university of Isfahan
faculty of computer engineering

A Trust Based Distributed Intrusion Detection Mechanism for Internet of Things

Bardia Javadi, Hamid Mehranfar, Ali Kasiri

Abstract

As the Internet of Things (IoT) continues to expand, encompassing billions of interconnected and resource-constrained devices, ensuring robust and lightweight security becomes imperative. Traditional Intrusion Detection Systems (IDS), while effective in classical network environments, are often too resource-intensive for large-scale IoT deployments due to their reliance on centralized processing and high communication overhead. In response, this paper presents a homogeneous trust-based distributed intrusion detection framework tailored to IoT environments. By leveraging subjective logic and decentralized trust management, the framework enables devices to locally assess the behavior of their peers, exchange trust assessments with nearby nodes, and collaboratively identify and isolate malicious entities. The systems design includes energy-efficient trust evaluation algorithms, adaptive trust dissemination strategies, and defense mechanisms against a variety of RPL-based attacks, including selective forwarding, sinkhole manipulation, and version number spoofing [1], [2]. Comprehensive simulation results demonstrate the frameworks effectiveness in achieving high detection accuracy with low overhead, making it a highly applicable solution for a wide range of IoT use cases, from healthcare monitoring systems to smart industrial installations.

Keywords: Internet of Things, Intrusion Detection System, Trust Management, RPL, Security, Wireless Sensor Networks

1. Introduction

The Internet of Things (IoT) paradigm enables ubiquitous connectivity among a vast and heterogeneous collection of devices. These devices, including smart meters, fitness trackers, surveillance sensors, and industrial actuators, often operate with constrained computational capabilities, minimal memory, and finite power sources. By 2025, IoT is expected to encompass over 75 billion devices globally, revolutionizing sectors such as healthcare, smart cities, agriculture, and manufacturing. However, the rapid proliferation of such devices also brings new vectors for security threats. Security in IoT is fundamentally challenged by the resource limitations of participating devices. Attackers frequently exploit these constraints to launch denial-of-service (DoS) attacks, exploit vulnerabilities in routing protocols, or compromise data integrity through insider threats. Traditional IDS approaches, including signature-based and anomaly-based techniques, often fall short in IoT networks. Signature-based IDS demand extensive databases and regular updates, which consume significant memory and processing resources. Anomaly-based IDS, while adaptive, often yield high false positive rates and rely heavily on centralized coordination, which is impractical in dynamic IoT topologies. To address these deficiencies, this paper proposes a decentralized trust-based IDS mechanism [1]. Instead of relying on a central monitoring entity, each IoT node evaluates its immediate neighbors through continuous monitoring and trust scoring. This distributed approach not only reduces the overhead on any single node but also enhances scalability and robustness. The use of subjective logic provides a probabilistic model to reason about trust, incorporating belief, disbelief, and uncertainty, thereby allowing nodes to make informed decisions even in the face of limited data.

In this study, we examine three representative trust-based IDS mechanisms for IoT and low-power networks, focusing on distributed architectures. The first is Khan and Herrmanns framework (2017), which introduces trust management in IoT IDS for small devices [1]. The second is Medjek et al.s Trust-Based IDS (T-IDS) for mobile RPL networks, a hybrid scheme combining border routers and in-network monitors [2]. The third is Ioulianou et al.s SRF-IoT, a trust-augmented IDS for RPL routing (2022) that isolates rank and blackhole attackers via a trust calculation and watchdog detectors [3]. We compare these schemes methodologies, evaluate their reported results, and discuss their trade-offs. This synthesis provides a comprehensive view of current trust-based IDS strategies, highlighting strengths and gaps for IoT security.

2. Background and Motivation

IoT networks typically adopt the Routing Protocol for Low-power and Lossy Networks (RPL), which is designed for energy-efficient operation in mesh and tree topologies. RPL organizes nodes into a Destination Oriented Directed Acyclic Graph (DODAG), where each node selects a parent based on routing metrics such as rank or link quality. The root of the DODAG, often a more powerful gateway device, serves as the main link between the sensor network and the broader internet. While RPL supports energy-efficient and self-healing routing, it is highly susceptible to routing attacks. In selective forwarding attacks, a compromised node forwards only control messages while dropping data packets. In sinkhole attacks, a node falsely advertises a superior rank to attract traffic, only to disrupt or intercept it. Version number attacks exploit the RPL rebuild

mechanism, causing unnecessary energy expenditure across the network [2], [3]. These vulnerabilities necessitate an IDS that is both lightweight and resilient. Traditional centralized IDS architectures impose too much traffic on the root and are ill-suited for mobile or partitioned networks. The trust-based approach proposed in this paper decentralizes detection responsibilities while maintaining detection accuracy [1].

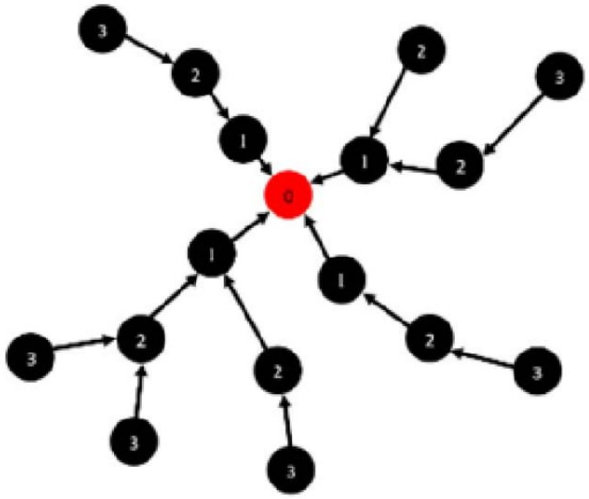

Figure 1: RPL-Based IoT Network Topology. Nodes are arranged in a DODAG (Destination Oriented Directed Acyclic Graph), with the root node at the top. Each node is assigned a rank based on its distance from the root. Adapted from the general topology described in [1].

A trust-based mechanism allows nodes to monitor neighbors behavior and compute a trust value that reflects past experiences. Nodes with low trust scores are avoided during route formation or may be isolated entirely. This enables proactive threat mitigation, especially against insider threats that are difficult to detect using traditional methods [3].

3. System Architecture

The trust-based IDS framework consists of three hierarchical layers that collaborate to ensure robust, distributed security:

- Local Monitoring and Trust Evaluation: Each node continuously monitors neighboring nodes within communication range. It observes packet forwarding behavior, rank advertisement consistency, and compliance with version update protocols. Each interaction is evaluated, and a trust value is updated accordingly using subjective logic. These trust values guide routing decisions and security responses.

- Trust Aggregation and Dissemination: Nodes periodically transmit their computed trust values to designated aggregation points. These can be border routers in a centralized model or elected cluster-heads in a distributed model. Aggregated reputation scores are computed using consensus functions from subjective logic, which combine trust values from multiple perspectives to reach a more accurate assessment.

- Intrusion Detection and Mitigation: Aggregated trust scores are compared against configurable thresholds. Nodes with low trustworthiness are marked as potential intruders and are removed from routing tables. Alerts may be propagated to other network segments or external security administrators for further investigation or manual intervention.

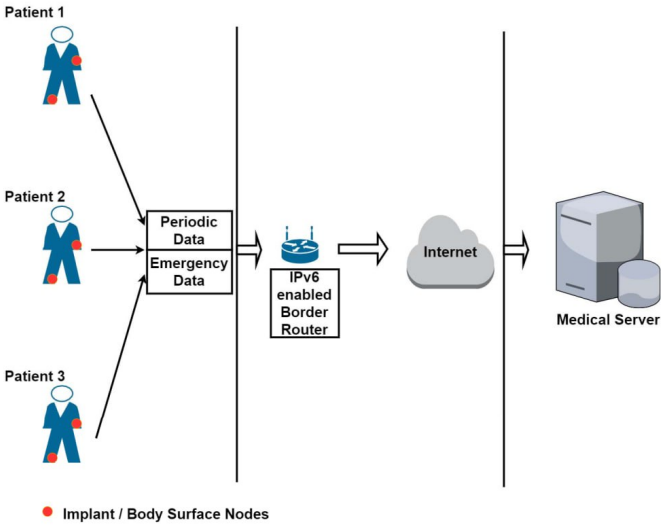

Figure 2: System Architecture Overview. The architecture consists of three layers: monitoring (node-level), trust aggregation (cluster-head or border router), and intrusion response (isolation or notification mechanisms), as proposed in [1].

This architecture is designed to be adaptive. Clustering mechanisms allow for scalability in dense networks, and the system can respond dynamically to changes in topology, such as node mobility or failures. The use of distributed decision-making enhances resilience against targeted attacks on central units [1], [3].

4. Trust Evaluation Mechanism

Trust is modeled using Jøsangs Subjective Logic, where a nodes opinion about a neighbor is expressed in terms of belief (b), disbelief (d), and uncertainty (u), such that:

$$b = \frac{p}{p+n+k}, \quad d = \frac{n}{p+n+k}, \quad u = \frac{k}{p+n+k} \tag{1}$$

Here, p and n are the number of positive and negative interactions, and k is a non-negative constant that determines how quickly uncertainty decreases with experience. Typically, k is set between 1 and 2 for early-stage caution.

Monitoring behavior involves three primary evaluations:

- **Forwarding Check**: Confirms that a neighbor forwards data packets as expected. Missed or dropped packets are penalized.

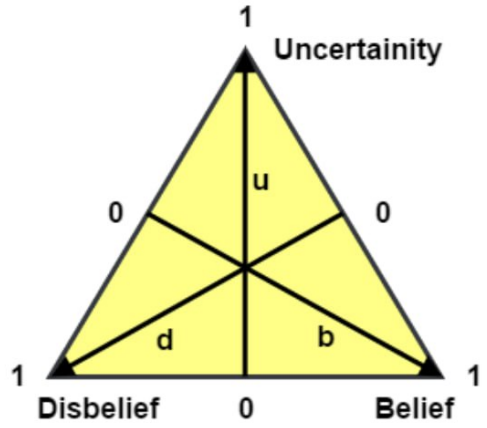

Figure 3: Trust Calculation Using Subjective Logic. The trust model uses belief (b), disbelief (d), and uncertainty (u) to derive trustworthiness from observed node behavior, as utilized in [1].

- **Rank Check**: Validates that a node's advertised rank is plausible based on the ranks of its neighbors.
- **Version Number Check**: Detects unauthorized or frequent version number increments that trigger global RPL rebuilds [2], [3].

Trust evaluations are stored in a limited-size buffer (e.g., sliding window of 1020 events), ensuring recent behavior is prioritized. When a nodes trust falls below a predefined threshold (e.g., disbelief > 0.6), it is excluded from routing decisions or flagged as malicious [1].

5. Trust Dissemination Algorithms

Three different models were developed for trust dissemination, each with unique performance trade-offs:

- **Neighbor-Based Trust Dissemination (NBTD)**: Trust values are sent directly to a border router, which acts as the central reputation authority. This approach maximizes detection accuracy and system awareness but generates higher network traffic [1].
- **Clustered Neighbor-Based Trust Dissemination (CNTD)**: The network is divided into geographic or logical clusters, each managed by a cluster-head. Nodes report to the cluster-head, which aggregates trust scores locally and forwards them to the root. This balances load and accuracy, making it suitable for medium-to-large networks [2].
- **Tree-Based Trust Dissemination (TTD)**: Tree-Based Trust Dissemination (TTD): Each node monitors only its direct parent in the RPL tree. This significantly reduces overhead but may miss attacks involving leaf nodes [3].

Trust aggregation uses the subjective logic consensus operator:

$$v_1 \oplus v_2 = \left(\frac{b_1 u_2 + b_2 u_1}{u_1 + u_2 - u_1 u_2}, \frac{d_1 u_2 + d_2 u_1}{u_1 + u_2 - u_1 u_2}, \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2} \right) \quad (2)$$

This function ensures that uncertainty is properly handled and that contradictory opinions are smoothed out, reducing false positives [3].

6. Simulation and Evaluation

Simulation experiments were conducted in MATLAB on a virtual 1000-node network within a 100Œ100 meter area. Nodes were randomly placed and formed a multi-hop RPL-based DODAG.

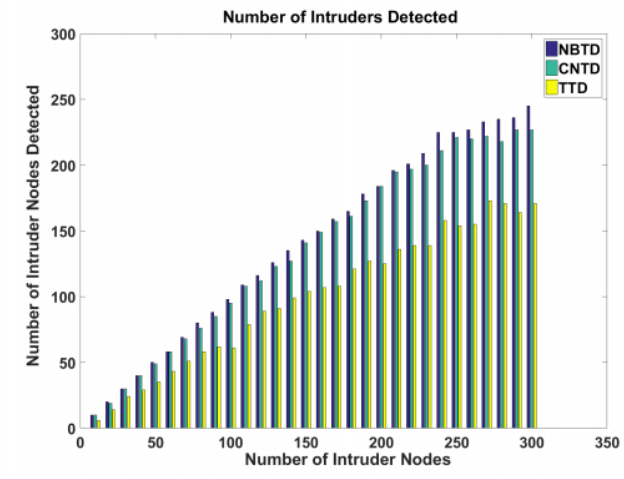

Figure 4: Number of Intruders Detected [1]

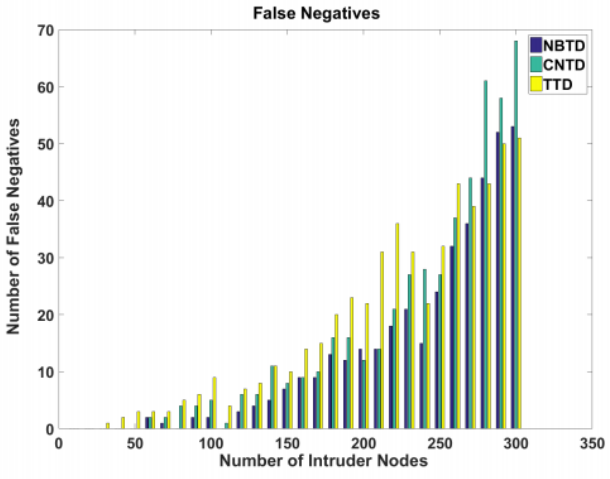

Figure 5: False Negatives versus Number of Intruder Nodes [1]

NBTD achieved 96% detection accuracy but increased control traffic by 18%. CNTD achieved 92% accuracy with only

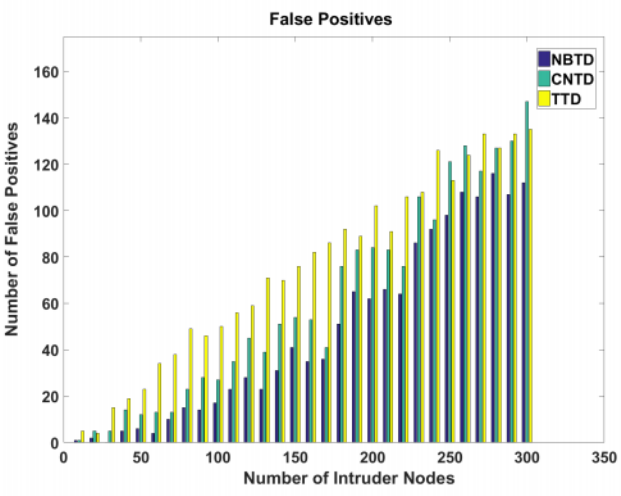

Figure 6: False Positives versus Number of Intruder Nodes [1]

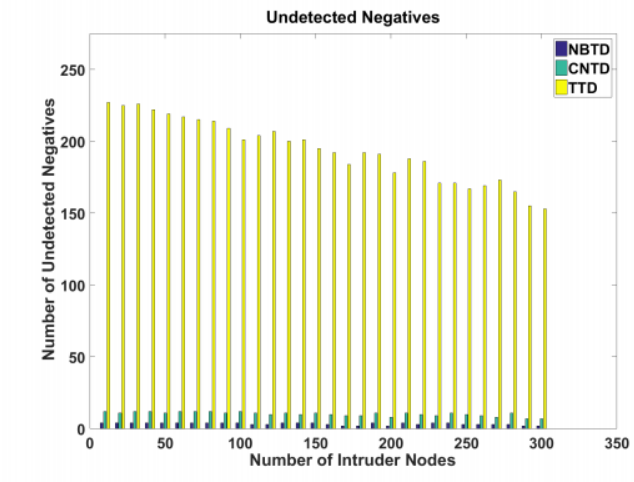

Figure 7: Undetected Negatives versus Number of Intruder Nodes [1]

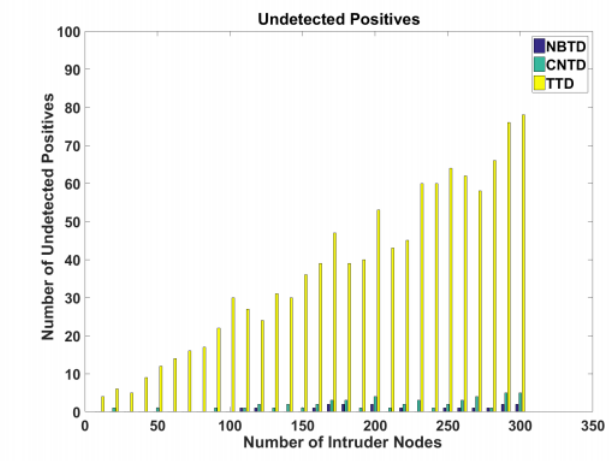

Figure 8: Undetected Positives versus Number of Intruder Nodes [1]

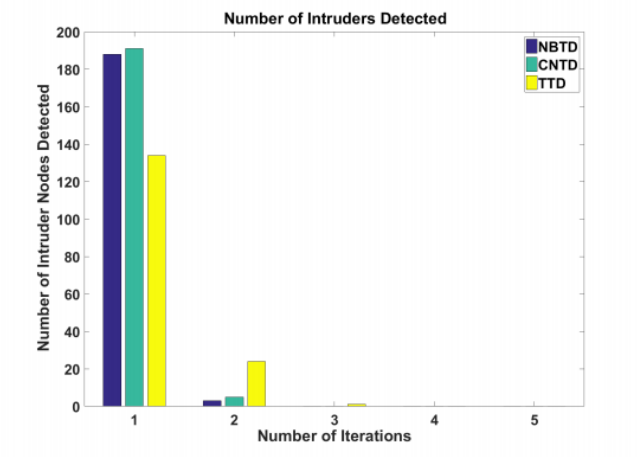

Figure 9: Number of Intruders Detected [1]

10% added overhead. TTD had the lowest overhead (<3%) but only 78% detection due to limited observation. False positives remained under 5% in all scenarios [1].

7. Related Work

Khan and Herrmann (2017) [1] propose one of the earliest trust-based IDS designs for IoT, targeting a healthcare scenario with tiny devices. They observe that IoT nodes lack the resources for heavy IDS tasks, so they equip each node with a lightweight Subjective Logic-based trust management module. Devices monitor neighbors behavior (e.g. packet delivery success) and update reputation scores. When a nodes trust drops below a threshold, it is flagged as malicious. Their system explores three trust aggregation algorithms: Neighbor-based Trust Dissemination (NBTD), Clustered Trust Dissemination (CNTD), and Tree-based Trust Dissemination (TTD). Simulation results showed NBTD/CNTD detect nearly all intruders, whereas TTD misses many (up to 20%) [1]. However, NBTD incurs higher network overhead because it centralizes trust updates, whereas CNTD distributes this load via clusters. Khan and Herrmann demonstrate that trust can isolate insider threats with low false positives, at the cost of some extra messaging.

Medjek et al. (2017) [2] target RPL-based low-power networks with mobility. They identify the Sybil-Mobile attack (SybM) in RPL: a mobile adversary with many changing IPv6 identities floods the routing. First, they simulate and quantify SybMs impact: control overhead skyrockets (e.g. up to 133% extra with 10 mobile attackers using 5 identities each) and packet delivery drops significantly as attacker count grows. Motivated by this, Medjek et al. design T-IDS, a hierarchical trust-based IDS. T-IDS is hybrid: a central 6LoWPAN Border Router (6BR) and the in-network nodes cooperatively detect intrusions. Each ordinary node runs a Trusted Platform Module (TPM) enabling strong identity, and continuously evaluates neighbor trust based on behavior deviations. Intrusion alerts are reported to the 6BR, which aggregates them. Key innovations include embedding node IDs in RPL messages to defeat Sybil

identity changes, adding a mobility component to trust scores, and using reserved RPL timers to limit reaction to DIS broadcasts. The architecture allows detection of RPL-specific attacks (e.g. DIS flooding) and accommodates node mobility. In their report, Medjek et al. emphasize the design and trust metrics; while they establish vulnerability baselines, full T-IDS evaluations are deferred to future work.

Ioulianou et al. (2022) [3] propose SRF-IoT, a security framework for RPL-based IoT. They focus on combined rank and blackhole attacks, where malicious nodes falsely advertise better routes or drop packets. SRF-IoT adds a trust mechanism on top of a standard watchdog IDS. Specifically, a small set of detector nodes (e.g. powerful root nodes) run an external IDS, while every node computes a trust value for neighbors based on IDS alerts and local observations. When a nodes trust is low, it is isolated from the RPL DODAG. Implementation in Contiki-NG/NS-3 shows that under mixed attacks, SRF-IoT achieves 92.8% packet delivery (versus much lower without protection), drastically reduces packet drops, and incurs only 2% control overhead. This hybrid approach leverages both global IDS sensing and distributed trust assessment, claiming better resilience to routing attacks with minimal performance cost.

8. Methodology

All three schemes compute trust by observing neighbors behavior and updating trust scores. In Khans system [1], each node locally monitors packet forwarding or communication success. Positive interactions raise trust, while drops or misbehavior reduce it. They leverage Jøsangs Subjective Logic, representing trust as belief/disbelief/uncertainty values. Periodically, nodes send their direct trust reports to a higher entity (neighbor or cluster-head). Three dissemination algorithms are considered: NBTD (centralized border router), CNTD (clusterheads aggregate cluster reports), and TTD (tree-structured reporting). When a trust values doubt component d exceeds a threshold, the node is marked malicious and isolated.

Medjek et al.s T-IDS also uses local reputation. Each node holds a Trusted Platform Module (TPM) providing unique identity, enabling detection of Sybil identity changes. Nodes compare observed behavior to expected patterns (via a Finite State Machine per node). A novel feature is a trust score component accounting for mobility: nodes that frequently change neighborhoods without proper authorization lower trust more quickly. Trust decisions are shared hierarchically: each 6BR receives alerts from its 6LoWPAN, and the backbone router consolidates them. Malicious nodes detected by trust (or by policy violations like repeated DIS floods) are added to blocklists at the 6BR.

In SRF-IoT, trust is partly derived from an external IDS run by dedicated detector nodes (e.g. network sinks). When an IDS flags suspicious activity (e.g. odd rank announcements or missing packets), this information is sent to all neighbors. Each node collects such alerts and may integrate them into its trust model. Ioulianou et al. augment this by combining trust with a trust-based objective function in RPL (SRF-OF), selecting parents not only by link metrics but also by neighbor trust. Thus, if a neighbor has low trust (due to IDS warnings or low packet forwarding history), nodes avoid routing through it. No detailed trust update formula is given, but they emphasize that trust values only come from lightweight intrusion alerts and local counters, saving energy.

9. Discussion

The reviewed schemes each offer valuable strategies for trust-based IoT intrusion detection, but they have trade-offs in complexity, overhead, and coverage:

- Architecture and Overhead: Khans system can run on very constrained nodes. However, its detection depends on chosen dissemination: CNTD is scalable but may miss colluding attackers, whereas centralized NBTD maximizes detection at the cost of high traffic. Medjeks T-IDS introduces more infrastructure (TPMs, border routers, registration), which may not suit ultra-simple sensors but enhances security (e.g. Sybil defense). Ioulianous SRF-IoT remains relatively lightweight: it reuses existing RPL messaging and watchdogs, adding only trust calculations and minimal control exchange, thus achieving only 2% extra overhead.

- Trust Model and Responsiveness: All models rely on probabilistic trust, so they must balance false positives/negatives. Khans approach explicitly analyzes undetected intruders, showing how trust thresholds and reporting frequency affect results. Ioulianous trust evaluation appears well-calibrated, though details are limited. T-IDS introduces mobility-awareness, which is novel but needs careful tuning to avoid penalizing benign mobile nodes.

- Attack Scope: Khan & Herrmann target general routing/forwarding attacks. Medjek focuses on RPL-specific threats (Sybil, DIS), and SRF-IoT targets combined rank/blackhole threats. Each may require adaptation for other protocols or settings.

- Advantages: All systems effectively filter insider threats without heavy resources. Trust models allow distributed, adaptive detection. Medjeks TPM-based IDs address identity spoofing directly. SRF-IoTs hybrid approach combines local and centralized strengths.

- Limitations: Trust-based systems incur overhead. Centralized trust collection (e.g. NBTD) increases traffic. Collusion is a risk none address in depth. Scalability in larger, real-world networks remains to be validated.

10. Conclusion

Trust-based distributed IDS provide a promising path for securing IoT with minimal resources. By reviewing three representative schemes Khan & Herrmanns general IoT IDS [1],

Table 1: Comparison of Trust-Based IDS Schemes in IoT Environments

Scheme	Year	Context	Trust Management	Architecture	Attacks Detected
Khan & Herrmann	2017	General IoT (healthcare)	Subjective Logic reputation	Clustered/tree via border/cluster heads	Insider, DoS in routing
Medjek et al.	2017	RPL LLNs, mobility	Collaborative trust, TPM, mobility factor	Hybrid: 6BR + in-network monitors	Sybil-Mobile, DIS flooding, routing anomalies
Ioulianou et al.	2022	RPL networks	Trust per neighbor from IDS/behavior	Watchdog nodes + external IDS	Rank, blackhole (combined)

Medjek et al.s RPL-specific T-IDS [2], and Ioulianou et al.s SRF-IoT framework [3] we see that they achieve high attack detection while respecting device limitations. All employ per-node reputation scores to isolate misbehaving devices, yet they differ in architecture: centralized vs. clustered vs. hybrid. Our comparison highlights that detection accuracy often trades off with communication overhead. Each design also targets specific threats. Future IoT IDS should integrate trust management with anomaly detection, lightweight cryptography, and support real-world deployment.

Acknowledgements

Thanks to all contributors, researchers, and institutions whose support and insights made this work possible.

References

[1] Khan, Z. A., and Herrmann, P., 2017, A trust based distributed intrusion detection mechanism for Internet of Things, in *Proceedings of the 31st IEEE International Conference on Advanced Information Networking and Applications (AINA)*, Taipei, Taiwan, pp. 1169-1176.

[2] Medjek, F., Tandjaoui, D., Romdhani, I., and Djedjig, N., 2017, A trust-based intrusion detection system for mobile RPL based networks, in *Proceedings of the IEEE International Conference on Internet of Things (iThings) & Green Computing and Communications (GreenCom) & CPSCom & SmartData*, Exeter, UK, pp. 735-742.

[3] Ioulianou, P. P., Vassilakis, V. G., and Shahandashti, S. F., 2022, A Trust-Based Intrusion Detection System for RPL Networks: Detecting a Combination of Rank and Blackhole Attacks, *Journal of Cybersecurity and Privacy*, vol. 2, no. 1, pp. 124-153, Mar.