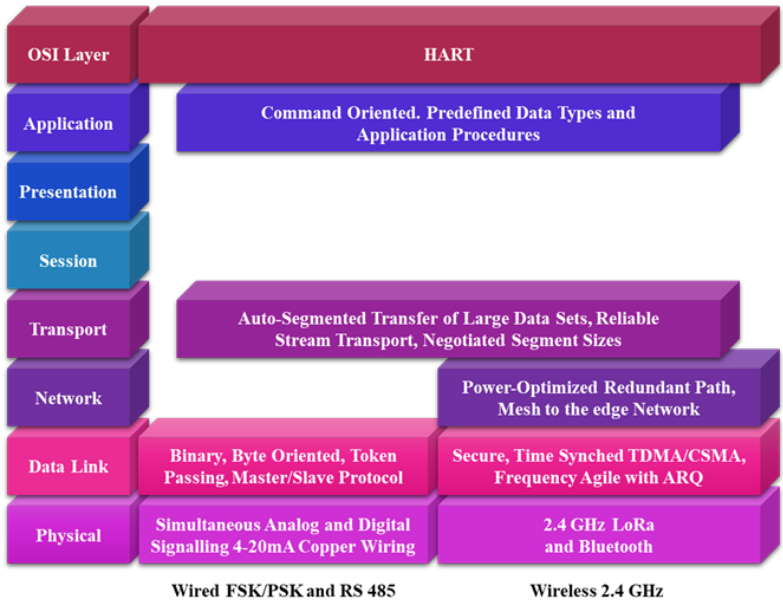

Figure 1.4 WirelessHART Protocol Architecture

WirelessHART uses mesh networking technology by design, where each device in a mesh network can act as a router for messages from other devices. This widens the range of the network and gives redundant communication routes to extend reliability in challenging radio environments encountered in process facilities [21][22][23]. Networks can scale up to 1000 nodes, but latency can be long and nondeterministic because transmissions occur only within an allocated time slot, and retransmissions are minimised.

Each WirelessHART network contains three major components:

- Wireless field devices that are connected to process or manufacturing equipment.
- Gateways that communicate among devices and on-premises host applications connected to high-speed backbone or other communications networks.
- A Network Manager configures the network, schedules communications between devices, monitors network health, and manages message routes. The Network Manager can be embedded into gateways, host applications, or process automation controllers.

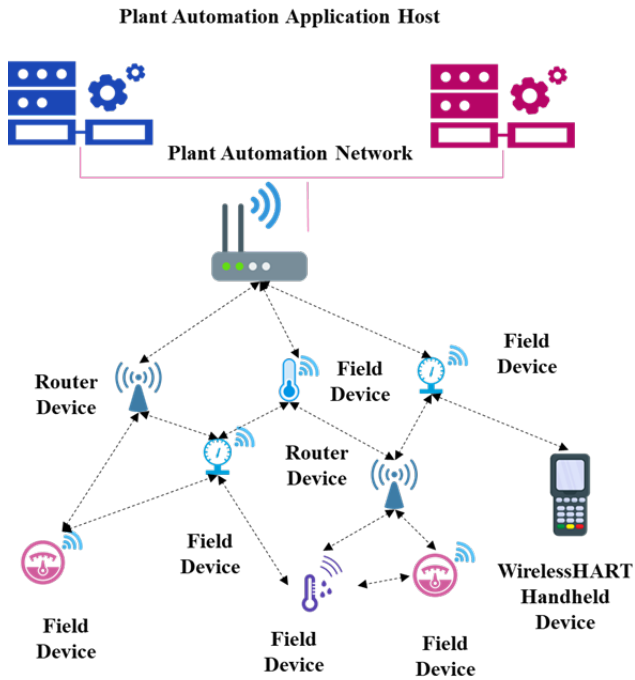

Figure 1.5 WirelessHART Mesh Networking

WirelessHART supports different messaging modes, such as one-way publishing of process and control values, spontaneous exception notification, ad-hoc request/response, and auto-segmented block transfers of large data sets, to provide flexibility to meet different application requirements. These features enable communications to be tailored to the application's needs, lowering power consumption and overhead.

The WirelessHART mesh networking topology applied to an industrial plant use case is illustrated in Figure 1.5.

WirelessHART is used in industrial environments that require security to provide the highest levels of protection to the network and data. Security includes encryption and authentication.

1.3.5.8 Z-WAVE

Z-Wave [24][25] is the wireless technology for secure, trusted home applications like home appliances, lighting control, security systems, garage door openers, thermostats, windows, locks, etc.

It is a mesh network low-energy wireless communications protocol used in systems controlled via the Internet and locally through devices or a Z-Wave gateway or central control device serving as hub controller and portal.

The Z-Wave Alliance [26] demands the mandatory implementation of the Security 2 (S2) framework on all devices receiving certification. Z-wave delivers packet encryption, integrity protection and device authentication services. End-to-end security is provided at the application level (communication using command classes). It has an in-band network key exchange and AES symmetric block cipher algorithm using a 128-bit key length.

Products using Z-Wave mesh protocol are interoperable and communicate with each other regardless of brand or platform, and the Z-Wave mesh networks become more reliable as more devices are added (e.g., a Z-Wave network with 100 devices is more reliable than a Z-Wave network with 30 devices). Z-Wave's interoperability at the application layer assures that Z-Wave devices share information and allows all Z-Wave hardware and software to work together.

Z-Wave uses the unlicensed industrial, scientific, and medical (ISM) band and operates at 868.42 MHz in Europe and 908.42 MHz in the US. Z-Wave delivers data rates of 9.6 kbps and 40 kbps, with output power at one mW.

Z-Wave range between two nodes is 100 m in an outdoor, unobstructed setting. For in-home applications, the range is 30 m for no obstructions and 15 m with walls in between.

1.3.5.9 6LoWPAN

6LoWPAN [27][28] itself is not a mesh protocol; it is an open standard defined in RFC 6282 by the Internet Engineering Task Force (IETF) for a network where every wireless network node is battery-powered and has a IPv6 address. Thus, a set of local nodes can make a wireless mesh network.

6LoWPAN defines how to run IP version 6 (IPv6) over low data rate, low power, and small footprint radio networks (LoWPAN) as typified by the IEEE 802.15.4 radio [28].

IP addresses may be static or dynamic if a network node that can issue IPv6 addresses is acting as or like a Dynamic Host Configuration Protocol (DHCP) server. For IoT networks, it is typical to have a node connected to both WLAN and LAN that performs the gateway functions to collect local data and control local nodes. If local 6LoWPAN demands such a functionality, it typically performs the DHCP server functions too.

1.4 LoRa and LoRaWAN Technology

LoRa and LoRaWAN are related but distinct technologies used together to create long-range, low-power wireless communication networks for the IoT and other edge applications.

1.4.1 LoRa physical layer

LoRa operates in the sub-GHz ISM bands, such as 433MHz, 868 MHz (Europe) or 915 MHz (North America).

Semtech has released a LoRa chipset operating at the 2.4 GHz frequency band, which is globally available with km-range capabilities, enabling region-independent hardware design chipsets [3][4].

LoRa, compared with other technologies operating in the 2.4 GHz band, such as Wi-Fi and Bluetooth, offers several significant advantages in range and power consumption in comparison with other existing techniques.

The BLE standard range is from 50 m indoors to 165 m outdoors, and the maximum range of 2.4 GHz Wi-Fi networks typically reaches around 100 m. LoRa's outdoor range is more than five times the outdoor range of BLE, and more than eight times typical IEEE 802.11 networks.

LoRa modulation is able to offer a higher receiver sensitivity and robustness against noise and interference. Some of the specific details will be explained in the next sub-chapters.

Chirp Spread Spectrum Modulation (CSS)

LoRa modulation uses a form of chirp spread spectrum modulation, where the transmit signal frequency varies continuously over time. Instead of transmitting data on a fixed carrier frequency, LoRa uses chirp signals that start at one frequency and sweep across the spectrum. The LoRa chirping signal sequence makes LoRa signals robust against narrowband interference because the signal energy is spread over a wider frequency range.

Symbols and Data Rate

LoRa allows to adapt the number of bits per symbol according to the signal-to-noise ratio available over the link. Long range is achieved by reducing the number of bits per symbol, increasing the amount of energy per bit, and thus reducing the resulting bit rate.

Spreading Factor (SF)

The spreading factor (SF) is a critical parameter in LoRa modulation that determines the signal's robustness and range. The SF defines the rate at which

the chirp signal spreads across the frequency spectrum and the amount of (potential) processing gain on receiver side.

Higher SF results in a lower data rate but better resistance to interference and an extended communication range. Conversely, a lower SF provides a higher data rate but with reduced range and increased susceptibility to noise.

Signal Bandwidth (BW)

The bandwidth of the LoRa signal also influences communication performance. LoRa modulation can operate in different bandwidths, typically 125 kHz, 250 kHz, or 500 kHz for sub-GHz LoRa.

A wider bandwidth allows for higher data rates but may reduce the communication range. Narrower bandwidths, on the other hand, result in lower data rates but offer increased range and better interference immunity.

Reception and Demodulation

On the receiver side, LoRa demodulation involves analysing the received chirp signal to decode the transmitted symbols. The receiver can determine the transmitted symbols and extract the original data by comparing the received signal with predefined chirp sequences.

Forward Error Correction (FEC)

In addition to the modulation scheme, a forward correction algorithm with several code rates can be applied, which enables the receiver to recover corrupted bits. This feature helps to decrease the number of packet retransmissions in noisy environments.

Sub-GHz Frequency Bands

The license-free sub-GHz ISM band allows transmitting within fixed defined frequency bands which vary depending on the region.

In this context, it is not possible to use the same type of radio hardware equipment because the used frequencies significantly impact the used chips, antenna matching circuits and the connected antennas.

The combination of a robust wireless transmission scheme with long-range capabilities and a low power footprint makes the LoRa technology ideal for battery powered IoT devices that can last up to 10 years.

The LoRa technology became public combined with the first LoRa radio modules and the so-called LoRaMAC protocol, today known as LoRaWAN protocol and defined within the LoRaWAN standard.

The following subchapters outline the most compelling aspects of the standard.

Table 1.1 Frequency Band Overview

No.	Region	Frequency Band
#1	Europe	863 MHz – 870 MHz
#2	Europe	433,05 MHz - 434,79 MHz
#3	North America	902 MHz– 928 MHz
#4	China	470 MHz – 510 MHz
#5	Korea	920 MHz – 925 MHz
#6	Japan	920 MHz – 925 MHz
#7	India	865 MHz – 867 MHz

1.4.2 LoRaWAN protocol

The LoRaWAN protocol defines methods, packet formats and LoRa physical layer radio parameters to ensure interoperability between IoT end devices and a given network infrastructure. The LoRaWAN standard itself is maintained by the non-profit association the LoRa Alliance [2].

The standard defines a system architecture consisting of at least three different component types with different roles and responsibilities.

The composition of end devices, gateways, and a central network server enables applications to create a star-of-star network topology.

LoRaWAN End Devices

These are typically sensors or actuators that need to communicate wirelessly over large distances through the LoRaWAN Link Layer protocol, formerly known as LoRaMAC protocol.

LoRaWAN Gateways

Gateways operate as intermediate devices with less intelligence. They relay the uplink and downlink messages between end devices and the network server using different TCP/IP-based protocols. A network can consist of several gateways.

LoRaWAN Network Server

The network server includes all the intelligence for controlling the radio network resources, e.g., network access, a security parameter, spreading factors (adaptive radio data rates) etc.

The network server is connected to all gateways and the application server, which hosts the application data and business logic. Suitable TCP/IP-based protocols typically handle these connections.

LoRaWAN allows IoT devices to transmit data over long distances to LoRaWAN gateways, which act as intermediaries between the end devices and the network server. LoRaWAN's key features are:

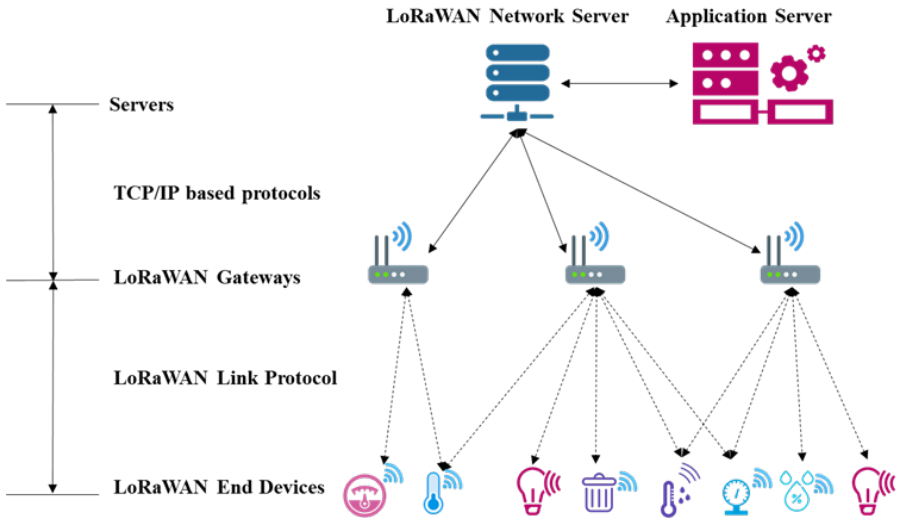

Figure 1.6 LoRaWAN Network Architecture

- **Low power** – LoRaWAN is designed to operate with low-power IoT devices, enabling long battery life for sensors and devices.
- **Wide area coverage** – LoRaWAN provides wide area coverage by leveraging the long-range capabilities of LoRa.
- **Public or private networks** – LoRaWAN can be deployed in public networks managed by network operators or in private networks operated by organisations.
- **Security** – LoRaWAN incorporates several security mechanisms, including end-to-end encryption and device authentication, to ensure secure data transmission.
- **Adaptive data rate** – LoRaWAN supports adaptive data rates, allowing devices to adjust their transmission speed based on the quality of the communication link, ensuring efficient data transfer.

LoRa and LoRaWAN form a powerful combination for creating efficient and scalable IoT communication networks. LoRaWAN defines a communication protocol and network architecture for IoT low-power wide area networks (LPWANs) and is designed to address the requirements for low power consumption (i.e., long battery life), long-range, and variable data rates (0.3 kbps – 50 kbps) while maintaining low operating and deployment costs.

1.4.3 2.4 GHz LoRa

In addition to sub-GHz LoRa, Semtech has developed a transceiver circuit with LoRa modulation for the 2.4 GHz ISM band. Compared to the sub-GHz solution this radio enables additional applications with diverse requirements.

The 2.4 GHz LoRa might be more suitable for applications operating in urban environments with higher device density, but covering shorter distances. On the other hand, sub-GHz LoRa is well-suited for applications needing extended range and better penetration of obstacles. Table 1.2 offers a brief comparison of the two radio technologies.

The integration of 2.4 GHz LoRa and a mesh protocol stack holds the potential to enhance the capabilities of edge AI-enabled IoT applications, particularly in terms of range coverage, network density, and robustness against single points of failure.

Table 1.2 Frequency Band Overview

Aspect	Sub-GHz LoRa	2.4 GHz LoRa
Frequency Band	433 MHz, 868 MHz, 915 MHz, depending on region / country	2.4 GHz Worldwide available
Range	Longer range	Shorter range
Penetration	Better penetration of obstacles	Lower penetration
Susceptibility to Interference	Lower	Higher due to higher signal channel bandwidth and multiple usage of the 2,4 GHz ISM band
Applications	Agriculture, rural areas, wide-area IoT networks	Smart Cities, densely populated areas, short-distance IoT networks
Interference Potential	Lower potential	Higher potential
Network Density	Lower density networks	Higher density networks
Tx Limits	Duty Cycle Limit 0.1%, 1%, 10% depending on sub-band	Unlimited
Bandwidth	125 kHz, 250 kHz, 500 kHz	203 kHz, 406 kHz, 812 kHz, 1625 kHz
Data rate	0.3 kbps – 0.9 kbps	0.2 kbps - 203 kbps

1.5 LoRa Mesh and Enabling AI Technologies

The convergence of technologies (including edge AI, IoT, distributed architectures, and federated learning) results in intelligent edge mesh communication systems performing efficient connectivity by optimising data rates, coverage, energy, and interference. LoRa when combined with edge AI and IoT, enhances connectivity and enables novel use cases:

- **Comprehensive area coverage** – LoRa's long-range capabilities allow devices to communicate over several kilometres, making it suitable for large-scale IoT deployments in smart agriculture, asset tracking, and environmental monitoring.
- **Energy efficiency** – LoRa devices consume very little power, making them ideal for battery-operated IoT sensors and devices, which can operate for extended periods without frequent battery replacements.
- **Low cost and scalability** – LoRa's low infrastructure cost and simple deployment enable cost-effective and scalable IoT solutions across diverse environments.

The Figure 1.7 illustrates a typical mesh topology with end nodes and gateways offering AI. For tasks like secure device enrolment, automatic firmware deployments or additional system monitoring a single or multiple application servers can be connected by wired or wireless IP based communication links. By combining edge AI, IoT, and LoRa, adopters can benefit from improved data rates, reduced latency, increased efficiency,

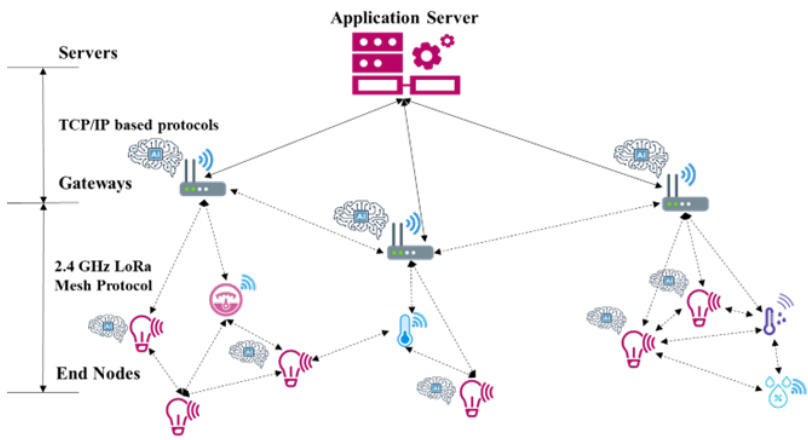

Figure 1.7 Edge AI Enabled LoRa Mesh Network

and cost-effectiveness. This convergence opens opportunities for innovation, automation, and optimisation across various sectors.

1.6 Applications for LoRa Mesh

LoRa mesh networks offer a versatile and reliable solution for applications that require low-power and extended-range wireless communication. LoRa mesh networks are suited for the following applications:

Industrial Automation: In industrial settings, LoRa mesh networks can be deployed for machine-to-machine (M2M) communication, asset tracking, and control systems. They enable monitoring and control of equipment and processes with extended-range.

Building Management Systems: LoRa mesh networks can be employed to optimise energy consumption in commercial buildings by managing lighting and other energy-related equipment more efficiently. However, it can be argued to what extent it remains energy efficient to reach indoor end nodes from an outdoor base station.

Smart Metering: LoRa-based intelligent metering systems can enable utilities to remotely monitor and manage energy, water, and gas consumption in residential and industrial settings.

Wireless Sensor Networks (WSNs): LoRa is a popular choice for creating WSNs, where many battery-powered sensors communicate with a gateway for data collection and analysis.

Smart Agriculture: LoRa mesh networks can be deployed in agricultural settings to monitor soil conditions, automate irrigation systems, and track livestock.

Lighting Control: LoRa can be used in wireless lighting control systems, enabling users to create adaptive and energy-efficient lighting environments.

Environmental Monitoring: LoRa mesh networks can be employed for monitoring environmental parameters, such as air quality, temperature, and humidity, in smart cities or remote areas. Furthermore, those networks can aid in predicting critical situations such as fires, floods, or earthquakes.

1.7 Conceptual Edge AI and LoRa Mesh Device Architecture

This chapter outlines a possible device architecture which integrates AI and 2.4 GHz LoRa Mesh technologies.

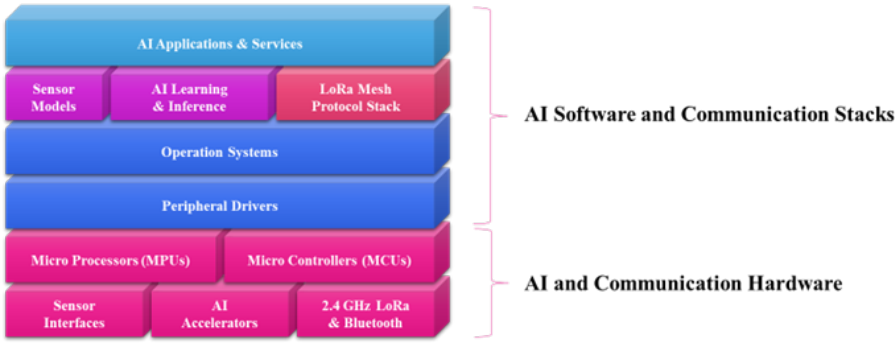

Figure 1.8 Conceptual Edge AI and LoRa Mesh Device Architecture

The purpose of this architecture is to provide a foundational framework for designing and implementing edge-devices with respect to the hardware and embedded software aspects.

The subsequent subsections should provide more detailed explanations of the provided building blocks, starting from the bottom, which includes the hardware-related units.

1.7.1 Sensor and interfaces

Typical IoT end devices include sensors (or actuators) that are connected via serial interfaces such as UART, SPI, or I2C to embedded microcontrollers running corresponding sensor drivers. More sophisticated devices may feature camera interfaces for image processing or display interfaces to connect displays that provide complex visual feedback to users. Consequently, the selection of microcontrollers/processors, sensor interfaces, power supplies, and connectors greatly depends on the specific application requirements. Designers and engineers must consider these factors when developing either a dedicated device or a multipurpose edge AI computing platform.

1.7.2 AI accelerators

Compared to pure software solutions, AI hardware accelerators offer better computational performance with a lower energy consumption footprint due to their parallel architecture. AI accelerators are designed for deep learning (DL) neural network computations and machine learning (ML) applications.

1.7.3 2.4 GHz LoRa and Bluetooth radios

The integration of 2.4 GHz LoRa and Bluetooth radio technologies can be achieved using modules that include their own microcontroller running the corresponding protocol stack. Such modules typically offer serial interfaces like UART or SPI for configuration, control, and data transfers. While 2.4 GHz LoRa is primarily used for long-range data exchanges within the application, short-range Bluetooth can be used for tasks such as single device maintenance and firmware updates. This can be accomplished through smartphones, tablets, or other portable computers that have Bluetooth available as a standard connectivity service.

1.7.4 Microcontrollers and microprocessors

These units are available from various manufacturers, offering a wide range of processing capabilities, including single-core and multi-core devices, as well as various memory and interface options. Microprocessor systems are typically capable of running embedded Linux, providing enhanced flexibility in choosing an appropriate programming language with higher abstraction and extensive library support. Microcontrollers are more likely to run smaller operating systems like FreeRTOS or proprietary ones, often with varying levels of real-time support and are directly connected to sensors and actuators.

Arm-based architectures with AI/ML-optimised cores support the development of lightweight microcontrollers with embedded coprocessing to optimise overall processing capability, local analytics, and power consumption. The edge AI methods, techniques, frameworks, and tools enable the embedded design to develop, train, optimise and deploy edge AI models on microcontroller-based hardware.

1.7.5 Peripheral driver

The connection between hardware and software is typically established through peripheral drivers. These drivers offer an interface for the higher layers of embedded software and ensure secure control and configuration of the underlying hardware units. In the case of operating systems like Linux, such drivers must adhere to specified interfaces and be implemented according to predefined rules. Additionally, in smaller microcontroller-based systems, similar driver software has been developed for the same purpose.

1.7.6 Operating systems

An operating system acts as an intermediary between hardware and embedded software applications. It manages and coordinates various hardware and software components to provide a stable and efficient environment for middleware and application software to run on a device. The choice of the operating system is, like hardware selection, significantly dependent on application requirements. Furthermore, it must be compatible with the selected hardware to support the lower-level peripheral drivers and interfaces.

1.7.7 Sensor models

A sensor model is a representation of how a sensor behaves and interacts with the environment it is monitoring. The model is a mathematical or computational description that helps understand and predict the relationship between the input (physical quantity being sensed) and the output (measurement or signal generated by the sensor).

Sensor models are used for various purposes, including:

- Simulation – they can be used to create virtual sensor behaviours in software simulations, allowing engineers to test systems before physical implementation.
- Calibration – sensor models help in calibrating real sensors by understanding how their measurements correspond to actual physical values.
- Data Fusion – when multiple sensors are used to gather information, their models can help combine and interpret the data accurately.
- System Design – in designing complex systems, sensor models aid in selecting appropriate sensors and understanding their integration.
- Fault Detection – deviations between actual sensor outputs and model predictions can indicate sensor malfunctions.

Sensor models can be as simple as linear equations or as complex as sophisticated computational simulations. They consider various factors that affect sensor behaviour, such as noise, sensitivity, non-linearity, temperature dependence, and more. By having an accurate model, engineers can improve the reliability and accuracy of systems relying on sensor data.

1.7.8 AI learning and inference

This building block includes the two fundamental aspects of an AI enabled edge device.

- AI Learning – this is the process in which AI systems gain knowledge and insights from data. It employs algorithms to identify patterns and learn how inputs relate to desired outputs. There are different types of AI learning, including supervised, unsupervised, and reinforcement approaches.
- AI Inference – this is the phase when a trained AI model is used to make predictions or decisions based on available data. This is the practical application of what the AI system has learned before.

1.7.9 2.4 GHz LoRa Mesh Protocol Stack

The LoRa Mesh Protocol Stack encompasses the functionalities to enable end-to-end communication within a wireless mesh topology. This includes tasks such as neighbour and route discovery, packet forwarding, route adaptation and maintenance, device management, and medium access control. Additionally, specific metrics and interfaces may be exposed to the embedded AI unit, enhancing adaptive routing algorithms through AI-based methods and techniques.

1.7.10 AI applications and services

This upper layer encompasses specific aspects and services tailored to a particular distributed edge AI-enabled application. The associated software components within this layer utilize middleware layer components at the highest available abstraction level to meet the application's specific functional and non-functional requirements.

1.8 Challenges and Future Research Directions

Built-in edge AI and wireless mesh connectivity capability that integrates processing units with AI-based capabilities, multiprotocol communication wireless modules for real-time monitoring and high-performance micro-electromechanical systems (MEMS) accelerometer sensors extend the functionalities and features of intelligent edge devices. This facilitates data aggregation, integration, and processing.

Building AI into wireless edge devices and sensors allows edge devices to learn and infer. Inference and decision making are performed within the edge device based on data collected through its sensors.

Long-range mesh network designs with edge AI capabilities enable effective monitoring through infrequent data updates communicated over long distances.

The LoRa mesh network can include security mechanisms while maintaining a low-energy profile for battery-powered edge sensors.

Lightweight authentication and encryption techniques can avert spoofing and provide confidentiality in message exchanges between edge nodes and the base station.

Updates can be performed using GPS-enabled time synchronisation and a concurrent transmission property inherent to LoRa.

An overview of the primary challenges and future research directions for edge AI and wireless LoRa mesh connectivity is presented in the next paragraphs.

Various wireless routing protocols, such as AODV, OLSR, DSR and RPL, face different challenges depending on the specific characteristics of the networks in which they are deployed. The following are some common challenges that these protocols frequently encounter:

- **Scalability** – all of these protocols need to scale with the increasing number of nodes in a network. As the network grows, more routing information must be managed and distributed. This can lead to increased overhead and longer route discovery times, especially for protocols based on proactive topology updates.
- **Mobility** – in wireless networks, devices can move frequently or follow unpredictable patterns. Protocols must be able to adapt to these changes and maintain efficient routes, even for mobile devices.
- **Connectivity** – fluctuations and interferences – Wireless networks are susceptible to connectivity fluctuations, interferences, and signal attenuations. Routing protocols must cope with these variations to provide stable and reliable routes.
- **Energy efficiency** – energy efficiency is crucial in IoT networks and battery-operated devices. Routing protocols should be designed to minimise energy consumption and maximise battery life.
- **Security** – wireless networks are vulnerable to security threats, such as man-in-the-middle attacks and routing manipulation. Routing protocols must rely on other mechanisms to secure communication and ensure the integrity of routing information.

- **Overhead and latency** – routing protocols generate additional overhead in the network to distribute and update routing information. This overhead can reduce the available bandwidth and lead to higher latency and increased energy consumption.
- **Complexity** – some routing protocols can be complex, especially when optimised for specific use cases. The implementation and management of such protocols can be challenging.
- **Interoperability** – in some cases, wireless networks must communicate with different devices and technologies from other vendors or protocols. Ensuring interoperability between different protocols can be a challenge.

It can be a challenge to find the appropriate edge AI learning techniques and AI input parameters when combining communication protocols with AI-based methods at the application level to enhance the overall performance of the wireless network itself.

These challenges are crucial when selecting and implementing a routing protocol for a wireless network. The routing protocol must meet the specific requirements of the network and the characteristics of the connected devices to ensure optimal performance and reliability.

The challenges for federated learning systems are potentially related to wireless communication efficiency, platform and sub-system heterogeneity, data heterogeneity, and protection of privacy [41][42][43]:

- **Wireless communication efficiency** – federated networks can include many edge nodes/devices, and the communication latency in the network may be larger than the time for computations carried out locally at the edge nodes/devices. As for all wireless networks, bandwidth is limited depending on the wireless technology solution used. Efficient communication strategies are needed to reduce the size and number of messages transmitted, such as the communication rounds constituting the training cycles, which are typically repeated iteratively until the global model converges and the targeted accuracy is achieved. To increase communication efficiency, local updates carried out in parallel on the nodes/devices for each communication round can reduce the total number of communication rounds. The size of messages transmitted can be reduced by using model compression methods, such as subsampling and quantification, and latency and bandwidth challenges can be reduced by decentralised topologies and training.

- **Platform and sub-system heterogeneity** – the federated network is a heterogeneous system typically without inherent seamless properties. The system may be challenged by different communication protocols, variations in hardware capabilities (e.g., processor units and memories) and various restrictions on energy consumption. When many edge nodes/devices are included in a system or its sub-systems, node/device fault tolerance properties are essential for the case of node loss (e.g., communication failure or power limitations) during a training/learning iteration. To reduce the possible adverse effects of heterogeneity, parallel iterative operations can be facilitated by asynchronous communication, and the number of nodes/devices participating in each communication round for training/learning can be increased and/or selected by active node/device sampling to maximise the aggregated node/device update within a defined timeframe; the effect of dropout can also be reduced/eliminated by implementing fault tolerance solutions that facilitate redundancy.
- **Data heterogeneity** – the data collected/generated from a potentially large number of edge nodes/devices in a federated network may be heterogeneous because of differences in populations, samples, and results. That is; the data used for modelling and analysing are usually not uniformly distributed across the edge nodes/devices, and variations in data types, attributes, data labelling, data points and data refresh rates, challenge the training/learning processes. Machine learning methods, such as meta-learning and multi-task learning, have been extended to modelling in federated infrastructure, but they have limitations in terms of scalability, robustness, and automation.
- **Protection of privacy** – the federated learning approach benefits privacy by keeping raw data (and possibly derived data) on each edge node/device in the federated network. However, sharing model updates in the network during the training/learning processes can expose sensitive information to a third party. Modular and differential approaches can enhance privacy in a federated infrastructure, but there may be trade-offs between privacy and model accuracy.

1.9 Discussion and Conclusions

LoRa is a wireless communication technology used in low-power, long-range communication applications. It provides low data rates to meet the

requirements of remote edge nodes, which periodically send small amounts of sensor data. The architecture of LoRaWAN builds on a star topology that creates a single hop between an edge node (sensor IoT node) and the gateway. LoRa mesh networks are available for various applications that cannot be sufficiently managed by LoRaWAN architecture. The work in [32] has demonstrated that LoRaWAN applications can be extended using multi-hop LoRa, in which intermediate nodes can operate as repeaters that broadcast traffic to other LoRa nodes to reach a gateway.

The advantage of a LoRa mesh network is that the network coverage area can be expanded without adding more base stations. Furthermore, mesh networks combined with LoRa technology and AI-based techniques for routing optimisation can bring advantages to the application of the wireless sensor network in terms of improving the coverage area and promoting low power consumption.

Different wireless technologies, such as ZigBee, Z-Wave, BLE, Wi-SUN and Wi-Fi, use mesh topologies in which each device can be a router relaying the packet of the other devices to the end node. The main difference between LoRa and these other technologies is the ability for long-range transmission. This advantage can assist in expanding the network model without the need for additional base stations. In addition, low bandwidth makes LoRa resistant to channel noise, long-term relative frequency drift, Doppler effects and fading.

The key parameters used to configure the LoRa radio module are the modulation method, frequency range, bandwidth (BW), spreading factor (SF), coding rate (CR) and transmission power (TP). Artificial intelligence-based ML methods applied to LoRa and LoRaWAN for efficient resource management (e.g., BW, SF, CR and TP) can enhance LoRaWAN network performance and efficiency.

Edge AI solutions can be used in the processing modules of IoT devices that transmit information packets via the LoRa mesh network.

As the communication bandwidth of a LoRa link is low, performing ML on the IoT device allows for sending classification results rather than sending a higher amount of raw sensor data for remote classification. This saves the bandwidth of the low-capacity LoRa communication link.

Communication in a LoRa mesh network must adopt bandwidth-saving strategies, considering the duty cycle limitations of sub-GHz LoRa. Long range lacks packet delivery guarantees; for instance, using federated learning will require additional protocols for reliable messaging.

Depending on the setup and operational conditions, messages in a LoRa mesh network are also delivered with delays, excluding applications with strict real-time requirements. Consequently, distributed intelligence within a LoRa mesh network must determine the trade-off between using local computation and using communication resources. In these cases, there is a need for network integration of the LoRa mesh layer with the internet in full-stack edge IoT applications.

Long-range mesh networks and ML techniques deployed on edge IoT nodes can become communication substrates for building distributed intelligence with tiny edge nodes. The application can be extended using federated ML over LoRa communication, which is performed by embedded devices at the IoT layer.

Long-range mesh topology combined with intelligent gateways, AI-based routing optimisation and ML algorithms implemented in the processing nodes can be used for applications, such as intelligent lighting systems that provide extended coverage with limited data rates. Compared with other protocols for controlling large numbers of light devices, this technology can be suitable for lighting control.

In addition, the presented technologies can offer several benefits that contribute to the improved efficiency, scalability, and reliability of agricultural sensor systems. Sensors placed further from the central control point can still communicate through intermediate nodes, extending the coverage range of the overall network.

Mesh networks are self-healing, meaning that if one sensor node fails or is disrupted, the network can dynamically reroute data through alternative paths. This is an important feature, for example, in agriculture, in which environmental factors, such as weather, crop growth and equipment malfunctions, can temporarily disrupt communication. Sensor nodes equipped with edge AI functionality can detect or even predict such situations to improve system reliability and maintainability and, as a result, reduce costs.

Mesh networks can easily accommodate the addition of new sensor nodes without requiring significant infrastructure changes. This scalability is crucial, for example, in agriculture, in which the number of sensors might need to increase as the plantation expands or as new monitoring needs arise.

Communication through intermediate nodes alleviates the requirement for additional infrastructure components, thereby decreasing overall system costs. Moreover, agricultural sensor nodes frequently function in areas that are remote or difficult to reach, underscoring the importance of battery

life. Leveraging low-power mesh protocols enables sensors to operate for extended durations without frequent battery replacements.

Acknowledgements

This research was conducted as part of the EdgeAI "Edge AI Technologies for Optimised Performance Embedded Processing" project, which has received funding from KDT JU under grant agreement No 101097300. The KDT JU receives support from the European Union's Horizon Europe research and innovation program and Austria, Belgium, France, Greece, Italy, Latvia, Luxembourg, Netherlands, and Norway.

The authors thank Fetze Pijlman and Jean-Paul Linnartz from Signify and the Eindhoven University of Technology, the Netherlands, for all their careful, constructive, and insightful comments and feedback about this work.

References

[1] LoRa (PHY). SEMTECH. Available at: https://www.semtech.com/lora/what-is-lora

[2] LoRa Alliance. https://lora-alliance.org/

[3] Semtech. Semtech SX128x Long Range Datasheet. 2019. Available online: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R000000HVET/HfcgiChyabtiPTh6EjcDM6ZEwAOQV7IirEmRULggMM

[4] Semtech. Application Note: Ranging with the SX1280 Transceiver. 2017. Available online: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R000000HVET/HfcgiChyabtiPTh6EjcDM6ZEwAOQV7IirEmRULgggMM

[5] OpenThread. Available online: https://openthread.io/

[6] Thread. Available online: Available online: https://www.threadgroup.org/

[7] ZigBee Mesh Network. Available online: https://www.emcu-homeautomation.org/zigbee-mesh-network-ver-3-introduction/

[8] Zigbee. The Full-Stack Solution for All Smart Devices. Connectivity Standards Alliance. Available online: https://csa-iot.org/all-solutions/zigbee/

[9] G. R. Hiertz et al., "IEEE 802.11s: The WLAN Mesh Standard," in IEEE Wireless Communications, vol. 17, no. 1, pp. 104-111, February 2010, doi: 10.1109/MWC.2010.5416357.

[10] B.A.T.M.A.N. protocol concept. Available online: https://www.open-mesh.org/projects/open-mesh/wiki/BATMANConcept

[11] D. Johnson, N. Ntlatlapa, and C. Aichele, Simple pragmatic approach to mesh routing using BATMAN. 2nd IFIP International Symposium on Wireless Communications and Information Technology in Developing Countries, CSIR, Pretoria, South Africa, 6-7 October, pp 10, 2008. Available at: http://hdl.handle.net/10204/3035

[12] S.M. Darroudi, and C. Gomez, Bluetooth low energy mesh networks: A survey. Sensors, 17(7), p.1467, 2017. Available at: https://doi.org/10.3390/s17071467

[13] R. Heydon, and N. Hunn, Bluetooth low energy. In CSR Presentation, Bluetooth SIG. 2012. Available at: https://www.Bluetooth.Org/DocMan/handlers/DownloadDoc.Ashx

[14] H. S. Kim, S. Kumar, and D. E. Culler, "Thread/Open thread: A compromise in low-power wireless multihop network architecture for the internet of things," IEEE Communications Magazine, vol.57, no.7, pp.55-61, 2019. Available at: https://par.nsf.gov/servlets/purl/10136090

[15] S.C. Ergen, ZigBee/IEEE 802.15. 4 Summary. UC Berkeley, September, 10(17), p.11, 2004. Available at: http://users.eecs.northwestern.edu/~peters/references/ZigtbeeIEEE802.pdf

[16] P. Lech, P. Włodarski, Analysis of the IoT WiFi Mesh Network. In: R. Silhavy, R. Senkerik, Z. Kominkova, Z. Oplatkova, Z. Prokopova, P. Silhavy, (eds) Cybernetics and Mathematics Applications in Intelligent Systems. CSOC 2017. Advances in Intelligent Systems and Computing, vol 574. Springer, Cham, 2017. Available at: https://doi.org/10.1007/978-3-319-57264-2_28

[17] W. S. Conner "IEEE 802.11 TGs Functional Requirements and Scope," IEEE802.11- 04/1174r13, Jan. 2005.

[18] H. Harada, K. Mizutani, J. Fujiwara, K. Mochizuki, K. Obata, and R. Okumura, IEEE 802.15. 4g based Wi-SUN communication systems. IEICE Transactions on Communications, 100(7), pp.1032-1043, 2017. Available at: https://www.jstage.jst.go.jp/article/transcom/E100.B/7/E100.B_2016SCI0002/_pdf/-char/en

[19] S. M. Hassan, R. Ibrahim, K. Bingi, T. D. Chung, N. Saad, Application of Wireless Technology for Control: A WirelessHART Perspective, Procedia Computer Science, Volume 105, pp. 240-247, 2017. ISSN 1877-0509. Available at: https://doi.org/10.1016/j.procs.2017.01.217

[20] J. Song, S. Han, A., Mok, D. Chen, M. Lucas, M. Nixon, and W. Pratt, WirelessHART: Applying wireless technology in real-time industrial

process control. In 2008 IEEE Real-Time and Embedded Technology and Applications Symposium, pp. 377-386, April 2008. Available at: https://www.cecs.uci.edu/~papers/cpsweek08/papers/rtas08/9B.pdf

[21] A. Saifullah, Y. Xu, C. Lu and Y. Chen, "Real-Time Scheduling for WirelessHART Networks," 2010 31st IEEE Real-Time Systems Symposium, San Diego, CA, USA, 2010, pp. 150-159. Available at: https://doi.org/10.1109/RTSS.2010.41

[22] T. Lennvall, S. Svensson and F. Hekland, "A comparison of WirelessHART and ZigBee for industrial applications," 2008 IEEE International Workshop on Factory Communication Systems, Dresden, Germany, 2008, pp. 85-88. Available at: https://doi.org/10.1109/WFCS.2008.4638746

[23] P.A.M. Devan, F.A. Hussin, R. Ibrahim, K. Bingi, F. A. Khanday, A Survey on the Application of WirelessHART for Industrial Process Monitoring and Control. Sensors 2021, 21, 4951. Available at: https://doi.org/10.3390/s21154951

[24] S. J. Danbatta and A. Varol, "Comparison of Zigbee, Z-Wave, Wi-Fi, and Bluetooth Wireless Technologies Used in Home Automation," 2019 7th International Symposium on Digital Forensics and Security (ISDFS), Barcelos, Portugal, 2019, pp. 1-5. Available at: https://doi.org/10.1109/ISDFS.2019.8757472

[25] M. Lilli, C. Braghin, and E. Riccobene, Formal Proof of a Vulnerability in Z-Wave IoT Protocol. In Proceedings of the 18th International Conference on Security and Cryptography (SECRYPT 2021), pages 198-209, 2021. ISBN: 978-989-758-524-1. doi: 10.5220/0010553301980209. Available at: https://www.scitepress.org/Papers/2021/105533/105533.pdf

[26] Z-Wave Alliance. Available at: https://z-wavealliance.org/

[27] J. Olsson, 6LoWPAN demystified. Available at: https://www.ti.com/lit/wp/swry013/swry013.pdf?ts=1691185461616&ref_url=https%253A%252F%252Fwww.google.com%252F

[28] G. Mulligan, The 6LoWPAN architecture. EmNets '07: Proceedings of the 4th workshop on Embedded networked sensors, June 2007, pp. 78–82, 2007. Available at: https://doi.org/10.1145/1278972.1278992

[29] A. Tønnesen, Implementing and Extending the Optimized Link State Routing Protocol, 2004. Available at: http://www.olsr.org/docs/report.pdf

[30] T. Clausen, and P. Jacquet, Optimized Link State Routing Protocol (OLSR), Oktober 2003. Available at: https://datatracker.ietf.org/doc/html/rfc3626

[31] P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, Optimized link state routing protocol for ad hoc networks, 2001. Available at: https://ieeexplore.ieee.org/document/995315

[32] J.R. Cotrim, and J.H. Kleinschmidt, "LoRaWAN mesh networks: A review and classification of multihop communication", Sensors, 20 (15) (2020). Available at: https://www.mdpi.com/1424-8220/20/15/4273

[33] J. Wang, M. Abolhasan, D. R. Franklin, and F. Safaei, OLSR-R∧3: Optimised link state routing with reactive route recovery, 2009. Available at: https://ro.uow.edu.au/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1791&context=infopapers

[34] C.E. Perkins and E.M. Royer, Ad hoc On Demand Distance Vector (AODV) Routing, 1999. Available at: https://ebelding.cs.ucsb.edu/sites/default/files/publications/wmcsa99.pdf

[35] C. Perkins, E. Belding-Royer, and S. Das, Ad hoc On-Demand Distance Vector (AODV) Routing, July 2003. Available at: https://datatracker.ietf.org/doc/html/rfc3561

[36] D. Johnson, D. Maltz, and J. Broch, DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks. Available at: https://cs.brown.edu/courses/cs295-1/dsr-chapter00.pdf

[37] D. Johnson, Y. Hu, and D. Maltz, The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4, 2007. Available at: https://datatracker.ietf.org/doc/html/rfc4728

[38] O. Iova, G. P. Picco, T. Istomin, and C. Kiraly, RPL, the Routing Standard for the Internet of Things . . . Or Is It?, 2017, Available at: https://hal.science/hal-01647152/document

[39] T. Tsvetko. RPL: IPv6 routing protocol for low power and lossy networks. Sensor nodes–operation, network and application (SN), 2011, 59. Jg., Nr. 2. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=59b65811b94ba2162a9083744aef83fe09d381b0#page=67

[40] T. Winter, and P. Thubert, RPL: IPv6 routing protocol for low-power and lossy networks, 2012. Available at: https://datatracker.ietf.org/doc/html/rfc6550

[41] Li, T., Sahu, A. K., Talwalker, A., and Smith, V. "Federated Learning: Challenges, Methods, and Future Directions," 2019. Available at: https://arxiv.org/pdf/1908.07873.pdf

[42] Iqbal, Z. and Chan, H.Y. "Concepts, Key Challenges and Open Problems of federated learning," International Journal of Engineering, 2021. Available at: https://www.ije.ir/article_132537.html

[43] Almanifi, O. R. A., Chow, C-O., Tham, M-L., Chuah, J. H.m and Kanesan, J. "Communication and computation efficiency in federated Learning: A survey," Elsevier, ScienceDirect, Internet of Things, Volume 22, 2023. Available at: https://www.sciencedirect.com/science/article/pii/S2542660523000653