VIKING

(Vital Infrastructure, NetworKs, INformation and Control System ManaGement)

REPORT D4.2 Robust and Secure Communication Protocols

PROJECT TITLE:	Vital Infrastructure, Networks, Information and Control Systems Management			
PROJECT ACRONYM:	VIKING			
GRANT AGREEMENT NUMBER:	225643			
PROJECT START DATE:	01.11.2008			
DURATION:	36 MONTHS			
PROJECT CO-ORDINATOR:	GUNNAR BJÖRKMAN, ABB		(ABB)	DE
PROJECT MEMBERS:	ABB	(1)	(ABB)	DE
	EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZÜRICH	(2)	(ETHZ)	CH
	E.ON	(4)	(E.ON)	DE
	ASTRON INFORMATICS LTD.	(5)	(Astron)	HU
	KUNGLIGA TEKNISKA HÖGSKOLAN	(6)	(KTH)	SE
	UNIVERSITY OF MARYLAND FOUNDATION	(8)	(USMF)	US
	MML ANALYS & STRATEGI	(8)	(MML)	SE

DOCUMENT IDENTIFIER:	D.4.2
ISSUE:	1.0
ISSUE DATE:	2011-10-31
PREPARED:	LABORATORY FOR COMMUNICATION NETWORKS, KTH
APPROVED	
DISSEMINATION STATUS:	PUBLIC

Project co-funded by the European Community within the 7th Framework Programme.

History Chart

Issue	Date	Changed Page (s)	Cause of Change	Implemented by
Original	2011.10.31			

Authorization

No.	Action	Name	Signature	Date
1	Prepared	György Dán Ognjen Vukovic Gunnar Karlsson Henrik Sandberg Kin Cheong Sou Gunnar Björkman Mathias Ekstedt		
2	Approved			

The information in this document is subject to change without notice.

Distribution List

This document is distributed as below.

Additional copies held by unnamed recipients will not be updated.

Electronic Copy Number	Name	Address
1	Cristian Olimid	EU, Brussels
2	Gunnar Björkman	ABB AG, Mannheim
3	Pontus Johnson	KTH, Stockholm
4 - 9	VIKING consortium members	

Contents

List of Abbreviations

BDD	Bad Data Detector
BIM	Bayesian Inference Method
BITW	Bump in the Wire
ICCC	Inter-control Center Communications
ICCP	Inter-control Center Communication Protocol
IP	Internet Protocol
MILP	Mixed Integer Linear Program
MPLS	Multiprotocol Label Switching
MPM	Maximum Posteriory Method
OPGW	Optical Ground Wire
PMU	Phasor Measurement Unit
RTU	Remote Terminal Unit
S2CC	Substation to Control Center Communications
SCADA	Supervisory Control and Data Acquisition
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Networking
TCP	Transmission Control Protocol
TLS	Transport Layer Security
TSO	Transmission System Operator
WAMS	Wide Area Monitoring System
WAN	Wide Area Network

0 Executive summary

The objective of WP4 in the VIKING project is to develop methodologies for the mitigation and protection of critical infrastructures. The focus is on threat detection in SCADA systems for power networks, and on attack mitigation through improved resource management, communication protocols and algorithms. WP4 builds upon the requirement studies of WP1, the models developed in WP2, and the threats described in WP3. This deliverable reports the work done in Task 4.2. The focus of Task 4.2 is on mitigating attacks against the communication infrastructure and the communication protocols used in the Supervisory Control and Data Acquisition (SCADA) systems in power systems. The two main areas of SCADA communications are substation to control-center and control-center to control-center communications. This deliverable describes contributions related to these two areas of SCADA communications.

The first contribution considers attack mitigation schemes for substation to control-center communications applied to the problem of power system state estimation. Based on realistic models of the communication infrastructure used to deliver measurement data from the substations to the state estimator, it is shown that the vulnerability of the power system state estimator is highly dependent on the communication infrastructure. To quantify the system's vulnerability two security metrics are introduced: the importance of individual substations and the cost of attacking individual measurements. Using the metrics it is shown how various network layer and application layer mitigation strategies, like single and multi-path routing and data authentication, can be used to decrease the vulnerability of the state estimator. The efficiency of the algorithms is illustrated on the IEEE 118 and 300 bus benchmark power systems.

The second contribution considers the problem of improving the availability of control-center to control-center communications by hiding the communication patterns from potential attackers. Hiding the communication patterns is achieved using relaying, which is often used in networked systems to achieve anonymous communication. Relaying introduces communication overhead and increased end-to-end message delivery delay, but in practice overhead and delay must often be kept reasonably low. To understand how to optimize anonymity for limited overhead and delay, the trade-off is analyzed between relationship anonymity and communication overhead under passive traffic analysis attacks. Two passive traffic analysis attacks and two anonymity networks are considered. The results show that, contrary to intuition, increased overhead does not always improve anonymity. For given system size and number of attackers the characteristics of the optimal overhead are studied. It is then shown that if the number of attackers is unknown it is best to optimize for a higher than expected number of attackers to make the system robust.

1 Introduction

The VIKING project is an EU financed Framework 7 Collaborative STREP Project and is part of themes 4, ICT, and 10, Security. VIKING stands for Vital Infrastructure, NetworKs, INformation and Control Systems ManaGement and will be executed between November 1, 2008 and October 31, 2011 by a consortium of industrial and academic partners.

1.1 SCADA Security and Power Transmission and Distribution

Society relies heavily on the proper operation of the electric power system. Many of the critical infrastructures could be able to operate without power for shorter periods of time, but longer power outages could have devastating economical and humanitarian consequences. The importance of the electric power system is illustrated, for instance, by the economic and social impacts of the 2003 North-East American blackout [1], which affected approximately 50 million people.

In order ensure efficient and reliable operation, large electric power systems are equipped with an IT infrastructure that allows for system-wide supervision and control. These industrial control systems are referred to as Supervisory Control And Data Acquisition (SCADA) systems. SCADA system collect measurement data from Remote Terminal Units (RTUs) located at the substations via a SCADA network, and typically store the measured data at one or more control centers. The data are processed at the control center, and serve as the basis for automatic or human decisions, which results in control commands sent back to the RTUs at the substations. SCADA systems greatly improve the coordination, monitoring, and utilization of resources in the power system. But by relying on the SCADA infrastructure, the power system becomes susceptible not only to operational errors but also to cyber-attacks.

Many reports discuss the vulnerabilities of SCADA systems to cyber attacks [2, 3, 4, 5, 6], but real incidents [7, 8, 9] also confirm the importance of this issue. [10] proposed a framework in order to clarify the interaction between the power system and the IT infrastructure and to identify the vulnerabilities and the malfunctions that could lead to an abnormal operation of the power network. Taking a different perspective, the authors of [11] attempted to quantify the impact of a cyber attack on the power market. In [12], the robustness of the US power network against cascading failures was tested.

1.2 VIKING Project

The VIKING research project [13] proposes a novel concept to address the challenges introduced by the interaction between the IT systems and the transmission and distribution systems. The main objective is to develop, test, and evaluate methodologies for the analysis, design and operation of resilient and secure industrial control systems for critical infrastructures. Specifically, the aim is to increase the understanding of vulnerabilities of integrated control systems, determine the impact on the electric power transmission and distribution system due to possible failures or attacks and develop strategies to eliminate or to mitigate these effects. In order to achieve this goal, VIKING has formulated four strategic objectives described below.

- Provide a holistic framework for identification and assessment of vulnerabilities for SCADA systems. This framework should provide computational support for the prediction of system failure impacts and security risks.
- Provide a reference model of potential consequences of misbehaving control systems in the power transmission and distribution network that can be used as a base for evaluating control system design solutions.
- Develop and demonstrate new technical security and robustness solutions able to meet the specific operational requirements that are posed on control systems for our target area.
- Increase the awareness of the dependencies and vulnerabilities of cyber-physical systems in the power industry.

These objectives are addressed by the following work packages:
WP1: Requirements study
The safety requirements imposed on the power transmission and distribution systems will be determined. This study will address the issue from three different perspectives, reflecting the three key players involved in the process: 1. The physical power transmission and distribution processes, 2. The IT infrastructure monitoring and regulating these processes, and 3. The users and operators that interact with the system. The results will serve as guidelines against which the performance of the risk assessment and mitigation methods developed in subsequent work packages will be evaluated.
WP2: Modeling
WP2 will provide the ingredients and modeling support for the risk assessment and mitigation studies. In particular, WP2 will: 1. Identify the system's intrinsic vulnerabilities and external threats, and 2. Develop models to capture the architecture of the IT infrastructure and the coupling of the physical power transmission

and distribution system and the IT infrastructure supervising it.

WP3: Risk assessment and evaluation methodologies

Methodologies for determining the effect of particular vulnerabilities and threats on the overall power transmission and distribution systems will be developed. The architecture and physical-IT models (developed in WP2) will be used to assess the impact that specific threats and system failures (identified in WP2) can have on the overall system. The results will then be coupled to models for estimating the consequences and costs of particular disruptions of service. This will allow us to determine the overall impact of threats and failures on the system, both in technical and in economic terms.

WP4: Mitigation and protection

WP4 will develop methodologies for the mitigation of the effects determined by WP3, and will be based on the modeling work of WP2. It will also draw on the WP3 studies, to identify the threats and vulnerabilities which can have the greatest potential impact on the system. WP4 will then develop strategies and countermeasures for reducing the risk of the most critical threats and vulnerabilities and/or ameliorating their impact.

WP5: Case study and test-bed

Proof of concept case studies to realistic size problems will be provided by the end users. The tools and methodologies developed in WP2-WP4 will be deployed on these case studies and their performance in terms of the requirements identified in WP1 will be established. Case studies will be evaluated through a test-bed consisting of a SCADA system integrated with a computer simulated physical infrastructure and communication networks.

1.3 Objective of WP4

The objective of WP4 is to develop methodologies for the mitigation and protection of critical infrastructures. The focus is on development of new methods for anomaly and threat detection and countermeasures through secure resource management and communication protocols. WP4 builds upon the requirement studies of WP1, the models developed in WP2, and threats found in WP3. These models are used as the foundations on which the threat assessment and mitigation methodologies in WP4 can be based. The work package of WP4 has three main tasks.

Task 4.1: Anomaly monitoring and resilient resource management

SCADA systems handle heterogeneous and complex data traffic. In this task, we study anomalies the adversaries can generate in this traffic, and how they can be detected using online monitoring. We also consider how resource management in the presence of anomalies can be improved by introducing network architectures capable of providing quality of service. How these problems have been addressed

are reported in this deliverable, as outlined below in Section 1.4.

Task 4.2: Robust and secure communication protocols

Critical infrastructures increasingly rely on data transmission over public telecommunication networks and on open communication standards, e.g., networking protocols. Networking protocols have extensions that make them resilient to traditional security attacks (such as IPSec for data integrity, identification, non-repudiation, man-in-the middle attack, etc), but a mission critical system can be rendered useless by other means as well. We will here investigate how mission-critical applications can be attacked in non-intrusive ways, and develop improved protocols that are resilient to non-intrusive attacks, but at the same time are fair to standard protocols.

Task 4.3: Mitigation and protection against threats

The final step in the process will be to develop mitigation strategies against threats that exploit the coupling between the IT infrastructure and the physical power transmission and distribution system. Based on the results of Task 3.2 we will first prioritize the threats and failures identified in Task 2.1 according to their potential impact on the system. We will then develop strategies to reduce the effects of the most prominent threats and failures.

1.4 Objectives of D4.2

The objective of D4.2 is to report the work done in Task 4.2, and thereby concludes the task. The work done and the approaches taken are summarized in the following.

Inputs from other tasks

- WP1 Attacks that corrupt measured data were listed in the requirements study in WP1. See, for example, RT-0002: Access to process communication channels, and RT-0002: Attack on ICCP connections, in [14].
- WP2 Simplified versions of the steady-state model of power systems obtained in WP2 are used here.
- WP3 In the impact analysis being carried out in Task 3.2, the stealthy false-data attacks have been identified as potentially important.
- WP4 Based on the anomaly detection carried out in Task 4.1, the mitigation against stealthy false-data attacks requires a decrease of the attack surface.

Output to other tasks

- WP3 The physical impact (Task 3.2) and society impact (Task 3.3) of the false-data attacks will be evaluated in WP3.

Related publications The general problem of securing power system IT infrastructures, and its relationship to system performance and reliability were discussed in [15]. The use of different networ-layer schemes and authentication mechanisms to mitigate attacks against the power system state estimator was investigated in [16]. The problem of improving availability through anonymous communications was considered in [17].

1.5 Document structure

The document is organized in four main sections. Section 1 provides information about the aim of the deliverable. Section 2 gives an overview of SCADA communications and the outlines the challenges in deploying security solutions. Section 3 describes a framework to assess the vulnerability of the state estimator against attacks as a function of the communication network topology, and discussess various network layer and application layer mitigation schemes. Section 4 investigates the use of anonymity networks to increase communication availability for the purpose of control center to control center communications.

2 SCADA Communications and Security

At the heart of the IT infrastructure for power system control and operation there are one or several Supervisory Control And Data Acquisition (SCADA) systems. Apart from the remote collection of vast amounts of real-time process measurements taken from the grid, e.g., in transformer stations, SCADA systems include functions for the remote control of process devices like breakers and tap changers. The acquired data are presented to the operators in the control center via an advanced graphical user interface, among others equipped with alarming features to alert the operators to changing operating conditions. Many SCADA systems include computerized models of the supervised process (i.e., the power system). The models enable simulation of alternative process states parallel to the physical process, which can be used for optimization and contingency analysis.

Reliability and performance have traditionally been the key design goals for the IT infrastructure used in power systems. A certain level of security was maintained by keeping the power system IT infrastructure isolated, and by using proprietary communication protocols. Power system IT infrastructures are, however, increasingly integrated with other IT infrastructures at the power utilities, including public infrastructures. At the same time, the proprietary protocols are being replaced by standardized communication protocols and interfaces to ensure interoperability between components from different vendors. The standardization of power system models, like the Common Information Model (CIM), is ongoing with the goal to ease the exchange of engineering data between and within utilities. These recent trends increase the exposure of the power system IT infrastructure to attacks, and are strong drivers towards using cryptographic protocols to secure power system communications. Nevertheless, a number of criteria have to be taken into account when designing secure communication solutions for power systems.

First, the communication and IT infrastructure of power systems have to satisfy very diverse application requirements. At one extreme, in the case of management information exchanged between utilities (e.g., control centers), data is transferred in batches with very loose delay constraints, and standard cryptographic protocols like TLS [18] can be used to provide authentication and confidentiality. At the other extreme, in the case of substation automation and inter-substation protection, the communication delays must be kept in the order of a few milliseconds, so that the delay introduced by encryption algorithms can already be critical for proper system operation. Thus, security solutions might have to be tailor-made for specific application scenarios.

Second, the power system's communication and IT infrastructure already consist of a vast number of components. The cost of securing the tens of thousands of components of a continent-wide infrastructure can be prohibitive, and therefore

Application			Performance - Typical values			Security - Importance		
			Distance [km]	Latency [s]	Throughput [1/sec]	Confidentiality	Integrity Authenticity	Availability
S2CC	SCADA	Data acquisition	1000	$1-10$	5000 meas	Medium	High	High
		Commands		1	0.1 command	High	High	High
		Alarms and events		1	500 events	Medium	High	High
	PMU data for WAMS			2×10^{-2}	18 meas/PMU	Medium	High	High
Substation automation (Intra-substation)			0.5	1	200 meas	Medium	High	High
Line protection (Inter-substation)			50	10^{-3}	2 meas	Low	High	High

Table 1: Approximate performance and security requirements of substation to control-center (S2CC), intra-substation and inter-substation communication for various applications. High level of integrity and availability has to be provided while satisfying very diverse performance requirements.

it is important to understand how the security of individual system components contributes to and affects the secure operation of the power grid. Also, in addition to the traditional IT and communication infrastructure security solutions and practices, in a cyber-physical system models of the physical process can often be leveraged to improve system security.

Achieving security in slowly evolving power system control and operation systems is a complex problem. Simply adding state-of-the-art security solutions and mechanisms to existing systems is often not feasible: security solutions can violate requirements on performance and reliability, which continue to have highest priority. Some security solutions would probably meet the requirements if completely new systems and architectures were deployed, but today as well as in the future we have to live with a large share of legacy equipment. Thus, in practice the challenge of security design in power system control and operation systems implies finding a proper level of trade-off between security, system properties like performance and reliability, and cost. Table 1 illustrates the heterogeneity of the performance and the security requirements of some power system applications.

3 Power System-aware Mitigation for RTU Communication

In general the SCADA communications related to power system state monitoring and estimation are referred to as substation to control center communication (S2CC). Nevertheless, the purpose of S2CC is not only to enable the monitoring of the state of the power system, but also to control the actuators located in the substations. In the following we give a brief overview of the communication protocols and security solutions available for S2CC, and then describe a framework to assess the vulnerability and to improve the security of the power system state estimator against attacks on the SCADA communication infrastrucrture.

3.1 Substation to Control Center Communications

Electric power transmission systems extend over large geographical areas, typically entire countries. The measurement data and status information taken at substations are collected by Remote Terminal Units (RTUs) located at the substations, and wide-area networks (WANs) are used to deliver the multiplexed measurement data from the RTUs to the control center of the transmission system operator (TSO).

Traditionally, S2CC was performed over low bitrate point-to-point transmission links, e.g., using power line communications, microwave, and leased lines. Point-to-point communication links are also present in modern S2CC, e.g., through the use of cellular and satellite communications. Fig. 1 shows a simple example of a power system and illustrates the corresponding S2CC communication infrastructure based on point-to-point communication links.

Nevertheless, modern WAN infrastructures are increasingly based on overhead ground wire (also called optical ground wire, OPGW) installations that run between the tops of the high voltage transmission towers or along underground cables. In the case of OPGW installations SONET or SDH is used to establish communication links (called virtual circuits) between the substations and the control center, but wide-area Ethernet is expected to become prevalent in the near future. As an effect the data sent from a remote substation to the control center might traverse several substations, where switches, multiplexers or cross connects multiplex the data from different substations onto a single OPGW link. The high capacity available in OPGW infrastructures also makes it possible to deliver voice, video and other data traffic multiplexed with the measurement and status information sent by the RTUs.

There is a large variety of application layer protocols for S2CC, and different protocols often coexist within a SCADA system. Legacy protocols (such as

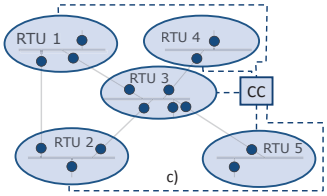

Figure 1: A simple power grid consisting of 5 buses. Each bar represents a substation (connected to a city distribution grid or a power plant, for instance) and the solid lines represent transmission lines. The circles indicate measurements of power flows, power injections, and voltages. There is an RTU at each substation, which collects the measurements taken at the substation and sends them over a point-to-point communication network (dashed lines) to the control-center.

Modbus, RP-570, and Profibus) and the proprietary protocols of equipment vendors, are slowly replaced by protocols standardized in the last decade, such as DNP3 and IEC 60870-5 for data acquisition and control, or IEEE C37.118-2005 for PMU data. Both the legacy and the standardized protocols were developed with performance and reliability in mind. To detect bit errors, they usually include an error detection code calculated by the RTU, which is sent along with the data. The error detection code is usually based on cyclic redundancy check (CRC).

Security extensions that provide confidentiality, integrity and authentication were standardized for these protocols recently, such as the IEC 62351-5 for IEC 60870-5. Communications with legacy equipment that does not support the security extensions can be secured using bump in the wire (BITW) solutions, like AGA-12 and YASIR [19]. A BITW solution consists of two devices that are inserted in the communication link near to the sender (the RTU) and the receiver (the control center), respectively. The sender side device encrypts and authenticates the output of the sender, which is then decrypted by the receiver side device. The sender and the receiver are not aware of the existence of the BITW devices. Nevertheless, authentication using a BITW device is not tamper-proof. If an attacker can access the communication link between the RTU and the BITW device, it can bypass the authentication and can modify all data. Newer RTUs are expected to contain tamper-proof authentication modules, and would hence be secure despite physical access.

Although security solutions are available for most communication protocols, deployment has been slow, not only because of the associated equipment costs, but also due to the overhead of managing encryption keys, and because of the potential impact of encryption on data availability in case of a lock out.

3.2 Power System State Estimation and Stealth Attacks

Measurements are taken and sent at a low frequency in SCADA systems, and therefore steady-state estimators are used for state estimation. We refer to [20, 21] for a complete treatment of state estimation, and provide a brief description in the following.

Consider a power system that has $n+1$ buses. We consider models of the active power flows P_{ij} (between bus i and j), active power injections P_i (at bus i), and bus phase angles δ_i, where $i, j = 1, \ldots, n+1$. (A negative P_i indicates a power load at bus i.) The state-estimation problem we consider consists of estimating n phase angles δ_i given M active power flow and injection measurement values z_m. One has to fix one (arbitrary) bus phase angle as reference angle, for example $\delta_0 := 0$, and therefore only n angles have to be estimated, i.e., the vector $\delta = (\delta_1, \delta_2, .., \delta_n)$. The active power flow measurements are denoted by $z = (z_1, \ldots, z_M)^T$, and are equal to the actual power flow plus independent random measurement noise e, which we assume has a Gaussian distribution of zero mean, $e = \left(e_1, \ldots, e_M\right)^T \in \mathcal{N}(0, R)$ where $R := \mathbf{E}ee^T$ is the diagonal measurement covariance matrix.

When the phase differences $\delta_i - \delta_j$ between the buses in the power system are all small, then a linear approximation, a so called DC power flow model, is accurate, and we can write

$$z = H\delta + e, \tag{1}$$

where $H \in \mathbb{R}^{M \times n}$ is a constant known Jacobian matrix that depends on the power system topology and the measurements, see [20, 21] for details. The state estimation problem can then be solved as

$$\hat{\delta} := (H^T R^{-1} H)^{-1} H^T R^{-1} z. \tag{2}$$

The phase-angle estimates $\hat{\delta}$ are used to estimate the active power flows by [21]

$$\hat{z} = H\hat{\delta} = H(H^T R^{-1} H)^{-1} H^T R^{-1} z. \tag{3}$$

The BDD system uses such estimates to identify faulty sensors and bad data by comparing the estimate $\hat{z}$ with z: if the elements $\hat{z}_m$ and z_m are very different, an alarm is triggered because the received measurement value z_m is not explained well by the model. For a more complete treatment of BDD we refer to [20, 21].

An attacker that wants to change measurement m (its *value* z_m) might have to change several other measurements m' to avoid a BDD alarm to be triggered. Consider that the attacker wants to change the measurements from z into $z_a := z + a$. The *attack vector* a is the corruption added to the real measurement vector z. As was shown in [22], an attack vector must satisfy

$$a = Hc, \quad \text{for some } c \in \mathbb{R}^n, \tag{4}$$

in order for it not to increase the risk of an alarm. The corresponding a is termed a *stealth attack* henceforth.

In the recent study [23] it was verified that, despite the simplifying assumptions, stealth attacks can be made large in real (nonlinear) SE software: in the example considered in [23], a power flow measurement was corrupted by 150 MW (57% of the nominal power flow) without triggering alarms.

3.3 Power System Communication Model

The $n+1$ buses of the power system are spread over a set of substations $\mathcal{S}$, $|\mathcal{S}| = S$. We denote the substation at which measurement m is taken by $S(m) \in \mathcal{S}$, and we denote the substation at which the control center is located by $s_{cc} \in \mathcal{S}$. We model the communication network by an undirected graph $\mathcal{G} = (\mathcal{S}, E)$; an edge between two substations corresponds to a communication link between the two substations (e.g., a point-to-point link from a substation to the control-center, or an OPGW link between two substations connected by a transmission line). The graph $\mathcal{G}$ is connected but is typically sparse. For each substation $s \in \mathcal{S}$ there is a set of established routes $\mathcal{R}_s = \{r_s^1, \ldots, r_s^{R(s)}\}$ from s to s_{cc} through $\mathcal{G}$. $\mathcal{R}$ denotes the collection of all $\mathcal{R}_s$. We represent a route by the set of substations it traverses including s itself and the control center s_{cc}, i.e., $r_s^i \subseteq \mathcal{S}$. The order in which the substations appear in the route is not relevant to the considered problem. If $R(s) = 1$ then all measurement data from substation s are sent over a single route to the control center. If $R(s) > 1$ then data is split equally among the routes such that unless the data sent over all routes get corrupted the control center can detect the data corruption using the error detection code.

We consider two forms of end-to-end authentication: non tamper-proof and tamper-proof. We denote the set of substations with *non tamper-proof* authentication (e.g., substations with a BITW device to *authenticate* the data sent to the control center, or an RTU with a non tamper-proof data authentication module) by $\mathcal{E}^N \subseteq \mathcal{S}$. For a route r_s^i we denote by $\sigma_{\mathcal{E}^N}(r_s^i)$ the set of substations in which the data is *susceptible* to attack despite non tamper-proof authentication. By definition $\sigma_{\mathcal{E}^N}(r_s^i) = \{s\}$ if $s \in \mathcal{E}^N$ and $\sigma_{\mathcal{E}^N}(r_s^i) = r_s^i$ otherwise, that is, non tamper-proof authenticated data can be modified at the substation where it originates from, if physical access is possible.

Similarly, we denote the set of substations with *tamper-proof* authentication (e.g., substations with a tamper-proof RTU that *authenticates* the data sent to the control center) by $\mathcal{E}^P \subseteq \mathcal{S}$. Data authenticated in a tamper-proof way is not susceptible to attack at any substation on the route, hence $\sigma_{\mathcal{E}^P}(r_s^i) = \emptyset$ for every route r_s^i.

Finally, a substation can be *protected* against attacks, e.g., by guards, video

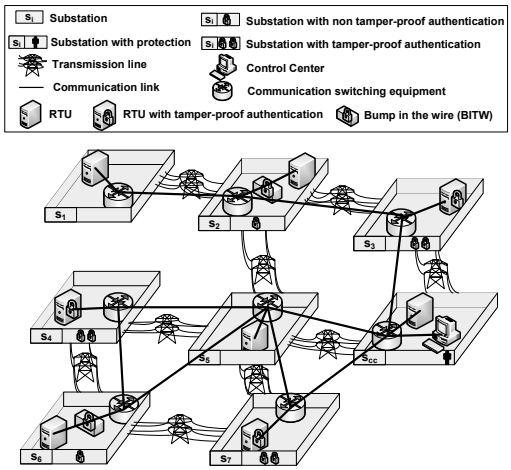

Figure 2: A simple example of a power system and its communication infrastructure. We have $\mathcal{E}^N = \{s_2, s_6\}$, $\mathcal{E}^P = \{s_3, s_4, s_7\}$, and $\mathcal{P} = \{s_{cc}\}$. A measurement taken at substation $s_1 \notin \mathcal{E}^P \cup \mathcal{E}^N$ is susceptible to attacks at substations s_1, s_2, and s_3. A measurement taken at substation $s_6 \in \mathcal{E}^N$ is only susceptible to attacks at substation s_6 ($\sigma_{\mathcal{E}^N}(r^1_{s_6}) = \{s_6\}$). A measurement taken at substation $s_4 \in \mathcal{E}^P$ is not susceptible to attacks ($\sigma_{\mathcal{E}^P}(r^1_{s_4}) = \emptyset$).

surveillance or using tamper-proof system components. We denote the set of protected substations by $\mathcal{P} \subseteq \mathcal{S}$. Protected substations are not susceptible to attacks, therefore $\sigma_{\mathcal{P}}(r^i_s) = r^i_s \setminus \mathcal{P}$. We assume that the substation where the control center is located is protected, that is, $s_{cc} \in \mathcal{P}$.

Fig. 2 illustrates a simple power system and its OPGW communication infrastructure. Some substations have applied mitigation schemes, such as non tamper-proof authentication, tamper-proof authentication, and protection.

3.4 Attack model and security metrics

We consider an attacker whose goal is to perform a *stealth attack* on some power flow or power injection measurement m. To perform the stealth attack, the attacker has to manipulate measurement data from several measurements to avoid a BDD alarm. To manipulate measurement data the attacker gets access to the communication equipment located at a subset of the substations. For example, the attacker could get physical access to the equipment in an unmanned substation or could remotely exploit the improper access configuration of the communication equipment. By gaining access to a substation $s \in \mathcal{S}$ (i.e., the switching equipment and the RTU) the attacker can potentially manipulate the measurement data that are *measured in* substation s and the data that are *routed through* substation s, un-

less multi-path routing, data authentication or protection make that impossible. To perform a *stealth attack* on a particular measurement m (its value z_m) the attacker might need to attack several substations simultaneously, which increases the cost of performing the attack.

In the following we propose two security metrics to characterize the vulnerability of the system with respect to the importance of individual substations and with respect to the vulnerability of individual measurements. Both metrics depend on the mitigation measures implemented by the operator. We also propose an approximation for each metric based on the communication graph topology.

3.4.1 Substation Attack Impact (I_s)

We quantify the importance of substation s by its *attack impact* I_s, which is the number of measurements on which an attacker can perform a *stealth* attack by getting access to a *single* substation s.

By definition $I_s = 0$ if the substation is protected ($s \in \mathcal{P}$). Otherwise, we define I_s as follows. A measurement m can be attacked if and only if the susceptible parts of all routes from $S(m)$ to the control center pass through substation s. Let us denote by $\mathcal{M}_s \subset \{1,\dots,M\}$ the index set of all such attackable measurements. Then measurement $m \in \mathcal{M}_s$ can be *stealthily* attacked if and only if the following system of equations has a solution with respect to unknowns $a \in \mathbb{R}^M$ and $c \in \mathbb{R}^n$

$$a = Hc, \quad a(m') = 0, \ \forall\, m' \notin \mathcal{M}_s, \quad \text{and } \ a(m) = 1. \tag{5}$$

The attack impact I_s is then the cardinality of the set of measurements for which (5) has a solution. That is,

$$I_s = \big|\, \{ m \mid \exists\, a \ \text{ satisfying (5)} \} \,\big|. \tag{6}$$

The attack impact of a substation depends on the routing $\mathcal{R}$, the set $\mathcal{E}^N$ of substations with non tamper-proof authentication, the set $\mathcal{E}^P$ of substations with tamper-proof authentication, and the set $\mathcal{P}$ of protected substations. I_s can be calculated with complexity $O(M^3)$, as shown in [17].

3.4.2 Measurement Attack Cost (Γ_m)

We quantify the vulnerability of measurement m by the minimum number of substations that have to be attacked in order to perform a stealth attack against the measurement, and denote it by Γ_m. If the substation at which the measurement is located is protected and uses non tamper-proof authentication ($S(m) \in \mathcal{P} \cap \mathcal{E}^N$), or it uses tamper-proof authentication ($S(m) \in \mathcal{E}^P$) then the measurement is not vulnerable and we define $\Gamma_m = \infty$.

Otherwise, for a measurement m we define Γ_m as the cardinality of the smallest set of substations $\omega \subseteq \mathcal{S}$ such that there is a stealth attack against m involving some measurements m' at substations $S(m')$ such that every route of the substations $S(m')$ involved in the stealth attack is susceptible to attack at least in one substation in ω. That is,

$$\begin{aligned} \Gamma_m = & \min_{\omega \subseteq \mathcal{S}; \omega \bigcap \mathcal{P} = \emptyset} |\omega| \;\; s.t. \;\; \exists \, a, c \;\; s.t. \;\; a = Hc, \;\; a(m) = 1 \text{ and} \\ & a(m') \neq 0 \implies \omega \bigcap \sigma_{\mathcal{E}}(r^i_{S(m')}) \neq \emptyset, \quad \forall \, r^i_{S(m')} \in \mathcal{R}_{S(m')}, \end{aligned} \tag{7}$$

where $\sigma_{\mathcal{E}}(r^i_{S(m')})$ denotes the substations in route $r^i_{S(m')}$ that are susceptible to attack despite the authentication applied at substation $S(m')$, i.e., $\sigma_{\mathcal{E}}(r^i_{S(m')}) = \sigma_{\mathcal{E}^P}(r^i_{S(m')}) \bigcap \sigma_{\mathcal{E}^N}(r^i_{S(m')})$.

The attack cost of a measurement depends on the routing $\mathcal{R}$, the set $\mathcal{E}^N$ of substations using non tamper-proof authentication, the set $\mathcal{E}^P$ of substations using tamper-proof authentication, and the set $\mathcal{P}$ of protected substations. The calculation of the attack cost Γ_m can be formulated as a mixed integer linear program (MILP), as shown in [16].

The following proposition establishes a relationship between the substation attack impact and the measurement attack cost.

Proposition 1. $I_s = 0 \; \forall s \in \mathcal{S} \iff \min_m \Gamma_m > 1$.

Proof. Follows directly from the definitions (6) and (7). If $\nexists s \; I_s > 0$ then a stealth attack against any measurement requires at least two substations to be attacked, $\Gamma_m \geq 2$. If $\exists s \; I_s > 0$ then attacking substation s is sufficient to attack some measurement m and hence $\Gamma_m = 1$. □

3.4.3 Numerical results

In the following we show numerical results obtained using the algorithms for two IEEE benchmark power systems: the IEEE 118 and 300 bus power systems. We use these IEEE systems for illustration because their size allows us to show a richer set of phenomena than using a 40 bus power network. Measurements are assumed to be taken at every power injection and power flow.

Network topologies Let us consider two communication network topologies. In the first topology every substation communicates directly to the control center, hence the communication network graph is a star graph of order $|S| + 1$: the control center has degree $|S|$ and all substations have degree 1. We refer to this communication network graph as the *star topology*. In the second topology there

is an edge between two substations s and s' in the communication network graph if there is a transmission line between any two buses in substations s and s'. The control center is located adjacent to the substation with highest degree s_{cc}. We refer to this communication network graph as the *mesh topology*.

Topology vs. Vulnerability without Mitigation We start with considering a baseline scenario, when no mitigation is used. Authentication is not used at any substation ($\mathcal{E}^N = \emptyset$, $\mathcal{E}^P = \emptyset$). For the mesh topology we consider that all substations use a *single shortest path* ($|\mathcal{R}_s| = 1$) to the control center s_{cc}, and the substation to which the control center is adjacent is protected ($\mathcal{P} = \{s_{cc}\}$). In the following we show the attack impact and the measurement attack cost for the star and for the mesh communication network topologies.

For the mesh topology Fig. 3 shows the attack impact I_s for the substations for which $I_s > 0$. The results show that there are several substations that would enable an attacker to perform a *stealth* attack on a significant fraction of the measurements in the power system, e.g., on about 1000 measurements for the 300 bus system (approx. 90% of all measurements). Almost 50% of the substations have non-zero attack impact, and the attack impact decreases slower than exponentially with the rank of the substation.

Table 2 shows the measurement attack costs for the star and the mesh topologies. For the star topology and the 118 bus power system there are no measurements with attack cost 1, and most of the measurements (more than 90%) have the attack cost of at least 3. Interestingly, for the 300 bus power system the attack costs are significantly lower. Almost 20% of the measurements have attack cost 1 and only around 45% of the measurements have an attack cost of at least 3. The reason is that in the 300 bus power system topology there are more substations with several buses, and an attacker can tamper with more measurements by accessing such substations.

The measurement attack costs for the mesh topology are significantly lower than those for the star topology; e.g., for the 118 bus power system more than 75% of the measurements have attack cost 1 for the mesh topology, while none for the star topology. The significant difference in terms of the attack costs shows the importance of considering the communication network topology when estimating the system security.

Motivated by the large substation attack impacts and low measurement attack costs in the case of shortest path routing, in the following we investigate how the operator can improve the system security by changing single-path routes, using multi-path routing, authentication and protection.

Table 2: Number of Measurements with Particular Measurement Attack Cost for the IEEE 118 and IEEE 300 systems

System	Topology	1	2	3	4	5	6
IEEE118	Star (Γ_m)	0	47	279	71	32	26
	Mesh (Γ_m)	374	78	11	0	0	0
IEEE300	Star (Γ_m)	209	251	378	188	41	2
	Mesh (Γ_m)	975	89	3	6	0	0

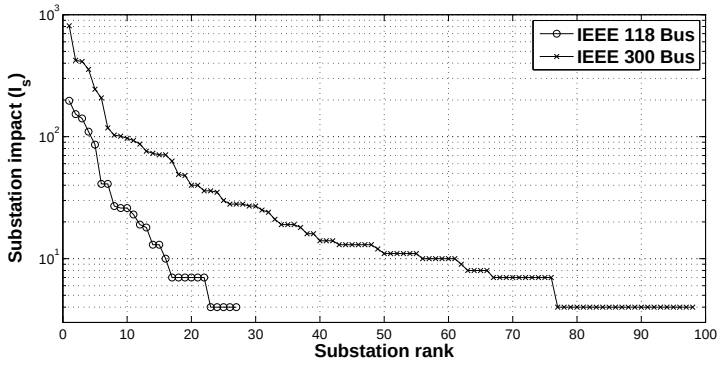

Figure 3: Attack impact I_s of the substations in the IEEE 118 and 300 bus systems in decreasing order of attack impact. The case of shortest path routing.

3.5 Mitigation measures against attacks

In the following we consider how an operator could improve the security of the system by (i) changing the routes used by the substations (ii) by using multipath routing (iii) and by using data authentication and/or protection.

First, we formulate a result regarding mitigation schemes that make stealth attacks impossible to perform, i.e., mitigation schemes such that $\Gamma_m = \infty$, $\forall m$. For this to hold, the minimum number of *measurements* z_m needed to be protected is the number of buses n [22, 24]. The straightforward way to protect this many measurements is to deploy tamper-proof authentication at all substations. The following result suggests that one can mitigate stealth attacks by deploying authentication in significantly less substations.

Proposition 2. *Consider the power system graph, i.e., the graph with vertex set $\mathcal{S}$, and edges the transmission lines. If $\Gamma_m = \infty$ $\forall m$ then $\mathcal{E}^P \cup \mathcal{P}$ is a dominating set of the power system graph.*

Proof. The dominating set of a graph is a subset of the graph's vertices such that every vertex is either a member of the subset or is adjacent to a vertex in the subset.

To prove the proposition, we show that if $\mathcal{E}^P \cup \mathcal{P}$ is not a dominating set of the power system graph then there is at least one measurement m with $\Gamma_m < \infty$.

Since $\mathcal{E}^P \cup \mathcal{P}$ is not a dominating set, there is at least one substation s that is unprotected and not authenticated, and is not adjacent to any substation $s' \in \mathcal{E}^P \cup \mathcal{P}$. Take a measurement m at a bus at substation s. This measurement can be attacked by using an attack vector $a = Hc$ for a vector c whose only non-zero component is that corresponding to a bus at substation s. a has nonzero components corresponding to measurements at adjacent buses, and these measurements are located at substations that do not use either authentication or protection. Hence $\Gamma_m < \infty$. This concludes the proof. □

The cardinality of the dominating set of connected graphs is typically much smaller than the number of vertices, hence perfect protection might be achievable without installing tamper-proof authentication at every substation. The numerical results in Section 3.6 validate this observation.

Next, we turn to the problem of decreasing the vulnerability of the system. A natural goal for the operator would be to improve the most vulnerable part of the system, that is, to minimize $\max_{s\in\mathcal{S}} I_s$ or to maximize $\min_{m\in\mathcal{M}} \Gamma_m$, potentially subject to some constraints on the feasible set of mitigation measures (e.g., due to financial reasons). Maximizing the cost of the least cost stealth attack can lead to increased average attack cost as well, compared to maximizing the average attack cost [24].

Instead of the above formulations, we formulate the operator's goal as a multi-objective optimization problem. As we show later, the solution to this problem formulation is a solution to the max-min formulation. We define the objective γ to be the minimization of the number of measurements with attack cost γ, $|\{m|\Gamma_m = \gamma\}|$. The objectives are ordered: objective γ has priority over objective $\gamma' > \gamma$. Formally, we define the objective vector $w \in \mathbb{N}^{S-1}$ whose γ^{th} component is $w_\gamma = |\{m|\Gamma_m = \gamma\}|$. The goal of the operator can then be expressed as

$$\underset{\mathcal{R},\mathcal{E}^N,\mathcal{E}^P,\mathcal{P}}{\text{l}exmin}\ w(\mathcal{R},\mathcal{E}^N,\mathcal{E}^P,\mathcal{P}), \tag{8}$$

where l*exmin* stands for lexicographical minimization [25], $w(\mathcal{R},\mathcal{E}^N,\mathcal{E}^P,\mathcal{P})$ is the objective vector calculated for the established routes $\mathcal{R}$, the sets $\mathcal{E}^N$ and $\mathcal{E}^P$ of authenticated substations, and the set $\mathcal{P}$ of protected substations, and the optimization is performed over all feasible mitigation schemes. The minimal objective vector w, $w_\gamma = 0$ $(1 \le \gamma \le S-1)$ corresponds the case when no measurement can be stealthily attacked, i.e., $\Gamma_m = \infty$ for all $m \in \mathcal{M}$.

Proposition 3. *The solution to (8) is a solution to* $\max_{\mathcal{P},\mathcal{E}^N,\mathcal{E}^P,\mathcal{R}} \min_{m\in\mathcal{M}} \Gamma_m$. *Furthermore, if* $\max_{\mathcal{P},\mathcal{E},\mathcal{R}} \min_{m\in\mathcal{M}} \Gamma_m > 1$ *the solution to (8) is a solution to* $\min_{\mathcal{P},\mathcal{E}^N,\mathcal{E}^P,\mathcal{R}} \max_{s\in\mathcal{S}} I_s$.

Table 3: CSF algorithm for given $\mathcal{R}$, $\mathcal{E}^N$, $\mathcal{E}^P$, $\mathcal{P}$ and γ^*

1. Set $\hat{\mathcal{S}} = \emptyset$
2. **for** $\forall m$ where $\Gamma_m = \gamma^*$ **do**
3. $X = \{x|\text{x is a valid stealth attack assuming } \mathcal{E}^N = \mathcal{S}\}$
4. $\exists X_{\gamma^*} \subseteq X$ s.t. $\forall x \in X_{\gamma^*}, \gamma^* = ||\omega||$
5. $\hat{\mathcal{S}} = \hat{\mathcal{S}} \cup \{\hat{s}|x(\hat{s}) = 1, \forall x \in X_{\gamma^*}\}$
6. **end for**
7. **for** $\forall \hat{s} \in \hat{\mathcal{S}}$
8. **create** $\mathcal{R}'_{\hat{s}}$ **and set** $\mathcal{R}'(\hat{s}) = (\mathcal{R} \setminus \mathcal{R}_{\hat{s}}) \cup \mathcal{R}'_{\hat{s}}$ or
9. **set** $\mathcal{E}^{N\prime}(\hat{s}) = \mathcal{E}^N \cup \hat{s}$ **or** $\mathcal{E}^{P\prime}(\hat{s}) = \mathcal{E}^P \cup \hat{s}$ **or** $\mathcal{P}'(\hat{s}) = \mathcal{P} \cup \hat{s}$
9. **calculate** $w^{\hat{s}}(\mathcal{R}'(\hat{s}), \mathcal{E}^{N\prime}(\hat{s}), \mathcal{E}^{P\prime}(\hat{s}), \mathcal{P}'(\hat{s}))$
10. **end for**
11. $\hat{s}^* = \arg\min_{\hat{s}} w^{\hat{s}}$
12. **if** $w^{\hat{s}^*} < w$
13. **return** $\mathcal{R}'(\hat{s}^*)$, $\mathcal{E}^{N\prime}(\hat{s}^*)$, $\mathcal{E}^{P\prime}(\hat{s}^*)$, $\mathcal{P}'(\hat{s}^*)$
14. **else if** $\gamma^* < S - 1$
15. **Set** $\gamma^* = \gamma^* + 1$ **and** GOTO (1)
16. **else**
17. **return** $\mathcal{R}$, $\mathcal{E}^N$, $\mathcal{E}^P$ and $\mathcal{P}$
18. **end if**

Proof. We prove the first part of the proposition by contradiction. Let w be the solution to (8), i.e., the lexicographically minimal objective vector, and denote by γ^* the smallest attack cost for which $w_{\gamma^*} > 0$, i.e., $\gamma^* = \min\{\gamma|w_\gamma > 0\}$. Let $\gamma' = max_{\mathcal{P},\mathcal{E}^N,\mathcal{E}^P,\mathcal{R}} \min_{m\in\mathcal{M}} \Gamma_m$ be the max-min solution and w' a corresponding objective vector. Assume now that $\gamma^* < \gamma'$. For $\gamma < \gamma'$ the objective vector has $w'_\gamma = 0$. Since $\gamma^* < \gamma'$, $w'_{\gamma^*} = 0$, and hence according to the definition of lexicographical ordering $w' < w$, which contradicts to the assumption that w is lexicographically minimal.

The second part of the proposition follows directly from Proposition 1 and from the first part of the proposition. □

We solve the lexicographical minimization in (8) in an iterative way [25]. Consider given $\mathcal{R}, \mathcal{E}^N, \mathcal{E}^P, \mathcal{P}$ and let $\gamma^* = \min\{\gamma|w_\gamma > 0\}$. If $\gamma^* = \infty$ the system is not vulnerable. Otherwise, we use the *critical substation first* (CSF) algorithm shown in Table 3 to decrease w_γ for some $\gamma \geq \gamma^*$ as long as that is possible.

The algorithm starts by calculating the set $\hat{\mathcal{S}}$ of *critical* substations. In order to find the *critical* substations, the algorithm identifies measurements with attack cost $\Gamma_m = \gamma^*$. Each such measurement has at least one stealth attack ω with attack

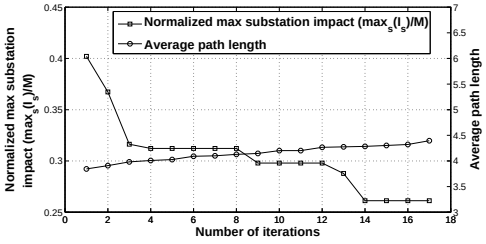

Figure 4: Maximum normalized attack impact and average path length vs. the number of single-path routes changed in the IEEE 118 bus system and mesh topology.

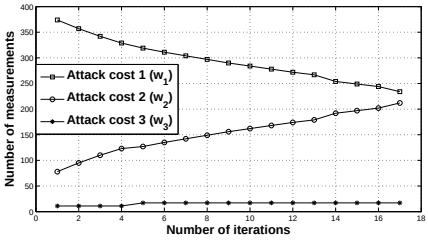

Figure 5: Number of measurements for various attack costs vs. the number of single-path routes changed in the IEEE 118 bus system and mesh topology.

cost $||\omega|| = \gamma^*$. The substations that are contained in ω for every such stealth attack are *critical* substations. There is at least one such substation, the substation $S(m)$. The critical substations are the candidates for route reconfiguration, authentication or protection.

For every *critical* substation $\hat{s}$ the algorithm considers an alternate mitigation scheme. The alternate mitigation scheme could contain a new set of routes $\mathcal{R}'_{\hat{s}}$ between substation $\hat{s}$ and the control center, or it could be the set of authenticated or protected substations augmented by $\hat{s}$ ($\mathcal{E}^{N\prime}(\hat{s}) = \mathcal{E}^N \cup \hat{s}$, $\mathcal{E}^{P\prime}(\hat{s}) = \mathcal{E}^P \cup \hat{s}$ or $\mathcal{P}'(\hat{s}) = \mathcal{P} \cup \hat{s}$). For every alternate mitigation scheme the algorithm calculates the objective vector $w^{\hat{s}}$ and selects the one with the minimal objective vector, $w^{\hat{s}}$. If the alternate mitigation scheme improves the system's level of protection, i.e., $w^{\hat{s}} < w$ then the algorithm terminates. Otherwise the algorithm considers a higher attack cost $\gamma^* = \gamma^* + 1$, and continues from Step 1.

3.6 Numerical results

In the following we illustrate on the IEEE 118 and the IEEE 300 bus networks how different mitigation schemes can be used to decrease the system's vulnerability to attacks. We use the same two communication topologies as in Section 3.4.3.

3.6.1 The case of single-path routing

Modifying single-path routes has the smallest complexity among the mitigation schemes we consider, hence we start with evaluating its potential to decrease the vulnerability of the system. For single-path routing the alternate mitigation schemes differ only in terms of routing. Consequently, $\mathcal{P}'(\hat{s}) = \mathcal{P}$, $\mathcal{E}^{P\prime}(\hat{s}) = \mathcal{E}^P$ and $\mathcal{E}^{N\prime}(\hat{s}) = \mathcal{E}^N$.

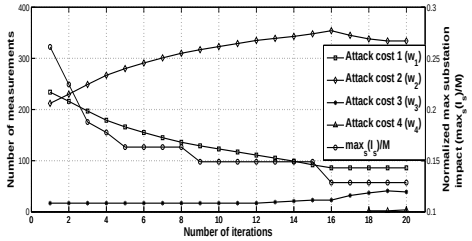

Figure 6: Maximum attack impact and number of measurements for various attack costs vs. the number of multi-path routes. IEEE 118 bus system, mesh topology.

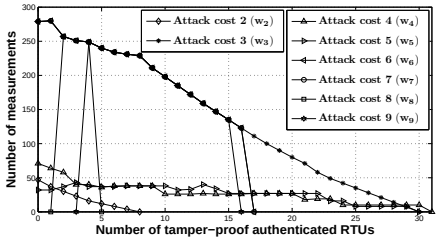

Figure 7: Maximum attack impact and number of measurements for various attack costs vs. the number of tamper-proof authenticated RTUs ($|\mathcal{E}^P|$). IEEE 118 bus system, star topology.

In the star topology, substations are directly connected to the control center. Hence, modifying single-path routes is not feasible. For the case of the mesh topology, in order to obtain $\mathcal{R}'(\hat{s})$ from $\mathcal{R}$ for a critical substation $\hat{s}$ we modify the only route $r_1^{\hat{s}}$ in $\mathcal{R}_{\hat{s}}$. For a route $r_1^{\hat{s}}$ we create the shortest alternate route $r_1^{\hat{s}\prime}$ that avoids the substation $s \in r_1^{\hat{s}}$ that appears in most substation attacks ω with cardinality γ^*.

Fig. 4 shows the maximum normalized substation attack impact, i.e., $\max_s I_s/M$, as a function of the number of single-path routes changed in the 118 bus system. The maximum attack impact shows a very fast decay, and decreases by almost a factor of two. At the same time the average path length to the control center increases by only 10%.

Fig. 5 shows the number of measurements that have attack cost 1, 2 and 3 (i.e., w_1, w_2 and w_3) as a function of the number of routes changed in the 118 bus system for the mesh topology. By changing single-path routes the algorithm could increase the attack cost for about 200 measurements from $\Gamma_m = 1$ to $\Gamma_m = 2$, and for some measurements to $\Gamma_m = 3$ (e.g., at iteration 5). The importance of increasing the attack cost from 1 to 2 for a particular measurement lies in the fact that while an attacker could gain access to a single substation by accident, simultaneous access to more than one substations would require significant coordinated effort. Thus it is encouraging that a scheme as simple as modifying single path routes is so efficient in mitigating attacks. After 16 iterations the algorithm could not find any single-path route that would lead to increased attack cost for any measurement. Hence, we turn to multi-path routing.

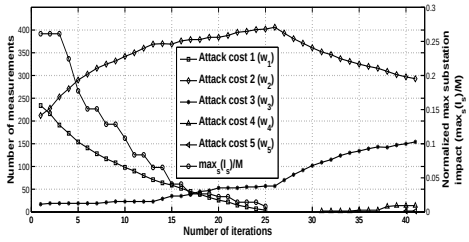

Figure 8: Maximum attack impact and number of measurements for various attack costs vs. the number of non tamper-proof authenticated RTUs ($|\mathcal{E}^N|$). IEEE 118 bus system, mesh topology.

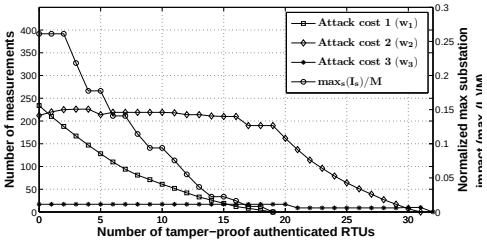

Figure 9: Maximum attack impact and number of measurements for various attack costs vs. the number of non tamper-proof authenticated RTUs ($|\mathcal{E}^N|$). IEEE 118 bus system, mesh topology.

3.6.2 The case of multi-path routing

In the case of multi-path routing the alternate mitigation schemes differ only in terms of routing, as for single-path routing. Consequently, $\mathcal{P}'(\hat{s}) = \mathcal{P}$, $\mathcal{E}^{P\prime}(\hat{s}) = \mathcal{E}^P$ and $\mathcal{E}^{N\prime}(\hat{s}) = \mathcal{E}^N$.

Since in the star topology substations are directly connected to the control center, multi-path routing can not decrease the vulnerability of the system. For the mesh topology, to obtain $\mathcal{R}'(\hat{s})$ from $\mathcal{R}$ for a critical substation $\hat{s}$, we consider the single route $r_1^{\hat{s}}$ in $\mathcal{R}_{\hat{s}}$, and construct the shortest route $r_2^{\hat{s}\prime}$ such that $r_2^{\hat{s}\prime}$ and $r_1^{\hat{s}}$ are node-disjoint. The routes in $\mathcal{R}_{\hat{s}}{}'$ are then $r_1^{\hat{s}\prime} = r_1^{\hat{s}}$ and $r_2^{\hat{s}\prime}$.

Multi-path routing introduces complexity in the management of the communication infrastructure. In the case of SDH at the link layer several virtual circuits have to be configured and maintained. In the case of Ethernet some form of traffic engineering is required (e.g., using MPLS). Hence the cost of establishing a multi-path route from a substation to the control center has a higher cost than changing a single-path route, considered in the previous subsection. We therefore take the set of routes $\mathcal{R}$ obtained in the last iteration of the algorithm in the previous subsection as the starting point for deploying multi-path routing.

Fig. 6 shows the maximum normalized substation attack impact and the number of measurements with attack costs 1 to 4 vs. the number of multi-path routes in the 118 bus system for the mesh topology. Multi-path routing could decrease the maximum attack impact by 50% through increasing the number of measurements with attack cost $\Gamma_m = 2$ and $\Gamma_m = 3$. There are 86 measurements with attack cost 1 when the algorithm terminates. We note, however, that dual-path routing provides a significant improvement compared to single path routing. The attack costs could further be increased by considering more than two simultaneous paths

whenever possible. Instead of exploring this direction, we turn to authentication.

3.6.3 The case of authentication

In the case of (non) tamper-proof authentication the alternate mitigation schemes differ in terms of the set of (non) tamper-proof authenticated substations $\mathcal{E}^P$ ($\mathcal{E}^N$). Consequently, $\mathcal{P}'(\hat{s}) = \mathcal{P}$ and $\mathcal{R}'(\hat{s}) = \mathcal{R}$.

To obtain $\mathcal{E}^{N\prime}(\hat{s})$ from $\mathcal{E}^N$ for a critical substation $\hat{s}$ we add substation $\hat{s}$ to the set of substations using non tamper-proof authentication, i.e., $\mathcal{E}^{N\prime}(\hat{s}) = \mathcal{E}^N \cup \hat{s}$. We follow a similar procedure to augment the set $\mathcal{E}^P$ of substations with tamper-proof authentication.

Apart from the deployment costs (e.g., new equipment), authentication requires that secret keys be protected and managed, which results in costs for the operator. The cost of introducing authentication is certainly higher than that of reconfiguring single-path routing, but it is difficult to compare its cost to that of introducing multi-path routing. We therefore take the set of routes $\mathcal{R}$ obtained in the last iteration of the algorithm for single-path routing as the starting point for deploying authentication.

Fig. 7 shows the number of measurements with attack cost 1 to 9 as a function of the number of tamper-proof authenticated RTUs in the 118 bus system for the star topology. Note that there are no measurements with attack cost 1. With 31 substations using tamper-proof authentication stealth attacks are impossible to perform. The 31 substations form a dominating set of the power system graph, in accordance with Proposition 2. Note that this number is less than one third of the number of substations in the system, which is $S = 109$.

Fig. 8 shows the maximum normalized substation attack impact and the number of measurements with attack cost 1 to 5 as a function of the number of non tamper-proof authenticated RTUs in the 118 bus system for the mesh topology. Authentication eliminates measurements with attack cost $\Gamma_m = 1$ after 25 substations are authenticated. Furthermoe, upon termination more measurements have attacks cost $\Gamma_m \geq 3$, than using multi-path routing.

Fig. 9 shows the maximum normalized substation attack impact and the number of measurements with attack cost 1 to 3 as a function of the number of tamper-proof authenticated RTUs in the 118 bus system for the mesh topology. Authentication eliminates measurements with attack cost $\Gamma_m = 1$ ($\Gamma_m = 2$, $\Gamma_m = 3$) after 19 (31,32) substations are authenticated. With 32 using tamper-proof authentication stealth attacks are impossible to perform. These 32 substations also form a dominating set of the power system graph, in accordance with 2. We note that authenticating the 31 substations found to make stealth attacks impossible for the star topology would also make stealth attacks impossible for the mesh topology.

3.7 Conclusion

The results obtained using the framework described in this section allow us to draw important conclusions regarding the vulnerability of the state estimator to attacks on the communication infrastructure and regarding the potential mitigation strategies. First, the topology of the communication infrastructure is of paramount importance. Second, the vulnerability of the state estimator can be significantly decreased by using very simple network layer mitigation schemes, which do not require the installation and management of secret keys in the substations. Finally, as expected, authentication is unavoidable to make attacks impossible, but it is important to note that the number of substations that has to be authenticated is significantly less than one would expect if not taking into account the communication infrastructure.

4 Increasing Availability through Anonymous Communications

Conceptually, the purpose of inter-control center communication (ICCC) is very similar to that of substation to control center communications: to enable monitoring of the state of the power system and to control actuators located in substations. Nevertheless, in the case of ICCC the exchange of information and control messages happens between different organizations, such as distribution service providers, transmission service providers, neighboring utilities, regional and national control centers, electricity producers and other electricity market participants. An example ICCC scenario is shown in Fig. 10 based on the communication infrastructure of the Italian transmission system [26].

The information obtained using ICCC is often used in the state estimation process, hence the security of ICCC affects the security of state estimation. In a smart grid environment, the importance of and the reliance on ICCC is expected to increase for a number of reasons. First, the number of independently managed electricity market participants is expected to increase, and their secure operation requires information about the state of other market participants. Second, due to its distributed nature the stability of the smart grid will rely increasingly on wide area measurement and control systems, which will span several, geographically distant market participants and might require real-time data delivery with stringent delay and throughput requirements.

4.1 Inter-Control Center Communication

In the past there were a number of proprietary protocols in use for inter-control center communications, but the predominant protocol in use today is the Inter-Control Center Communication Protocol (ICCP, IEC 60870-6/TASE.2). ICCP provides a point-to-point connection, called an association, between a pair of nodes, that is, two control centers. Two nodes can maintain several ICCP associations with each other simultaneously, and can use different associations to exchange data with different priorities. The rationale for maintaining several associations is that the service level requirements of the information exchanged between two nodes spans a wide range, from real-time data exchange with stringent delay requirements to the bulk exchange of planning data and schedules. ICCP can operate on top of a variety of transport layer protocols, both connectionless and connection oriented, but most often it is used on top of TCP/IP.

Although ICCP was standardized only a few years ago, it does not include either confidentiality, or integrity, or authentication. It only provides access control via so called bilateral tables. Bilateral tables specify the access rights between

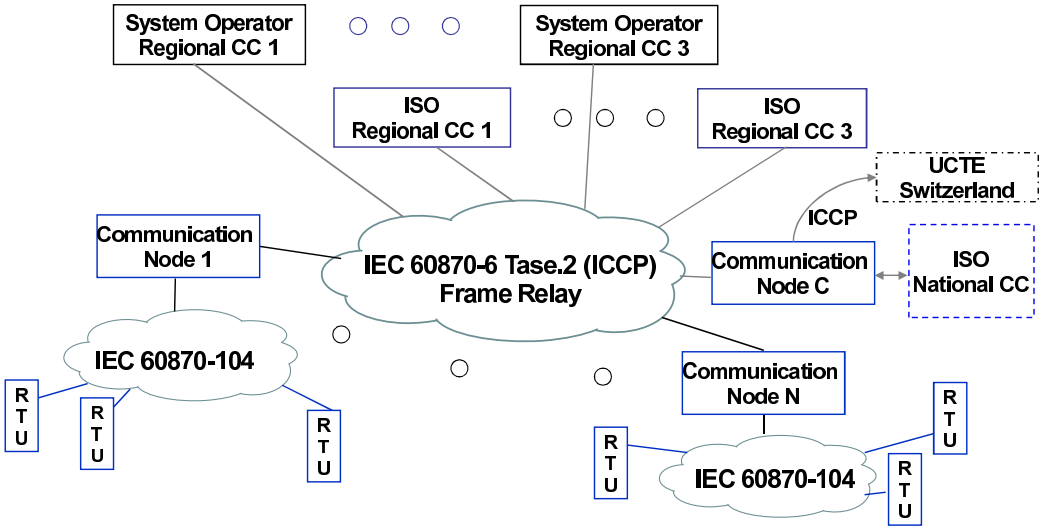

Figure 10: Schematic view of the SCADA communication infrastructure of the Italian transmission system, which used to be the largest ICCP deployment worldwide [26]. In the actual system measurement data from around 250 RTUs are delivered to one of 22 communication nodes using IEC 60870-5-104 over TCP/IP. ICCP is used to deliver data from the communication nodes to the 3 regional control centers of the transmission system operator (TSO), and to the 3 regional control centers and to the national control center of the independent system operator (ISO). ICCP is used also to communicate to the control center of the Union for the Co-ordination of Transmission of Electricity (UCTE). For simplicity, the figure does not show ICCP connections to power generation control centers.

two control centers that have an ICCP association. Confidentiality, integrity, and authentication can be provided by lower layer protocols, for example, TLS [18] when ICCP is used over the TCP/IP protocol stack. The number of nodes connected by ICCP associations in today's power systems is relatively low, in the order of tens, like in the case of the system shown in Fig. 10. As the number of nodes is low and control centers are relatively well protected key management is not an issue in practice today.

4.2 Beyond Cryptographic Security: Information Availability

With the problems of integrity, authentication, confidentiality and access control solved to a large extent, the most important issue in ICCC is information availability. Unlike in the past, when ICCC was performed mainly over dedicated point-to-point connections, such as leased lines, communication is shifting to public wide

area networks, such as the Internet. The use of public network infrastructures might be cheaper, but it poses stringent requirements on network availability and it exposes ICCC to denial of service attacks. The resilience to network failures can be improved by maintaining multiple independent communication paths between the nodes, at the price of increased costs. If the ICCP connection is established over a TCP/IP network, multi-homing and redundant routers can be used to provide fast failover. The ICCP standard enables such solutions to be implemented over TCP, but there is no standardized solution, and hence the interoperability between the products of different vendors can be an issue in practice. Alternatively, ICCP can be used over the Stream Control Transmission Protocol (SCTP) and can rely on the multipath and failover capabilities of the transport layer protocol. In general there is, however, a trade-off between the frequency of path bouncing and the speed of path failover in the case of a network failure [27].

Denial of service attacks are even more difficult to mitigate. An attacker that monitors the data traffic of encrypted ICCP associations can use traffic analysis to extract information from the traffic patterns, e.g., it can detect the increase of data rates, which is typically a sign of abnormal system state, and can disable communications when it is most needed. Traffic analysis attacks can either be mitigated through masking (i.e., continuously transmitting at the peak rate) or through relaying over mix networks [28], which delay every message at random upon relaying. In principle ICCP allows associations to be relayed over control centers, so that a mixing network can be used to hide the identity of the sender and the receiver of an ICCP association from an outside attacker [29].

Nevertheless, a mix network introduces the possibility of inside attacks: due to the long life-cycles of SCADA systems software corruption is a threat, and the complexity of the code-base makes corruption hard to detect. A compromised control center can reveal the routing information of the mix network and thereby it can enable attacks despite using a mix network. One possible solution to mitigate the attacks even in the presence of compromised control centers is to use anonymity networks to establish overlay routing paths among the control centers. An anonymity network hides the sender and/or the receiver of the messages routed through the overlay from the relaying nodes, and thereby it makes it difficult for an inside attacker to identify the associations between the nodes [29]. Depending on whether it is the sender, the receiver or the association between a pair of nodes that is to be hidden, an anonymity network can be designed to provide sender, receiver or relationship anonymity, respectively [30]. In the following we focus on relationship anonymity, i.e., on hiding the fact that there is an association between two nodes.

Anonymity networks can provide some level of relationship anonymity against inside attackers (e.g., [31], [32]) by hiding the sender or the receiver from the relay nodes. Nevertheless, good sender or good receiver anonymity in itself does

not necessarily lead to the best possible relationship anonymity [33]. Furthermore, mix networks and anonymity networks come at the price of increased data rates and end-to-end delay. Increased data rates lead to increased communication costs, while long delays are undesirable for time-sensitive data; hence the mix and anonymity networks have to be configured appropriately if they are to be used for ICCC. The results presented in the following address, among others, the above design issues.

4.3 System Model

We consider an anonymity network that consists of a set $\mathscr{N}$ of nodes, $N = ||\mathscr{N}||$. The nodes (e.g., control-centers communication with each other using ICCP) act as *sources*, *destinations* and as *relay* nodes for each others' messages. The underlying communication network is a complete graph.

The *inside attacker* is in control of a set $\mathscr{C} \subset \mathscr{N}$ ($C = ||\mathscr{C}||$) of compromised nodes, and its goal is to learn the communication patterns. To achieve its goal, the attacker can observe the messages traversing the nodes in $\mathscr{C}$ and the protocol specific information contained in the messages. It can recognize if the same message visits several compromised nodes. The attacker has some *a priori* belief of the system traffic matrix. For every message that the attacker observes, it calculates the probability $P(\hat{S}(a), \hat{R}(b))$ for every pair of nodes $(a,b) : a \in \mathscr{N}, b \in \mathscr{N} \setminus \{a\}$ that it is the sender-receiver pair (s,r) of the message. The attacker maintains a counter for every pair of nodes (*a,b*), and it increases the counters with the calculated probabilities for every observed message. The attacker uses the counters to estimate the number of exchanged messages between every pair of nodes in a given time interval.

We consider two metrics: the *overhead* of the anonymity network and the *relationship anonymity*. We define the *overhead* as the average number of nodes $E[K]$ that an arbitrary message visits. We quantify the *relationship anonymity* by the average increase of the counter corresponding to the real sender-receiver pair (s,r) for every message sent by (s,r), including the ones not observed by the attacker. In general, the relationship anonymity depends on two factors. First, on the probability of having an attacker node on the path. Second, on the probability assigned to the sender-receiver pair $P(\hat{S}(s), \hat{R}(r))$ by an attacker node on the path. Both factors are functions of the anonymity protocol, the number of nodes N and the number of inside attacker nodes C. Furthermore, the probability $P(\hat{S}(s), \hat{R}(r))$ depends on the method used by the attacker for counting. We consider two counting methods.

4.3.1 Bayesian inference method

Using the Bayesian inference method, when the attacker intercepts a message, it considers every pair of nodes (a,b) as a possible sender-receiver pair of the message regardless of how likely they are. Let us denote by $P(H_{1+}|S(s),R(r))$ the probability that an attacker node occurs on the path given that (s,r) is the sender-receiver pair, and by $P(\hat{S}(s),\hat{R}(r)|H_{1+},S(s),R(r))$ the probability that the attacker identifies (s,r) as the sender-receiver pair given its occurrence on the path. Then we can express the relationship anonymity under the BI method as

$$P_{relB}(s,r) = P(\hat{S}(s),\hat{R}(r)|H_{1+},S(s),R(r)) \cdot P(H_{1+}|S(s),R(r)), \tag{9}$$

4.3.2 Maximum posteriori method

Using the Maximum posteriori method method, when the attacker intercepts a message, it identifies the set $\mathscr{Q}$ of the most likely sender-receiver pairs. In the worst case the set $\mathscr{Q}$ is a singleton, $||\mathscr{Q}|| = 1$, and the anonymity is likely to be low. At the other extreme, $\mathscr{Q}$ can contain all possible send-receiver pairs, $||\mathscr{Q}|| = (N-C)\cdot(N-C-1)$, which corresponds to perfect relationship anonymity. In general $(a,b) \in \mathscr{Q}$ does not imply that (a,b) is the actual sender-receiver pair, not even when $||\mathscr{Q}|| = 1$. Nevertheless, intuitively, we can say that $(s,r) \in \mathscr{Q}$ is more likely than $(s,r) \notin \mathscr{Q}$.

Let us denote by $P((s,r) \in \mathscr{Q}|H_{1+},S(s),R(r))$ the probability that the sender-receiver pair is one of the most likely sender-receiver pairs, i.e., $(s,r) \in \mathscr{Q}$, and by $P(\hat{S}(s),\hat{R}(r)|(s,r) \in \mathscr{Q},H_{1+},S(s),R(r))$ the probability that the attacker identifies (s,r) as the sender-receiver pair given its occurrence on the path and $(s,r) \in \mathscr{Q}$. Using this notation we can express the relationship anonymity under the MP method as

$$\begin{aligned} P_{relM}(s,r) = & P(\hat{S}(s),\hat{R}(r)|(s,r) \in \mathscr{Q},H_{1+},S(s),R(r)) \cdot \\ & P((s,r) \in \mathscr{Q}|H_{1+},S(s),R(r)) \cdot P(H_{1+}|S(s),R(r)). \end{aligned} \tag{10}$$

4.4 Anonymity networks

We consider two anonymity networks, MCrowds and Minstrels. The two anonymity networks are designed to provide relationship anonymity through hiding both the sender and the receiver, in two different ways.

4.4.1 MCrowds

MCrowds is an anonymity network inspired by Crowds [32], which was proven to provide optimal sender anonymity [34]. In MCrowds the sender specifies a set

$\mathscr{M}$ of nodes as receiver for a message. The number $M = ||\mathscr{M}||$ of receiver nodes is a system parameter. Nodes specified in the set $\mathscr{M}$ are not used for relaying. For a message to reach its intended receiver r it must be that $r \in \mathscr{M}$; the other $M-1$ nodes are chosen uniformly at random. The sender then sends the message to one of the $\mathscr{N} \setminus \mathscr{M}$ nodes for relaying. The relay node sends the message for further relaying with probability p_f to one of the $\mathscr{N} \setminus \mathscr{M}$ nodes, and with probability $1-p_f$ the message is sent as a multicast message to all receiver nodes specified in $\mathscr{M}$. Upon multicasting, the receiver set is removed from the message. Node r recognizes that it is the receiver while the other $\mathscr{M} \setminus \{r\}$ nodes discard the message. For $M = 1$ MCrowds is equivalent to Crowds, except for that the receiver node $r \in \mathscr{N}$.

4.4.2 Minstrels

Minstrels uses nodes as message relays in the same way as Crowds with the difference that the number of nodes visited by a message is bounded.

When a node s wants to send a message to a node r it picks a node uniformly at random among the other $N-1$ nodes (excluding s) and forwards the message. The next node forwards the message to one of the other $N-2$ nodes (excluding itself and the sender node s) chosen uniformly at random. Every subsequent forwarder picks one of the non-visited nodes to forward the message. When node r receives the message, it will send the message further in order to improve the receiver anonymity. The path ends when all N nodes have been visited.

The message, or part of it, is encrypted with the receiver's public key. When a node receives the message, it checks if it is the receiver by trying to decrypt the encrypted part of the message. If the decrypted part of the message represents valid data, the node is the receiver. Note that a node does not know who is the receiver, it can only check whether it is the receiver itself.

To bound the path length, every message records the set $\mathscr{V}$ of the visited nodes in its header. The set can be implemented, for example, using a Bloom filter, to keep its size small. When a relaying node receives a message, it will relay the message only to non-visited nodes. To control the maximum path length (i.e., delay) the sender can initialize the set $\mathscr{V}$ of visited nodes with a number $f \in \{0,\ldots,N-1\}$ of the nodes in the system. These initialized nodes are considered as visited so that the message can not be relayed to them. A message traverses all nodes except for the initialized nodes in the set $\mathscr{V}$ and hence the sender must not include the receiver in the set $\mathscr{V}$. The sender picks the number of initialized nodes at random: it initializes the set with f nodes with probability $P(F=f)$, where $\sum_{f=0}^{N-1} P(F=f) = 1$. For $f=0$ the set is empty, for $f=1$ the set is initialized only with the sender and for $f>1$ the set is initialized with the sender and $f-1$ other nodes. The distribution of F is a system parameter, and we use it to explore

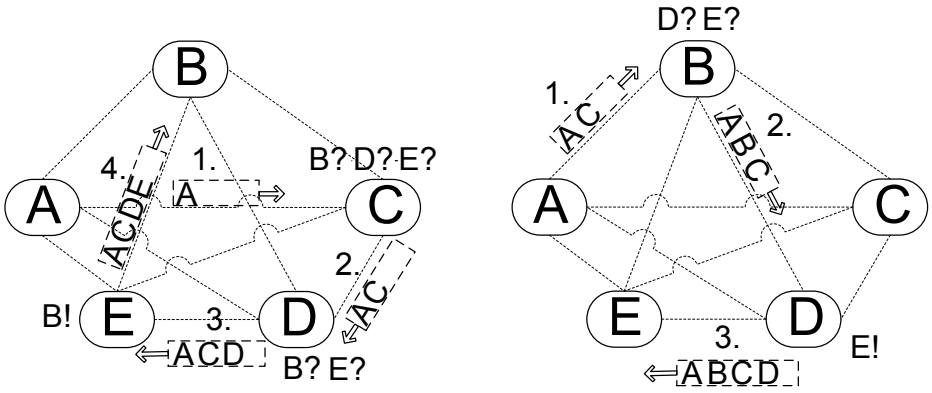

Figure 11: A simple example of Minstrels with five nodes.

the anonymity-overhead trade-off.

Fig. 11 shows two simple examples with five nodes, node A as sender and node D as receiver. Fig. 11 (left) shows a case when the set $\mathscr{V}$ is initialized with the sender node A and the message is forwarded to node C. Node C checks if it is the receiver, puts itself in the set and chooses the next hop uniformly at random among nodes (B,D,E). The next hop, node D, follows the same procedure with only two forwarding options (B,E). Fig. 11 (right) shows another case when the set $\mathscr{V}$ is initialized with the sender and node C, and the message is forwarded to node B. Node B adds itself to the set and decides to which of the remaining nodes (D,E) to forward the message. Node C is considered as already visited.

4.5 Overhead and Anonymity

In the following we derive expressions for the communication overhead and the relationship anonymity provided against inside attackers for MCrowds and for Minstrels.

4.5.1 Communication Overhead

We start with calculating the communication overhead of MCrowds and Minstrels. For MCrowds, the mean number of nodes visited by a message is the expected value of a geometric distribution with success probability $1-p_f$ plus the multicast messages, i.e.,

$$E[K] = \frac{p_f}{1-p_f} + 1 + M. \tag{11}$$

For Minstrels and for a given number f of initialized nodes in the set $\mathscr{V}$, the number of nodes visited by a message is equal to $K = N - f$. The mean number

of visited nodes depends on the distribution of F and it can be expressed as

$$E[K] = \sum_{f=0}^{N-1} P(F = f) \cdot (N - f). \tag{12}$$

4.5.2 Relationship Anonymity Against Inside Attackers

In the following we derive the relationship anonymity expressions for MCrowds and for Minstrels.

MCrowds We start the calculation of the relationship anonymity with expressing the probability of having an attacker node on the path. This probability depends on the number of receiver nodes M, because M influences the number M_C of attacker nodes in the receiver set $\mathscr{M}$. Let us denote by $\mathscr{M}_C$ the set of attacker nodes in the receiver set, $M_C = ||\mathscr{M}_C|| \in \{max(0, M-(N-C-1)), ..., min(M-1, C)\}$. For $M = 1$ there cannot be attacker nodes in the receiver set, $P(M_C = 0) = 1$. For $M > 1$ we have

$$P(M_C = m_C) = \left(\begin{array}{c} M-1 \\ m_C \end{array} \right) \frac{\prod_{k=2}^{M-m_C}(N-C-k)\prod_{k=0}^{m_C-1}(C-k)}{\prod_{k=2}^{M}(N-k)}.$$

Let us denote the event that the position of the first attacker node is i by H_i. H_i happens if the message is first relayed $i-1$ times through trusted nodes but the i^{th} hop is an attacker node. Conditioned on m_C we have

$$P(H_i | M_C = m_C, S(s), R(r)) = \frac{C-m_C}{N-M}\left(p_f \frac{N-C-M+m_C}{N-M}\right)^{i-1}.$$

If the message is again relayed over an attacker node on any position after i, the attacker does not gain any additional information about the sender-receiver pair (s,r) of the message, and hence the probability assigned to the sender-receiver pair does not change. Therefore, we focus only on the position of the first attacker node on the path. Let us now denote by H_{1+} the event that there is an attacker on the path as a relay. Conditioned on m_C this event happens with probability

$$\begin{aligned} P(H_{1+} | M_C = m_C, S(s), R(r)) &= \sum_{i=1}^{\infty} P(H_i | M_C = m_C, S(s), R(r)) \\ &= \frac{C-m_C}{N-M-p_f(N-C-M+m_C)}. \end{aligned}$$

Let us denote by I the event that the first attacker node on the path is immediately preceded on the path by the sender. Note that $H_1 \Rightarrow I$ but the opposite is not

true since the sender may appear multiple times on the path. Since the receiver set size is M, there are $N-M$ possible relays and so the predecessor is the sender with probability $P(I|H_{1+},M_C=m_C)=\frac{N-M-p_f(N-C-1-M+m_C)}{N-M}$. Similarly, we denote by $\overline{I}$ the event that a particular node other than the predecessor is the sender, this happens with probability $P(\overline{I}|H_{1+},M_C=m_C)=\frac{1-P(I|H_{1+},M_C=m_C)}{N-C-1-M+m_C}$.

Next, we calculate the probability $P(I|H_i,M_C=m_C,S(s),R(r))$ that the sender is the predecessor of a relaying attacker. For $i=1$ we have $P(I|H_1,M_C=m_C,S(s),R(r))=1$ while for $i>1$ we have $P(I|H_i,M_C=m_C,S(s),R(r))=\frac{1}{N-C-M+m_C}$. The probability $P(\overline{I}|H_i,M_C=m_C,S(s),R(r))$ of the complement events follows.

Then, for the BI method the probability that a *relaying* attacker assigns to the actual sender of the message is

$$\begin{aligned}P(\hat{S}(s)|H_i,M_C=m_C,S(s),R(r))=&\\ &P(I|H_i,M_C=m_C,S(s),R(r))\cdot P(I|H_{1+},M_C=m_C)\\ &+P(\overline{I}|H_i,M_C=m_C,S(s),R(r))\cdot P(\overline{I}|H_{1+},M_C=m_C),\end{aligned}$$

The probability assigned to the receiver is $P(\hat{R}(r)|H_i,M_C=m_C,S(s),R(r))=\frac{1}{M-m_C}$. The events are conditionally independent, hence the probability assigned to the sender-receiver pair (s,r) is the product of the two.

What remains is to calculate the probability for a non-relaying attacker. Let us denote by $\overline{H}_{1+}$ the event that a message does not visit any attacker node as a relay, the complement event of H_{1+}. If $M_C=0$ and $\overline{H}_{1+}$ happens then the attacker does not observe the message. Otherwise, if $\overline{H}_{1+}$ happens but $M_C>0$ then the attacker nodes in the receiver set $\mathscr{M}$ get the multicast message from the last relay node (the one that decides to send the message to the receivers, with probability $1-p_f$). The sender is the last relay node with probability $P(I|\overline{H}_{1+},M_C=m_C,S(s),R(r))=\frac{1}{N-C-M+m_C}$. At the same time the probability that the last relay node is the sender is $P(I|\overline{H}_{1+},M_C=m_C)=\frac{1}{N-C-M+m_C}$. The probability $P(\hat{S}(s)|I,\overline{H}_{1+},M_C=m_C,S(s),R(r))$ is the product of these two probabilities. Since the last relay node removes the receiver set $\mathscr{M}$ from the message, the receiver is hidden among the $N-C$ trusted nodes, and it cannot be the last relay. Therefore, $P(\hat{R}(r)|I,\overline{H}_{1+},M_C=m_C,S(s),R(r))=\frac{1}{N-C-1}$. Again, the two events are independent.

Finally, the sender is not the last relay node with probability $P(\overline{I}|\overline{H}_{1+},M_C=m_C,S(s),R(r))$. In this case the attacker identifies the sender with probability $P(\hat{S}(s)|\overline{I},\overline{H}_{1+},M_C=m_C,S(s),R(r))=\frac{1-P(I|\overline{H}_{1+},M_C=m_C,S(s),R(r))}{N-C-1}$. The receiver is hidden among $N-C-2$ nodes, so that $P(\hat{R}(r)|\overline{I},\overline{H}_{1+},M_C=m_C,S(s),R(r))=\frac{1}{N-C-2}$. The two events are independent.

It can happen that there is an attacker node on the path as a relay (H_{1+}) and there is at least one attacker node specified in the receiver set ($M_C>0$). In this

case it is clear that the attacker assigns higher probability to the actual sender-receiver pair (*s*,*r*) when the message is observed by the relaying attacker node than when it is observed by the attacker node in the receiver set. Since the attacker can recognize if the same message visits several compromised nodes, the attacker does not recalculate the assigned probability when it gets the multicast message. Therefore, we do not provide separate expressions for this case.

Finally, we express $P_{relB}(s,r)$, using the law of total probability conditioning on the first attacker position on the path and on the number M_C of attacker nodes in the receiver set $\mathscr{M}$

$$\begin{aligned} P_{relB}(s,r) = \sum_{m_C} \sum_{i=1}^{\infty} & P(\hat{S}(s),\hat{R}(r)|H_i, M_C = m_C, S(s), R(r)) \\ & \cdot P(H_i|M_C = m_C, S(s), R(r)) \cdot P(M_C = m_C) \\ & + \sum_{m_C \neq 0} P(\hat{S}(s),\hat{R}(r)|\overline{H}_{1+}, M_C = m_C, S(s), R(r)) \\ & \cdot P(\overline{H}_{1+}|M_C = m_C, S(s), R(r)) \cdot P(M_C = m_C). \end{aligned} \tag{13}$$

The relationship anonymity $P_{relM}(s,r)$ for the MP method can be expressed similar to (13). If an attacker node is a relay at position i of the path then the attacker identifies the sender-receiver pair (s,r) as one of the most likely sender-receiver pairs $\mathscr{Q}$ if the the sender appears as the predecessor (on position $i-1$). This happens with probability $P((s,r) \in \mathscr{Q}|H_i, M_C = m_C, S(s), R(r)) = P(I|H_i, M_C = m_C, S(s), R(r))$. The cardinality of the set $\mathscr{Q}$ is equal to $M - m_C$, the number of trusted nodes in the receiver set $\mathscr{M}$, so the probability assigned to the sender-receiver pair (s,r) is $P(\hat{S}(s),\hat{R}(r)|(s,r) \in \mathscr{Q}, H_i, M_C = m_C, S(s), R(r)) = \frac{1}{M-m_C}$.

Otherwise, if the attacker is not a relay node (event $\overline{H}_{1+}$) but is in the receiver set, then it identifies the last relay node as the most likely sender. Hence, the sender-receiver pair is in the set $\mathscr{Q}$ if the sender is the last relay. Thus $P((s,r) \in \mathscr{Q}|\overline{H}_{1+}, M_C = m_C, S(s), R(r)) = P(I|\overline{H}_{1+}, M_C = m_C, S(s), R(r))$. The cardinality of the set $\mathscr{Q}$ equals the number of potential receivers $N-C-1$, so that $P(\hat{S}(s),\hat{R}(r)|(s,r) \in \mathscr{Q}, H_i, M_C = m_C, S(s), R(r)) = \frac{1}{N-C-1}$.

Minstrels When the first attacker node on the path gets the message, the attacker knows the number f_C of attacker nodes that the set of visited nodes was initialized with by the sender. f_C is a realization of the random variable F_C, whose distribution depends on the number f of initialized nodes in $\mathscr{V}$.

In Minstrels the probability that the attacker assigns to a sender-receiver pair does not only depend on the node that the message is received from, i.e., the predecessor p, but also on the contents of the set $\mathscr{V}$ of visited nodes that the message carries. Consequently, the attacker distinguishes between three disjoint

sets of nodes: the predecessor node ($\{p\}$), nodes in the set of visited nodes except the predecessor ($\mathscr{V} \setminus \{p\}$), and nodes not in the set of visited nodes ($\overline{\mathscr{V} \cup \{p\}}$). These sets form a partition of the set of all trusted nodes in the system, and nodes belonging to the same set are equally likely to be the sender (and the receiver). As a shorthand for the universe of distinguishable events we use the notation $\Omega_s = \{s = p, s \in \mathscr{V} \setminus \{p\}, s \in \overline{\mathscr{V} \cup \{p\}}\}$, where, for example, $s = p$ is the event that the predecessor is the sender. Similarly, we define $\Omega_r = \{r = p, r \in \mathscr{V} \setminus \{p\}, r \in \overline{\mathscr{V} \cup \{p\}}\}$ for the distinguishable events regarding the receiver.

If the message visits multiple attacker nodes on its path, the attacker does not gain more information that could increase the probability assigned to the sender-receiver pair (*s*,*r*). Hence, we only consider the first attacker node on the path that gets the message. Given the information on $\mathscr{V}$, f_C, and p available to the attacker, we can use the law of total probability to expand (9) and (10) conditional on the size $||\mathscr{V}|| = v$ of the set of visited nodes, $\omega_s \in \Omega_s$, $\omega_r \in \Omega_r$, and $F_C = f_C$,

$$P_{relB}(s,r) = \sum_{f_C}\sum_{v}\sum_{\omega_s}\sum_{\omega_r} P(\hat{S}(s),\hat{R}(r)|\omega_r,\omega_s,f_C,H_{1+},v,S(s),R(r)) \tag{14}$$

$$\cdot P(\omega_r,\omega_s,f_C,H_{1+},v|S(s),R(r)), \tag{15}$$

$$P_{relM}(s,r) = \sum_{f_C}\sum_{v}\sum_{\omega_s}\sum_{\omega_r} P(\hat{S}(s),\hat{R}(r)|(s,r)\in\mathscr{Q},\omega_r,\omega_s,f_C,H_{1+},v,S(s),R(r)) \tag{16}$$

$$\cdot P((s,r)\in\mathscr{Q}|\omega_r,\omega_s,f_C,H_{1+},v,S(s),R(r)) \tag{17}$$

$$\cdot P(\omega_r,\omega_s,f_C,H_{1+},v|S(s),R(r)). \tag{18}$$

Note that (15) and (18) represent the probability that a message with (s,r) as sender-receiver pair is received by an attacker node and carries particular information. Eq. (17) is the probability that the sender-receiver pair (s,r) is in the set $\mathscr{Q}$, while (14) and (16) are the probabilities that the attacker correctly identifies the sender-receiver pair (s,r).

Before we turn to the calculation of the probability $P(\omega_r,\omega_s,v,f_C,H_{1+}|S(s),R(r))$ we introduce the notation $H(v,f_C|F=f)$ for the joint event $||\mathscr{V}|| = v$, H_{1+}, and $F_C = f_C$ for a given number of initialized nodes f. Clearly, $v \geq f$. The probability of this event can be expressed as

$$P(H(v,f_C|F=f)) = \begin{cases} \frac{C}{N-1} & v=0, F=0 \\ P(F_C=0|F=f)\frac{N-C-1}{N-1}\frac{C}{N-v}\prod_{z=1}^{v-1}\frac{N-C-z}{N-z} & v\geq 1, f=0 \\ P(F_C=f_C|F=f)\frac{C-f_C}{N-v}\prod_{z=f}^{v-1}\frac{N-C+f_C-z}{N-z} & v\geq 1, f>0, \end{cases} \tag{19}$$

where $P(F_C|F=f)$ is the probability that the set of visited nodes is initialized with f_C attacker nodes, given that it is initialized with f nodes by the sender. Due

to the rules of initialization, $f_C \in \{max(0, f-1-(N-2-C)), min(f-1, C)\}$. For $F = 0$ and $F = 1$ there cannot be any initialized attackers, hence $P(F_C = 0|F \in \{0,1\}) = 1$ and $P(F_C > 0|F \in \{0,1\}) = 0$. For $F > 1$ we have

$$P(F_C|F = f) = \begin{pmatrix} f-1 \\ f_C \end{pmatrix} \frac{\prod_{k=2}^{f-f_C}(N-C-k)\prod_{k=0}^{f_C-1}(C-k)}{\prod_{k=2}^{f}(N-k)}. \quad (20)$$

We now turn to the calculation of the probability $P(\omega_r, \omega_s, v, f_C, H_{1+}|S(s), R(r))$, i.e., the probability that the attacker would receive a particular message sent by s to r. If the sender is the predecessor ($s = p$) the receiver cannot be the predecessor, hence $P(r = p, s = p, v, f_C, H_{1+}|S(s), R(r)) = 0$. For the rest of the cases we show the probabilities in a tabular form to improve readability.

For $||\mathcal{V}|| = 0$ and $||\mathcal{V}|| = 1$ there can be no attackers in the set of visited nodes (when received by the first attacker), because if the sender initializes the set of visited nodes it has to include itself in the set. Hence, for $||\mathcal{V}|| = 0$ and $||\mathcal{V}|| = 1$ we have $F_C > 0$ with probability 0. Furthermore, for $||\mathcal{V}|| = 0$ the sender must be the predecessor ($s = p$) and the receiver cannot be in the set of visited nodes ($r \in \overline{\mathcal{V} \cup \{p\}}$). Every other tuple in $\{(\omega_s, \omega_r) : \omega_s \in \Omega_s, \omega_r \in \Omega_r\}$ has probability 0. Table 4 shows the corresponding probability, i.e., the probability that the sender initializes the message with an empty set, and chooses the attacker as next hop. For $||\mathcal{V}|| = 1$ the sender and the receiver cannot both be in the set of visited nodes. Furthermore, if the sender or the receiver is in the set of visited nodes, it must be the predecessor, hence $s \in \mathcal{V} \setminus \{p\}$ and $r \in \mathcal{V} \setminus \{p\}$ have probability 0. Table 5 shows the probabilities for the remaining cases for $||\mathcal{V}|| = 1$. As an example, the second row in the table is the probability that the sender initializes the set empty, forwards the message to the receiver, which then forwards the message to the attacker.

For $||\mathcal{V}|| > 1$ there may or may not be attackers in the set of initialized nodes. Table 6 shows the probabilities for $||\mathcal{V}|| > 1$ when there are no attackers in the set of initialized nodes ($F_C = 0$). When there are attackers in the set of initialized nodes ($F_C > 0$), the sender has to be in the set of visited nodes. Furthermore, if the sender is the predecessor ($s = p$) then the receiver cannot be in the set of visited nodes ($r \in \mathcal{V} \setminus \{p\}$), because this could only happen if the sender had initialized the set of visited nodes with the receiver, but then the receiver would never receive the message. The corresponding probabilities for $||\mathcal{V}|| > 1$ and $F_C > 0$ are shown in Table 7.

Let us now turn to the calculation of the probabilities that the attacker correctly identifies the sender-receiver pair (s, r) used in the Bayesian inference method (14). Given a message received by an attacker node that contains information ($||\mathcal{V}|| = v$, $\omega_s \in \Omega_s$, $\omega_r \in \Omega_r$, and $F_C = f_C$) the attacker would identify (s, r) as the

Table 4: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 0, F_C = 0, H_{1+}|S(s), R(r))$

Ω_s, Ω_r	
$s = p,\ r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=0)P(H(0,0\|F=0))$

Table 5: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 1, F_C = 0, H_{1+}|S(s), R(r))$

Ω_s, Ω_r	
$s = p,\ r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=1)P(H(1,0\|F=1))$
$s \in \overline{\mathscr{V} \cup \{p\}},\ r = p$	$P(F=0)P(H(1,0\|F=0))\frac{1}{N-C-1}$
$s \in \overline{\mathscr{V} \cup \{p\}},\ r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=0)P(H(1,0\|F=0))\frac{N-C-2}{N-C-1}$

Table 6: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| > 1, F_C = 0, H_{1+}|S(s), R(r))$

Ω_s, Ω_r	
$s = p,\ r \in \mathscr{V} \setminus \{p\}$	$P(F=0)P(H(v,0\|F=0))\frac{v-1}{(N-C-1)^2}$
$s = p,$ $r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=0)P(H(v,0\|F=0))\frac{(N-C-v)}{(N-C-1)^2}$ $+P(F=v)P(H(v,0\|F=v))$
$s \in \mathscr{V} \setminus \{p\},$ $r = p$	$P(F=0)P(H(v,0\|F=0))\frac{v-2}{(N-C-1)^2}$ $+\sum_{k=1}^{v-1}P(F=k)P(H(v,0\|F=k))\frac{1}{N-C-k}$
$s \in \mathscr{V} \setminus \{p\},$ $r \in \mathscr{V} \setminus \{p\}$	$P(F=0)P(H(v,0\|F=0))\frac{(v-2)^2}{(N-C-1)^2}$ $+\sum_{k=1}^{v-2}P(F=k)P(H(v,0\|F=k))\frac{v-k-1}{N-C-k}$
$s \in \mathscr{V} \setminus \{p\},$ $r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=0)P(H(v,0\|F=0))\frac{(N-C-v)(v-2)}{(N-C-1)^2}$ $+\sum_{k=1}^{v-1}P(F=k)P(H(v,0\|F=k))\frac{N-C-v}{N-C-k}$
$s \in \overline{\mathscr{V} \cup \{p\}},\ r = p$	$P(F=0)P(H(v,0\|F=0))\frac{(N-C-v)}{(N-C-1)^2}$
$s \in \overline{\mathscr{V} \cup \{p\}},\ r \in \mathscr{V} \setminus \{p\}$	$P(F=0)P(H(v,0\|F=0))\frac{(v-1)(N-C-v)}{(N-C-1)^2}$
$s \in \overline{\mathscr{V} \cup \{p\}},\ r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F=0)P(H(v,0\|F=0))\frac{(N-C-v)(N-C-v-1)}{(N-C-1)^2}$

sender-receiver pair with probability

$$P(\hat{R}(r), \hat{S}(s)|\omega_r, \omega_s, f_C, H_{1+}, v) = \frac{P(\omega_r, \omega_s, v, f_C, H_{1+}|S(s), R(r)) \cdot P(R(r)|S(s)) \cdot P(S(s))}{\sum_{(a,b)} P(\omega_r, \omega_s, v, f_C, H_{1+}|S(a), R(b)) \cdot P(R(b)|S(a)) \cdot P(S(a))} \quad (21)$$

where the summation in the denominator is over all possible non-attacker sender-

Table 7: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| > 1, F_C > 0, H_{1+}|S(s), R(r))$

Ω_s, Ω_r	
$s = p$, $r \in \overline{\mathscr{V} \cup \{p\}}$	$P(F = v)P(H(v, f_C|F = v))$
$s \in \mathscr{V} \setminus \{p\}$, $r = p$	$\sum_{k=f_C+1}^{v-1} P(F = k)P(H(v, f_C|F = k)) \frac{1}{N-C+f_C-k}$
$s \in \mathscr{V} \setminus \{p\}$, $r \in \mathscr{V} \setminus \{p\}$	$\sum_{k=f_C+1}^{v-2} P(F = k)P(H(v, f_C|F = k)) \frac{v-k-1}{N-C+f_C-k}$
$s \in \mathscr{V} \setminus \{p\}$, $r \in \overline{\mathscr{V} \cup \{p\}}$	$\sum_{k=f_C+1}^{v-1} P(F = k)P(H(v, f_C|F = k)) \frac{N-C+f_C-v}{N-C+f_C-k}$

Table 8: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 0, F_C = 0, H_{1+}|S(a), R(b))$

Ω_s, Ω_r, a, b	
$s = p$, $r \in \overline{\mathscr{V} \cup \{p\}}$, $a = s$, $\forall b$	$P(F = 0)P(H(0, 0|F = 0))$

receiver pairs (a, b). $P(S(s))$ is the (a priory) probability that node s sends a message, and $P(R(r)|S(s))$ is the probability that node s selects node r as the receiver of a message. Since the a-priori traffic matrix is homogeneous and attackers are informed about each other, all trusted nodes are believed to be equally likely the sender, $P(S(s)) = \frac{1}{N-C}$, and any trusted node (except the sender) is believed to be chosen equally likely as the receiver, i.e., with probability $P(R(r)|S(s)) = \frac{1}{N-C-1}$. The same observation holds for $P(S(a))$ and $P(R(b))$, so that these probabilities cancel out each other in (21).

We already calculated the numerator of (21), so in order to finish our calculations we only have to express $P(\omega_r, \omega_s, v, f_C, H_{1+}|S(a), R(b))$ and only for the cases when the numerator of (21) is non-zero, and when $a \neq s$ or $b \neq r$.

The attacker can receive a message with an empty set of visited nodes ($||\mathscr{V}|| = 0, F_C = 0$) only if the sender is the predecessor, hence, $P(\omega_r, \omega_s, ||\mathscr{V}|| = 0, F_C = 0, H_{1+}|S(a), R(b)) > 0$ only for $a = s$. Nevertheless, the receiver of the message can be any trusted node $b \neq s$ (we use $\forall b$ as a shorthand notation). The corresponding probability $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 0, F_C = 0, H_{1+}|S(a), R(b))$ is given in Table 8.

The attacker can receive a message with only one node in the set of visited nodes ($||\mathscr{V}|| = 1$), in which case the node in the set is the predecessor. The set could have been sent by the predecessor ($a = p$) or by a node not in the set ($a \in \overline{\mathscr{V} \cup \{p\}}$), but in either case there cannot be any attacker node initialized in the set ($F_C = 0$). The receiver could be any other node ($\forall b$). The probability of receiving such a message $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 1, F_C = 0, H_{1+}|S(a), R(b))$ is given in Table 9.

The probabilities for $||\mathscr{V}|| > 1$ can be obtained following a similar reasoning.

Table 9: $P(\Omega_r, \Omega_s, ||\mathscr{V}|| = 1, F_C = 0, H_{1+}|S(a), R(b))$

Ω_s, Ω_r, a, b	
$s = p$, $r \in \overline{\mathscr{V} \cup \{p\}}$, $a = s$, $\forall b$	$P(F=1)P(H(1,0\|F=1))$
$s = p$, $r \in \overline{\mathscr{V} \cup \{p\}}$, $a \neq s$, $\forall b$	$P(F=0)P(H(1,0\|F=0))\frac{1}{N-C-1}$
$s \in \overline{\mathscr{V} \cup \{p\}}$, $r = p$, $a = r$, $\forall b$	$P(F=1)P(H(1,0\|F=1))$
$s \in \overline{\mathscr{V} \cup \{p\}}$, $r = p$, $a \neq r$, $\forall b$	$P(F=0)P(H(1,0\|F=0))\frac{1}{N-C-1}$
$s \in \overline{\mathscr{V} \cup \{p\}}$, $r \in \overline{\mathscr{V} \cup \{p\}}$, $a \in \{s,r\}$, $\forall b$	$P(F=0)P(H(1,0\|F=0))\frac{N-C-2}{N-C-1}$
$s \in \overline{\mathscr{V} \cup \{p\}}$, $r \in \overline{\mathscr{V} \cup \{p\}}$, $a \notin \{s,r\}$, $\forall b$	$P(F=0)P(H(1,0\|F=0))\frac{N-C-3}{N-C-1}$ $+P(F=1)P(H(1,0\|F=1))$

We refer to [17] for a complete description of the probabilities in order to ease readability.

We now turn to the calculation of the probability (17) that the sender-receiver pair (s,r) is one of the most likely sender-receiver pairs, i.e. $(s,r) \in \mathscr{Q}$, used in the Maximum posteriori method. The sender-receiver pair (s,r) is in the set $\mathscr{Q}$ if the probability $P(\omega_r, \omega_s, v, f_C, H_{1+}|S(a), R(b))$ for every sender-receiver pair $\forall(a,b)$ is less than or equal to $P(\omega_r, \omega_s, v, f_C, H_{1+}|S(s), R(r))$.

The sender-receiver pairs in the set $\mathscr{Q}$ are node pairs that have the same a-posteriori probability of being the actual sender-receiver pair. Hence, when the sender-receiver pair (s,r) is in the set $\mathscr{Q}$ of most likely sender-receiver pairs the probability (16) that the attacker assigns to (s,r) is equal to

$$P(\hat{S}(s), \hat{R}(r)|(s,r) \in \mathscr{Q}, \omega_r, \omega_s, f_C, H_{1+}, v, S(s), R(r)) = \frac{1}{||\mathscr{Q}||}. \qquad (22)$$

4.5.3 Bounds For Relationship Anonymity

In order to have a better understanding of the relationship anonymity provided by the described anonymity networks, we define upper and lower bounds for the relationship anonymity. To obtain the upper bound, we consider that whenever the attacker intercepts a message, it knows the sender-receiver pair with probability $P(\hat{S}(s), \hat{R}(r)|H_{1+}, S(s), R(r)) = 1$. Hence, the bound is equivalent to the probability of having an attacker node on the path $P(H_{1+}|S(s), R(r))$. To obtain the lower bound, we consider that whenever the attacker intercepts a message, it assumes that any trusted pair of nodes is equally likely to be the sender-receiver pair with probability $P(\hat{S}(s), \hat{R}(r)|H_{1+}, S(s), R(r)) = \frac{1}{(N-C)(N-C-1)}$.

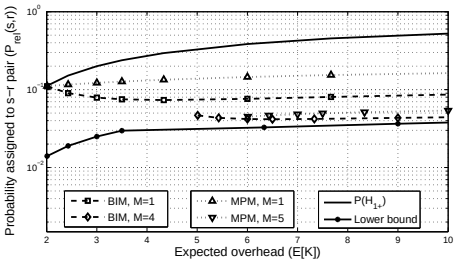

Figure 12: Relationship anonymity vs. overhead for MCrowds, $N = 10$, $C = 1$

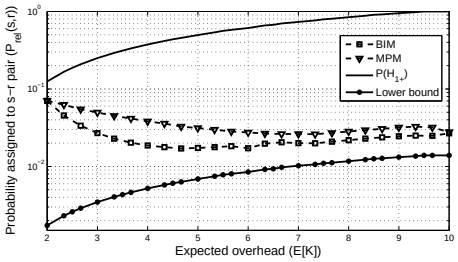

Figure 13: Relationship anonymity vs. overhead for Minstrels, $N = 10$, $C = 1$

4.6 Numerical Results

In the following we use the expressions described above for the Bayesian inference method (BIM) and for the Maximum posteriori method (MPM) to get insight into the relationship anonymity-overhead trade-off provided by MCrowds and by Minstrels. To explore the trade-off, for MCrowds we use the relaying probability $p_f \in (0,1)$ and $M \in \{1,\ldots,N-2\}$, and for Minstrels we use various uniform, binomial, and triangular distributions for the number F of initialized nodes. The attacker's a-priory belief is that the traffic matrix is homogeneous.

Fig. 12 and Fig. 13 show the probabilities $P_{relB}(s,r)$ and $P_{relM}(s,r)$ assigned to the sender-receiver pair as a function of the expected overhead for $C = 1$ attacker node in a system of $N = 10$ nodes. Fig. 12 shows results for MCrowds, and Fig. 13 shows results for Minstrels. A higher value of $P_{relB}(s,r)$ means that the sender-receiver pair is more exposed, i.e., has worse relationship anonymity. The bounds are obtained by finding the distribution of F for Minstrels, and the receiver set size M for MCrowds, that has the lowest $P(H_{1+}|S(s),R(r))$ for a given overhead.

An expected overhead of $E[K] = 2$ corresponds to one relay on average, while $E[K] = N$ is the maximum expected overhead for Minstrels. One would expect that higher overhead always provides better relationship anonymity (i.e., low assigned probability), but surprisingly this is not the case.

For the Bayesian inference method (denoted by *BIM* in the figures), above a certain level of overhead a further increase of the overhead (more relaying) has a negative effect on the anonymity for both anonymity networks. The reason is that as the expected number of relays increases, the probability $P(H_{1+}|S(s),R(r))$ of having an attacker node on the path increases faster than the certainty of the attacker about the identity of the sender-receiver pair decreases.

For the Maximum posteriori method (denoted by *MPM* in the figures) increased overhead improves the anonymity for Minstrels up to a certain level. Interestingly, for MCrowds increased overhead always results in worse anonymity. We also observe that both Minstrels and MCrowds provide worse relationship

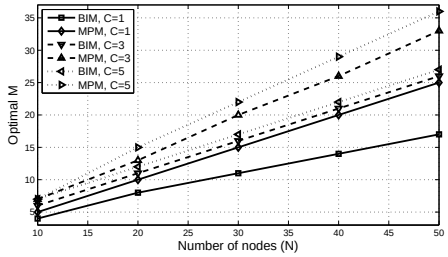

Figure 14: Optimal M vs. number of nodes in the system

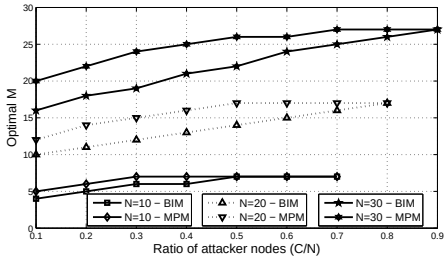

Figure 15: Optimal M vs. ratio of attacker nodes

anonymity for the Maximum posteriori traffic analysis method than for the Bayesian inference traffic analysis method.

For high overhead, the anonymity provided by both anonymity networks approaches its lower bound. Despite the fact that for Minstrels the probability $P(H_{1+}|S(s),R(r))$ of having an attacker node on the path is higher than for MCrowds, Minstrels provides better relationship anonymity. The reason is that Minstrels hides the sender and the receiver among a bigger subset of nodes.

Fig. 12 suggests that MCrowds performs better for larger values of the receiver set size M. This is not true in general. For a larger M the receiver is better hidden but, at the same time, the sender is more exposed because there are fewer potential relays. Hence there should be an optimal receiver set size M. Fig. 14 shows the optimal value of M as a function of the number N of nodes in the system. The optimal receiver set size M increases both with the number of nodes in the system (almost linearly) and with the ratio $\frac{C}{N}$ of attacker nodes. The value of M used in Fig. 12 ($M = 4$ for both BIM and MPM) is in fact optimal for $N = 10$ and $C = 1$.

Fig. 15 shows the optimal receiver set size M as a function of the ratio $\frac{C}{N}$ of attacker nodes in the system. We can see that the optimal value of M is a non-decreasing function of the ratio of attacker nodes. For a given ratio of attacker nodes the optimal receiver set size M for MPM is always greater or equal than the optimal M for BIM. The optimal M for MPM and the optimal M for BIM have the same maximal value. As the system gets larger, the highest optimal value of M for MPM and for BIM is reached at higher values of the ratio of attacker nodes. Hence, with more attacker nodes in the system it is better to increase the receiver set size M if it is lower than the highest optimal value.

Fig. 16 and Fig. 17 show the optimal overhead (where the probabilities $P_{relB}(s,r)$ or $P_{relM}(s,r)$ are the lowest) as a function of the ratio of attacker nodes $(\frac{C}{N})$ for MCrowds and for Minstrels, respectively. For MCrowds, the optimal overhead for both BIM and MPM increases with the system size N. For a given ratio of attacker nodes $\frac{C}{N}$ the optimal overhead for BIM is greater than or equal to the op-

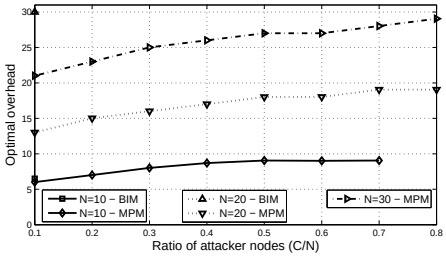

Figure 16: Optimal overhead vs. ratio of attacker nodes for MCrowds

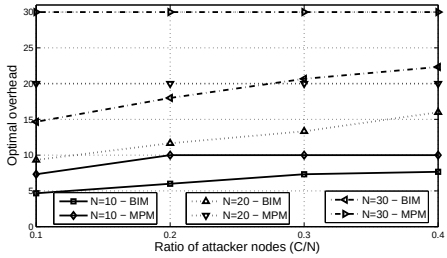

Figure 17: Optimal overhead vs. ratio of attacker nodes for Minstrels

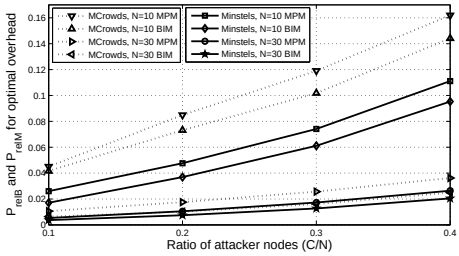

Figure 18: Relationship anonymity for optimal overhead vs. ratio of attacker nodes

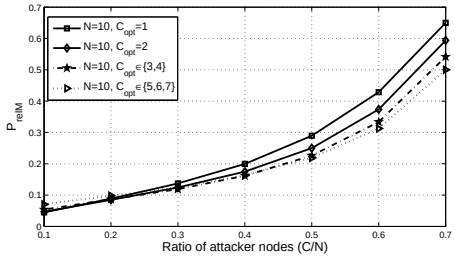

Figure 19: Best relationship anonymity ($P_{relB}(s,r)$) vs. ratio of attacker nodes for MCrowds

timal overhead for MPM. It is interesting to note that for the considered system sizes N the optimal overhead is in the interval $\{2..N\}$ and it increases with the ratio of attacker nodes. For Minstrels, the optimal overhead for BIM increases with the system size N and it is lower than the optimal overhead for MPM. The optimal overhead for MPM is equal to the maximum overhead for Minstrels ($E[K] = N$) except for $N = 10$ and $\frac{C}{N} = 0.1$.

Fig. 18 shows the probabilities $P_{relB}(s,r)$ and $P_{relM}(s,r)$ at the optimal overhead as a function of the ratio of attacker nodes ($\frac{C}{N}$). As the ratio of attacker nodes increases, the probabilities $P_{relB}(s,r)$ and $P_{relM}(s,r)$ increase almost linearly. However, for larger systems the probabilities are lower for the same ratio of attacker nodes. Consequently, with an increase in the system size the attacker needs to corrupt more than proportional number of nodes in order to achieve the same values of $P_{relB}(s,r)$ and $P_{relM}(s,r)$. Hence, both for Minstrels and for MCrowds, it is always beneficial to have more nodes in the network for the same ratio of attacker nodes $\frac{C}{N}$.

In practice the ratio of the attacker nodes is not known by the system designer, hence the anonymity network must be inevitably optimized for an unknown pa-

rameter. In Fig. 19 we investigate the sensitivity of the relationship anonymity to misestimating the ratio of attacker nodes. Fig. 19 shows the probability $P_{relM}(s,r)$ as a function of the actual ratio $\frac{C}{N}$ of attacker nodes for a system size of $N = 10$ nodes. The expected overhead is selected to be optimal for various ratios of attacker nodes, from $\frac{C}{N} = 0.1$ to $\frac{C}{N} = 0.7$. Interestingly, $P_{relM}(s,r)$ is less sensitive to the actual ratio of attacker nodes when the anonymity network is optimized for a higher ratio of attacker nodes. The anonymity network optimized for a lower ratio of attacker nodes performs worse for higher $\frac{C}{N}$ ratios than the anonymity network optimized for a higher ratio of attacker nodes for lower $\frac{C}{N}$ ratios. Therefore, it is better to optimize the anonymity network for a higher ratio of attacker nodes than the actual ratio. We observed the same behavior for bigger system sizes N, while for BIM it is always better to have higher overhead (more relaying). For Minstrels, we observed the same behavior for BIM, and for MPM and system size $N = 10$. In the case of MPM and a larger system size, the optimal overhead is always the highest overhead for Minstrels ($E[K] = N$).

4.7 Conclusion

The presented results lead us to the following interesting conclusions. First, best relationship anonymity might not be achieved at the highest possible overhead. The optimal overhead depends on the anonymity network, traffic analysis method, system size, and the number of attacker nodes. Second, for an attacker it is always better to use the Maximum posteriori method than the Bayesian inference method for traffic analysis in case of the MCrowds and the Minstrels anonymity networks. Third, MCrowds and Minstrels can achieve better relationship anonymity in bigger systems, but at the price of higher overhead. Fourth, when the number of attacker nodes is unknown MCrowds and Minstrels are less sensitive if they are optimized for a high ratio of attacker nodes. Fifth, for MCrowds it always beneficial to have more than one node specified as the receiver of the message ($M > 1$). Finally, for the considered system sizes N and ratios of attacker nodes ($\frac{C}{N}$), Minstrels achieves better relationship anonymity than MCrowds.

It is also important to mention that for the purpose of ICCC already a limited amount of overhead could provide a level of relationship anonymity that would make traffic analysis attacks by compromized control center software very difficult.

References

[1] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.

[2] "Comptuterworld, DHS to review report on vulnerability in west coast power grid." [Online]. Available: http://www.computerworld.com/s/article/9138017

[3] "NewScientist, How to short-circuit the US power grid." [Online]. Available: http://www.newscientist.com/article/mg20327255.900-how-toshortcircuit-the-us-power-grid.html

[4] "ITWORLD, Power grid is found susceptible to cyberattack." [Online]. Available: http://www.itworld.com/security/64770/power-grid-found-susceptiblecyberattack

[5] S. Spoonamore and R. Krutz, "Smart grid and cyber challenges: National security risks and concerns of smart grid," 2009.

[6] S. M. Amin, "Smart grid: Opportunities and challenges toward a stronger and smarter grid," in *MIT Energy Conference Accelerating Change in Global Energy*, 2009.

[7] "Forbes, congress alarmed at cyber-vulnerability of power grid." [Online]. Available: http://www.forbes.com/2008/05/22/cyberwar-breachgovernment-tech-security_cx_ag_0521cyber.html

[8] "Cnn, sources: Staged cyber attack reveals vulnerability in power grid." [Online]. Available: http://www.cnn.com/2007//US/09/26/power.at.risk/index.html

[9] "The wall street journal, electricity grid in u.s. penetrated by spies." [Online]. Available: http://online.wsj.com/article/SB123914805204099085.html

[10] D. Kirschen and F. Bouffard, "Keep the lights on and the information flowing," *IEEE Power and Energy Magazine*, vol. 7, no. 1, p. 5060, 2009.

[11] M. Negrete-Pincetic, F. Yoshida, and G. Gross, "Towards quantifying the impacts of cyber attacks in the competitive electricity market environment," in *IEEE Power Tech Conference*, 2009.

[12] J. W. Wang and L. L. Ronga, "Cascade-based attack vulnerability on the us power grid," p. 13321336, 2009.

[13] "Viking Project." [Online]. Available: http://www.vikingproject.eu

[14] "D1.1: SCADA system security requirements specification, VIKING deliverable,."

[15] G. Dán, H. Sandberg, G. Björkman, and M. Ekstedt, "Challenges in power system information security," *IEEE Security and Privacy Magazine, to appear*.

[16] O. Vuković, K. Sou, G. Dán, and H. Sandberg, "Network-layer protection schemes against stealth attacks on state estimators in power systems," in *Proc. of IEEE SmartGridComm*, Oct. 2011.

[17] O. Vuković, G. Dán, and G. Karlsson, "On the trade-off between relationship anonymity and communication overhead in anonymity networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Jun. 2011.

[18] T. Dierks and E. Rescorla, "rfc 5246, The Transport Layer Security (TLS) protocol, Version 1.2," Aug. 2008. [Online]. Available: http://www.ietf.org

[19] P. Tsang and S. Smith, "YASIR: A low-latency, high-integrity security retrofit for legacy scada systems," in *Proc. of IFIP/TC11 International Information Security Conference*, 2008.

[20] A. Monticelli, "Electric power system state estimation," *Proc. of the IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[21] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Marcel Dekker, Inc., 2004.

[22] Y. Liu, P. Ning, and M. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. of the 16th ACM conference on Computer and Communications Security (CCS)*, 2009, pp. 21–32.

[23] A. Teixeira, G. Dán, H. Sandberg, and K. H. Johansson, "A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator," in *Proc. IFAC World Congress*, Aug. 2011.

[24] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Proc. of IEEE SmartGridComm*, Oct. 2010.

[25] J. Ignizio and T. Cavalier, *Linear Programming*. Prentice Hall, Englewood Cliffs, NJ, 1994.

[26] H. Mueller, "Outage analysis: Italy," *Network Manager News, News and Information for Users of Network Manager Worldwide*, vol. 2, no. 1, pp. 1–3, 2004.

[27] P. Natarajan, N. Ekiz, P. Amer, and R. Stewart, "Concurrent multipath transfer during path failure," *Computer Communications*, vol. 32, no. 15, 2009.

[28] D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," *Commun. of the ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[29] K. Sampigethaya and R. Poovendran, "A survey on mix networks and their secure applications," *Proc. of the IEEE*, vol. 94, no. 12, pp. 2142–2181, 2006.

[30] A. Pfitzman and M. Köhntopp, "Anonymity, unobservability, and pseudonymity - a proposal for terminology," in *Anonymity*, 2000, pp. 1–9.

[31] P. Syverson, D. Goldschlag, , and M. Reed, "Anonymous connections and onion routing," in *Proc. IEEE Symp. on Security and Privacy*, May 1997, pp. 44–54.

[32] M. Reiter and A. Rubin, "Crowds: Anonymity for web transactions," *ACM Trans. Inform. Syst. Security*, pp. 66–92, 1998.

[33] V. Shmatikov and M. H. Wang, "Measuring relationship anonymity in mix networks," in *Proc. of WPES*, 2006.

[34] G. Danezis, C. Díaz, E. Käsper, and C. Troncoso, "The wisdom of crowds: attacks and optimal constructions," in *Proc. of ESORICS*, 2009.