Practical Security for Cooperative Ad Hoc Systems

Hans Walter Behrens,
supervised by K. Selçuk Candan.
Arizona State University
Tempe, AZ
{hwb, candan}@asu.edu

***Abstract*—Existing consumer devices represent the most pervasive computational platform available, but their inherently decentralized nature poses significant challenges for distributed computing adoption. In particular, device owners must willingly cooperate in collective deployments even while others may intentionally work to maliciously disrupt that cooperation. Public, cooperative systems benefit from low barriers to entry improving scalability and adoption, but simultaneously increase risk exposure to adversarial threats via promiscuous participant adoption. In this work, I aim to facilitate widespread adoption of cooperative systems by discussing the unique security and operational challenges of these systems, and highlighting several novel approaches that mitigate these disadvantages.**

***Index Terms*—ad hoc networks, cooperative communication, network security, homomorphic hashing, pervasive computing**

I. Introduction and Motivation

Pervasive computing, including IoT deployments and edge computing, offers transformative potential compared to current computing paradigms. Most existing systems rely on centralized authorities to mediate interactions between participants, reducing security risks from malicious agents while also forming a communication bottleneck that acts as a central point of failure. Ad hoc systems, in contrast, increase topological dynamicity in exchange for higher latency and susceptibility to adversarial disruption.

Extensive literature has evaluated the efficiency of ad hoc systems, but relatively little focuses on adversarial mitigation in an ad hoc context. In particular, the node capture assumption in which an adversary knows the content of a participants' memory forms a key security requirement for "bring your own device" (BYOD) cooperative systems. Operational resilience also improves as system centralization decreases, since fewer points of failure exist for network disruption or surveillance.

These types of resilient low-trust systems suggest many potential practical applications, such as wide-area monitoring in developing or disaster-stricken areas, resilient mesh communication independent of infrastructure, or community-driven computing with donated devices working towards common computational goals.

In my PhD thesis, I explore several weaknesses of low-trust ad hoc systems, propose solutions to these drawbacks through the lens of communication security, and contextualize them in concrete and practical applications.

This research is supported by NSF#1610282 "DataStorm: A Data Enabled System for End-to-End Disaster Planning and Response" and NSF#1610282 "BIGDATA: Discovering Context-Sensitive Impact in Complex Systems".

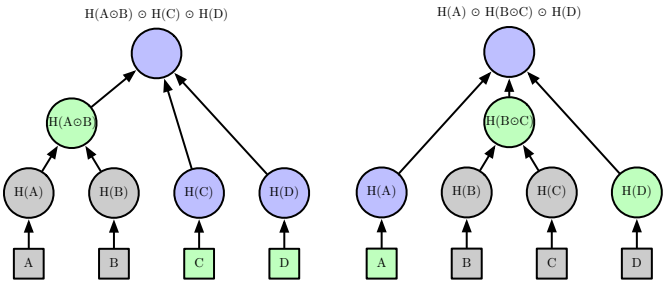

Fig. 1. Two equivalent HMTs, with provided information in green, computed information in blue, and hidden information in gray. Note that neither tree knows value B, nor shares a known intermediate ancestor as input.

II. Encryption, Hashing, and Homomorphism

Defining practicality changes significantly when shifting from powerful traditional servers to low-powered, untrusted nodes. Malicious participants undermine symmetric encryption solutions, while the loosely-coupled and extremely numerous nature of ad hoc networks discourage asymmetric encryption. Hash-based approaches offer efficiency gains over cryptographic techniques, as sacrificing data secrecy permits improved computational efficiency [5] for low-powered devices.

One potential alternative, homomorphic hashing [4], offers security and efficiency advantages over these traditional approaches. Specifically, I proposed a novel related structure, the homomorphic Merkle tree [2], which allows the compartmentalization of shared secrets (see Fig. 1). Thus, node capture attacks can be mitigated without requiring impractical per-node key ledgers. The following works utilize this style of tree as a component for ensuring message validity and authenticity.

III. Addressability

Routing plays a key role in ad hoc systems. Delivering a message between two participants who have not previously communicated, over an unknown topology populated with potentially malicious nodes, poses several unique challenges relative to centralized approaches. For example, as the number of participating nodes grows, maintaining a ledgers and topologies becomes increasingly impractical for low-powered pervasive devices. Oblivious routing, in which each node makes a local determination about message delivery, offers a promising method to sidestep this scaling issue at the expense of optimality. The functional effects of such a change differ based on the application; I will examine two potential cases, *categorical* and *spatial* addressability.

A. Categorical Addressability

Although fully-decentralized systems offer many benefits, a small number of secure, higher-powered devices may be practical in some scenarios. A wireless sensor network represents an excellent example, where inexpensive devices deployed in a large, hard-to-secure area must report sensor readings to a small number of coordinating systems.

As adversaries may compromise any device which they can physically access, the network operator aims to obtain the most accurate observation possible without prematurely degrading the network with expensive flood-based messaging. In [1], I propose a mechanism by which the network operator maintains categorical rosters (called "topic channels") of different nodes to permit communication while maintaining node capture resilience and response integrity. However, the approach used flood-based delivery, a relatively expensive choice. To mitigate this cost, I shift the structure of responses from publish-subscribe, a common choice for wireless sensing, to an "on-demand" query-response model. Thus, only nodes from the target topic channel participate in the routing and validation, reducing network propagation and computational costs.

B. Spatial Addressability

By leveraging the inherently spatial nature of wireless sensor network deployments, I am able to increase both efficiency and functionality relative to the previous approach. In a recent work [2], I build upon these flexible addressability concepts to *(1)* remove the need for a centralized secure ledger, *(2)* improve routing efficiency without sacrificing resilience or efficacy, and *(3)* maintain the privacy of the addressed nodes' locations. Rather than relying on encryption of location data, hash comparisons provide an efficient and effective mechanism to compare both equality and membership of spatial information without requiring enrollment or channel establishment as was required for categorical addressability. At the same time, angular oblivious routing provides an effective mechanism for heuristic-driven delivery without revealing the precise location targeted – only its general trajectory from the source.

IV. Distributed Byzantine Sensor Fusion

One disadvantage of the previous approach derives from the impracticality of masking respondents' locations, especially over the span of several query-response interactions. Additionally, carefully planned node capture attacks can potentially partition the network topology (a so-called *eclipse attack*).

To prevent topological attacks, I have explored the use of shared randomness as a way of providing probabilistic bounds on message delivery [3]. By structuring the query-response cycle as a form of agreement between a sender and an unknown number of recipients, and maintaining the use of spatial metadata, I can provide a robust form of distributed sensor fusion that returns correct results when the the rate of adversarial compromise $f < 50\%$.

Conceptually, nodes in the network use the shared nonce from a query as a seed, to establish a synchronized random generator which produces several (virtual) spatial points. From these, the participating nodes collectively create a shared ordering of nodes I term a Trémaux forest, using only node-local decisions. This feature makes the tree-like structures adversarially resilient, simple to agree upon, and easy to store.

Using these decentralized virtual structures, nodes generate one reply per tree, and route it up to their immediate parent if one exists. Although adversarial nodes higher in the tree exert more weight on the replies of that tree, validation restricts their options to reply pruning, rather than falsification or modification. Collectively, over all trees, almost every node will see at least a few responses reach a root.

Finally, the querier collects the responses from the nodes near the virtual roots, and intersects the validation trees to evaluate the behavior of participating nodes. Using a behavioral heuristic, the querier performs classification boosting to ignore nodes which appear to have improperly pruned subtrees from their responses, and adjusting the final results accordingly.

As before, this approach leverages the inherently-spatial nature of a deployment to improve efficiency, security, and resilience of a cooperative system that contains an arbitrary number of adversarial participants.

V. Conclusions and Future Work

Pervasive computing encompasses a much broader spectrum of capabilities beyond sensing. For example, distributed machine learning represents an area experiencing recent growth. As the role of parallelism increases, shifting computation from specialized datacenters to pervasive edge computing nodes holds the potential to further increase scalability.

In such a system, even a single adversarial node could shift algorithm results through the inclusion of a malicious input or output. In future work, I aim to create a generalized representation of these effects in high-dimensional space, describe potential attack models that exploit this representation, and propose algorithmic and protocol revisions for affected entries in the literature to mitigate these risks.

Computational tasks amenable to an edge computing paradigm generally are both small and self-contained, making them better suited for redundant execution than other motivating scenarios in trusted third-party computation literature. Furthermore, by analyzing behavioral changes from a graph topology perspective, malicious interference becomes both easier to detect and much more difficult to mask.

By generalizing my work from distributed sensing to a broader collaborative computation context, I hope to broaden its practical applications, such as secure decentralized communication networks, adversarially-resilient crowd computing platforms, and other community-driven systems.

References

[1] H. W. Behrens and K. S. Candan, "Adversarially-Resistant On-Demand Topic Channels for Wireless Sensor Networks," in *SRDS*, 2018.

[2] ——, "WindRose: Adversarially-resistant oblivious routing with masked geographic targeting," in *DSN*, 2020, in review.

[3] ——, "Pando: Efficient Byzantine-Tolerant Distributed Sensor Fusion Using Forest Ensembles," in *ICC*, 2020, in review.

[4] M. Bellare, O. Goldreich *et al.*, "Incremental Cryptography: The Case of Hashing and Signing," in *CRYPTO*, 1994.

[5] G. Pereira, R. Alves *et al.*, "Performance Evaluation of Cryptographic Algorithms over IoT Platforms and Operating Systems," *Security and Communication Networks*, 2017.