UNIT III : Aspects of communication using IoT Protocol

Syllabus:

1. **WSN protocol**
2. **SCADA and RFID**
3. **6LowPan, COAP protocol**
4. **Sensor deployment and Node discovery**
5. **Data Aggregation and dissemination**

Wireless Sensor Network

Wireless Sensor Network (WSN), is an infrastructure-less wireless network that is deployed in a large number of wireless sensors in an ad-hoc manner that is used to monitor the system, physical, or environmental conditions.
Sensor nodes are used in WSN with the on board processor that manages and monitors the environment in a particular area. They are connected to the Base Station which acts as a processing unit in the WSN System. The base Station in a WSN System is connected through the Internet to share data. WSN can be used for processing, analysis, storage, and mining of the data.

FEATURES WIRELESS SENSOR NETWORKS

 A network configuration where every participant can communicate with one another without going through a centralized point.

 WSN comprises a large number of distributed, self-directed, tiny, low powered devices called sensor nodes

 WSN has embedded devices that are networked to supportively collect, process, and convey data to the users.

 WSN has restricted computing and processing capabilities.

 WSN can be defined as a self-configured and infrastructure-less wireless networks

 It is used to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants

 It pass their data through the network to a main location or sink where the data can be observed and analysed

 A sink or base station acts like an interface between users and the network.

 One can retrieve required information from the network by injecting queries and gathering results from the sink.

 A wireless sensor network contains hundreds of thousands of sensor nodes. The sensor nodes can communicate among themselves using radio signals.

 A wireless sensor node is equipped with sensing and computing devices, radio transceivers and power components.

 The individual nodes in a wireless sensor network (WSN) are inherently resource constrained.

 They have limited ¬ processing speed ¬ storage capacity ¬ communication bandwidth

 After the sensor nodes are deployed, they are responsible for selforganizing an appropriate network infrastructure often with multi-hop communication with them.

 Then the onboard sensors start collecting information of interest

 Wireless sensor devices also respond to queries sent from a ―control site‖ to perform specific instructions or provide sensing samples.

 The working mode of the sensor nodes may be either continuous or event driven.

 Global Positioning System (GPS) and local positioning algorithms can be used to obtain location and positioning information

Advantages

- **Low cost:** WSNs consist of small, low-cost sensors that are easy to deploy, making them a cost-effective solution for many applications.
- **Wireless communication:** WSNs eliminate the need for wired connections, which can be costly and difficult to install. Wireless communication also enables flexible deployment and reconfiguration of the network.
- **Energy efficiency:** WSNs use low-power devices and protocols to conserve energy, enabling long-term operation without the need for frequent battery replacements.
- **Scalability:** WSNs can be scaled up or down easily by adding or removing sensors, making them suitable for a range of applications and environments.
- **Real-time monitoring:** WSNs enable real-time monitoring of physical phenomena in the environment, providing timely information for decision making and control.

Disadvantages

- **Limited range:** The range of wireless communication in WSNs is limited, which can be a challenge for large-scale deployments or in environments with obstacles that obstruct radio signals.
- **Limited processing power:** WSNs use low-power devices, which may have limited processing power and memory, making it difficult to perform complex computations or support advanced applications.
- **Data security:** WSNs are vulnerable to security threats, such as eavesdropping, tampering, and denial of service attacks, which can compromise the confidentiality, integrity, and availability of data.

- **Interference:** Wireless communication in WSNs can be susceptible to interference from other wireless devices or radio signals, which can degrade the quality of data transmission.
- **Deployment challenges:** Deploying WSNs can be challenging due to the need for proper sensor placement, power management, and network configuration, which can require significant time and resources.
- while WSNs offer many benefits, they also have limitations and challenges that must be considered when deploying and using them in real-world applications.

Applications of wireless sensor network

1.Military applications 2. Area monitoring 3. Transportation 4. Health applications 5. Environmental sensing 6. Structural monitoring 7. Industrial monitoring 8. Agricultural sector 9. Neighbor node discovery 10. Smart sensing 11. Data storage and processing 12. Data aggregation 13. Target tracking 14. Control and monitoring

In Detail:

1. **Military Applications:** WSNs are crucial for battlefield surveillance, intrusion detection, and reconnaissance. They facilitate real-time tracking of enemy movements, monitoring of critical infrastructure, and early warning systems for potential attacks.
2. **Environmental Monitoring:** WSNs are used to monitor air and water quality, track forest fires, detect landslides, and monitor other environmental conditions. This enables proactive measures for disaster prevention and mitigation.
3. **Health Applications:** WSNs play a vital role in patient monitoring, enabling continuous tracking of vital signs, remote diagnosis, and personalized healthcare solutions. They can be used for fall detection, elderly care, and managing chronic diseases.
4. **Industrial Monitoring:** WSNs facilitate real-time monitoring of industrial processes, equipment health, and environmental conditions. This enables predictive maintenance, improved efficiency, and reduced downtime.
5. **Transportation Applications:** WSNs are used for intelligent transportation systems, including smart parking, traffic management, and vehicle tracking. They help optimize traffic flow, reduce congestion, and improve road safety.
6. **Agriculture:** WSNs are employed in precision agriculture for monitoring soil conditions, optimizing irrigation, tracking livestock, and managing greenhouses. This leads to increased crop yields and efficient resource utilization.
7. **Smart Homes and Buildings:** WSNs automate various functions in smart homes, including lighting, temperature control, and security systems. They also monitor structural health and energy consumption in buildings.

8. **Disaster Management:** WSNs are used for early warning systems for earthquakes, tsunamis, and other natural disasters. They also aid in rescue operations and post-disaster recovery efforts.
9. **Other Applications:** WSNs are also utilized in areas like wildlife tracking, structural health monitoring, and underwater exploration.

SCADA

SCADA stands for supervisory control and data acquisition. It is a type of software application program for process control. SCADA is a central control system that consists of controllers network interfaces, input/output, communication equipment, and software. SCADA systems are used to monitor and control the equipment in the industrial process which includes manufacturing, production, development, and fabrication. The infrastructural processes include gas and oil distribution, electrical power, water distribution. Public utilities include a bus traffic system, airport. The SCADA system takes the reading of the meters and checks the status of sensors in regular intervals so that it requires minimal interference of humans.

SCADA:

SCADA (supervisory control and data acquisition) is a category of industrial control systems (ICS) that remotely gathers data in real time from industrial processes in order to supervise and control equipment and conditions. SCADA tools enable organizations to control and monitor their industrial devices and make data-driven decisions regarding their industrial processes. SCADA can be used to manage almost any type of industrial process via a graphical user interface (GUI). Both SCADA and ICS fall under the broader category of operational technology (OT).

Components of a SCADA system

SCADA systems gather real-time data to help with the remote management of industrial sites in order to enhance industrial efficiency and output. A SCADA system requires both hardware and software components. The most common hardware devices in a SCADA system include the following:

- **Field devices (sensors and actuators).** A sensor is a device that detects inputs from industrial processes. It functions like a gauge or meter, displaying the status of a machine. An actuator controls the mechanisms of the various processes. It acts like a switch, dial or control valve to control a device. Both sensors and actuators are controlled and monitored by SCADA field controllers.
- **SCADA field controllers.** These elements are microcomputers that interface directly with sensors and actuators and send data to the human-machine interface (HMI). They also send control commands to the field devices to which they are attached. Field controllers come in two varieties:
- **Remote telemetry units.** RTUs, also called remote terminal units or remote telecontrol units, interface with field devices such as sensors, actuators and valves to collect telemetry data and transform it into useful information for human consumption. RTUs are often placed in remote locations, and can be programmed to suit different environmental conditions and applications. They can then initiate specific actions based on specific triggers.
- **Programmable logic controllers.** A PLC is a small digital computer typically used to control industrial processes based on certain inputs, as well as the standards and instructions provided for a specific process or application. PLCs can monitor and control many types of complex, automated and repeatable processes on a constant basis.
- **SCADA supervisory computers.** These control all SCADA processes and are used to gather data from field devices. They also send commands to those devices in order to control industrial processes.
- **Communication infrastructure.** Communication infrastructure enables SCADA supervisory systems to communicate with field devices and field controllers. This infrastructure enables SCADA systems to collect data from field devices and to control those devices.

SCADA software is HMI software**.** It consolidates and presents data from SCADA field devices, PLCs and RTUs for human consumption. HMI software usually includes a GUI that enables operators to understand, control, analyze and modify the status of SCADA-controlled equipment and processes.

The HMI provides a human-friendly way to interpret data, react to alarms and make data-driven decisions. For example, the SCADA system might send a notification to a user's phone that a device is not working properly. The operator can then decide to stop the device via the HMI. The HMI sends the appropriate command to the PLC or RTU, which then forwards the command to the malfunctioning device and causes it to stop.

Features of SCADA systems

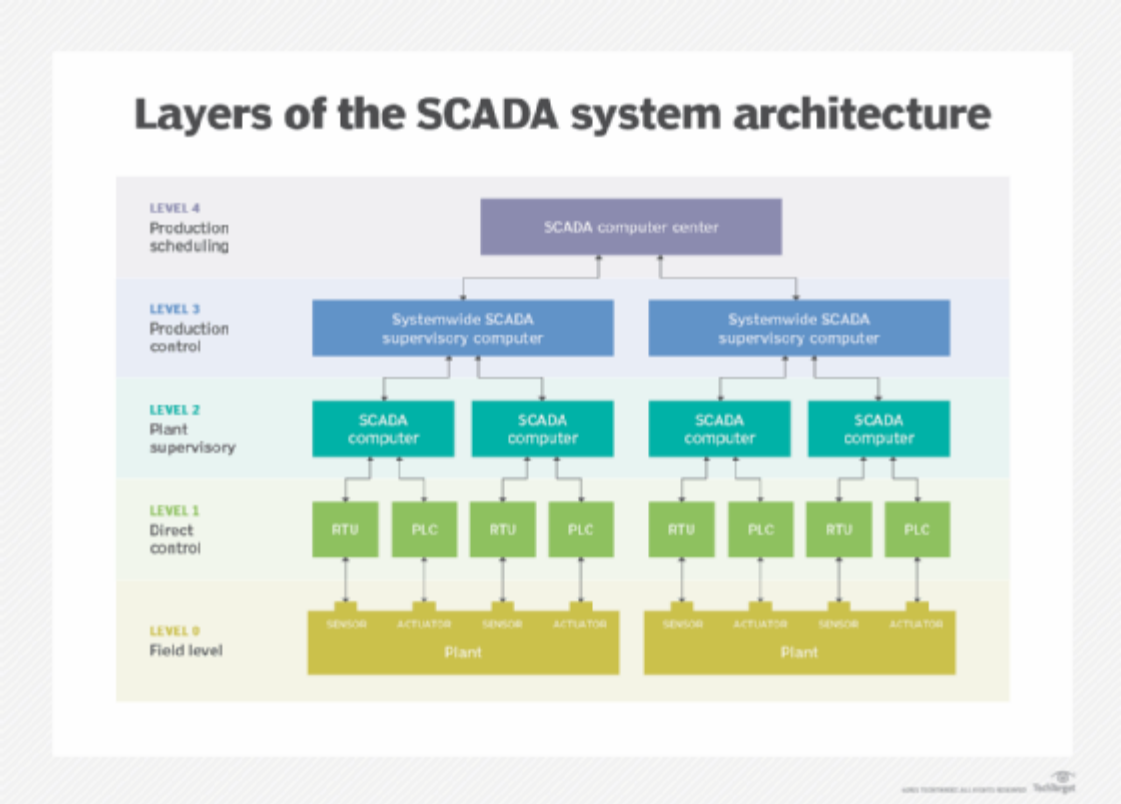
Most SCADA systems include the following features:

- **Data acquisition** is a foundation of SCADA systems. Sensors collect data and deliver it to field controllers, which, in turn, feed data to SCADA computers. The information is then delivered to an HMI device to support human analysis and decision-making.
- **Remote control** through the control of field actuators is based on the data acquired from field sensors.
- **Networked data communication** enables all SCADA functions. Data collected from sensors must be transmitted to SCADA field controllers that communicate with the SCADA supervisory computers. Remote control commands are transmitted back to actuators from the SCADA supervisory computers.
- **Data presentation** is achieved through HMIs, which represent current and historical data to the operators running the SCADA system.
- **Real-time and historical data** are both important parts of the SCADA system. Users can use live data for real-time equipment tracking and maintenance, and historical data to track current performance against historical trends and determine what -- if any -- improvements are required.
- **Alarms** alert SCADA operators to potential problems or significant conditions. Alerts can be configured to notify operators when processes are blocked, when systems fail or when other aspects of SCADA processes must be stopped, started or adjusted.
- **Reporting** on SCADA system operations can include reports on system status, process performance and reports customized to specific uses.

SCADA architecture

SCADA systems operate at five of the six levels defined in the Purdue Enterprise Reference Architecture (PERA), a conceptual model that shows the relationships between OT and information technology (IT):

- **Level 0.** The field level includes field devices, such as sensors -- used to forward data relating to field processes and actuators used to control processes.
- **Level 1.** The direct control level includes local controllers, such as PLCs and RTUs, that interface directly with field devices in ways that include accepting data inputs from sensors and sending commands to actuators. Control processors and process-specific microcontrollers are also placed at this level.
- **Level 2.** The plant supervisory level includes local supervisory systems that aggregate data from level controllers and issue commands for those controllers to carry out. Examples of such systems include alarm servers and process analytic systems.
- **Level 3.** The production or site-wide supervisory/control level includes systemwide supervisory systems that aggregate data from Level 2 systems to produce ongoing reporting to the production scheduling level, as well as other site or regionwide functions, such as alerts, reporting and HMIs.
- **Level 4.** The production scheduling level includes business systems and networks used to manage ongoing processes.
- **Level 5.** The level supporting individual business units and users with systems typically situated in corporate data centers for the likes of servers supporting human resources systems, internal email, CRM, document management, backup, Active directory and security operations.



SCADA use cases and industry examples

SCADA automates and manages many kinds of industrial processes that are too complex or cumbersome for human monitoring and control. SCADA is particularly useful for monitoring and controlling remote processes, where it can help reduce waste and improve overall production efficiency.

SCADA is most commonly used in these industrial settings:

- Electricity generation and distribution.
- Oil and gas refining operations.
- Telecommunications infrastructure.
- Transportation and shipping infrastructure.
- Fabrication and other industrial processing.
- Food and beverage processing.
- Recycling operations.

- Chemical manufacturing.
- Water and waste control.

RFID:

RFID full form is Radio Frequency Identification. It is an automatic identification method that uses wireless non-contact radio frequency waves in which data is digitally encoded in RFID tags or smart labels which can be read by reader through radio waves. The transfer of data takes place between a reader and a movable thing that can be identified & track. RFID can be considered similar to barcodes as data read from tags are stored in database or you can say device that captures the label's data, and stores data in a database.

What is RFID?

Radio Recurrence Distinguishing proof (RFID) is an abbreviation for Radio Recurrence ID. The guideline is the non-contact information correspondence between the RFID peruser and the RFID tag to accomplish the motivation behind recognizing the objective, which is a sort of programmed distinguishing proof innovation.

Direct non-contact two-way information correspondence through radio recurrence, and utilize radio recurrence to peruse and compose recording media (electronic labels or radio recurrence cards), to accomplish the reason for recognizing targets and information trade, which is viewed as the most possible improvement in the 21st 100 years. One of the data innovation.

Components of RFID System

An RFID (Radio-Frequency Identification) system typically consists of the following components:

- **RFID Tag:** Also known as a transponder, this is a small wireless device that contains a microchip

  and an antenna. The tag stores information about the item it is attached to, such as its unique identification number and other relevant data.
    - **Integrated Circuit :** It is used for storing & processing data.
    - **Antenna :** It is used for transmitting receiving signal.
- **RFID Reader:** Also known as an interrogator, this is a device that emits radio waves and can read the information stored on an RFID tag. It is connected to a computer or other device that processes the data received from the tag.

- **Antenna:** This is a device that transmits and receives radio waves between the RFID tag and reader. It can be integrated into the RFID tag or reader, or it can be a separate component.
- **Middleware:** This is software that sits between the RFID reader and the back-end computer system. It helps to manage and process the data received from the RFID tags, and can also provide additional functionality such as filtering and error checking.
- **Back-end System:** This is the computer system that stores and processes the data received from the RFID tags. It can be a simple database or a more complex system, depending on the application.
- **Power supply:** Some RFID systems require an external power supply to function, while others use batteries or are powered by the reader through a process known as backscatter.
- **Accessories:** Depending on the application, additional accessories such as mounts, cases, and antennas may be required to optimize the performance of the RFID system
- **Active Tag:** These have their own power supply and allows a read range of about 100 feet.
- **Passive Tag:** A reader inductively gives power to Passive Tags as they don't have their own power supply. Passive Tags are most widely used Tag and their read range is approximately 30 feet

Applications of RFID

- Document tracking.
- Controlling access to restricted areas
- Asset tracking
- Personnel tracking
- Inventory management
- ID badging
- Supply chain management
- Manufacturing
- Healthcare

Features of RFID

1. Applicability: RFID innovation depends on electromagnetic waves and doesn't need actual contact between the two gatherings.
2. Proficiency: The perusing and composing velocity of the RFID framework is incredibly quick, and a commonplace RFID transmission process is normally under 100 milliseconds. High-recurrence RFID reader might recognize and peruse the substance of

various labels simultaneously, which significantly works on the proficiency of data transmission.

3. Uniqueness: Each RFID tag is novel. Through the balanced correspondence between RFID labels and items, the ensuing course of every item can be plainly followed.
4. Effortlessness: RFID labels have a straightforward construction, high acknowledgment rate and basic understanding hardware. Particularly as NFC innovation is continuously promoted on cell phones, every client's cell phone will turn into the least complex RFID reader.

6LoWPAN

6LoWPAN is an IPv6 protocol, and It's extended from IPv6 over Low Power Personal Area Network. As the name itself explains the meaning of this protocol is that this protocol works on Wireless Personal Area Network. WPAN is a Personal Area Network (PAN) where the interconnected devices are centered around a person's workspace and connected through a wireless medium. You can read more about WPAN at WPAN. 6LoWPAN allows communication using the IPv6 protocol. IPv6 is Internet Protocol Version 6 is a network layer protocol that allows communication to take place over the network. It is faster and more reliable and provides a large number of addresses.
6LoWPAN initially came into existence to overcome the conventional methodologies that were adapted to transmit information. But still, it is not so efficient as it only allows for the smaller devices with minimal processing ability to establish communication using one of the Internet Protocols, i.e., IPv6. It has very low cost, short-range, low memory usage, and low 1 bit rate. It comprises an Edge Router and Sensor Nodes. Even the smallest of the IoT devices can now be part of the network, and the information can be transmitted to the outside world as well. For example, LED Streetlights.
It is a technology that makes the individual nodes IP-enabled. 6LoWPAN can interact with 802.15.4 devices and also other types of devices on an IP Network. For example, Wi-Fi. It uses AES 128 link layer security, which AES is a block cipher having key size of 128/192/256 bits and encrypts data in blocks of 128 bits each. This is defined in IEEE 802.15.4 and provides link authentication and encryption.

Basic Requirements of 6LoWPAN

- The device should be having sleep mode in order to support the battery saving.
- Minimal memory requirement.
- Routing overhead should be lowered.

Features of 6LoWPAN

- It is used with IEEE 802.15,.4 in the 2.4 GHz band.
- Outdoor range: ~200 m (maximum)
- Data rate: 200kbps (maximum)
- Maximum number of nodes: ~100

Advantages of 6LoWPAN

- 6LoWPAN is a mesh network that is robust, scalable, and can heal on its own.
- It delivers low-cost and secure communication in IoT devices.
- It uses IPv6 protocol and so it can be directly routed to cloud platforms.
- It offers one-to-many and many-to-one routing.
- In the network, leaf nodes can be in sleep mode for a longer duration of time.

Disadvantages of 6LoWPAN

- It is comparatively less secure than Zigbee.
- It has lesser immunity to interference than that Wi-Fi and Bluetooth.
- Without the mesh topology, it supports a short range.

Applications of 6LoWPAN

- It is a wireless sensor network.
- It is used in home-automation,
- It is used in smart agricultural techniques, and industrial monitoring.
- It is utilised to make IPv6 packet transmission on networks with constrained power and reliability resources possible.

CoAP

CoAP or Constrained Application Protocol, as the name suggests, is an application layer protocol that was introduced by the Internet Engineering Task Force in the year 2014. CoAP is designed for the constrained environment. It is a web-based protocol that resembles HTTP. It is also based on the request-response model. Based on the REST-style architecture, this protocol considers the various objects in the network as resources. These resources are uniquely assigned a URI or Uniform Resource Identifier. The data from one resource to another resource is transferred in the form of CoAP message packets whose format is briefly described later. The Client requests for some resources and in response to that, the server sends some response over which the client sends an acknowledgement. Although, some types of CoAP do not involve the receiver sending acknowledgments for the information received.

How Does CoAP Works

Constrained Application Protocol (CoAP) is an application layer protocol designed for resource-constrained devices and networks, particularly in the context of the Internet of Things (IoT).

- **Client-Server Model**: CoAP model is essentially a client/server model enabling the client to request for service from server as needed and the server responds to client's request.
- **Resource-Oriented**: CoAP treats various objects in the network as resources, each uniquely identified by a URI (Uniform Resource Identifier). Clients can request information about these resources, and servers provide responses2.
- **Methods**: CoAP supports several methods similar to HTTP:
    - GET
    - POST
    - DELETE
    - PUT

- **Asynchronous Messaging**: CoAP messages are asynchronous because it uses the User Datagram Protocol (UDP). Unlike TCP-based protocols, CoAP does not require acknowledgments for every message, which helps conserve energy in resource-constrained devices.
- **Energy Efficiency**: CoAP is designed to minimize energy consumption while simplifying communication between clients and devices. It achieves this by managing resources, providing device descriptions, and supporting mechanisms to determine if a device is powered on or off.

Methods in CoAP

- **GET** - The get method is used to retrieve resource information identified by the request URI. In response to GET method success a 200(OK) response is sent.
- **POST** - The post method creates a new subordinate resource under the parent URI requested by it to the server. On successful resource creation on the server, a 201 (Created) response is sent while on failure a 200 (OK) response code is sent.
- **DELETE** - The delete method deletes the resource identified by the requested URI and a 200 (OK) response code is sent on successful operation.
- **PUT** - The PUT method updates or creates the resource identified by the request URI with the enclosed message body. The message body is considered as modified version of a resource if it already exists at the specified URI otherwise a new resource with that URI is created. A 200 (OK) response is received in former case whereas a 201 (Created) response is received in later case. If the resource is neither created nor modified then an error response code is sent.

CoAP Features

- Lightweight and Simple
- RESTful Architecture
- UDP-Based
- Asynchronous Communication
- Low Header Overhead
- Multicast Communication
- Proxy and Caching

Applications of CoAP

- **Real Time Monitoring in Grid** - Smart cities can monitor the distribution and generation of power remotely. The CoAP sensors could be embedded inside the transformers and the data could be transferred over GPRS or 6LowPAN.
- **Defense utilities** - The armory and tanks are now-a-days fitted with sensors so that information could be communicated remotely without any interference. The CoAP sensors could detect any intrusion. This makes them capable to transfer more data even under low bandwidth network.
- **Aircraft utilities** - The Aircraft sensors and actuators could be connected with other sensors and communication can take place using smart CoAP based sensors and actuators.

Advantages of CoAP Protocol

Following are some of the key benefits of CoAP protocol.

1. **Simplicity and Low Overhead:** CoAP is a simple protocol with minimal overhead due to its operation over UDP.
2. **Power Efficiency:** It allows for short wake-up times and long sleep states, contributing to extended battery life for IoT and M2M devices.
3. **Secure Communication:** CoAP utilizes IPSEC or DTLS for secure communication.
4. **Flexible Communication:** Synchronous communication is not mandatory in CoAP.
5. **Low Latency:** It boasts lower latency compared to HTTP.
6. **Low Power Consumption:** CoAP consumes less power than HTTP.
7. **Reliable Communication:** Through the use of ACK messages, CoAP achieves reliability similar to HTTP while avoiding unnecessary retransmissions.
8. **Ideal for Home Networks:** CoAP is a suitable protocol choice for home communication networks, finding applications in information appliances, communication equipment, and control equipment in smart homes.

   LoRaWAN devices
   Telecom equipment

Disadvantages of CoAP Protocol

Following are some of the drawbacks of CoAP protocol.

1. **Unreliability (UDP-Based):** Due to its reliance on UDP, CoAP is inherently unreliable. Messages may arrive out of order or get lost during transmission. To mitigate this, CoAP implements a stop-and-wait mechanism with exponential backoff retransmission and duplicate detection.
2. **Increased Processing Time:** Acknowledging each received message increases processing time. Furthermore, CoAP does not verify if the received message has been decoded correctly.
3. **Security Overhead:** While it can be secured using DTLS, CoAP is inherently unencrypted (like MQTT). Implementing DTLS adds implementation overhead.
4. **NAT Traversal Issues:** CoAP can face communication challenges for devices behind Network Address Translation (NAT).

Sensor deployment and node discovery:

Sensor deployment and node discovery are crucial aspects of Internet of Things (IoT) systems, especially within Wireless Sensor Networks (WSNs). Sensor deployment refers to the physical placement of sensor nodes in an area of interest, while node discovery is the process by which these nodes locate and communicate with each other to form a network. Effective deployment and discovery are essential for achieving desired coverage, connectivity, and energy efficiency.

Sensor Deployment:

Random Deployment:

Sensor nodes are scattered randomly, often by methods like dropping from a helicopter, which can be suitable for large, inaccessible areas.

Planned Deployment:

Sensor nodes are placed according to a predefined pattern or strategy, potentially optimizing for coverage, connectivity, or energy consumption.

Deterministic Deployment:

Sensors are placed at specific, predetermined locations, often in accessible areas with known conditions.

Stochastic Deployment:

Deployment relies on probabilistic methods, potentially balancing coverage and energy efficiency.

Multi-Objective Deployment:

Deployment strategies may aim to optimize multiple factors simultaneously, such as coverage and energy consumption.

Homogeneous vs. Heterogeneous Deployment:
Sensors can be of the same type (homogeneous) or different types with varying capabilities (heterogeneous).
Node Discovery:

Initial Discovery:

After deployment, nodes need to discover each other and form a network.

Localization:

Nodes may need to determine their own location within the network, which can be achieved through various localization algorithms.

Coverage Hole Detection:

After localization, nodes can identify areas with insufficient coverage (coverage holes) and potentially trigger actions like reconfiguring the network or deploying additional nodes.

Connectivity Maintenance:
Once a network is established, nodes need to maintain connectivity and adapt to changes in the environment or node failures.
Challenges in Sensor Deployment and Node Discovery:

Energy Constraints:

Sensor nodes have limited power, requiring careful energy management strategies.

Limited Resources:

Sensor nodes also have limited processing power, memory, and communication bandwidth.

Scalability:

IoT networks can involve a large number of sensors, requiring scalable solutions for deployment and discovery.

Security:

Ensuring secure communication and data transmission is crucial in IoT deployments.

Coverage and Connectivity:
Achieving adequate coverage and ensuring reliable communication between nodes is essential.
Solutions and Advancements:

Clustering:

Clustering techniques can help reduce energy consumption and improve network performance by organizing nodes into groups.

Advanced Deployment Algorithms:

Algorithms like EDTD-SC (Enhanced Delaunay Triangulation-based Deployment with Sensor Clustering) aim to optimize coverage and energy efficiency.

5G Technologies:

5G technologies are being explored to address scalability, latency, and bandwidth challenges in IoT.

AI and Machine Learning:
AI and ML techniques can be applied to optimize sensor deployment, node discovery, and data analysis.
By addressing these challenges and leveraging advancements in technology, IoT systems can be deployed effectively to monitor and manage diverse environments, leading to significant benefits in various applications.

DATA AGGREGATION:

Data aggregation is collecting and efficiently gathering data from various sources into a more digestible and insightful format. The goal is to generate meaningful patterns and actionable insights to drive strategic decision-making. In the context of IoT, the idea remains the same but with a much broader and more complex scale of operation.

Implementing data aggregation in the Internet of Things (IoT) is an intricate process involving several steps:

- **Step 1: Data collection**

The process begins with data collection, which involves gathering data from multiple IoT devices. This data can be varied, both in terms of type and source. However, its raw and unstructured nature at this stage makes it challenging to directly derive meaningful insights.

- **Step 2: Data pre-processing**

Once the data is collected, it moves on to the pre-processing stage. Here, the raw data is cleaned and standardised. It involves removing inconsistencies, filling in missing values, and resolving conflicts to ensure the data is fit for statistical
analysis.

- **Step 3: Data aggregation**

The next step is data aggregation, where the pre-processed data is compiled and summarised into a more digestible format. This process involves transforming the data by calculation, sorting, or grouping to provide a summary that can offer useful insights.

- **Step 4: Data analysis**

Finally, the aggregated data is analysed. This is where the data is examined in-depth to uncover trends, patterns, and correlations. With the help of data analysis tools, valuable insights are extracted from the aggregated data, which can then be utilised for decision-making.

Data Dissemination:

Data Dissemination is a procedure where the server initiates and manages transfer of data as well as updates. It also helps in maintaining data consistency and cache management. It is defined as "Pushing data to mobile devices from a server or some other computer." Mobile devices can select time and cache required data. In ad-hoc network, traffic is peer to peer. Multi-hop routing is used to communicate data. In wireless sensor network, other traffic models are possible which are as follows:

- Data Collection Model: The source sends data to a collection periodically on demand
- Data Diffusion Model: A sensor node that generates data based on its sensing mechanism's observation.

Data dissemination has two different entities:

- Source: Generating data.
- Event: Something that needs to be reported for example, in target detection, some abnormal activity.

- Sink: A node randomly located in the field, that is interested in events and seeks such information.

- Event in the below diagram indicates the information to be reported or sent. After source receives an interest from the sink, the event is transferred from the source to the sink. Data dissemination is a two step process.

First, the node that is interested in some events, broadcasts its interests to its neighbours periodically. Interests are then propagated through the whole sensor network.

In the second step, nodes that have requested data, send back data after receiving the request. Intermediate nodes in the sensor network also keep a cache of received interests and data