Cooperative IDS for Detecting Collaborative Attacks in RPL-AODV Protocol in Internet of Everything

Erukala Suresh Babu, National Institute of Technology, Warangal, India
Bhukya Padma, National Institute of Technology, Warangal, India
Soumya Ranjan Nayak, School of Computer Engineering, KIIT (Deemed University), Bhubaneswar, India
Nazeeruddin Mohammad, Prince Mohammad Bin Fahd University, Saudi Arabia
Uttam Ghosh, Meharry Medical College, USA*

ABSTRACT

Internet of everything (IoET) is one of the key integrators in Industry 4.0, which contributes to large-scale deployment of low-power and lossy (LLN) networks to connecting people, processes, data, and things. The RPL is one of the unique standardized routing protocols that enable efficient use of smart devices energy, compute resources to address the properties and constraints of LLN networks. The authors investigate the RPL-AODV routing protocol’s performance in combining the advantages of both RPL and AODV routing protocol, which works together in a low power resource-constrained network. The main challenging issue is collaborating the AODV and RPL routing protocol in the LLN network. This paper also models the collaborative attacks such as wormhole, blackhole attack for AODV, and rank and sinkhole attacks to exploit the vulnerability of RPL protocol. Finally, the cooperative IDS combining specification-based and signature-based IDS is proposed to detect the collaborative attacks against the RPL-AODV routing protocol that effectively monitors and provides security to the LLN networks.

KEYWORDS

LLN Network, RPL-AODV Protocol, Collaborative Attacks, IDS

INTRODUCTION

The advancement of technological capabilities bringing in ubiquitous connectivity and low-powered Devices in the Internet of Things (IoT) are a flourishing phenomenon. There has been a meteoric rise in the number of internet-connected devices around us in recent years. Global tech advisory firm Gartner has predicted that by the end of 2021, there will be more than 50 billion connected devices, which is approximately three times the current human population and 70 billion in the next five
years. The Internet of Everything (IoE) is one of the key technologies. integrators in industry-4.0, which contributes to the large-scale deployment of low-power and lossy (LLN) networks (Ghaleb et al., 2018) that enable networked connection among People, processes, data, and things here Things are physical devices and objects connected to the internet for intelligent decision-making; Real-time insight Data is mainly used to leverage these data into more useful information for decision-making, and the process is used to deliver the right information to the right person (or machine) at the right time and finally, People involved in more Relevant, valuable ways. Despite the rise, one of the major inconveniences slowing rapid adaption is the ‘security’ (Mutchler & Warkentin, 2020; Zhou & Jing, 2020) of these devices (Suresh Babu et al., 2016). The myriad of companies that manufacture them and the multiple protocols from the many standards that existing out there is one of the major reasons why there is still a non-standardized approach to solving these problems. In particular, we look at the Mirai malware family, which became notoriously popular in 2016 when it was first used to coordinate a massive denial of service (DoS) (Babu et al., 2016; Van Kerkhoven et al., 2019; Vishwakarma & Jain, 2020) attack using an army of innocent IoT devices (Babu, Dadi, Singh et al, 2022; Hosen et al., 2020; Nagarajan et al., 2021). This attack is part of the bigger picture indicating the rise of attacks using a legion of “compromised devices that the end-users are unaware of. In addition to the security issue, LLN Networks also possess some other challenging issues, such as (1) the maximum The size of the packet at the physical layer is 127 bytes, which results in the maximum The size of the frame at the data link layer is 102 bytes (Babu, Kavati, Nayak et al, 2022; Babu et al., 2015). If we security parameter, it includes security overhead, which is still reduced to 81 bytes on the link layer. (2) Low bandwidth for such a constrained network includes data rates of 20 kbps, 40kbps, and 250 kbps for each physical layer defined at 868 MHz, 915 MHz, and 2.4 GHz, respectively, (3) Device location is not predefined. Sometimes devices move to a new location, and (4) Devices in LLN may go to sleep for energy conservation. Such devices can’t communicate when they are in sleep mode. (5) It consists of a large number of restricted devices with low power and processing, limited memory, and energy when the devices are battery-operated. (6) All the nodes in LLN are connected through lossy links, which are generally not stable and supports low data rates.

The routing protocol for low-power and lossy networks (RPL) (Dall’Ora et al., 2014; Idris Khan et al., 2014) is one of the unique standardized routing protocols that enable efficient use of smart devices’ energy and compute resources, build flexible topology and data routing that are used to address the above properties and constraints of the LLN networks. This protocol defines IPv6 routing protocol for low-power and lossy networks (RPL), which provides the Multi-point-to-point (MP2P) and point-to-multi-point (P2MP) traffic operation. In addition, this RPL protocol uses an IPv6 protocol stack with additional features like- optimizing the IPv6 over IEEE 802.15.4, enabling header compression, fragmentation, and & reassembling, etc. However, this routing protocol had certain limitations such as Single Path Routing, Implicit Hop Count Impact, Downward Routing Incompatible Modes, and Memory Constraints in Storing Mode, Non-Storing Mode contains Long Source Header, This paper addresses the challenging issue: (1) The AODV (Patel et al., 2014) and RPL routing protocol in the LLN network by resolving the above limitations of AODV and RPL protocol where AODV protocol uses a flooding mechanism and while RPL protocol is mainly used in the restricted network. (2) The AODV and RPL protocol itself has many security risks and possibilities of attacks (Chang et al., 2014; Kumar et al., 2015; Van Kerkhoven et al., 2019). Most such efforts have been put into a mechanism to defend against individual attacks (Foley et al., 2020; Pongle & Chavan, 2015) against AODV and RPL routing protocol. (3) Intrusion detection (Shen et al., 2020; Wang et al., 2021) fails to handle high detection rates, early detection of known attacks, the ability to detect novel, unknown attacks, and a low false-positive rate.

Motivation
There has been a meteoric rise in internet-connected devices around us in recent years. The Internet of Everything (IoET) is one of the key technology integrators in industry 4.0, contributing to the
large-scale deployment of low-power and lossy (LLN) networks that enable networked connection among people, processes, data, and things. However, all the nodes in LLN are connected through lossy links, which are generally not stable and support low data rates. Many restricted devices have low power and processing limited memory and energy when battery-operated. In addition, the major inconveniences of these LLN networks mainly depend upon the “security” of the devices. The myriad of companies that manufacture them and the multiple Protocols from the many standards that exist out there are one of the major reasons why there is still a non-standardized approach to solving these problems. The routing protocol for low-power and lossy networks (RPL) is one of the unique standardized routing protocols that enable efficient use of smart devices’ energy, compute resources, build flexible topologies, and data routing used to address the above properties and constraints of the LLN networks. The RPL protocol itself has many security risks and possibilities for attacks. Most such efforts have been put into a mechanism to defend against individual attacks against the RPL routing protocol. Intrusion detection fails to handle a high detection rate, early detection of known attacks, and the ability to detect novel, unknown attacks, and a low false-positive rate.

Our Contribution

This paper presents the proposed work that detects collaborative attacks against the RPL-AODV routing protocol using the cooperative IDS mechanism in LLN networks. This paper provides the following solutions for the above challenging issues:

1. We investigate the RPL-AODV routing protocol’s performance in combining the advantages of RPL and AODV routing protocols, which work together in low-power resource-constrained LLN networks. This RPL-AODV routing protocol LLN network establishes the path from the origin node to the target node only on an on-demand basis.
2. We modeled the collaborative attacks such as wormhole, and blackhole attack, which exploits the vulnerability of the AODV protocol, and rank attack and sinkhole attack which exploits the vulnerability of the RPL protocol. This collaborative attack may cause a more devastating impact on LLN networks than an uncoordinated attack. The collaborative model was developed to investigate the weakness of AODV and RPL protocol in LLN networks that exploit the LLN environment’s vulnerabilities. These collaborative attacks use the combined efforts of more than one attacker against the target victim from a security perspective.
3. A hybrid IDS has been proposed by combining signature and specification-based techniques to overcome the limitations of signature and anomaly-based approaches. This combination enables the hybrid IDS to detect both cases, either signature or specification attacks, and consumes less energy. The proposed cooperative IDS is a hybrid-based intrusion detection system using Ensemble Machine Learning Approach that combines specification-based and signature-based IDS as cooperative IDS to detect collaborative attacks against the RPL-AODV routing protocol that effectively monitors the LLN network.

PRELIMINARIES

Overview of LLN Networks

Low-power and lossy (LLN) networks (Ghaleb et al., 2018) enable networked connections among people, processes, data, and things. However, resource constraint for devices in LLN includes low power, low processing, and low storing capabilities; the communication technologies are subject to highly asymmetric characteristics of the link, high loss of data, low data rates, and variable data loss on lossy links and communication in short range. The nodes in LLN usually have similar characteristics, although there may be differences in node storing and computing capabilities. In this matter, IETF has defined sensor nodes depending upon the capabilities of nodes into several classes, i.e., class 0,
class 1, and class 2. Devices in class 0 are highly constrained in processing and memory and are not capable of communicating without a gateway node. Devices in class 1 are less restrictive than class 0 devices and can communicate without a gateway node. Devices in class 2 are the least restrictive and can support a protocol stack similar to that used in traditional computers. LLN has the following main characteristics:

- The maximum size of the packet at the physical layer is 127 bytes; this results in the maximum size of the frame at the data link layer being 102 bytes. Other than this, there can be security overhead on the link layer; thus, for data packets, there can be a maximum possible size is 81 bytes.
- Low bandwidth for such a constrained network includes data rates of 20 kbps, 40 kbps, and 250 kbps for each physical layer, defined at 868 MHz, 915 MHz, and 2.4 GHz, respectively.
- Device location is not predefined, and sometimes devices move to a new location.
- Devices in LLN may go to sleep for energy conservation, and such devices can’t communicate when they are in sleep mode.
- It consists of many restricted devices with low power and processing, limited memory, and energy when the devices are battery-operated.
- All the nodes in LLN are connected through lossy links, which are generally not stable and support low data rates.
- It supports different patterns of traffic, not only point-to-point (P2P) but point-to-multi-point (P2MP) or multi-point-to-point (MP2P) also in many cases.

Challenges in LLN

Some of the challenges of the LLN networks are presented below:

- **Duty Cycle and Power:** Battery-operated wireless devices must keep the percentage of time active low. In IP, the assumption is device is always connected.
- **Multicast:** IEEE 802.15.4, which is embedded wireless radio technology, does not support multicasting, and in such a constrained network, flooding is a waste of bandwidth and power.
- **Frame Size and Bandwidth:** Generally, embedded wireless radio technologies have a limited bandwidth range of 20-250 kbps while the frame size is 40-200 Bytes. In the case of IEEE 802.15.4 frame size is 127 bytes. In standard IPv6, the minimum size of the frame is 1280 bytes and therefore requires fragmentation.
- **Reliability:** In a wireless embedded network, unreliability problem occurs due to low energy or energy exhaustion, node failure, and sleep duty cycle.
- **Limited Management and Configuration:** LLN devices have limited capabilities for input, and it is hard to reach the location of such devices. Therefore, the protocols used in LLN must have minimized configuration and are easy for bootstrapping.
- **Fragmentation and Reassembly:** In IEEE 802.15.4, the maximum frame length is 127 bytes at the data link layer, which does not match the maximum transfer unit of 1280 bytes in IPv6. So to transmit IPv6 frames over the wireless radio links in IEEE 802.15.4, the frames are required to divide into different small segments. For this work, the extra overhead is generated in the header to reassemble the data packets at the end in the correct sequence. When there is the reassembly of the data packets, the extra overhead is removed, which was added earlier, and the data packet is restored to its original IPv6 format. Based on the routing used, there can be different fragmentation sequences. When it meshes under routing, then at the final destination, only other fragments are reassembled, while when it is a route over the network, packets are reassembled at every hop. Therefore, every node needs sufficient storage to route over the network for the fragments. More traffic is generated since all the fragments pass instantly in the mesh under the system. In mesh under the system, if a single fragment is missing at the time of reassembling, there is a need to
retransmission—the whole packet. Since when the devices are battery-operated, fragmentation needs to be avoided. Memory need is a major factor since the reassembly of all fragments is done at the final destination. Therefore, header compression and keeping the payload low are of utmost importance.

- **Header Compression:** In the most pessimistic scenario, the greatest size accessible for transmitting IP parcels over an IEEE 802.15.4 wireless frame is 81 B, and without optional headers, the header in IPv6 is 40 bytes. After this, only 41 bytes are left for the upper layer protocols like TCP and UDP. 8 bytes are used in the UDP header, while 20 bytes are for the TCP header. This leaves data over UDP of 33 bytes and 21 bytes over TCP. Fragmentation and reassembly are also required, consuming more bytes, and leaving only a few data bytes. Hence, if one somehow managed to utilize the protocols as may be, it leads to more fragmentation and reassembly; this happens even when the packet size is just 10s of bytes. This point requires header compression.

**Overview of Routing Protocols**

- **Ad-Hoc On-Demand Distance Vector (AODV):** Distance vector routing protocols include AODV. In a distance vector, every node knows the distance to reach other nodes their neighbors. AODV supports both multicast and single cast. AODV follows three mechanisms: route discovery, generation of route messages, and route maintenance. When there is an on-demand route request from a node, the route discovery operation is performed by AODV, which is a similar process in the DSR protocol. Following are the four various AODV messages:
  - **RREQ (Route Request Message):** When a new route is needed from source to destination, RREQ is used.
  - **RREP (Route Reply Message):** It acknowledges the RREQ.
  - **RERR (Route Error Message):** It is a message for route error.
  - **To check the presence of active neighbors, broadcasting Hello messages is done periodically.**

AODV forms a path from the source to the destination using the RREQ-RREP cycle. Whenever any source node wants a route to send a packet to its destination, it floods RREQ in the network. For the starting node and nodes which receive the packet, the routing table contains the backward pointer for those nodes and thus updates the route information. For the destination, in addition to the IP address of the source, the current sequence number, and a broadcast id, RREQ contains the updated sequence number. A node that is either the destination node (D) or an intermediate that knows the path to the D with a higher sequence number that received the route request message will transmit the route reply message. Nodes can maintain both RREQ’s broadcast id and the IP address of the source.

**Overview of Routing Protocol in LL (1) Network**

The routing protocol for low-power and lossy networks in IPv6 infrastructure, i.e., RPL, is suitable for resource-constrained devices. The main focus of RPL is to give IPv6 infrastructure to the wireless embedded devices, which are battery operated that communicate using low-power radios, and those devices can deliver the data over a number of hops. This routing protocol was useful in different applications in the wireless sensor network and IoT domain. It is treated as a critical component that connects the IETF protocol application layer for LLN to the low-power network. The resource-constrained devices include power restrictions, storage, and processing restrictions. The communication between different nodes is subject to low data rates, high packet loss, a small communication range, limitation on frame size and dynamically changing network topology Routing Requirements in LLN:
• **Traffic Support**: LLN routing protocol should be able to yield two-direction connectivity between two arbitrary nodes and can support multicast and unicast.

• **Resource Constraint**: Resource constraints must be implemented in resource-limited nodes (e.g., the memory of 128kB (host) for 8-bit devices and memory of 256kB (router) for 16-bit devices).

• **Path Diversity**: For reliable delivery of packets, there must be alternate routes for the transmission of packets. The delivery ratio of packets should not be more than 3-retransmission over unreliable lossy links.

• **Convergence Time**: Routing must converge when a new node joins the network within several minutes.

• **Node Property Awareness**: It must consider node characteristics, for example, sleep interval, memory, and power budget for routing.

• **Heterogeneous Routing**: It must be likely to create several ways under various functionalities for various flows to guarantee that the critical application can’t be conceded and less critical applications can access the network.

• **Security**: Protocols must support message integrity to avoid malicious nodes or attackers from changing routing functions or participating in routing decisions.

**RPL’s Key Feature**

RPL key highlights a type of DVR (distance vector routing protocol) that fabricate DAG (directed acyclic graph) based on selected routing constraints and metrics. The DAG is built-in support of efficient upstream traffic pattern support with resource-constrained nodes. RPL basis is to construct Destination Oriented Directed Acyclic Graph, i.e., DODAG, which is rooted in at least a single DODAG root and supports IPv6 bidirectional communication between the nodes in the network. The RPL’s process of parent selection and control message with the simple structure of a network in which there is only one DODAG. Every node in the graph advertises its routing constraints and metrics via DIO (DODAG Information Object). When a node receives DIO from its neighbor, it chooses its preferred parent for routing based on the objective function (OF), and it collects path information (e.g., DODAG ID, RANK) from DIO and creates a route topology (DODAG). In RPL, there can be several parent nodes for a single node to accomplish reliable delivery of packets via path diversity. Based on the Trickle Timer, DIO messages are sent to accomplish a balance between fast recovery/convergence and control overhead (consumption of energy). As a request, upon receiving DIS, i.e., DODAG Information Solicitation, DIO is sent to the node from which DIS was received. Rank can be defined by objective function to show route distance from a particular device to LLN Border Router. Metrics of node and link are useful in parent selection and rank calculation.

**Overview of an IDS System**

This subsection presents the signature-based, anomaly-based, and specification-based IDS used to detect the various attacks (Seyfollahi & Ghaffari, 2021; Xu et al., 2019):

**Signature-Based Detector**

In this misuse-based detection mechanism, ecollaborative attack attacks can be encoded in advance and used to match against user behaviour. These collaborative attacks are followed well-defined patterns and signatures that exploit system weaknesses and application software. The proposed misuse-based detection mechanism uses **snort**, which is a popular lightweight signature-based IDS that can analyze real-time traffic analysis and data flow in the network; it uses a basic analysis and security engine (BASE) to generate the alerts, performs the protocol analysis, and finally detects different types of attack. In the LLN network, packets flow from the edge node to the application passing through various network technology, internet technology, and service discovery processes. Snort works as a network packet sniffer that inspects the packet contents with known virus signatures encapsulated as
rules and detects abnormal connections, records events, initiates action and stores the information in a log file/database. In other words, whenever any packet comes into the network, the snort checks the behavior of the network. If the pattern matches and performance degrades, then snort stops the processing of the packet, discards the packet, and stores its detail in the signature database. Finally, compares those packets with the database of known attack signatures, and Warnings will be generated with various attacks occurring in the network, as shown in Figure 2.

Snort IDS provides flexible and quiet powerful rule of descriptive language. These rules are created from known intrusions that occurred in the past and can be applied to detect multiple types of intrusion activity. Every rule consists of two logical parts: the Rule Header and Rule Body. The rule header parts are comprised of five sections; (1) Rule Actions- These actions to be taken when an intrusion is detected, (2) End-to-End source and destination information that contains source and destination IP addresses, and (4) port numbers depending on the protocol used. (3) Direction of traffic (5) Protocol type used such as RPL, AODV, TCP, UDP, or ICMP. While the Rule Body part contains keywords and content. The Keywords are separated by a colon (:) and the content is separated by a semicolon (;). Rule options consist of various conditions that help decide whether the mentioned misuse operation has occurred or not. Options fields are available for all rule types and may be used to generate complex behaviors from the program. The option part of the snort rule structure starts from the parenthesis. It consists of alert (message), strings (content), load length (size), type (class type), priority (priority value), version (rev), and other important data used for mode matching, i.e., offset, depth, etc.

The proposed Snort IDS detects volume-based DDoS (Khader & Eleyan, 2021; Suresh Babu et al., 2019; Van Kerkhoven et al., 2019) attacks that include the combination of HTTP, ICMP, UDP, RPL, AODV, RPL, and other spoofed packet floods to target the victim for the resources. The attacker will randomly spoof the IP source combine UDP packets with port 80 for the destination, and send ICMP echo request packets that target the victim machine. The proposed Snort captures the RREQ, RREP control packets, UDP packets, and ICMP packets recorded in the database and alerts the detected intrusions to the user.

**Anomaly-Based Detector**

Normal data models are built based on normal traffic in the anomaly-based approach. The deviation from the normal model will be considered an attack or anomaly. The main advantage of using this approach is detecting attempts to exploit new and unforeseen vulnerabilities. However, it suffers from a high false alarm rate; this happens mainly because of a lack of a training data set that covers all the legitimate areas. Moreover, it detects and suspects every abnormal behaviour as an intrusion, which is not always an indicator of intrusions. The abnormal behaviour can also happen due to policy changes or the offering of new services by a site. Figure 1 depicts the architecture of anomaly-based detection that contains the network traffic, Feature Generation, Feature Selection, Classification Models, and Alert/Warnings.

**RELATED WORK**

This section presents the literature review that provides the basis for the proposed work. The review of existing techniques aims to detect, prevent, and mitigate IoT routing attacks (Yavuz et al., 2018). The issue of routing attacks mainly arises because of the lack of standards across the domain. It is important to understand that majority of the IoT users are non-technical people using them in the form of CCTV cameras, smartwatches, smart homes, etc. Analyzing the system’s internal workings like networking is not a recommended approach for such people. Therefore, it would make sense if the device manufacturing companies put in measures like closing ports that would normally not be used by end-users and setting up non-trivial credentials. We also studied a prevention mechanism in which unused ports are closed and default login credentials are changed to thwart brute-force DoS
attacks (Almusaylim et al., 2020; Kumar et al., 2015) which are currently the most happening. These attacks can also be stopped if standards are maintained across devices and companies manufacturing them. Some referred papers have described different security attacks (Babu et al., 2016) against the RPL protocol. The different attacks include attacks on topology, attacks against resources, and attacks on traffic. The major consequences of these attacks are denial of service, network congestion, and network instability, leading to performance degradation (Ioulianou & Vassilakis, 2020). Some protocols were mentioned to detect the attacks or general solutions like heart beat protocol, rank authentication, IDS system-based building global view of the network, etc.

The Interconnection of IoT Devices with the IPv6 or 6LoWPAN protocol are very useful for enabling low-powered IoT devices to achieve scalability. Amit Kumar Sikder et al. (2018) proposed an overview of the SLS. They reviewed different IoT-enabled communication protocols, which can be used to realize the SLS in a smart city context. Moreover, the author analyzed different usage scenarios for IoT-enabled indoor and outdoor SLS and provided an analysis of the power consumption. The authors have developed IoT-enabled smart lighting systems that can reduce power consumption by up to 33.33% in indoor and outdoor settings. Chia-Wen Lu et al. (2013) have proposed a SIP-based protocol for effective communication in smart grids. Many existing IP-based services can thus be re-used to monitor WSN’s real-time status. From a perspective on network management service, the author compares the advantages and disadvantages of ZigBee and IP protocols. Since ZigBee is only appropriate for small-scale networks and suffers from the scope expansion of a sensor network.

The 6LoWPAN protocol provides connectivity, compatibility, and coverage using IoT devices. A 6LoWPAN-based neighborhood area network for a smart grid communication infrastructure is proposed in (Abdel Hakeem et al., 2019; Chen et al., 2013; McDermott et al., 2018). A NAN is a key component of a smart grid communication network infrastructure that enables the communication between end devices and various controllers within a smart grid. It can be created to cover a vast geographic area by using infrastructure-based access networks such as WiMAX or LTE-based systems.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoET</td>
<td>Internet of Everything</td>
</tr>
<tr>
<td>LLN</td>
<td>Low-Power and Lossy networks</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>RPL</td>
<td>Routing protocol for low power and lossy network</td>
</tr>
<tr>
<td>IDS</td>
<td>Intrusion Detection System</td>
</tr>
<tr>
<td>AODV</td>
<td>Ad-hoc On-demand Distance Vector</td>
</tr>
<tr>
<td>DoS</td>
<td>Denial of Service</td>
</tr>
<tr>
<td>NS3</td>
<td>Network Simulator-3</td>
</tr>
<tr>
<td>MP2P</td>
<td>Multi-point-to-point</td>
</tr>
<tr>
<td>P2MP</td>
<td>Point-to- multi-point</td>
</tr>
<tr>
<td>WSN</td>
<td>Wireless Sensor Networks</td>
</tr>
<tr>
<td>DODAG</td>
<td>Destination-oriented directed acyclic Graph</td>
</tr>
<tr>
<td>MoP</td>
<td>Mode of operation</td>
</tr>
<tr>
<td>DIO</td>
<td>DODAG Information Object</td>
</tr>
<tr>
<td>RREQ</td>
<td>Route Request</td>
</tr>
<tr>
<td>FPR</td>
<td>False Positive Rate</td>
</tr>
</tbody>
</table>
The author developed a 6LoWPAN-based NAN architecture that can handle all smart meters in a NAN coverage area while meeting the QoS requirements of various applications inside NANs. A protocol of 6LoWPAN and its application in smart lighting and healthcare are proposed in (Le et al., 2011; Verma & Ranga, 2020). These smart lights based on Power Line Communication (PLC) are short on low-data rate and inappropriate communication protocol. They have updated the smart lighting system from PLC to 6LoWPAN. 6LoWPAN nodes replace the PLC nodes, and 6LoWPAN routers replace the controllers. From these implementations, they have gained more advantages in transmission rate, signal range, and compatibility compared to PLC. 6LoWPAN with IP-Standard interconnection makes it easier to integrate various types of sensors for monitoring.

We further investigated various attacks and mechanisms to ensure secure routing against security attacks in the RPL protocol that presents BLSTM-RNN detection, performed at the packet level, focusing on text recognition within features otherwise normally discarded by flow-based techniques (Zaminkar et al., 2021). The BLSTM introduced two independent layers to accumulate contextual information from the past and the future. The authors choose four attack vectors used by Mirai-User Datagram. Protocol (UDP) flood, Acknowledgement (ACK) flood, Domain Name System (DNS) flood, and Synchronize (SYN) flood. Messages between the C&C server and the infected device were captured as the device’s normal data. All the analysis was done using the .pcap files after converting them into the CSV format. The dataset included features such as No., Time, Source, Destination, Protocol, Length, info, and has about 4,000,000 data points. In (Almusaylim et al., 2020), the authors feel that Blockchains can address the problem through centralized DDoS mitigation systems by introducing a distributed database relying on a P2P network that provides reliability and a high level of trust. Blockchain works because when a member node receives a new block, it is verified, and then broadcasted to the rest of the network. The miners are the only nodes allowed to add the block to the blockchain. The miners are selected randomly using a consensus algorithm such as the PoW proof of work. In (Rghiou et al., 2014; Soe et al., 2020; Vishwakarma & Jain, 2020), traditional attack detection systems cannot be located in IoT environments because of the diverse architecture of the underlying network methodologies and the different natures of such devices.

### Table 2. Summary of related work on routing attacks in LLN networks

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ambili et al. (2021)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>Anhtuan et al. (2011)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>Semih Cakir et al. (2020)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Chin-Yang et al. (2003)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>John Foley et al. (2020)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>Jian-Ming et al. (2014)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>M. Zhang et al. ()</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td>M. Napiah et al. (2018)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>9</td>
<td>Ioulionou et al. (2022)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>10</td>
<td>Mina Zaminkar et al. (2021)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>Junqi Duan et al. (2014)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td>Van Kerkhoven et al. (2019)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td><strong>Our Contribution</strong></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>
Additionally, new attacks can be distinct from those already on traditional network devices. Heavy encryption methods cannot be deployed on these resource-hungry devices. Rule-based detection systems are comparatively easier to circumvent, and machine learning-based systems can somewhat detect the variances of many kinds of attacks. Furthermore, the ML classifier training process is tough to implement on these low-resource devices. Authors’ Model - The authors’ model uses ANN, the J48 Algorithm (also called C4.5 and is a descendant of ID3), Naive Bayes, and Correlation-Based Feature Selection. Using multiple different Machine Learning algorithms, select the best matching one according to the detection accuracy obtained for each sub-engine. In this way, the authors successfully create a hybrid detection architecture. In (Al-Shargabi & Aleswid, 2020; Napiah et al., 2018) proposed the Merkle tree-based wormhole attack avoidance mechanism against the DAG-based structure of RPL protocol that generates the hash for the information stored in the tree. The author
has proposed an authentication mechanism for avoiding the promotion of routes but increasing the cost of communication with the root. If an entry fails to be discovered, the authentication element is authenticated with the hashed security element at the root and the public key of the new node which discussed the TRAIL - Trust Anchor Interconnection Loop generic approach to detect and prevent topological inconsistencies (Perrey et al., 2013). Each node is enabled to validate its upward path to the root and detect rank spoofing on it. The key idea of TRAIL is to validate upward paths to the root using a round-trip message. Without relying on encryption chains like VeRA++, a node can conclude rank integrity from a recursively intact upward path. The Summary of related work and comparative study is shown in Table 2.

PROPOSED WORK

This section presents the proposed work that detects collaborative attacks against the RPL-AODV routing protocol using the Cooperative IDS Mechanism in LLN Networks of the Internet of Everything (IoET). First, we investigate the RPL-AODV routing protocol’s performance in combining the advantages of both RPL and AODV routing protocols, which works together in low-power resource-constrained network (LLN) networks (Ambili & Jose, 2020). The main challenging issue in collaborating the AODV and RPL routing protocol in the LLN network of the Internet of Everything (IoET). AODV protocol uses a flooding mechanism, while RPL protocol is only used in the restricted network. Next, we modeled the collaborative attacks (Duan et al., 2014; Suresh Babu et al., 2019; Tseng et al., 2003), such as wormhole, and black attack, which exploits the vulnerability of the AODV protocol, and rank attack and sinkhole attack exploits the vulnerability of the RPL protocol. Specifically, we modeled the collaborative attacks-wormhole, black attacks, rank attacks, and sinkhole attacks against the RPL-AODV routing protocol. Finally, we proposed cooperative IDS with a combination of specification-based and signature-based IDS as cooperative IDS to detect collaborative attacks against the RPL-AODV routing protocol that effectively monitors the LLN network of the Internet of Everything (IoET). The overall work is presented in the proposed architecture as shown in Figure 1.
Overview of RPL-AODV Routing Protocol

The Routing Protocol for Low Power and Lossy Networks (RPL) is the IPv6-based distance vector routing protocol for LLNs of the Internet of Everything (IoET) that supports multipoint-to-point, point-to-multipoint, and point-to-point traffic flows from the root in the destination-oriented directed acyclic Graph (DODAG). This traffic will be happening between different routers within the DODAG. However, the routers do not contain the information of another router of the network. Therefore, the traffic flows information in these networks is operated in two modes. One is the non-storing mode and the storing mode. In the non-storing mode, the root of the DODAG will receive every data packet from the routers or the edge nodes. While in the storing mode, the data packets will be received by the common predecessor node. But, the data packets need to flow over the longer path, resulting in congestion at the root level of the DODAG. In the RPL network, to discover a better path, the originator acts as a local root in the temporary destination-oriented DODAG that introduces the DIO control message. Once the neighbor routers receive the DIO message from the originator node, it adds its IPv6 addresses and then multicast this DIO message to the target node. The whole process is encapsulated using point-to-point route discovery P2P-RDO options in DODAG, either hop-by-hop or source routing mode. However, both the hop-by-hop and source routing mode adds the extra overhead of the address vector that restricts satisfying the objective function constraint. The RPL-AODV protocol uses point-to-point route discovery features of the RPL protocol with different operation modes. The RPL-AODV protocol uses two different multicast messages to find the possible asymmetric routes to achieve high route diversity. RPL-AODV eliminates the need to address overhead in the case of hop-by-hop mode. A significant reduction in control packet size is useful for restricted LLN low-power networks (Khanuja & Adane, 2020).

Modeling of RPL-AODV Routing Protocol for Internet of Everything (IoET)

This paper proposes an RPL-AODV routing protocol in low power and lossy network (LLN) that establishes the path from the origin node to the target node only on-demand basis. The mechanism of the route discovery process in RPL-AODV protocol performs reactively when the origin node wants to transmit the data packet to the target node, when is no route, or the existing route doesn’t satisfy the requirement. This route discovery process of the RPL-AODV protocol achieves high route diversity with the help of asymmetric communication using bidirectional links for finding the route from the origin node to the target node and from the target node to the origin node and also eliminates the constraints of traversing a common predecessor node, which is there in original RPL protocol. Further, RPL-AODV enables route discovery for symmetric communication along paired DODAG, as shown in Figure 2. In discovering the routes, the RPL-AODV uses a route discovery process containing two control messages, route request (RREQ) and route reply (RREP). The discovery of routes is achieved by forming a temporary DODAG, which is rooted at the origin node. The paired DODAG Instances are constructed based on the mode of operation (MoP) in RPL-AODV. The RREQ control messages instance is sent from the origin node to the target node, and the RREP control message instance is sent back from the target node to the origin node. The rank calculated from DODAG Information Object (DIO) helps intermediate routers join the DODAG instance. The transmission of data from the origin node to the target node is based on the route found in the RREP instance, and the transmission of acknowledgment from the target node to the origin node is based on the route found in the RREP instance.

The following Figure 3 shows the RPL-AODV DIO option, which contains the RREQ-DIO message that comprises the DODAG ID field filled by the IPv6 address of the origin node. RREQ-DIO must carry out only one RREQ option in RPL-AODV MoP. In the RREQ option, the origin node forwards the following information: 

- **Type**: The type assigned to the RREQ option.
- **Option Length**: The option length in octets, omitting the length field and type field. Due to the number of octets and address vector, option length is variable. ‘S’ indicates the symmetric bit, representing a symmetric path from the origin node to the router transmitting the RREQ-DIO. ‘H’ is the value of H set to
zero, which indicates source routing, and one represents hop-by-hop routing. This flag controls both upstream and downstream routes. ‘X’ is reserved. ‘Compr’ is an unsigned integer of 4 bits, which is the useful field when H=0, i.e., in the case of source routing, and if it is hop-by-hop routing, i.e., H=1, upon reception, it is ignored and set to zero. ‘L’ is an unsigned integer of 2 bits indicating the time duration for a node in the RREQ instance for which it belongs to a temporary DAG, including both the target node and the origin node. Once the time is over, the node has to leave the directed acyclic graph (DAG), and for temporary DODAG, the node has to stop receiving or sending DIO. ‘MaxRank’ represents the upper limit of the rank. Orig SeqNo is the sequence number of the origin node, defined the same as in the AODV protocol. ‘Address Vector’ is an IPv6 address vector indicating the path that RREQ-DIO has passed. It is present only if the value of H is zero. If a node rank is higher or equal to the max rank, that node should not join the RREQ instance. When the rank of a target node is equal to the max rank, that node can join the RREQ instance. Upon receiving RREQ, if the rank is higher or equal to the max rank, the router must discard it.

Option Length is the option length in octets, omitting the length field and type field. Due to the number of octets and address vector, option length is variable. G is Gratuitous route. H’s value is set to zero, which indicates source routing, and one represents hop-by-hop routing for the downstream route. The value if H here is set the same as in the RREQ option. X is reserved. Compr is an unsigned integer of 4 bits. This field is useful when H=0, i.e., in the case of source routing, and if it is hop-by-hop routing, i.e., H=1, it is ignored and set to zero upon reception. L is an unsigned integer of 2 bits indicating the time duration for a node in the RREQ instance for which it belongs to a temporary DAG, including both the target node and the origin node. Once the time is over, the node has to leave the directed acyclic graph (DAG), and for temporary DODAG, the node has to stop receiving or sending DIO. Rsv: Initialization of Rsv is set to zero, and upon reception, it is ignored. Orig SeqNo: The sequence number of the origin node is defined the same as in AODV. Address Vector: is an IPv6 address vector indicating the path RREP-DIO has passed for the asymmetric path. It is present only if the value of H is zero.

Figure 3. Packet format of DIO RREQ instance option

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------------------------------------------+
<table>
<thead>
<tr>
<th>TYPE</th>
<th>Option Length</th>
<th>S</th>
<th>H</th>
<th>X</th>
<th>Compr</th>
<th>L</th>
<th>Max Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orig Sequence No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Address Vector (Optional, Variable Length)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Figure 4. Format of DIO RREP option

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------------------------------------------+
<table>
<thead>
<tr>
<th>TYPE</th>
<th>Option Length</th>
<th>G</th>
<th>H</th>
<th>X</th>
<th>Compr</th>
<th>L</th>
<th>Max Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shift</td>
<td>Rsv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Address Vector (Optional, Variable Length)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Problem Formulation

Consider an LLN network that consists of both non-storing and storing nodes. In such a network, a large message is further divided into multiple segments, and the packet which is traveling from the node \( N_i \) (origin) to \( N_j \) (destination) is cached in a queue at the intermediate locations \( N_k \) and then transmitted to \( N_h \), which is the next hop of the node. The intermediate nodes act as storing nodes. Requirements for packet flow between nodes are raised at a random time, and there may be packets of various lengths. In such networks, random variables such as packet average delay that average flow in the channel, from origin to the destination node is represented as:

\[
AD = \frac{i}{\gamma} \sum_{i=1}^{NA} \frac{f_i}{(c_i - f_i)}
\]

where:

\[
\gamma = \sum_{i=1}^{NN} \sum_{j=1}^{NN} r_{ij}
\]

\[
AD = \frac{\text{Per packet total average delay (sec / packet)}}{\text{NA = No. of edges}}
\]

\[
r_{ij} = \text{It is the average rate of a packet from i to j (packet/second)}
\]

\[
f_i = \text{Total bit rate on the channel i (bits / second)}
\]

\[
c_i = \text{Channel(i) capacity (bits / second)}
\]

The RPL-AODV protocol is a routing problem in which packets are transmitted from an origin node to some target node through symmetric or asymmetric links in a resource-constrained network. Mathematically, we will represent the RPL-AODV routing problem. Consider a graph \( G \in (V, E) \), where \( V \) is the set of nodes and \( E \) is the edges in the LLN network. The cost \( C_{ij} \) is associated with each edge \( (i, j) \in E \), and each edge has some constraint \( U_{ij} \) on capacity. Let the decision variable be \( V_{ij} \), which is defined per edge \( (i, j) \in E \). Each of \( V_{ij} \) denotes a packet distribution from i to j. \( C_{ij} \cdot V_{ij} \) is the cost of a flow \( V_{ij} \). Every node \( j \) in graph \( G \) satisfies the flow constraint:

\[
\sum_{\{i, j, k \mid E \}} V_{j,k} - \sum_{\{i, j, k \mid E \}} V_{i,j} = b_j
\]

\( b_j \) is the flow amount generated by node \( j \). And to find the flow from the source node to the target node and minimize overall cost subjected to capacity and flow conservation constraints:

\[
\text{Minimize} \sum_{\{(i,j) \mid E \}} C_{ij} V_{ij}
\]

subject to:
\[ \sum_{\{i,(j,k) \in E\}} V_{j,k} - \sum_{\{i,(i,j) \in E\}} V_{i,j} = b_j \quad \forall j \in V(Vertices) \]

\[ 0 \leq V_{ij} \leq U_{ij}; \quad \forall (i,j) \in E(Edges) \]

Due to the restricted LLN network, there should be minimization in the overall link cost and the number of packets distributed, i.e., overall transportation cost should be minimized. Mathematically, we will represent the RPL-AODV routing problem, a flow model for a network with intermediate nodes. Let \( R \) be the RPL-AODV routing algorithm, and \( R(p) \) be the packet probability that uses path \( p \) from the origin node to the destination node. The RPL-AODV routing algorithm \( R \) can be represented as:

\[ \sum_{p \in R_{sd}} R(p) = 1; \forall s,d \in N \]

\[ \sum_{p \in R_{sd}} R(p) \geq 0; \forall p \in P \]

where \( P \) is the set of all paths of the RPL-AODV routing algorithm, and the Cost Function \( C(R) \) should be minimized:

Subject to \( \sum_{p \in R_{sd}} R(p) = 1 \)

\[ \sum_{p \in R_{sd}} R(p) \geq 0 \]

In the RPL-AODV routing protocol, the total number of packets sent by the source node is equivalent to the total packet received and is represented as:

\[ \sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \]

To minimize the total distribution cost from source node \( i \) to target node \( j \):

\[ \text{Minimize } y = \sum_{i=1}^{n} \sum_{j=1}^{m} C_{ij} x_{ij} \quad (1) \]

such that:

\[ x_{ij} > 0 \]

The packets are sent from source node \( i \) to all possible target nodes with available routes at that source:
subject to $\sum_{j=1}^{n} x_{ij} = \alpha_i$ \hspace{1cm} (2)

The packets that are sent to the target node $j$ from all possible source nodes ought to be equivalent to the received at that target node $j$:

$$\sum_{i=1}^{m} - x_{ij} = -b_j$$ \hspace{1cm} (3)

where $\alpha_i =$ No. of packets sent from the source; $b_j =$ No. of packets received at the destination; $c_{ij} =$ cost from source node $i$ to target node $j$, where $i = 1, 2, 3, \ldots, m - 1, m$ and $j = 1, 2, 3, \ldots, n - 1, n$ and $x_{ij} =$ no. of packets to be distributed from source node $i$ to target node $j$ where $i = 1, 2, 3, \ldots, m - 1, m$ and $j = 1, 2, 3, \ldots, n - 1, n$.

Modeling of DDoS Attacks Against the RPL-AODV Routing Protocol

This section collaboratively presents the most prominent well-known attacks such as sinkholes, rank attacks, black holes, and wormholes against the AODV and RPL routing protocols (Babu et al., 2016). Most such efforts have been put into mechanisms to defend against individual attacks in the existing work. The attack can be described as any process, method, or means used to attempt maliciously to compromise the nodes in the LLN network of the Internet of Everything (IoET). Different attacks are based on the AODV and RPL routing protocols used in the LLN network. The attackers’ malicious activities performed in the network include data stealing, data damage, and data modification, denial of services, or depleting network bandwidth and resources. First, we present the attacks against AODV and RPL routing protocol. Existing literature shows wormhole attacks and blackhole attacks are more vulnerable to the AODV routing protocol (Ioulianou et al., 2022), as shown in Figure 5. The wormhole attack is caused by two or more attackers or malicious nodes that can communicate with one another. One node is kept around the router, and the other is kept somewhere else in the LLN network. When one node receives a data packet without sending it on its normal path, it sends it out to
the other node directly through the tunnel. So in this way, the attacker can do a packet manipulation and cause network routing disruption since routing data did not reach every node. The wormhole attack can also be possible using one malicious node that can transmit incorrect information to two legitimate nodes at different locations, convincing them that they are neighbors. The other one is a black hole attack that will work with the perception of intercepting all the messages; the attacker can broadcast false information to all other nodes in the LLN network that has the shortest path. Using a forged RREP packet (route reply), all the packets are attracted by the black hole node, which falsely claims that it has the shortest path to the destination and drops all the data packets without sending them out to the destination. Once the attacker tries to access all the packets and can drop a few all the packets, this can be done either according to the data present in the packets. They assume that the malicious node has complete information about the data content transmitted in the LLN network.

While the rank attack and sinkhole attack will exploit the vulnerability of the RPL routing protocol, the rank attack will exploit the RPL protocol's weakness, which is mainly used to provide optimized routing topology and prevention of loops in LLN networks, as shown in figure 6. The mechanism used for the prevention of loops is based on the concept of rank to show the relationship of the nodes. Every node in the network needs to calculate its rank depending on the data collected from neighboring nodes. Each node requires selecting a preferred parent except that of the sink node, and the parent node rank should not be greater than the children's rank. The node's rank is calculated based on the RPL rule, which states that “in downward direction rank of a node is strictly increasing and in upward direction rank of a node is strictly decreasing.” The attackers will exploit this security flaw. Specifically, when a source node-1 transmits a packet through intermediate nodes (i.e., are the intermediate nodes) to the target node N. Consider be the rank of nodes from respectively. According to the rank rule, if node-1 transmits a packet in the upstream direction to node N, the condition must be satisfied, or if the packet travels in a downward direction, then must be satisfied. Every sender and receiver of the packet must check these conditions along the route. If the condition is not satisfied, the node must inform by setting the bit rank-error in the information of the RPL packet. The attackers will exploit this issue easily by omitting the checking function for rank in the compromised nodes. This attack is hard to disclose because it does not require anything to spoof and looks normal in the behavior of the compromised nodes. The attackers will exploit the following (1) creation of an unoptimized path, (2) disruption in the optimized paths, which results in an undiscoverable path, and (3) undetected creation of loops. These issues degrade the performance of the network, like delay and throughput. The other attack is the Sinkhole Attack against the RPL routing protocol. Intruders launch a sinkhole attack with the help of their rank. Intruder sends out better rank to the other nodes in the network to make them a neighbor in the destination-oriented directed acyclic graph for choosing it as a preferred parent node. This attack focuses on controlling packet traffic through the compromised or malicious node to a great extent in the network. The attacker cheated the authorized node to establish the link with the unauthorized or malicious node by showing that it has the optimal routes. The attacker forwards incorrect information to some legitimate node using a wormhole attack. A sinkhole attack behaves like a wormhole, blackhole, and selective forwarding attack.

In this proposed work, we mainly address collaborative attacks, which may cause a more devastating impact on LLN networks than uncoordinated attacks. The A collaborative model was developed to investigate the weaknesses of AODV and RPL protocol in LLN networks that exploits the LLN environment’s vulnerabilities. These Collaborative attacks use the combined efforts of more than one attacker against the target victim.

**Detection of Collaborative Attacks Against RPL-AODV Using Hybrid Based Intrusion Detection System**

This section presents the proposed hybrid-based intrusion detection system that solves the above challenging issues by combining the advantages of signature-based IDS that provides a high detection rate and early detection of known attacks, a low false-positive rate, and the ability of the anomaly
detection system to detect novel, unknown attacks. Hence, the sequence-based fusion of these two approaches should theoretically provide an effective intrusion detection system that enhances the overall performance of collaborative attack detection, shortening the detection delay, increasing detection accuracy, and reducing false-alarm rates. As shown in Figure 1, the proposed work is the hybrid intrusion detection system that deploys the two IDS methods in a two-staged manner to identify both known and novel attacks. In the first stage, misuse detection is employed; we choose Snort IDS, a lightweight signature-based detector. A database of known detection behaviours has been developed and updated over time. In this stage, the system compares the network traffic with an intrusion behaviour database in real-time. If any intrusion is detected, the system will start to produce alerts according to the event handling information present in the rules matched. So, the attacks are detected early without passing through further learning stages. The second stage is the anomaly detection system. This stage is used to overcome the first stage’s shortcomings and can detect novel attacks. After passing through signature-based detection, the remaining unknown network traffic is directed to the feature extraction stage to extract robust network features. The extracted non-redundant features are essential and selected to discriminate abnormal behaviour from normal network activities. This is achieved using machine learning classification such as Decision Tree, Naive-Bayes, and Support Vector Machine with the Adaboost Ensemble technique for increased accuracy in detecting malicious packets.

The proposed method measures various tokens for detecting the misbehaving aspect resulting from these collaborative attacks. The measures include the strength of the signal, sending rate of the packet, receiving rate of the packet, delivery ratio of packets, packet acknowledgment, sending ratio of packets, dropping rate of packet, forwarding rate of the packet, and channel sensing time. Some of the Parameters were considered while simulating the proposed RPL-AODV protocol.

- The received strength of the signal is the power measure enclosed in the radio signal received.
- Sending rate of the packet is the number of packets transmitted in a predefined duration.
- Receiving rate of the packet is the number of packets received in a predefined duration.
- The delivery ratio of the packet is the packet ratio that is delivered successfully and based on the number of packets that the sender transmits.
- Packet acknowledging rates are numbers defined for a node’s acknowledgments sent to another node.
- Sending ratio of the packet is defined as the ratio of the number of packets sent successfully and the number of packets that must be transmitted.
- The dropping rate of a packet is the number of packets dropped in a predefined duration.
- The forwarding rate of the packet is the number of packets received by some node to forward it in a predefined duration.
The duration for which a node has to wait to access the channel is called channel sensing time.

Hybrid-Based Intrusion Detection System Using Ensemble Machine Learning Approach

To overcome signature and anomaly-based weaknesses, a hybrid IDS has been proposed by combining signature and specification-based techniques. This combination enables the hybrid IDS to detect both cases, either signature or specification attacks, and consumes less energy. To classify the legitimate and anomalous feature vectors from the dataset. These datasets contain flow-based and packet-based content of network traffic. There is a need to perform the numerical statistical characteristics such as mean (for every 100 packets) and the size of the packet, protocol information with the direction identifiers. The AODV-RPL protocol performs leveling of the network by tree created using the child-parent relationship, and at a high level, it places a border router that acts as the tree base. To decrease the use of node resources, rather than the monitoring node monitoring all its neighbors, it only monitors the node which has an immediate relationship with the nodes, i.e., its child node or its parent node. The monitored node data result can be forwarded to the operator present at the border router for the comparison between them and outputs the final result to revoke the suspicious node or not, based on the analyzed and statistical data.

The proposed method uses ensemble machine learning techniques, which combine Support Vector Machine, Naive-Bayes, and Decision Tree classifier techniques used to classify network records with moderate variations between regular and malicious observations and achieve better performance. To achieve high accuracy, the proposed method uses Adaptive Boosting (AdaBoost) classifier is used to improve classifier accuracy, which is an iterative ensemble method as shown in Figure 1. This classifier combines decision trees, Naive Bayes, and Support Vector Machine to make them strong classifiers and improve their accuracy, create high precision models, and be less affected by the overfitting problem. The primary objective of this method is to allocate weight in the training set for each instance. Initially, all weights are considered to be equal. Still, the weights are raised for all the cases predicted wrongly in each iteration such that in the next epoch these instances are given a high likelihood of classification.

In contrast, the weights of correctly classified instances are reduced. The iterations are repeated until a good classifier with a low error rate is reached or until we exceed the defined a maximum number of estimators. Each iteration reduces training errors and tries to ensure a good fit for the data given. As shown in Figure 7, Voting Classifier is used to get the final prediction from the above three

Figure 7. Ensemble machine learning technique
strong classifiers. Soft voting is applied to the outcomes of the classifiers. Soft voting is achieved by averaging the probability distributions estimated by the individual techniques to get the best result. In soft voting, we predict the class labels based on the calculated classifier probabilities and the assigned weight to the classifier Alert/Warnings: When malicious instances are classified, alerts are generated, and the false positive rate is found for evaluating the performance.

The main advantage of the proposed hybrid IDS is that (1) it is efficient in spaces with great dimensions. (2) It is effective in memory as it selects a set of training samples known as support vectors in the decision-making function. (3) It makes predictions using probabilities. (4) It can handle both discrete and continuous data. (5) It can easily deal with missing values. (6) Easy to update as new data arrives (6) it decreases complexity and confusion and improves clarity. (3) It considers any possible outcome of a decision. Consequently, it tracks each node to the conclusion (7) it provides a high detection rate, early detection of known attacks, and a low false-positive rate.

RESULTS AND PERFORMANCE EVALUATION

This section presents results and performance analysis for studying the feasibility of our proposed work; we have implemented the proposed AODV-RPL protocol along with its attacks in the network simulators NS3 (Perrey et al., 2013), Whitefield (2013), and Contiki-Cooja (2014) have been utilized, and conducted a series of experiments ranging from 10 to 1000 nodes on AODV, RPL, AODV-RPL Protocol, and proposed collaborative attacks against AODV-RPL Protocol for generation of high-trustworthiness attack data within LLN networks. Table 3 shows the simulation parameters of the proposed.

PARAMETERS

Performance Metric

The proposed work uses the following performance metrics for an RPL-AODV routing protocol, signature-based and anomaly-based detectors using machine learning techniques to assess and measure the efficiency. The differing metrics used include packet delivery ratio (PDR), average end-to-end delay, routing overhead, precision, accuracy, recall, or detection rate, false-positive rate, and F1-score. The assessment of the The proposed intrusion detection system (IDS) is assessed based on the following: measurement:

1. **Packet Delivery Ratio (PDR)**: The ratio between the data packets received correctly by the target node and data packets sent by the origin node.

<table>
<thead>
<tr>
<th>Operating System</th>
<th>Ubuntu-18.08, Contiki-3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Nodes</td>
<td>10-1000 Nodes</td>
</tr>
<tr>
<td>Simulator</td>
<td>NS-3, Cooja</td>
</tr>
<tr>
<td>Transmission Range</td>
<td>250m</td>
</tr>
<tr>
<td>Simulation Duration</td>
<td>Variable</td>
</tr>
<tr>
<td>Physical Topologies</td>
<td>Grid-Center Topology and Random Topology</td>
</tr>
<tr>
<td>Traffic Type</td>
<td>UDP</td>
</tr>
<tr>
<td>Data Payload Size</td>
<td>127 Bytes/Packet</td>
</tr>
<tr>
<td>Routing Protocol</td>
<td>RPL-AODV Protocol</td>
</tr>
</tbody>
</table>

Table 3. Simulation
2. **Average End-to-End-Delay:** Average time taken by the packet correctly to deliver from the origin node to the target node.

3. **Routing Overhead:** Number of data packets received correctly by the target node within the time duration.

4. **Accuracy:** Calculated as the proportion of the adequately classified samples to the total samples.

5. **Precision:** Determined as the fraction of true positive samples to predict positive samples. It is the assurance of detection of a DDOS attack.

6. **Recall:** Expressed as the fraction of true positive samples to total positive samples and referred to as Detection Rate (DR) or True Positive Rate (TPR).

7. **The False-Positive Rate (FPR):** Determined as the percentage of false-positive samples to predict positive samples.

8. **F1_Score (F1):** Described as the precision and the recall harmonic average.

Figure 8 shows the packet delivery ratio with various LLN network nodes. It shows that the PDR for AODV is less than that for RPL and the RPL-AODV protocol. However, the proposed RPL-AODV protocol achieves better PDR than the RPL protocol with varying nodes. Figure 9 shows the packet delivery ratio with the collaborative attacks in the LLN network. The PDR is decreasing because the attacks become active at 132.1ms. It is observed that collaborative attacks will affect high packet drop that incurs the denial of service to the application layer. It sends only 12.5% of the application data to the border router—figure 10 shows routing overhead with varying times. The proposed RPL-AODV has less routing load than the AODV and RPL protocols. It reduces the size of four DIO requests and DIO replies, RREP, and RREQ control messages to two control DIO-RREQ instances and two DIO-RREP instances. Hence, the proposed routing protocol suits various low-power applications. Figure 11 shows the throughput of RPL, AODV, and RPL-AODV routing protocols with varying pause times. The throughput of RPL-AODV is decreasing because the attacks become active at this time, 15.1ms. It also observed that collaborative attacks affect high throughput loss, which degrades the performance of the LLN network. Figures 12 and 13 show the end-to-end delay of RPL, AODV, and RPL-AODV routing protocol with and without attacks in the LLN network. The proposed RPL-AODV protocol has a lower average end-end delay due to the shortest path and achieves higher route diversity than the AODV and RPL protocol. Further, figure-13 proposed RPL-AODV protocol incurs more end-to-end delay in the presence of collaborative attacks.

The proposed method also uses an ensemble machine learning-based attack detection methodology that uses various tokens for detecting the misbehaving aspect of these collaborative attacks with high

![Figure 8. Packet delivery ratio for different routing protocols](image)
accuracy and precision. We evaluated and compared the performance of the experimental results of the proposed hybrid IDS that combines the signature-based and Anomaly-based detectors using
various parameters such as Accuracy, precision, detection rate, false positive rate (FPR), F1_score, etc. To evaluate the performance of the experimental results of our proposed framework, we utilize the dataset generated by the simulation NS-3 and Cooja in the form of PCAP raw packet capture and Trace (tr) files. This data source has a hybrid network traffic’s actual normal activities and contemporary collaborative attacks. The labeling of normal vectors is given class as Normal, and for malicious vectors, the class is labeled as an attack. The tool tcpdump was also used to acquire all the raw network packet data, and features were created with the tools Argus, Bro-IDS, and twelve algorithms to generate 43 features with the class label. The nominal data type of class label of each record of the data is assigned with numerical values such as for normal instances as 0 and attack as 1. The data were separated into two sets: a training set and a test set in the 7:3 ratio, with approximately 2,76,232 training records and 122,332 test records. Each classifier was trained using a train set and validated using the test set. Table 4 gives normal and attack records in the dataset used for training the model.

From the dataset, the categorical features with nominal Data types include the following attributes:

- (HTTPS, HTTP, CoAP), Protocols – (UDP, RPL, AODV, RPL-AODV, and ICMP). These categorical
fields are encoded into numerical data types, with each being of unique value, such as TST=1, URN=2, RTA=3, etc. But the dataset is not ready to fit into the chosen classification models. It contains both quantitative and qualitative features that may have unwanted and redundant features, which cannot be used for the statistical techniques models. However, the proposed machine learning techniques use numerical statistics to classify the given input data affected by the qualitative attributes present in the data source. The quality of input data plays an essential role in obtaining a well-trained machine-learning model. Hence, data needs to be visualized for its quality and pre-processed before training the proposed model; the data should be cleaned and prepared for fitting into the used classification models. The features in the dataset are visualized for correlation. If any two attributes are highly correlated, they produce the same effect on the dependent variable. To reduce unnecessary computation or other costs, we can discard one of the two attributes. To perform this, we have used matplotlib and pandas packages for plotting the data correlation. The features are ranked within the [-1, 1] interval.

The above performance measures are obtained from the confusion matrix based on the predicted class calculated versus the actual class (ground truth). The confusion matrix is the process of presenting the result of binary classification. There are four possible outcomes as follows based on the two-class nature of the prediction:

- **True-Positive (TP):** Number of Attacks/anomalies that are successfully detected as attacks.
- **False-Positive (FP):** Number of Normal records incorrectly classified as attacks.
- **True-Negative (TN):** Number of Normal records successfully identified as normal.
- **False-Negative (FN):** Number of Attacks/anomalies classified as normal.

In phase one, we evaluated the performance of the experimental results of the signature-based use of Snort IDS. We tested 122,332 packets of UDP, ICMP, HTTP packets, RPL, AODV, RPL-AODV control, and data messages These UDP, ICMP, HTTP RPL, AODV, RPL-AODV control, and data packets are captured by snort based on the rules or signs. These rules and signs are written as collaborative attacks to detect intrusions. All the alerts are generated from snort IDS logged into the output module. All these modules are tested, and results are achieved.

Table 5 shows the significant in Detection Rate 89.13% for 25,000 packets, 92.34% for 97332, and, 89.67% for 122,332 packets respectively. It can also be observed from the table-5 that the proposed IDS provides a low false-positive rate of 0.7% for 25,000 packets, 0.9% for 97332, and, 1.3% for

<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Training</th>
<th>Testing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>174,217</td>
<td>75,332</td>
<td>2,49,549</td>
</tr>
<tr>
<td>Collaborative attacks</td>
<td>102,015</td>
<td>47,000</td>
<td>93,000</td>
</tr>
<tr>
<td>Total</td>
<td>2,76,232</td>
<td>122,332</td>
<td>3,42,549</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>Measures (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Packets</td>
</tr>
<tr>
<td>Snort IDS</td>
<td>25,000</td>
</tr>
<tr>
<td></td>
<td>97332</td>
</tr>
<tr>
<td></td>
<td>122,332</td>
</tr>
</tbody>
</table>
122,332 packets respectively, and finally, the proposed IDS achieves high accuracy of 99.12% for 25,000 packets, 97.65% for 97332 and, 98.70% for 122,332 packets respectively.

The dataset is divided into training and testing subsets to evaluate each classifier and ensemble. Method. Figures 17 and 18, and Figure 19, show that the DT technique produces a 93.7% accuracy, 95.0% detection rate, and 8.9% FPR; the SVM technique achieves a 74.1% accuracy, a 65.3% detection rate, and a 13.5% FPR. Lastly, the NB technique achieves an accuracy rate of 76.6%, 86.0% DR, and 37.9% FPR. When DT and NB techniques are combined with an accuracy rate of 94.2% and a detection rate of 95.7% and 9.3% FPR are achieved. NB and SVM techniques combined achieved 76.7%, 92.4% DR, and 54.5% FPR. Similarly, when DT and SVM techniques are combined, with an accuracy rate of 94.1%, the detection rate is 95.9%, and the FPR is 9.4% achieved. The Accuracy and Detection of
the Ensemble Method of the three techniques achieve 94.3% and 96.3%, respectively, while the FPR produced is 9.0%, which outperforms the DT, NB, and SVM performance techniques. Figure 17, Figure 18, and Figure 19 also show the significance in Detection Rate 88.56% for 25,000 packets, 90.72% for 97332, and, 89.43% for 82,332 packets respectively. It can also be observed from Table 5 that the proposed IDS provides a low false-positive rate of 0.7% for 25,000 packets, 0.9% for 97332, and, 1.3% for 122,332 packets respectively, and finally, the proposed IDS achieves high accuracy of 98.67% for 25,000 packets, 98.70% for 97332 and, 98.17% for 122,332 packets respectively.
Figure 20. Comparative analysis of throughput with varying zigbee nodes

Figure 21. Comparative analysis of throughput with varying wireless sensor network nodes

Figure 22. Comparative analysis of throughput with varying wireless adhoc network nodes

Figure 20, Figure 21, and Figure 22 show the throughput with various LLN network nodes. It shows that the throughput for AODV and RPL is less than the proposed RPL-AODV protocol. Moreover, the proposed RPL-AODV protocol achieves better throughput than the RPL protocol with varying nodes, and is compared with other existing methods presented (Halder et al., 2018; Lu et
al., 2011; Toscano & Bello, 2012). Figure 23 shows the average end-to-end delay of RPL, AODV, and RPL-AODV routing protocols in the LLN network. The proposed RPL-AODV protocol has a lower average end-end delay due to the shortest path and achieves higher route diversity than the AODV and RPL protocol. Moreover, the proposed RPL-AODV protocol achieves less latency than the RPL protocol with varying nodes and compared with other existing methods presented (Halder et al., 2018; Lu et al., 2011; Toscano & Bello, 2012) However, the Figure 13 proposed RPL-AODV protocol incurs more end-to-end delay in the presence of collaborative attacks.

RESEARCH DISCUSSIONS

Recently, there has been a meteoric rise in internet-connected devices around us in recent years. The Internet of Everything (IoET) is one of the key technologies integrators in industry-4.0, contributing to the large-scale deployment of Low-power and lossy (LLN) networks that enable networked connections among people, processes, data, and things. However, the major inconveniences of these LLN networks mainly depend on the ‘security’ of the devices. The routing protocol for low-power and lossy networks (RPL) is one of the unique standardized routing protocols that enable efficient use of smart devices’ energy, compute resources, build flexible topologies, and use data routing in LLN networks. The RPL protocol itself has many security risks and possibilities of attacks. Most such efforts have been put into a mechanism to defend against individual attacks against the RPL routing protocol. Intrusion detection fails to handle a high detection rate, early detection of known attacks, and the ability to detect novel, unknown attacks and have a low false-positive rate. This paper presents the proposed work that detects collaborative attacks against the RPL-AODV routing protocol using the cooperative IDS mechanism in LLN Networks. Initially, we investigated the RPL-AODV routing protocol’s performance by combining the advantages of RPL and AODV routing protocols, which work together in low-power, resource-constrained LLN networks. Next, we modeled the collaborative attacks that exploit the vulnerability of AODV and RPL routing protocol. These collaborative attacks use the combined efforts of more than one attacker against the target victim from a security perspective. Finally, A hybrid IDS has been proposed using an ensemble machine learning approach that combines specification-based and signature-based IDS as cooperative IDS to detect the collaborative attacks against the RPL-AODV routing protocol that effectively monitors the LLN network.

To study the feasibility of our proposed work, we have implemented the proposed AODV-RPL protocol along with its attacks in the network Simulators NS3, Whitefield, and Contiki-Cooja have been utilized and introduced the collaborative attacks against the AODV-RPL Protocol for generation of high-trustworthiness attack data within LLN networks. Further, we evaluated the proposed work with various performance metrics to assess and measure the efficiency.
CONCLUSION AND FUTURE WORK

This paper proposes a cooperative IDS mechanism that detects collaborative attacks against the RPL-AODV routing protocol in LLN networks of the Internet of Everything (IoET). First, we investigated the RPL-AODV routing protocol, which combines the advantages of both RPL and AODV routing protocols, which work together in low-power resource-constrained LLN networks of the Internet of Everything (IoET). This RPL-AODV routing protocol The LLN network establishes the path from the origin node to the target node only on an on-demand basis. Next, we modeled the collaborative attacks that cause a more devastating impact on LLN networks than uncoordinated attacks. The collaborative
model was developed to investigate the weakness of AODV and RPL protocols in LLN networks that exploit the LLN environment’s vulnerabilities. These collaborative attacks use the combined efforts of more than one attacker against the target victim. Finally, a hybrid IDS has been proposed using an ensemble machine learning approach that combines specification-based and signature-based IDS as cooperative IDS to detect the collaborative attacks against the RPL-AODV routing protocol that effectively monitors the LLN network of the Internet of Everything (IoET). As there is still a lot of scope in this area to come up with innovative solutions for achieving security in LLN networks.

REFERENCES


E. Suressh Babu is working as Assistant Professor, Department of Computer Science and Engineering, National Institute of Technology, Warangal. He obtained his Ph.D. from JNTU Kakinada, specializing in Networking and Security. He secured his M.Tech in Computer Science and Engineering from Visvesvaraya Technological University, Belgaum, and B.Tech in Computer Science and Engineering from JNTU Hyderabad. He has 16 years of Research, and Teaching in different levels of Professors, Associate, and Assistant professors. Currently, He works in the area of Blockchain Technology, Internet of Things, IoT Security, Wireless Networks, and Wireless Adhoc Network Security. Some of the areas where he published 40+ refereed Journals, 20+ Conferences, and 10 Book chapters.

Bhukya Padma is currently pursuing her Ph.D. from the National Institute of Technology Warangal, India. She has completed her M.tech in CSE from Kakatiya University College of Engineering And Technology, Hanamkonda, Telangana, India. She has completed her B.Tech from Kakatiya University College Of Engineering And Technology, India. Her areas of interest are the Internet of Things (IoT) and Blockchain Technologies.

Soumya Ranjan Nayak is currently working as Assistant Professor at Amity School of Engineering and Technology, Amity University, Noida, India. He received his Ph.D. degree in Computer Science and Engineering under MHRD Govt. of India fellowship from CET, BPUT Rourkela, India; with a preceded degree of M. Tech and B. Tech in Computer Science and Engineering. He has published over 90 articles in peer-reviewed journals and conferences of international repute like Elsevier, Springer, World Scientific, IQS Press, Taylor & Francis, Hindawi, MDPI, Inderscience, IGI Global, etc. Apart from that, 12 Book Chapter, 6 Books and Six Indian patents (two patents granted) and Two International patents (two patents granted) under his credit. His current research interests include medical image analysis and classification, machine learning, deep learning, pattern recognition, fractal graphics and computer vision. His publications have more than 700 citations, of h index of 15, and i10 index of 26 (Google Scholar). He serves as a reviewer of many peer-reviewed journals such as Applied Mathematics and Computation, Journal of Applied Remote Sensing, Mathematical Problems in Engineering, International Journal of Light and Electron optics, Journal of Intelligent and Fuzzy Systems, Future Generation Computer Systems, Pattern Recognition Letters, etc. He has also served as Technical Program Committee Member of several conferences of international repute.

Nazeeruddin Mohammad is the director of cybersecurity center at Prince Mohammad Bin Fahd University in the Kingdom of Saudi Arabia (KSA). Earlier, he completed his PhD from the School of Computing and Information Engineering at the University of Ulster, Coleraine, UK in 2007. He received his Bachelor of Engineering (B.E) degree in Electronics and Communications Engineering from Osmania University, India in 1996. He has an M.S. degree in Systems Engineering from King Fahd University of Petroleum & Minerals (KFUPM), KSA in 1999. He also received practical training and an Honours Diploma in Software Development from BDPS, India in 1997. He is a recipient of M.S. Research Scholarship from KFUPM (1997), Vice Chancellors Research Scholarship from University of Ulster (2004) and first prize in the Faculty of Engineering Business Plan Competition at the University of Ulster (2006). He is actively involved in the TPC/Review committees of renowned conferences and journals.

Uttam Ghosh is working as an Associate Professor of Cybersecurity in the Department of Data Science & Computer Science of School of Advanced Computational Sciences, Meharry Medical College, USA. Dr. Ghosh obtained his PhD in Electronics and Electrical Engineering from the Indian Institute of Technology Kharagpur, India in 2013, and has Post-doctoral experience at the University of Illinois in Urbana-Champaign, Fordham University, and Tennessee State University. Dr. Ghosh has published seventy papers at reputed international journals also top international conferences by IEEE, ACM, and Springer. His main research interests include Machine learning, Cybersecurity, Blockchain, Computer Networks, Smart Power grid and Software-Defined Networking.