Orbital Trust and Privacy: SoK on PKI and Location Privacy Challenges in Space Networks

David Koisser, Sanctuary; Richard Mitev, Technische Universität Darmstadt; Nikita Yadav, Indian Institute of Science, Bangalore; Franziska Vollmer and Ahmad-Reza Sadeghi, Technische Universität Darmstadt

https://www.usenix.org/conference/usenixsecurity24/presentation/koisser

This paper is included in the Proceedings of the 33rd USENIX Security Symposium.
August 14–16, 2024 • Philadelphia, PA, USA
978-1-939133-44-1

Open access to the Proceedings of the 33rd USENIX Security Symposium is sponsored by USENIX.
Orbital Trust and Privacy: 
SoK on PKI and Location Privacy Challenges in Space Networks

David Koisser  
Sanctuary  
david.koisser@sanctuary.dev

Richard Mitev  
Technical University of Darmstadt  
richard.mitev@trust.tu-darmstadt.de

Nikita Yadav  
Indian Institute of Science  
nikitayadav@iisc.ac.in

Franziska Vollmer  
Technical University of Darmstadt  
franziska.vollmer@stud.tu-darmstadt.de

Ahmad-Reza Sadeghi  
Technical University of Darmstadt  
ahmad.sadeghi@trust.tu-darmstadt.de

Abstract

The dynamic evolution of the space sector, often referred to as "New Space," has led to increased commercialization and innovation. This transformation is characterized by a surge in satellite numbers, the emergence of small, cost-effective satellites like CubeSats, and the development of space networks. As satellite networks play an increasingly vital role in providing essential services and supporting various activities, ensuring their security is crucial, especially concerning trust relationships among satellites and the protection of satellite service users.

Satellite networks possess unique characteristics, such as orbital dynamics, delays, and limited bandwidth, posing challenges to trust and privacy. While prior research has explored various aspects of space network security, this paper systematically investigates two crucial yet unexplored dimensions: (i) The integrity of PKI components directly impacts the security and privacy of satellite communications and data transmission, with orbital delays and disruptions potentially hindering timely certificate revocation checks. (ii) Conversely, transmitting user signals to satellites requires careful consideration to prevent location tracking and unauthorized surveillance. By drawing on insights from terrestrial studies, we aim to provide a comprehensive understanding of these intertwined security aspects, identify research gaps, and stimulate further exploration to tackle these research challenges in the evolving domain of space network security.

1 Introduction

The space sector has witnessed a remarkable surge in commercialization and innovation over the past decade. The number of operational satellites has quintupled to almost 5,500 since 2013 

1 https://www.ucusa.org/resources/satellite-database

since 2013 


3 https://www.spacefoundation.org/2012/04/05/space-foundations-2012-report-reveals-12-2-percent-global-space-industry-growth-in-2011/

4 https://www.interorbital.com/Cubesat420Kits.php

5 https://aws.amazon.com/ground-station/

6 https://www.isispace.nl/satellite-as-a-service/

7 https://www.esa.int/Enabling_Support/Operations/OPS-SAT

8 https://space-inventor.com/

33rd USENIX Security Symposium 6093
network refers to a system of interconnected satellites, ground stations, and other space-based assets designed to facilitate communication, data transmission, navigation, and other functions in space and between space and Earth. These networked satellites enhance global coverage, reduce latency, and support various applications, including broadband internet services. External users can tap into satellite-based communication services to access long-distance voice calls, data transmission, and internet connectivity.

A recent popular space network architecture is the mega-constellation, such as SpaceX’s Starlink satellite Internet project, aiming to deploy a total of 42,000 satellites\(^\text{10}\). The traditional space agency aims to exploit this development by designing communication protocols that support multi-hop communications for the sparsely connected space network. Namely, IETF’s Bundle Protocol\(^\text{11}\) [10] aims to enable Delay/Disruption Tolerant Networking (DTN)\(^\text{12}\) for satellite networks. NASA and their partners plan to use this technology for the Artemis program to establish an outpost on the Moon\(^\text{13}\).

However, the unique characteristics of satellite networks, including orbital dynamics and limited bandwidth, give rise to various security and privacy challenges, exacerbated by the diversity among satellite owners. Moreover, the interconnected nature of satellite systems means that an attack on one satellite can potentially impact others within the same network or constellation, amplifying the consequences of successful attacks. In the worst cases, satellite collisions may occur if satellites deviate from their orbits, generating dangerous space debris that poses risks to other satellites and spacecraft.

The existing literature on satellite security has focused on various aspects, such as satellite platform (e.g., potential attacks, design issues, access control)\([23,31,46,79,96]\), communication (e.g., cryptographic schemes, authentication protocols)\([53,65,87,93,97]\) and satellite networks (e.g., Global Navigation Satellite Systems (GNSS), key management, secure routing, network attacks)\([17,28,38,54,55,60,62,66,73,85,91,101,105,107]\).

Our literature investigation identified two critical security and privacy dimensions within satellite networks that have not yet received sufficient attention. The first dimension revolves around Public Key Infrastructure (PKI) complexity. PKI has been proposed for use in terminal to satellite encryption\([16,29]\) as a trust framework that provides the necessary key management, cryptographic protocols, and infrastructure to establish and maintain secure communication channels. In the case of Public Key protected Satellite-to-Satellite (SS) connections, there exist few works considering this scenario explicitly in the first place, whereas Ground-to-Satellite (GS)\([35,70,99]\) link security works exist in plathora. Unfortunately, works on SS security only rely on symmetric key cryptography\([34]\) or do not address revocation\([75]\), maintaining an open gap in the literature.

The usage of PKI in space or Delay Tolerant Network (DTN) scenarios was proposed for key distribution\([5,6,39,82]\) establishment\([6]\) or agreement\([69]\).

However, deploying PKI in satellite networks faces unique challenges. Orbital dynamics, limited bandwidth, and computational resources impose constraints on certificate management, with satellite mobility and intermittent connectivity complicating certificate revocation procedures.\(^\text{14}\)

The second aspect pertains to the security implications of satellite networks for users. As satellite networks become increasingly popular, ensuring user location privacy is crucial, especially with the rising adoption of satellite-based services like satellite internet for terrestrial applications. Signals transmitted from ground users to satellites can be intercepted by other satellites, enabling user location triangulation.

The interconnection between PKI complexity and location privacy within satellite networks is multifaceted, underscoring the intricate balance needed to uphold security and user autonomy. Compromised PKI elements, such as certificates, can lead to unauthorized access and data breaches, including location-related data. Conversely, robust mechanisms for preserving user location privacy are essential for strengthening trust in satellite networks. Advanced privacy-enhancing techniques, like location obfuscation, can be integrated into PKI frameworks to mitigate privacy risks while maintaining security assurances. Moreover, orbital delays and disruptions may impede timely certificate revocation checks, while transmitting user signals to satellites requires careful consideration to prevent location tracking and unauthorized surveillance.

We examine these challenges and explore research conducted in these areas, assessing their applicability and adaptability to the space domain. Given the limited research in space network security, we also incorporate insights from terrestrial studies as they can offer valuable perspectives for addressing security concerns in space networks.

This paper is structured as follows. Section 2 presents the background on space networks. Section 3 describes the challenges in space networks. Sections 4 and 5 present a comprehensive literature overview on revocation checks and user location privacy, respectively. We also provide insights into the prior works and highlight the open research challenges.

### 2 Background on Space Networks

Space-based communications architectures are historically quite simple. In the past, satellite communication required

---

\(^\text{10}\)https://spacenews.com/spacex-submits-paperwork-for-30000-more-starlink-satellites/

\(^\text{11}\)https://public.ccsds.org/Pubs/734x2b1.pdf

\(^\text{12}\)https://www.nasa.gov/directorates/heo/scan/engineering/technology/disruption_tolerant_networking_overview

\(^\text{13}\)https://www.nasa.gov/gateway/overview

\(^\text{14}\)Other space-related aspects concern the exposure of satellites to harsh environmental conditions such as radiation, temperature fluctuations, and mechanical stress, which can affect the reliability and integrity of PKI components onboard.
a direct connection from the satellite dish to the satellite. However, the future paradigm for space communications involves leveraging multi-hop communications among many satellites, i.e., space networks. Space networks refer to the interconnected system of satellites and other spacecraft that communicate with each other and with ground-based control stations. These networks play an important role in various applications, including telecommunications, Earth observation, navigation, and scientific research. Concrete examples can already be seen, such as the already-deployed European Data Relay Satellite System (EDRS)\textsuperscript{15}, which enables other Earth observation satellites to transmit data continuously instead of waiting to orbit over a satellite dish. Another example is the concept of creating an Interplanetary Internet\textsuperscript{16} in the solar system using Delay-/Disruption Tolerant Networking (DTN)\textsuperscript{17}. This concept is currently being standardized by both the IETF\textsuperscript{10} and CCSDS\textsuperscript{12}. As technology advances, space networks are expected to play an increasingly vital role in our interconnected world.

Figure 1 shows an example of how a space network may be structured. The initial layer of the architecture consists of the ground station, which includes a network of satellite dishes to establish the connection between Earth and space. At the next level, satellites are in different orbits\textsuperscript{17}. A distinction is made between three types of orbits: Geostationary orbit (GEO), Low Earth orbit (LEO), and Medium Earth orbit (MEO). The geostationary orbit (GEO) is positioned above the equator, allowing satellites to remain stationary over a fixed point on Earth. This orbit is primarily used for telecommunication and weather monitoring satellites. In contrast, low Earth orbit (LEO) is relatively close to Earth, and LEO satellites can be placed in different orbital planes as they do not have fixed orbital paths. An example of this is Starlink LEO\textsuperscript{18}. Satellites in LEO are interconnected, allowing them to downlink data via multi-hop connections, even if they are not in the line of sight to a satellite dish. Medium Earth Orbit (MEO) covers a wide range of orbits between LEO and GEO. Satellites in MEO do not require fixed orbital paths and are used for constellations such as GPS. It is important to note that satellites are in constant motion, resulting in a highly dynamic network topology. However, it is possible to predict the topology based on their regular orbits.

Satellite networks encounter various challenges that affect their efficiency, reliability, security, and privacy. One of the primary concerns is latency, as signals traveling between Earth and satellites experience delays due to the distances involved. Interference from atmospheric conditions, electromagnetic signals, and physical obstacles can degrade signal quality and disrupt communication. Adverse weather conditions such as rain, snow, and fog can further attenuate signals, affecting network performance. Further, satellite networks exhibit a highly dynamic topology due to the constant movement of satellites orbiting Earth. Consequently, any communication scheme should be able to tolerate potentially long delays. This may also lead to entire disruptions of some nodes, as they may lose line of sight to the rest of the network. Dealing with such delays is challenging, especially for security approaches that rely on reliable connections to other nodes. Limited bandwidth is another significant issue, as satellite communication links have finite capacity that can become congested in high-demand areas. Security threats, including unauthorized access and cyberattacks, also jeopardize the integrity of satellite communication systems. Hence, space networks require careful planning, coordination, and advanced technology to ensure reliable, efficient, and secure communication and data transfer in the challenging environment of outer space.

## 3 New Challenges in Space Network

Satellites progressively play an important role in numerous areas of modern life, underscoring the imperative of prioritizing the security of satellites and their users. Our investigations dive into two primary concerns within satellite networks: PKI integrity concerning handling certificate revocation and user location privacy. We concentrate on these security dimensions, which are emerging as indispensable yet remain largely unexplored in the existing literature. The former pertains to the security and trust relations among satellites, while the latter focuses on the security and privacy dynamics between the satellite network and its users. Subsequent subsections offer a comprehensive elucidation of these issues.

Revocation checks are fundamental to the integrity of PKI for space networks. They involve verifying and enforcing access control policies, ensuring that only authorized entities can interact with the satellite network. Revocation checks help mitigate the risks posed by compromised or unauthorized access attempts by regularly evaluating and updating the credentials of users and devices.

On the other hand, user location privacy is concerned with safeguarding the confidentiality and anonymity of users’ geographic location in a network. In the space scenario, this involves mechanisms to prevent malicious actors from track-
ing or deducing users’ whereabouts, thereby preserving the users’ privacy and safety.

The interconnection between revocation checks and user location privacy in the context of satellite networks lies in their complementary roles in ensuring the overall security and privacy of the system and provided services. Compromising either aspect can have cascading effects on the overall security posture of the satellite network.

For instance, if user location privacy is compromised, it could potentially aid malicious actors in circumventing access controls or targeting specific users. Similarly, weaknesses in PKI integrity caused by failing revocation checks could lead to unauthorized access and exposure of sensitive user location information.

3.1 Revocation Checks

In recent years, the utilization of public key cryptography within Space Delay Tolerant Networks (DTNs) [68] and specifically Satellite Networks [93] has been on the rise. This trend emphasizes the growing importance of PKI in guaranteeing the security and integrity of communication within space networks [29]. Hence, a well-defined and practical PKI for satellite networks is becoming increasingly desirable. PKI is a comprehensive system managing cryptographic keys and certificates, which is crucial for securing communication, ensuring data integrity, and facilitating secure access to online services. The European Space Agency (ESA) recently organized an open competition to develop a scalable PKI concept tailored for large constellation networks.

While many studies propose the integration of public key cryptography in space or between ground stations and satellites, a functional PKI is often lacking [16, 29, 36, 61]. Others suggest PKI utilization in space or DTN scenarios for purposes such as key distribution [5, 6, 39, 82], establishment [6], or agreement [69]. However, the critical and challenging issue of key revocation in a space environment has not been adequately addressed without imposing impractical resource requirements, such as the unmanageable memory usage associated with Certificate Revocation Lists (CRLs) [4, 20]. Ensuring revocation checks for keys and certificates is crucial when employing PKI. Failing to revoke compromised keys promptly poses severe security risks, enabling adversaries to exploit compromised keys to deceive other network nodes or take control of affected satellites.

As we detail in Section 4.2, traditional and popular revocation check approaches, like the Online Certificate Status Protocol (OCSP) [8], are unsuitable for space networks due to their reliance on uninterrupted connectivity and the potential for large communication overheads. OCSP requires direct contact with issuing certificate authorities (CAs), which is infeasible in space networks because of the inherent delays and disruptions. Similarly, CRLs pre-cache revocation information and depend on regular updates, which creates vulnerability windows and adds communication overhead, particularly in partially disconnected space networks.

In conclusion, revocation checks in space networks necessitate innovative methods and strategies. Addressing this challenge is a focal point of ongoing research efforts.

3.2 User Location Privacy

Satellite-based services encompass a diverse range of applications utilizing satellites orbiting Earth to provide essential functions such as communication, navigation, Earth observation, and scientific research. These services, facilitated by constellations like GPS and GLONASS, offer precise positioning, navigation, and timing data, serving applications from personal navigation devices to aviation and maritime navigation.

Preserving location privacy in space networks is evident due to the potential implications for individual privacy, security, and safety. The revelation of users’ geolocations can facilitate targeted surveillance, tracking, or even physical harm in sensitive or hostile environments. Furthermore, compromised location privacy may undermine users’ trust and confidence in satellite-based services, hindering their adoption and utilization.

While existing literature has extensively discussed threats such as eavesdropping, jamming, and spoofing on communication channels between ground users and satellites [43, 93], another critical but less explored class of attacks concerns compromising the location privacy of satellite-based service users. Signals transmitted from ground to satellite can be triangulated by neighboring satellites, enabling adversaries to estimate users’ geolocations. For instance, if a satellite internet user uploads a large file, adversarial satellites nearby can track the signal, measure its strength at multiple points, and estimate the user’s location. With the increasing access to satellites, this issue poses significant concerns for user privacy and security.

Addressing the challenges of realizing location privacy in space networks requires novel solutions tailored to satellite-based communication systems’ unique characteristics and constraints. Location privacy in space networks is usually achieved by spanning a mesh network over the already deployed base stations. Unfortunately, solutions to mitigate users’ geolocation in other types of networks (e.g., the internet) do not apply to partially connected networks, such as mesh networks, e.g., terrestrial ground-based satellite network users. We will discuss different approaches in Section 5.2. Solutions like TOR [19] are overlay networks and require a well-connected network (c.f., [49]) and, therefore, will be omitted from the following comparison and analysis.

4 Revocation Checks

This section dives into the literature on revocation checks applicable to the space networks, focusing on several core selec-
tion criteria. Firstly, many works in this domain assume that replacing an old key with a new one inherently results in revocation. However, this assumption proves insufficient when keys are re-issued before expiration, leaving both old and new keys considered benign despite potential compromises. Therefore, our focus lies on literature explicitly addressing revocation, aiming to address this limitation.

Secondly, a noteworthy aspect of our paper is the inclusion of literature addressing revocation within Delay/Disruption Tolerant Networking (DTN), as the system model in these environments applies to space networks. Given that space networks are among the primary use cases for DTN, incorporating this literature enriches our understanding of revocation mechanisms in relevant contexts.

Additionally, our selection criteria emphasize works that address mutual authentication scenarios. We exclude prior works assuming a client-server architecture (e.g., traditional Certificate Revocation Lists (CRLs)), where only one party needs to assure authentication and the validating party maintains a reliable connection to the issuing certificate authority (CA). Such scenarios differ significantly from those applicable to space networks. We opt not to include the cases where a simple on-demand revocation check mechanism suffices.

### 4.1 Literature Overview

This section offers details on revocation approaches applicable to the space network, which are presented in Table 1. The **Authentication Method** column treats both Public Key Cryptography (PKC) and PKI as distinct methods. PKI implies a chain of certificate authorities that issue certificates. PKC only means that the public cryptography algorithms (e.g., elliptic curves) are used without certificates and sometimes without a certificate authority.

We also distinguish between broadcast and gossip. Numerous papers generally leverage an implicitly reliable broadcast channel. However, often, there are no details on how the broadcast works practically. Other papers specifically mention that the nodes spread information by sending data to their neighbors, who, in turn, send it to their neighbors, and so on. While a varying degree of detail describes how this process works, we classify such methods as gossip.

Further, we analyze the mutual trust between the parties, as shown in the **Mutual Trust** column. The ∼ symbol indicates a trust imbalance between the parties, and their mutual authentication methods may differ. One party may require a timely revocation check while the other may merely rely on certificate expiration without a revocation check. The ✓ symbol indicates that both parties check each other’s revocation status.

Finally, the last column shows the evaluation approaches. **Theoretic** evaluation indicates that the work presents mathematical calculations for the evaluation. For example, equations to calculate expected storage overheads when using certain data structures. **Experimental** evaluation indicates a deployment on a real device to measure the run-time overheads in a representative environment. **Simulation** refers to large-scale evaluation using a network simulator to measure network overheads. We categorize the works based on revocation methods for the following discussion.

#### Short-lived credentials

Short-lived credentials implicitly revoke credentials if they are not renewed within a specified timeframe. Seth and Keshav [88] propose an authentication scheme for Delay-Tolerant Networks (DTNs) based on hierarchical identity-based cryptography (IBC) [27]. IBC [9] uses the identity of a node (e.g., an email address) as its public key. The certificate authority (CA) typically combines a node’s general public parameter with its identity. The hierarchical IBC approach in this work is based on regions, with an overarching root certificate authority (CA) and each region having its sub-certificate authority (CA). On registration, nodes authenticate via a PKI-based certificate towards the IBC-based CA. The CA then generates the respective private key for the node’s identity. Users may register region-specific keys with region CAs or get general keys with the root CA. Each IBC-based certificate is only valid for a defined time frame. Thus, it is argued that revocation is implicit as the certificate will eventually become invalid if not renewed. This also addresses the typical problem of re-issuing all nodes’ keys in IBC if one key is compromised.

Chen et al. [14] propose that each node generates its private key. It then sends a challenge signature to the CA, which uses the signature to calculate the user’s master key. The master key is used to derive symmetric session keys for the node. These session keys have a short lifetime; thus, each node must request a new one regularly. If the node is revoked, the CA does not issue a new session key and deletes the node’s identity material.

Meng et al. [70] target mutual authentication between a user and a satellite service. The CA on the ground delegates partial trust for authentication checks to the satellites to reduce delays. The satellites can use proxy signatures based on elliptic curves to provide proxy authentication for the CA. This proxy authority is temporary. Users first register with the CA directly and get temporary anonymous credentials and information on the proxy authorities. The revocation is thus implicit with the expiry of either side’s certificate. On certificate expiration, an online direct check is triggered with CA to get up-to-date validation information. The paper proposes a scheme to handle handovers, as long user sessions involve multiple satellites and thus proxy authorities.

#### Broadcast or gossip-based revocation dissemination

Another class of works distributes revocations in the network via broadcast or gossip. Roy-Chowdhury et al. propose to use both PKI and TESLA [86]. Timed Efficient Stream Loss-tolerant Authentication (TESLA) [80] is a symmetric broadcast authentication protocol. It uses a hash chain in which each hash represents a time slice, with a secret key at the end.
Table 1: Overview of revocation literature applicable to space networks. In the mutual trust column, the ∼ symbol denotes varying mutual authentication methods between parties and ✓ symbol indicates that both parties verify each other’s revocation status. The × symbol indicates lack of security analysis and evaluation.

<table>
<thead>
<tr>
<th>Authentication Method</th>
<th>Revocation Method</th>
<th>Trust Assumptions</th>
<th>Key Assumptions</th>
<th>Mutual Trust</th>
<th>Security Analysis</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[88] IBC</td>
<td>Short-lived keys</td>
<td>CAAs know node secrets, CAs are always trusted,</td>
<td>Secure PKI in place</td>
<td>∼ ✓ ×</td>
<td>✓</td>
<td>Informal Simulation</td>
</tr>
<tr>
<td>[14] PKC for registration, rest SKC</td>
<td>Short-lived session keys</td>
<td>CA can derive all secrets</td>
<td>—</td>
<td>— ✓</td>
<td>Informal</td>
<td>Experiment, theoretic</td>
</tr>
<tr>
<td>[70] PKC</td>
<td>Short-lived trust, online check on expiry</td>
<td>CA can derive all secrets</td>
<td>— ✓</td>
<td>✓ Formal</td>
<td>Experiment, theoretic</td>
<td></td>
</tr>
<tr>
<td>[86] PKI for registration, rest TESLA</td>
<td>Individual broadcast, short-lived certificates</td>
<td>CA can derive all secrets</td>
<td>CA has reliable broadcast</td>
<td>✓ Informal</td>
<td>Experiment</td>
<td></td>
</tr>
<tr>
<td>[81] PKI</td>
<td>Hierarchical Bloom filter, individual gossip</td>
<td>Majority benign nodes</td>
<td>Reputation-based adversary detection system</td>
<td>✓ Informal</td>
<td>Theoretic, simulation</td>
<td></td>
</tr>
<tr>
<td>[37] PKC</td>
<td>Individual gossip</td>
<td>Inherently secure second channel (confirms identity)</td>
<td>Revocation spreads faster than re-issued keys</td>
<td>✓ Informal</td>
<td>Simulation</td>
<td></td>
</tr>
<tr>
<td>[26] Distributed PKI</td>
<td>Implicit, individual broadcast</td>
<td>Group of reputable nodes as distributed CA</td>
<td>Reputation-based adversary detection system, reliable broadcast</td>
<td>✓ Informal</td>
<td>Theoretic, simulation</td>
<td></td>
</tr>
<tr>
<td>[16] PKC</td>
<td>CRL</td>
<td>—</td>
<td>—</td>
<td>✓ ✓</td>
<td>✓</td>
<td>Simulation</td>
</tr>
<tr>
<td>[35] PKC</td>
<td>Only CA can check for revocation</td>
<td>CA can reveal anonymous nodes</td>
<td>Satellites are not revoked</td>
<td>— Informal</td>
<td>Theoretic, experiment</td>
<td></td>
</tr>
<tr>
<td>[103] PKC</td>
<td>CRL contains parameters to compute new group key</td>
<td>CA can derive all secrets</td>
<td>Nodes sign as a group</td>
<td>— Informal</td>
<td>Experiment, theoretic</td>
<td></td>
</tr>
<tr>
<td>[30] IBC</td>
<td>CRL stored on blockchain</td>
<td>Registration authority group (blockchain consensus)</td>
<td>Reliable access to blockchain</td>
<td>∼ ✓</td>
<td>Simulation</td>
<td></td>
</tr>
<tr>
<td>[4] NOVOMODO</td>
<td>Time slice hashes for CA, hash table for nodes</td>
<td>—</td>
<td>Efficiency / vulnerability window tradeoff</td>
<td>✓ Informal</td>
<td>Experiment, simulation</td>
<td></td>
</tr>
<tr>
<td>[47] Distributed PKI</td>
<td>Terrestrial schemes</td>
<td>Multiple CAs via consensus</td>
<td>—</td>
<td>✓ ✓</td>
<td>In-orbit experiment</td>
<td></td>
</tr>
<tr>
<td>[48] PKI</td>
<td>Sparse Merkle Tree</td>
<td>—</td>
<td>Subset of nodes use more storage space</td>
<td>✓ Informal</td>
<td>In-orbit experiment, simulation</td>
<td></td>
</tr>
<tr>
<td>[20] PGP</td>
<td>Neighbors vouch for node’s revocation (and new key)</td>
<td>Neighbors with direct connection establish trust</td>
<td>Only protects against impersonating adversary</td>
<td>✓ Simulation</td>
<td>Simulation</td>
<td></td>
</tr>
<tr>
<td>[59] PKC</td>
<td>Time-specific certificates and CRL</td>
<td>CA can derive all secrets</td>
<td>Users know the satellites along the route</td>
<td>— Informal</td>
<td>Experiment, theoretic</td>
<td></td>
</tr>
</tbody>
</table>

of the chain. The hashes, concerning the current time, are used to calculate MACs over the sent messages. The hashes—and finally, the secret key—are gradually disclosed, which allows the receiving parties to confirm the authenticity of the message, given the arrival time of the message is before the hash disclosure. First, a node registers with the CA by providing a PKI-based certificate for identification. Then, the CA sends the node a TESLA certificate back, allowing the node to start its own authenticated TESLA chain. As nodes are assumed to be sufficiently time-synchronized, two nodes can use the current hash in the TESLA chain to authenticate each other. The nodes can confirm each other’s certificate with the respective hash disclosure in the next time slice. When a certificate needs to be revoked, the CA broadcasts this information to all nodes. The paper further argues that the TESLA certificates have a short lifetime, and thus, an eventual implicit revocation exists.

In Qian et al.’s revocation scheme for space networks [81], the revocation information is stored in a hierarchical set of Bloom filters [7]. This structure improves the false positive rate inherent in Bloom filters. To justify revocations in the first place, the paper assumes an adversary detection scheme is in place, which allows nodes to check each other for misbehavior. Each misbehavior is reported, and a reputation system tracks each node. In case the reputation of a node falls below a defined threshold, the CA revokes the node and issues a new bloom filter.

Jia et al. propose a public key distribution scheme for DTNs based on two-channel cryptography [37]. Aside from the normal broadband communication channel, these systems assume an inherently secure second low bandwidth channel and an established identity. Nodes exchange public keys and other identifying information on the normal channel, while a hash of the transferred data is sent via the second channel. Due to the inherent security of the channel, it is sufficient to authenticate the received credentials. In some cases, the two carriers (with many nodes) exchange credentials of an entire group of nodes, meaning an entire set of keys is exchanged.
For revocation, the affected node or carrier starts the epidemic routing and spreads the revocation. As this requires indirect forwarding by other nodes to reach the entire network, each node checks if there is a direct chain of trusted nodes to the receiver. Further, as revocation implies a re-issued public key and DTNs expect a dynamic topology, the paper assumes a quality of service mechanism to ensure the revocation spreads faster than the new key.

Fang et al. propose a distributed approach to PKI [26], in which a distributed certificate authority (CA) is established, i.e., a subset of nodes are chosen to operate as the CA cooperatively. The nodes are chosen based on a reputation system, and the most reputable nodes become part of the CA group. The paper assumes that an intrusion detection scheme is deployed, which allows the nodes to measure each others’ trustworthiness. If a node detects suspicious activities, it broadcasts the information to the network, and each node individually tracks the reputation of all the nodes in the network. A node can be revoked in three cases: (1) if a node’s reputation is low, (2) if a node’s certificate is expired, or (3) if a node revokes its certificate. The revocation is implicit in the first two cases. The paper does not address how a node proves its identity to others in the latter case.

Certificate Revocation List (CRL) Some works use straightforward revocation techniques, such as traditional CRLs in space networks. Cruickshank [16] is one of the first works to address revocation for space networks. The approach is essentially a standard hybrid for channel security between the ground users and the satellites. Public keys are exchanged via signed certificates, and a symmetric key is exchanged to establish a secure channel through symmetric encryption. They propose CRLs for revocation. However, the paper does not address the distribution of the CRL in the space environment.

Ibrahim et al. [35] propose a method where ground-based users and satellite networks authenticate each other while keeping user identity hidden. The user’s authentication data needs to be anonymous and untraceable. A Certification Authority (CA) uses elliptic curves to generate and share a set of blinded secrets to achieve this. When users register, they pick a random blinded secret from this set, compute a session key, and use it for communication protection. By changing the blinded secret regularly, the user stays anonymous. Only the original generator of the secrets can calculate the mutual session key so that the user can authenticate the service provider. The scheme resists jamming as authentication isn’t interactive, so the jamming adversary doesn’t know when the user communicates. For revocation, the service provider can determine the user’s identity and delete their registration, preventing them from establishing a valid session key. However, it’s unclear how this revocation information is sent to the satellites.

Yang et al. propose another approach for the same setting as above which leverages group signatures for roaming users [103]. A set of users sharing the same domain are grouped. The CA provides the necessary information to the users to calculate group signatures, and the satellites use traditional certificates. For user authentication, the satellites check if the group signature used by any of the group users is valid. They use a special type of CRL for revocation, which contains the necessary parameters to update the group certificate information. When a user is revoked, the satellites and the users update their certificates.

Guan et al. propose another approach based on IBC which leverages blockchain technology [30]. In this scheme, nodes register with one of the registration authorities, the registration authorities are blockchain nodes. The authorities come together to establish a permissioned blockchain. A permissioned blockchain contains a limited set of consensus nodes, as opposed to a permissionless blockchain, which anyone can join, and consensus happens amongst all the network’s participants. Upon registration, the registration authorities generate and encrypt their key shares, which the registering node uses to retrieve its private key. As the scheme uses IBC, the node’s identity acts as its public key, which is stored on the blockchain, along with the encrypted private key. Satellites act as proxy authorities, storing their identities on the blockchain. The blockchain is also used to store and distribute a CRL for revocation. The key assumption is that the nodes keep an up-to-date version of the chain and thus have reliable access to the blockchain.

Efficient data structures Bhutta et al. [4] and Silvio Micali [71] propose NOVOMODO scheme, which involves splitting the time until certificate expiration into discrete time slices, each represented as a hash in a hash chain spanning the entire validity period. When a node wants to authenticate itself before others, the Certificate Authority (CA) provides the current hash to append to the certificate, enabling a single hash instead of a complete proof akin to OCSP. Moreover, the validity of these time slices can be adjusted, allowing for handling delays with granularity. For instance, if each time slice is an hour—including a new validity hash per hour—nodes also accept hashes that are five hours old. This approach allows for handling delays with a selected granularity.

Multiple CAs within the network are responsible for their nodes, and the NOVOMODO scheme is employed explicitly for these CAs. The respective CA issues a hash table of revoked nodes alongside the current NOVOMODO hash for node revocation, facilitating validation even with delayed arrival. However, this efficiency-security trade-off entails more considerable periods, increasing vulnerability windows while reducing overhead.

Koisser et al. propose to use multiple CAs to form a distributed CA [47]. They explore various revocation schemes based on terrestrial methods. Since space networks involve many distrustful parties, agreeing on a single CA is challenging. To address this, the paper proposes a consensus protocol among CAs using a signature aggregation scheme, keeping communication overhead low. Revocation schemes are di-
vided into two categories: burden shifting to the prover and data structures containing all revocations. The first category includes OCSP Stapling, where the prover regularly fetches short-lived OCSP responses, and Merkle hash trees, providing a Proof-of-Inclusion (PoI). The second category proposes CR-Lite, using cascading Bloom filters, and Let’s Revoke, using incremental numbers and compressed bitvectors. The paper evaluates the runtime overhead of both signature aggregation and revocation schemes on an in-orbit satellite.

Koissier et al. [48] build on using a Merkle hash tree to aggregate certificate validation information as a revocation scheme for space networks. Specifically, the paper proposes to use the Sparse Merkle Tree (SMT) [50] to store all certificates’ validation information. The SMT builds a complete hash tree over the entire search space, i.e., all possible outputs of the used hash function. In the beginning, all leaves of the SMT are empty and are consequently populated by the hashes of all active certificates in the network. As the tree mostly contains empty leaves and branches, large parts of the tree can be omitted to make it calculable. Since the position of any certificate (or rather its hash) is deterministic in the tree, leaves can be deleted without recalculating the entire hash tree. This results in properties that allow up-to-date nodes to share their valid PoI to repair outdated nodes’ PoIs on any changes (e.g., a revocation that deletes a leaf in the SMT). This allows nodes to spread revocation information epidemically throughout the network without directly contacting the CA or inducing large communication overheads. The paper argues that the scheme enforces fast revocations due to the inherently epidemic spread and is independent of the network topology. This work also evaluated the proposed scheme on an in-orbit satellite.

Other approaches Djamaludin et al. [20] propose an approach to leverage a PGP-style trust network for revocation. In Pretty Good Privacy (PGP) [108], each node self-signs its certificate and trusts the certificates of only certain nodes (e.g., nodes it already met). If a new node, and thus an unknown certificate, is encountered, the node checks if anyone knows the new node of its trusted nodes. Thus, a chain of trust is established such that the node can indirectly establish trust in the new node. In the presented work, neighboring nodes (i.e., nodes that have a direct connection to each other) can inherently establish trust in each other’s identity. The PGP approach is then used to communicate with nodes that do not have a direct connection. If a node needs to revoke its certificate, it collects signatures from its neighbors. Thus, these neighbors essentially vouch for the revoked node without a certificate. The number of signatures required is defined as a fraction of the total number of nodes in the network. If the revoking node has collected enough signatures, it broadcasts its revocation. This typically also includes a new certificate, which is then vouched for in the usual PGP way. As trust is established by direct contact, the approach only protects against an adversary who tries to impersonate another node.

In the work of Liu et al. [59], satellite service users move and encounter different access points, i.e., satellites. The proposed scheme establishes a mutual authentication scheme between the users and the encountered access points. First, both users and access points register with the CA. Users anticipate the times they will encounter the various access points before they start their trip and send a list of this information to the CA. The CA generates anonymous credentials for the user based on group signatures, which are only valid for a specific period for each access point (according to the user’s trip plan). A list of the respective public keys is also retrieved to authenticate the access points. These credentials (per access point) are short-lived for revocation. Additionally, if a user is found malicious, the CA reveals the user’s identity and distributes a special CRL to the satellites, which contains parameters to subtract the user’s key from the group’s credentials. For the revocation of access points, only short-lived certificates are implied.

4.2 Analysis

In this section, we analyze the literature on revocation schemes in the light of their applicability to the space scenario and give intuitions about their shortcomings. The short-lived credential method implicitly revokes the certificate if it is not renewed in time [14, 70, 88]. Even though the assumptions required for these approaches are not impractical (c.f., Table 1), these straightforward approaches have a clear downside. To enforce a quick revocation in case of a compromise, the lifetime of the credentials needs to be kept short. However, the shorter the lifetime, the more network overhead is implied, as the credentials need to be renewed often. Further, some satellites may be disconnected at the time of renewal, leading to issues regarding either missed updates or how these credentials are distributed. For example, DTN uses a store-and-forward approach, and network disruptions imply that the forwarding nodes spend resources to store the message until they can successfully forward it. This strains the network and may even lead to situations in which a node may miss several renewals.

The next approach simply informs the network nodes of revocations directly via broadcasts or gossip-based dissemination [26, 37, 81, 86]. While this is a valid strategy in principle, we highlight two issues with this approach. First, these works do not consider how the respective dissemination is affected by the topology of satellites. In works that use broadcast [26, 86], there is no discussion on how to broadcast in a space network. As shown in Table 1 column Key Assumptions, a reliable broadcast technique is assumed. However, the effectiveness and swiftness of the revocation directly depend on the implicit properties of the used dissemination. While the broadcasting technique itself is an orthogonal issue, it needs to be considered to hint at the guarantees of the revocation. Another issue is that the broadcast/gossip messages may not arrive at all the nodes in the space networks, e.g., when
parts of the network are temporarily disconnected. One of the works assumes that a share of the network is malicious [81], and it can also affect the gossip protocol. This has crucial implications on the effectiveness of the revocation.

The straightforward revocation techniques [16, 30, 35, 103], such as traditional CRLs [8], pre-cache revocation information on virtually all certificates in the network, and the absence of a revocation indicates the certificate’s validity. Usually, a CA distributes such a CRL to all nodes in the network and regularly updates it. Note that even if there are no revocations since the previous update, it is important for the CA to distribute still an update such that nodes can be sure that they have the newest CRL. However, there are two intertwined problems with this approach. Updating the CRL in a defined interval also implies an inherent vulnerability window. For example, if a certificate is revoked just after a CRL update, nodes in the network will not know about the revocation until the next CRL update. While a revocation could also trigger an exceptional update to be distributed earlier, nodes in a space network may experience disruptions and not receive the exceptional update. Due to a recently received regular update, they will assume a compromised certificate to be benign. This problem can be avoided by shortening the time interval between expected CRL updates, yet such a strategy ties into the second problem. A CRL may grow to large sizes, and regularly sending out updates may induce large communication overheads. Combined with a partially disconnected space network due to delays and disruptions, this results in a non-trivial challenge when using CRLs. Further, so far, we assumed that the updates reach all nodes in the network, which may not be the case in an actual deployment. They also do not discuss the implications of how the data is disseminated. Further, there is a lack of discussion and analysis on updating this data, such as delta CRLs. Delta CRLs may significantly reduce the overhead of revocation information dissemination but may also cause problems if nodes are disconnected.

Other works use efficient data structures to reduce the network overhead [4, 47, 48]. These approaches alleviate the dissemination problem by removing the need to send distinct messages for revocation. However, these works also do not consider how to disseminate the revocation information by relying on terrestrial schemes (c.f., Table 1) column Revocation Method; thus, there can be issues in real-world deployment of these approaches. One exception is the work by Koisser et al. [48], whose revocation approach is topology-independent. They define exchange protocols relying on small amounts of data that allow nodes to ascertain if they’re outdated. If nodes notice that they are not up-to-date, they can reject credentials and try to update their revocation information without contacting the authorities. Djamaludin et al. [20] relies on a PGP-style revocation system in which neighbors vouch for each other. It assumes that the involved parties are willing to establish an entirely self-organized trust network, which may be difficult in a real-world scenario. Moreover, the neighbor-based trust and dissemination approach is prone to scalability issues, especially in the highly dynamic environment of a space network. Liu et al. [59] rely on credentials only valid for a specific time frame. This work assumes that the nodes know in advance exactly which other nodes they will encounter and interact with. This assumption may be impractical in many scenarios, especially for a large-scale space network. There is a trade-off between either using a pre-computed long list of encounters or regularly acquiring a new list. Both cases imply a significant communication overhead, with irregular large amounts of data or regular updates of small lists.

4.3 Open Research Challenges

In this section, we outline the identified research challenges based on our analysis of the prior works.

Multiple CAs In a practical deployment, PKI can include multiple stakeholders such as governmental space agencies and commercial players; relying on a single CA is unfeasible, as the parties mutually distrust and may not agree to trust a single authority. The use of multiple CAs may cause additional challenges in terms of scalability, which needs to be considered explicitly.

Some existing works do consider multiple CAs. In Fang et al. [26], a reputable group of nodes inside the network is elected to do the decision-making as a CA for the entire network. However, it assumes all parties agree to place their trust in a distributed approach. The work of Bhutta et al. [4] also considers multiple CAs, which work independently. If a node belonging to one CA wants to check the validity of another CA, the leveraged NOVOMODO approach supports efficient checks. Guan et al. [30] leverages a blockchain to let a group of authorities create a consensus-based repository, which stores all agreements among the group, e.g., issuance or revocation of certificates. However, this approach, therefore, assumes a reliable blockchain in place. Similarly, Koisser et al. [47] proposes a consensus protocol for secure agreement among CAs. However, this solution works only if the involved parties accept quorum-based decision-making, i.e., the others may override a single authority. Multiple CAs in the unique topology of space networks constitute a challenging research problem.

Topology Optimizations Network topology is another aspect that has received much attention. Unlike other dynamic networks, the periodic nature of orbits creates a predictable topology in space networks, which can be exploited to build solutions that efficiently spread revocation information. Yet, in our investigation of the prior works, we did not encounter any work that explicitly considered the network topology.

Practical Evaluation To thoroughly evaluate schemes designed for space networks, deploying and testing them on in-orbit satellites is essential. To our knowledge, only Koisser et al. [47, 48] have
conducted such tests for their proof-of-concept software on real-world satellites. Evaluations on actual space networks provide representative results that can lead to new insights and reveal opportunities for optimization.

5 Location Privacy for Space Networks

This section presents approaches for achieving location privacy for users of ground-based satellite services. Only one paper directly addresses the issue of location privacy for space networks [49]. However, the issue of location privacy arises in the security of the physical layer in satellite communications. Furthermore, we found that the issue of location privacy has been extensively researched in Wireless Sensor Networks (WSNs). Although originally designed for terrestrial networks, the system model applies to space networks. In WSNs it is assumed that nodes form a mesh network and use multi-hop routing to reach a sink, which is connected to the WAN. If we apply this to, e.g., a satellite-based Internet scenario, the sink would be the satellite terminal and the service user needs location privacy. Unlike WSNs, it is assumed that there are many sinks, which may form the mesh network instead of the users themselves.

5.1 Literature Overview

The section analyzes the approaches applicable to location privacy for ground-based satellite network users.

Table 2 depicts various location privacy approaches and their corresponding attributes. The Adversary Model concerns one of three adversary models - local, global, or hunter - related to them. The Global adversary passively eavesdrops on the entire network, representing the strongest adversary model. The Local adversary has a limited view, observing only specific network traffic. A Hunter adversary has a localized view, seeing traffic from nearby nodes, and can trace a user’s connection back to the root, assuming the adversary can move. In the section Phantom Routing, we offer a detailed definition of this adversary model as introduced in the paper [77]. Furthermore, Table 2 provides a general indication of each approach’s overhead, including latency, computational, and communication overheads. It is important to note that these schemes were chosen optimistically, as the overheads vary. The Latency overhead indicates the average delay of messages with each approach; e.g., taking a slightly longer route may induce less latency than storing messages before forwarding them at each hop. The Computational overhead refers to the additional computing resources or effort required to perform specific tasks or processes, e.g., apply complex cryptographic primitives on each node along the route. Finally, the Communication overhead considers the entire network rather than just the sending node, referring to the additional payload data.

Anonymous Satellite Internet A recent work that specifically addresses location privacy for satellite service users, such as satellite Internet, is AnonSat introduced by Koisser et al. [49]. Their scheme aims to protect location of satellite Internet users from triangulation by other satellites. The objective is to utilize cost-effective, off-the-shelf devices to facilitate the implementation of this system. This will enable the system to be used directly from consumer devices, such as common smartphones, without needing a bespoke device or application. AnonSat employs long-range wireless communication to establish a local mesh network among dispersed base stations in a designated area. It redirects users’ communication through distant base stations to prevent geolocation of the satellite Internet base station and nearby users. Consequently, an adversary monitoring base stations to find the user’s traffic and geolocate the user cannot assume that the up-linking station is closest to the user.

Physical Security for Satellite Communication While physical security approaches aim to enable satellite communication resilience against eavesdropping, they also facilitate location privacy properties. Various approaches propose the utilization of wireless communication characteristics to achieve communication confidentiality, which will be introduced below.

The goal is, therefore, to achieve a signal-to-noise ratio (SNR) above a specified threshold at the location of the authorized receiver, while maintaining the SNR below the threshold in all other areas. The low SNR of the transmission ensures that the message cannot be decoded by an unauthorized listener [2, 52, 100]. Lei et al. [52] use beamforming and power modulation to negate the SNR of the eavesdropper thereby preventing proper reception of the signal. This technique also safeguards location privacy by preventing geolocation when signals aren’t received. It necessitates the use of specialized satellite hardware and the acquisition of channel state information, which includes details about communication link properties such as scattering, fading, and power decay with distance. Additionally, some approaches use artificial or natural noise to disrupt eavesdroppings, like those by Yan et al [100] or An et al. [2]. Importantly, these approaches may potentially result in unintended consequences about location privacy. An attacker could potentially estimate the location by continuously testing the SNR and observing the direction of increase or decrease.

Below we describe additional approaches, not utilizing physical security but relying on software- or network-based approaches, developed for WSNs to address location privacy, in addition to those previously discussed.

Phantom Routing The first paper to propose the Phantom-Routing approach was by Ozturk et al. [77], which also introduced the hunter-adversary model. In this model, it is assumed that the adversary knows the sink’s location. If the target node routes messages to the sink, the adversary can trace the node used as the last hop of the route. This process could potentially be repeated until the adversary reaches the target node. In the context of a satellite Internet connection,
Table 2: Overview of approaches for user location privacy. The latency, computational and communication overhead are approximated.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Adversary Model</th>
<th>Key Assumption</th>
<th>Latency Overhead</th>
<th>Computational Overhead</th>
<th>Communication Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymous Satellite Internet</td>
<td>Local</td>
<td>No compromises of nodes</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Physical Security [52]</td>
<td>Global</td>
<td>Channel state information must be known</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Phantom Routing [42, 57, 58, 77, 92, 95, 98, 106]</td>
<td>Hunter</td>
<td>Single hunter</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Fake Traffic [1, 11, 63, 64, 67, 89]</td>
<td>Global</td>
<td>—</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Ring Routing [13,21,44,56,83,104]</td>
<td>Hunter</td>
<td>High capacity node form ring</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Multi-Path Routing [15,33,45,94]</td>
<td>Hunter</td>
<td>Single hunter</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Network Coding [24,25]</td>
<td>Global</td>
<td>Feasible overhead for homomorphic encryption</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Avoid Adversary [22,84]</td>
<td>Local</td>
<td>Adversary can be detected</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Artificial Delay [32,40,41]</td>
<td>Hunter</td>
<td>Messages may arrive out-of-order</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Pseudonymity [18,72,76,90]</td>
<td>Global</td>
<td>Pre-shared secret in place</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

the term 'sink' refers to the satellite dish that is connected to the WAN. The hunter can trace the terrestrial route until they locate the actual user. To counter this, the scheme operates in two phases. In the initial phase, the target node employs a random walk scheme to regularly route to various nodes, with the final node designated as the phantom node. Subsequently, in the second phase, messages are transmitted towards the sink, which may result in a flood of messages in that direction. This way, the adversary will be misled to different directions, never reaching the target node. Another work by Yan et al. [42] modifies this scheme to use a direct route to the sink instead of flooding. Xi et al. [98] propose an adaptation of the hunter-adversary model to consider several malicious nodes monitoring the network. Their approach involves the target node routing to multiple phantom nodes simultaneously, while the sink regularly transmits discovery messages. When a phantom node receives such a message, it forwards it to the sink, thereby utilizing multiple concurrent routes, which presents a challenge for any potential adversaries attempting to track them. Another proposal by Wang et al. [95] assumes that the adversary has a limited and known visibility range. The presented scheme incorporates this consideration to perform a directed random walk in a manner that ensures phantom nodes are dispersed sufficiently to prevent potential adversaries from gaining an advantage. To guarantee that the phantom nodes are always a minimum distance away from the target node, Lightfoot et al. [58] propose the selection of phantom nodes within an annular area surrounding the target node. This ensures that phantom nodes are sufficiently far away from the target, without including significant delays due to their proximity. To implement this successfully, it is necessary to have an understanding of the network topology. Similarly, in the RRIN scheme proposed by Li et al. [57], nodes only require knowledge about their neighbors. Note that this scheme is primarily designed for small networks and may not be suitable for larger networks. An alternative approach proposed by Zhou et al. [106] is to utilize an ant colony optimization instead. Here, reasonably direct routes are selected from target to sink, yet the algorithm is modified to select sufficiently different routes each time. The PEM scheme proposed by Tan et al. [92] starts by identifying the shortest route between the target and sink. Then, phantom nodes are selected from nodes that are near this shortest path.

**Fake Traffic** This context addresses works that consider global attackers, i.e. attackers who can monitor the entire network. While some works use fake traffic to address a local or hunter adversary [12,42], this is not the primary focus of this consideration. Mehta et al. [67] proposed a periodic collection scheme to address a Global adversary. In this scheme,
nodes collect messages that have been received, and then forward a fixed number of these messages during a given period, regardless of whether they have collected any messages. This also applies to the forwarding of artificially generated messages. A similar approach entails nodes regularly transmitting fake messages, as proposed by Shao et al. [89]. The target node can take the transmission interval into account to cover its messages. To reduce communication overhead and delays, the scheme further proposes to use variable and probabilistic transmission times. Similar to the previous approach, Majeed et al. [64] propose a scheme that lets nodes regularly send messages. The method has been extended to use multiple paths from source to sink, which increases privacy and dynamically adjusts the rate of fake news in the network to account for the load of real news. The idea of using a probabilistic distribution for the fake message rate is extended by Alomair et al. [1]. Instead of only using the distribution for the fake messages, the scheme considers the real messages as well and fits them into the distribution. The proposal by Mahmoud et al. [63] combines phantom routing with fake traffic generation. In order to enhance efficiency, the generation of fake traffic is only activated when genuine traffic is in operation. According to Bushnag et al. [11], it is recommended to use a low average rate for transmitting both fake and real messages. For implementation, the transmission rate for all messages is constant. Alternatively, real messages can be transmitted at a high rate while subsequent fake messages are transmitted at a low rate to maintain an unchanged overall average.

Ring Routing The ring routing scheme [44] establishes topological rings for each cluster, with nodes clustered together and forwarding in one direction. Messages are efficiently routed through multiple clusters, each with a designated head node that reliably transfers messages to other ring clusters. As messages go along similar routes and form cyclic paths, it becomes harder for a hunter adversary to trace messages back to the actual target node. According to Li and Ren [56] the ring routing approach can be effectively combined with the fake traffic approach. Here, a single ring is used with relay nodes breaking cycles and forwarding messages to a sink, supplemented by fake messages. A similar approach uses multiple rings, layered around the sink Yao et al. [104]. The target node then chooses an adjacent ring as its next hop. Afterward, the message is forwarded towards the sink, with a chance to be routed along any ring on its way. Another approach similar to the initial ring routing methodology [44] involves generating fake traffic within each cluster Ren et al. [83]. Dong et al. [21] also suggest implementing random walking within the ring cluster before performing the forwarding along the ring. Another proposal is to use flooding inside the ring cluster Chen et al. [13].

Multi-Path Routing Wang et al. [94] propose a multi-path scheme to protect against a hunter adversary. Multiple paths from the target to the sink are selected with different general angles towards the sink. Each node randomly chooses a new angle to establish diverse routes such that the adversary cannot trace the route back to the target. The work proposes another scheme that includes message aggregation and fake traffic generation to provide additional protection against a global adversary. Koh et al. [45] optimize the previous approach by leveraging a Bayesian strategy for route selection. The approach by Huang et al. [33] tries to optimize multiple paths globally for the network and merges redundant paths to significantly reduce network overhead. Chen et al. [15] optimize path selection by selecting next-hop nodes through an annular area around the target, combined with the general angle towards the sink. This approach establishes routes that are not only short but also sufficiently diverse.

Network Coding To protect against a global adversary, messages are divided into smaller segments and transmitted through multiple routes to the sink, ensuring a highly effective method Fan et al. [25]. Nodes combine parts from different messages and forward them collectively. The division of messages provides clear instructions on how to reassemble the message at the sink. The adversary’s ability to trace the message back through traffic analysis is directly linked to their ability to obtain this information. To avoid this problem, simply encrypting the aggregation information is the recommended solution. It should be noted that intermediate nodes will need to modify this information during transmission. The proposed scheme leverages homomorphic encryption to prevent intermediary nodes from accessing and adding their own aggregation information. This ensures that only the sink can decrypt the information. Another work by the same authors extends this approach with fake traffic to further exacerbate the adversary’s efforts [24]. In this case, the artificially generated message components are absorbed to increase efficiency. If any intermediary node processes the fake message, the homomorphic function will filter out any non-real parts. This approach effectively reduces the induced network overhead, without requiring the knowledge of any intermediary node.

Avoid Adversary To ensure network security, it is crucial to guard against internal adversaries. These adversaries may control specific nodes in the network, often by compromising them. The approach proposed by Dutta et al. [22] assumes that the adversary can be detected and geolocated by the network, thus enabling the protection of the network against this adversary. The network is divided into a grid. When any node detects an adversary within a grid, all benign nodes in the grid will immediately stop communicating until the adversary has left. To detect an arriving or leaving compromised node in a grid, the special grid nodes will detect the adversary’s movement and immediately inform the respective grid to deactivate or activate. Rios et al. [84] replaces the grid with a hop-based distance measurement. Benign nodes can detect adversarial nodes as they cannot authenticate themselves. The detecting node broadcasts a message to inform the rest of the
network, incrementing a counter for each hop of the message. This gives an estimate of the adversary’s location and allows the other benign nodes to circumvent the adversary when choosing a route for their packets. Nodes near to the adversary will buffer messages until the adversary has departed.

**Artificial Delay** According to Hong et al. [32], encrypting and padding all messages can make them indistinguishable from one another. In addition, they suggest the addition of a random delay before the forwarding of messages. This prevents a local adversary to trace messages back over multiple hops, even if they arrive out of order. The same work also suggests an extension to handle networks with low traffic. In this case, if a node receives only one message for a while, it re-triggers the random delay for the message until a set maximum of delay periods or another message arrives. The work by Kamat et al. [40] focuses on analyzing the effects of different distributions and how the intermediary nodes handle the buffering of messages. To avoid analyzing the chosen distribution of the delays, message delays can be mixed Kamat et al. [41]. For this, intermediary nodes can adopt strategies in which they skip waiting for the delay, e.g., just sending the message with the longest delay contained in the node’s buffer.

**Pseudonymity** One of the first works that protect against a global adversary via pseudonyms is by Misra et al. [72]. In this approach, all nodes have a shared secret key pair with their neighbors. Additionally, nodes form clusters, which select individual nodes as cluster heads. Each cluster establishes a cluster key and all traffic is forwarded via the cluster head. To protect against the adversary, the paper proposes two solutions. One leverages a global pseudonym space, in which every node gets its sub-space. Nodes then randomly select a pseudonym and the sink has an attribution table to know which node belongs to the used pseudonym. Forwarding nodes also keep a table that indicates the direction of nodes for certain pseudonym ranges. The other solution aims to reduce the storage overhead of using large pseudonym attribution tables. Instead of doing a lookup in a table, nodes use a keyed hash function that directly generates the pseudonym. The key for this hash function is a shared seed among the nodes. The work by Ouyang et al. [76] proposes two similar strategies to the previous ones. One strategy extends the keyed hashing for resolving pseudonyms to protect against an internal attacker. Briefly, any node will create two hashes, one for communicating with the sink and one for communicating with its next neighbor. The alternative strategy generates a complete hash chain to transmit to the sink, which effectively reverses the order of the other strategy. Another work leverages three mechanisms to achieve anonymous routing Sheu et al. [90]. One mechanism is an anonymous one-hop communication based on encryption keys established with the node’s neighbors. The second mechanism enables multi-hop routing, by using a pseudonym next hop routing table. Lastly, the last two mechanisms are combined to get a per-hop encryption along the entire route. The work of Di Pietro et al. [18] organizes the network in a tree-like structure with a sink as the root. First, the sink and all nodes will send random values down the tree for i rounds to create shared key values of i bits, which are used for per-hop encryption. We omitted the works specifically focusing on protecting the sink, as the location of the sink will be revealed through triangulation anyway.

### 5.2 Analysis

This section compares the individual approaches, highlighting their drawbacks and applicability to location privacy for ground-based satellite network users. The analysis is predominantly conducted under the key factors of latency, computational, and communication overheads as displayed in Table 2. This is because location privacy in network approaches is often a trade-off between throughput and privacy guarantees. Due to missing implementation, open-source code, or even evaluation of the specific works, the overheads were approximated. The work on Anonymous Satellite Internet [49] directly addresses the issue of location data protection in the specific context of satellite services and provides.

However, at the same time, a customized approach may offer several advantages over the presented scheme. Therefore, in the following, we provide an analysis on proposed mechanisms.

**Physical Security** may be able to inherently provide location privacy. However, they are designed for confidentiality and not for location privacy of high bandwidth services. Manipulating the SNR in this way will impair the data rates. Additionally, they assume the knowledge of all Channel State Information (CSI), which may be impractical.

Due to the random (and flooding) nature of Phantom Routing approaches, they incur large communication overheads. The random walk may also direct the message unfavorably, causing further delays. In contrast, other works rely on complete knowledge of the network topology to avoid this issue. In the highly dynamic environment of a space network, this is non-trivial to achieve and to consider for routing.

**Fake Traffic** approaches provide strong location privacy guarantees against an adversary with a global view. However, generating fake traffic across the entire network results in insignificant communication overheads. Additionally, these approaches also employ a form of aggregation with regular release, which also implies delays. Some of the works propose an efficient method for generating traffic that does not require knowledge of when a benign node is sending, ensuring correct timing of the generation (e.g., [89]). This is quite impractical and if this information leaks to an adversary, it could be abused to defeat the entire approach by exactly knowing the timing of the benign transmissions.

**Ring Routing** assumes a set of powerful nodes that establish the ring, which implies large bottleneck overheads in terms of communication and might not be achievable in many scenarios. Also, the one-way routing along the ring may imply
significantly longer routes, leading to delays.

**Multi-Path Routing** implies communication overhead proportional to the number of distinct routes. This approach also requires in-depth knowledge about the topology of the network to properly and efficiently set up the multiple routes.

**Network Coding** is a powerful approach against a global adversary. The scheme’s computational intensity results from homomorphic encryption at every hop. This is especially pronounced when numerous parties are communicating, as the holistic overhead is significant. Also, the setup of pre-shared secrets is inadequately addressed.

**Avoid Adversary** is effective if the network knows which nodes are compromised and their locations. One effective approach is to assume the presence of an adversary detection scheme that enables all nodes to validate each other and accurately identify any malicious nodes. This assumption is based on the fact that without such schemes, it is not feasible to ensure the security of the network.

**Artificial Delays** cause significant latency, detrimental in many use-cases like satellite Internet. Incorrectly received messages can cause further complications.

**Pseudonymity** While this approach protects against a global adversary, there are two caveats to accomplish this. Firstly, these approaches rely on a bootstrapping phase to set up shared secrets, which in some cases requires numerous all-to-all communication rounds. This may be impractical, especially for large networks. Secondly, the pseudonymity of an individual node depends on how many other nodes use the mechanism and if they are close to the node. Otherwise, it will be easy for an adversary to identify the node despite the pseudonym, if there are no other messages in the network.

**Tor network** As stated earlier, the Tor network [19] is not an applicable location privacy approach in satellite internet networks. In addition, satellite communications can be monitored by any satellite and, more concerning, can be triangulated to pinpoint the user’s location. For instance, many attacks on Tor rely on an adversary conducting entry-point monitoring [74,102]. While this is usually a challenging position for the adversary, it becomes much simpler with satellite-based Internet since the connection first goes to the satellite before reaching the Tor network. Additionally, some attacks on Tor assume the adversary can monitor both the entry and exit points [3,51,78]. However, as described in Section 5.3 onion routing could be utilized in future research.

### 5.3 Open Research Challenges

This section will outline some open research challenges we identified. As space networks are not commonly addressed in most works, numerous research questions remain unanswered in this unique environment.

**Physical Security** While we covered physical security as a potential way to achieve location privacy, the respective works were not designed with this goal in mind. Thus, we think there is an opportunity to optimize the general idea of this approach to either conceal the ground-based user’s location or maybe even prevent the triangulation entirely.

**Compromised Nodes** Internal adversaries are often overlooked in many works, despite their significant impact on network security. These attackers have control over a node within the network, usually as a result of a compromise. Thus, we think this problem can be addressed more explicitly and in combination with other approaches. Moreover, it is imperative to analyze the aspect of adversary detection schemes more practically, including scenarios of imperfect detection or dissemination of information about a detected compromise.

**Optimized Fake Traffic** There are opportunities to optimize fake traffic methods for ensuring location privacy by distributing generated traffic. The predictable and regular orbits of typical space networks can be leveraged to design an effective fake traffic generation strategy tailored for space networks.

**Onion Routing** One of the practical ways to achieve anonymity on the Internet is onion routing, e.g., as used by the Tor network [19]. As an overlay network, it can be utilized to provide better location privacy guarantees for satellite service users. Particularly, it can prevent an internal attacker from accessing the routes of specific messages, thereby ensuring location privacy protection.

### 6 Conclusion

In this SoK, we look into two important security areas that are crucial for robust and secure space networks, yet haven’t been explored enough in satellite network literature. Firstly, as collaborative efforts among satellites from diverse and distrusting entities gain popularity, setting up PKI becomes important for secure communication. However, we find that checking certificate revocations in space networks is hard due to long delays and disruptions. Secondly, with more people using satellite internet and other services, there’s concern about potential disclosure of users’ locations. This is a noteworthy consideration in safeguarding user privacy and security. Throughout this paper, we analyze existing research in these fields and point out the challenges that need more attention. Our examination revealed numerous open research challenges, emphasizing the need for further investigation and innovation in these domains.

### References


