Data Security and Privacy in Transactive Energy Markets

by

Daniel Sousa-Dias

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the degree of

Master of Computer Science

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Examining Committee

The following served on the Examining Committee for this thesis.

Internal Members: Javad Fattahi
Assistant Professor, School of Electrical Engineering & Computer Science
University of Ottawa

Melike Erol-Kantarci
Professor, School of Electrical Engineering & Computer Science
University of Ottawa

Supervisors: Daniel Amyot
Professor, School of Electrical Engineering & Computer Science
University of Ottawa

John Mylopoulos
Adjunct Professor, School of Electrical Engineering & Computer Science
University of Ottawa
Professor Emeritus, University of Toronto, Canada
Professor Emeritus, University of Trento, Italy

Ashkan Rahimi Kian
Adjunct Professor, School of Electrical Engineering & Computer Science
University of Ottawa
Founder & CTO, IEMS Solution Ltd

Declaration of Authorship

I hereby certify that I, Daniel Sousa-Dias, am the primary author of this thesis and of the three papers contained within it. I am aware of the University of Ottawa regulations concerning plagiarism, including those regarding consequent disciplinary actions. Any use of the works of any other author, in any form, is properly acknowledged at their point of use.

The concept for the literature review *A Review of Cybersecurity Concerns for Transactive Energy Markets* (Chapter 3) was developed with my supervisors Dr. Daniel Amyot, Dr. John Mylopoulos, and Dr. Ashkan Rahimi-Kian. On the basis of the review, I selected the topics to be addressed in the second and third papers of my thesis. My supervisors contributed editing and feedback for all three papers and the thesis. Dr. Masoud Bashari, an industrial partner who also co-authored two papers, reviewed and helped edit the papers *Cyclic Homomorphic Encryption Aggregation (CHEA) – a Novel Approach to Data Aggregation in the Smart Grid* (Chapter 4) and *Enhancing Trust in Transactive Energy with Individually Linkable Pseudonymous Trading Using Smart Contracts* (Chapter 5).

My paper co-authors agree with this declaration of authorship.

Abstract

Innovation in power generation, storage, and information technology have created new opportunities in grid management. Advanced metering equipment and smart grid infrastructure enable energy providers to operate more efficiently and cost effectively through better reporting and prediction. While these improvements to the traditional energy management model are important, opportunities for further gains in economic and energy efficiency have been discovered in distributed systems. The next generation of energy market will be *transactive*, enabling prosumers, consumers equipped with energy generation or storage devices, to trade energy directly between each other. This capacity will result in increased price efficiency, reduced transmission distances, and better integration of renewable energy sources into the grid.

The incentive to modernize the grid is clear, but increased information flow demands heightened security measures to protect consumer safety and trust. Many proposed transactive energy market solutions use distributed ledger technology, or blockchain, along with smart contracts to underpin their energy auctions. Blockchain technology has many desirable security properties that make it suitable for handling trades. However, security gaps remain in the processes not managed by the blockchain.

The goals of this thesis are to discover the cybersecurity gaps present in transactive energy market systems, identify the areas most in need of improvement, and propose solutions to some of those areas.

A thorough review of the literature led us to identify fourteen cybersecurity threat categories. We selected two processes that were significantly affected by these threats and that had few solutions addressing them. These processes are: energy usage data collection and market anonymity.

In addition to the literature review, this thesis contributes secure, privacy-preserving schemes for each of these processes, namely *Cyclic Homomorphic Encryption Aggregation* (CHEA) and *Individually Linkable Pseudonymous Trading Scheme* (ILPTS). Both schemes improve security and efficiency by reducing infrastructural requirements and increasing decentralization. Formal analysis found that both solutions successfully achieve their security design goals, while performance simulations found that CHEA performs well compared to similar data aggregation schemes from the literature.

Acknowledgements

I would first like to thank my supervisors, Dr. Daniel Amyot, Dr. John Mylopoulos, and Dr. Ashkan Rahimi-Kian, all of whose guidance and expertise were instrumental in the production of this thesis.

Second, the wonderful research team curated for the Symboleo project: Dr. Luigi Logrippo, Dr. Marco Roveri, Dr. Amal Anda, Regan Meloche, Sofana Alfuhaid, Aidin Rasti, and Alireza Parvizimosaed. Their feedback and patient attentiveness each week were well appreciated as I worked towards this goal, not to mention the camaraderie shared during the times we were happily able to meet in person.

Finally, thanks to Dr. Javad Fattahi and Dr. Melike Erol-Kantarci of uOttawa, and the talented team at IEMS, Dr. Masoud Bashari and Ehsan Saradar, for contributing their expertise to help shape my work.

This work was funded by the Ontario Research Fund project *CyPreSS: Software Techniques for the Engineering of Cyber-Physical Systems*, with support from NSERC Discovery Grant, fund 610877, titled *Engineering Requirements for Cyber-Physical Systems*.

Dedication

I dedicate this thesis to my loving family, without whose support this would not have been possible.

To my brother, Antonio; my sister, Jordan; and my parents Donna and Pedro.

Special thanks to Sarah, Bo, Maddie, Anika, Than, Darby, Dani, Emily, and Dr. B. Patrick. For believing in me, and everything else.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMI Advanced Metering Infrastructure

BBPS Blockchain-Based Patch System

BESS Battery Energy Storage System

CHEA Cyclic Homomorphic Encryption Aggregation

DER Distributed Energy Resource

DoS Denial of Service

DSO Distributed System Operator

ETSE Energy Trading and Security Enhancement

EV Electric Vehicle

EVE Electron Volt Exchange

FDI False Data Injection

FDIA False Data Injection Attack

FN Fog Node

HE Homomorphic Encryption

HEM Home Energy Management

ICT Information and Communication Technology

ID Identifier

ILPTS Individually Linkable Pseudonymous Trading Scheme

KGC Key Generation Center

MC Microgrid Controller

P2P Peer-to-Peer

PID Pseudonym ID

RES Renewable Energy Sources

SC Smart Contract

SEL Security Enhancement Layer

SG Smart Grid

SM Smart Meter

SPOF Single Point of Failure

TA Trusted Authority

TE Transactive Energy

TEM Transactive Energy Market

TTP Trusted Third Party

VPP Virtual Power Plant

Chapter 1

Introduction

1.1 Overview and Context

Several concurrent technological trends serve as the groundwork for the technology that is the focal point of this thesis, *Data Security and Privacy in Transactive Energy Markets*.

The first of these is the democratization of computational processing. Computer chips are getting cheaper, smaller, and more powerful, all at once. This trend has stayed consistent for the past 25 years, and is expected to do so for the coming decades.

Cheaper, smaller, and more powerful chips have enabled manufacturers and industrial designers to imbue a broad array of household products with computational abilities that would previously have been infeasible or not worth the expense. These "smart" devices include smart refrigerators, smart televisions, smart appliances (dishwashers, laundry machines, etc.), as well as other home automation devices like robotic vacuums, smart blinds, and smart lights. In combination with the rise in internet access and wireless communication technology, these new devices can be leveraged to produce powerful home automation workflows that enhance consumer convenience and free time.

Simultaneously, the energy distribution landscape is changing. The proliferation of consumer-driven power generation using renewable energy sources, such as solar panels and wind turbines, has necessitated novel approaches to energy management [11]. The computerization of metering technology – formally known as *advanced metering infrastructure* (AMI) – and the development of the smart grid have enabled more advanced energy management techniques, including higher-frequency billing, predictive generation, and improved automated demand response [14, 5].

Another significant revolution in grid management is the development of microgrids and transactive energy (TE) [3]. These dynamic energy markets enable distributed grid management and trading directly between consumers who can produce or store excess energy – hereafter referred to as *prosumers* [6].

Transactive energy markets (TEMs) promise improved economic efficiency, robustness to centralized generation failure, decreased transmission distances, and better integration of renewable energy sources and battery energy storage systems into the power grid [2].

These advances come with a price. Increased information density and flow create new privacy and security concerns relative to traditional energy markets [1]. A common solution proposed by researchers is to build the market infrastructure on *distributed ledger technology* (DLT), including blockchain technology [8]. Building on distributed applications is a sensible choice since it aligns with the principles of microgrids and transactive energy, which aim to improve robustness, security, privacy, and reliability through decentralization [7].

Despite its many benefits, distributed ledger technology is not a panacea. TEM operations outside of prosumer trading can still be vulnerable, and DLT can even introduce new attack surfaces.

It was our goal with this thesis to enhance the security posture of transactive energy markets across the two cyber layers: network and software. The aim is to improve security such that transactive energy technology can be deployed widely, with (enhanced) consumer trust and regulatory approval, so that the benefits of this technology can be fully realized.

We selected two specific, high-priority security items [12, 13, 4, 10, 9] and designed systems that aimed to address them within the transactive energy environment.

The products of this thesis, then (described in more detail in Section 1.3), are:

- A literature review of existing security concerns, solutions that have been proposed, and an evaluation of which concerns should be of highest priority going forward.
- Two solution blueprints, each aiming to address a high priority concern from the previously described review.

The goal with these solutions was to create robust designs that addressed security gaps in the existing transactive energy literature. The papers outlining these designs serve as blueprints that should facilitate real-world implementation by either the industry or researchers. The solutions are designed to function as modules within broader TEM implementations, specifically addressing the security gaps they are designed for.

This process will be guided by the research questions outlined in the following section. As this is a thesis by article, results will be presented in the articles contained within, with a more formal discussion of the research questions and thesis objective to follow in the final discussion chapter.

1.2 Research Questions

Four thesis-level research question are identified here. As each of the articles has its own local research questions, the thesis-level ones are labelled TRQ1 to TRQ4 to avoid confusion.

1.2.1 Thesis Research Question TRQ1

What privacy and security concerns exist in transactive energy implementations that are not currently being adequately addressed?

This question motivates:

- The discovery of security and privacy vulnerabilities in the transactive energy landscape.
- Assembly of proposed and existing solutions for these vulnerabilities.
- Analysis of the gaps that are found to exist, especially security or privacy vulnerabilities that are not well addressed.

The analysis should consider remaining concerns from the perspective of coverage as well as severity. A concern with few or no proposed solutions would be considered to be low coverage and therefore more deserving of attention, and vice versa. Severity is determined by the impact that an attack may have if it is carried out successfully, with consideration of its likelihood to occur – that is, a highly-impactful attack that requires unrealistic coordination is not as severe as a less serious attack that can be carried out easily.

1.2.2 Thesis Research Question TRQ2

Can systems be created to improve a selection of these concerns and what properties would they need to have?

This question motivates the development of novel proposals to address concerns that fall in the aforementioned gap, ideally ones with few proposed solutions and/or particularly severe consequences.

This is done first by analyzing the potential solution space, mapping the field of threat vectors associated with a selected strategy, and defining a set of properties that must be adhered to by a valid solution.

1.2.3 Thesis Research Question TRQ3

How well do the systems proposed align with their design and security goals?

In this question, we presume that an attempt is made to design a system that meets the specifications outlined in TRQ2. Here, then, we will evaluate the success of the system in achieving these design and security goals. This will involve outlining the details of the system and demonstrating that these processes do in fact satisfy desired privacy and security properties.

1.2.4 Thesis Research Question TRQ4

Are these proposed systems satisfactory in terms of performance and compared to competing solutions?

This research question addresses the efficacy of the solutions proposed. In order to answer this question we perform a number of evaluative measures such as:

- Mathematical performance analysis.
- Empirical performance analysis by simulation.
- Comparison to similar works.

1.3 Contributions

The contributions of this thesis are structured along three papers:

1. Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, and John Mylopoulos. A review of cybersecurity concerns for transactive energy markets. *Energies*, 16(13), 2023. doi:http://doi.org/10.3390/en16134838.

2. Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, Masoud Bashari, and John Mylopoulos. Cyclic Homomorphic Encryption Aggregation (CHEA) – a Novel Approach to Data Aggregation in the Smart Grid. Submitted to *Energies*, 2023. doi:https://doi.org/10.3390/en17040878

3. Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, Masoud Bashari, and John Mylopoulos. Enhancing Trust in Transactive Energy with Individually Linkable Pseudonymous Trading Using Smart Contracts. *Review pending.*

The first and second papers have been published, with the first receiving the Editor's Choice distinction. The third paper is currently under review.

1.3.1 A Review of Cybersecurity Concerns for Transactive Energy Markets

We contribute a rigorous analysis of DLT-based TEM literature in this paper, published in the journal *Energies* [12]. This paper draws evidence from 28 peer-reviewed papers to develop a coalition of security and privacy threats particular to the research domain.

We provide a thorough description and explanation of each of the security and privacy threats, along with potential methods of execution and the impact they may have on prosumers. We also provide a comprehensive compilation of existing solutions to each of these threats (where available), along with resources and descriptions. Finally, we perform an analysis of the gaps in threat coverage, offering priority evaluations for each potential area of research.

With this paper, we contribute a useful reference for TEM researchers and security researchers who may find threats and solutions they had not considered, or pursue a novel research direction based on our priority analysis and recommendations.

TEM industrialists may also consult it to discover solutions to problems they encounter or threats that may have been overlooked in order to improve their product security stature and value proposition.

1.3.2 Cyclic Homomorphic Encryption Aggregation (CHEA) – a Novel Approach to Data Aggregation in the Smart Grid

In this paper, we present CHEA, a solution for aggregating energy usage data in a TE environment.

For TE researchers, CHEA can provide a solution to a component of a TEM system or a jumping off point for further research (we discuss potential improvements in the paper). There is also potential for CHEA to be adapted to other domains, a possible use case for researchers in other fields.

For industrialists, CHEA is a fully-specified aggregation solution that can be implemented without additional network components, enabling security and privacy enhancement in concert with lowered costs and increased efficiency.

1.3.3 Enhancing Trust in Transactive Energy with Individually Linkable Pseudonymous Trading Using Smart Contracts

In this paper, we present the *Individually Linkable Pseudonymous Trading Scheme*, or ILPTS. This scheme aims to solve the fundamental incongruity between TE market privacy and consumer safety.

Like CHEA, ILPTS presents a novel architecture that can likely be iterated upon to refine it beyond its initial design. This offers an interesting research direction for TE and distributed system researchers. TE researchers can also implement ILPTS into a larger TEM framework to enhance market security and reduce the need for automated punishment protocols such as disconnecting power.

Similarly, industrial TE designers can use ILPTS to offer enhanced security and privacy to their customers and superior adherence to regulations. As well, ILPTS offers peace of mind that defrauded customers will be made whole and that access to power will be reliably maintained.

1.4 Thesis Structure

The structure of the thesis is as follows:

- Chapter 2 provides a description of the methodology employed at each stage of the thesis and as a whole.

- Chapter 3 contains a reformatted (but otherwise unchanged) version of the paper *A Review of Cybersecurity Concerns for Transactive Energy Markets*, which provides a review of transactive energy literature and analysis of the security concerns raised throughout.

- Chapter 4 presents the paper *Cyclic Homomorphic Encryption Aggregation (CHEA) – a Novel Approach to Data Aggregation in the Smart Grid*, a proposed solution to the energy usage data privacy concern raised in the literature review.

- Chapter 5 presents the paper *Enhancing Trust in Transactive Energy with Individually Linkable Pseudonymous Trading Using Smart Contracts*. This paper presents a proposed solution to the market privacy problem also found in the literature review. This paper has not yet been submitted for publication.

- Chapter 6 is a discussion, summarization, and analysis of the three papers, the conceptual connections between them, and an evaluation of the research questions outlined earlier in the Introduction.

- Finally, Chapter 7 concludes the thesis with a reflection on the research and contributions, and discussion of future research directions.

Bibliography

[1] Mario Baptista, Nuno Silva, Nicola Nostro, Tommaso Zoppi, and Andrea Ceccarelli. *STECA – Security Threats, Effects and Criticality Analysis: Definition and Application to Smart Grids*, pages 167–182. River Publishers, 09 2022. ISBN 9781003337485. doi:10.1201/9781003337485-8.

[2] Tao Chen, Qais Alsafasfeh, Hajir Pourbabak, and Wencong Su. The next-generation U.S. retail electricity market with customers and prosumers–a bibliographical survey. *Energies*, 11(1), 2018. doi:10.3390/en11010008.

[3] William Cox and Toby Considine. Structured energy: Microgrids and autonomous transactive operation. In *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6, Feb 2013. doi:10.1109/ISGT.2013.6497919.

[4] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[5] Ateeb Hassan, Hadi Nabipour Afrouzi, Chua Hong Siang, Jubaer Ahmed, Kamyar Mehranzamir, and Chin-Leong Wooi. A survey and bibliometric analysis of different communication technologies available for smart meters. *Cleaner Engineering and Technology*, 7:100424, 2022. ISSN 2666-7908. doi:10.1016/j.clet.2022.100424.

[6] M. Jayachandran, K. Prasada Rao, Ranjith Kumar Gatla, C. Kalaivani, C. Kalaiarasy, and C. Logasabarirajan. Operational concerns and solutions in smart electricity distribution systems. *Utilities Policy*, 74:101329, 2022. ISSN 0957-1787. doi:10.1016/j.jup.2021.101329.

[7] Mohsen Khorasany, Yateendra Mishra, and Gerard Ledwich. A decentralised bilateral energy trading system for peer-to-peer electricity markets. *IEEE Transactions on Industrial Electronics*, 06 2019. doi:10.1109/TIE.2019.2931229.

[8] Desen Kirli, Benoit Couraud, Valentin Robu, Marcelo Salgado-Bravo, Sonam Norbu, Merlinda Andoni, Ioannis Antonopoulos, Matias Negrete-Pincetic, David Flynn, and Aristides Kiprakis. Smart contracts in energy systems: A systematic review of fundamental approaches and implementations. *Renewable and Sustainable Energy Reviews*, 158:112013, 2022. ISSN 1364-0321. doi:10.1016/j.rser.2021.112013.

[9] Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker. Inferring personal information from demand-response systems. *IEEE Security & Privacy*, 8(1):11–20, 2010. doi:10.1109/MSP.2010.40.

[10] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[11] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23, 09 2017. doi:10.1109/RWEEK.2017.8088642.

[12] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, and John Mylopoulos. A review of cybersecurity concerns for transactive energy markets. *Energies*, 16(13), 2023. ISSN 1996-1073. doi:10.3390/en16134838.

[13] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, Masoud Bashari, and John Mylopoulos. Cyclic homomorphic encryption aggregation (chea)—a novel approach to data aggregation in the smart grid. *Energies*, 17(4), 2024. ISSN 1996-1073. doi:10.3390/en17040878.

[14] Noelia Uribe-Pérez, Luis Hernández, David De la Vega, and Itziar Angulo. State of the art and trends review of smart metering in electricity grids. *Applied Sciences*, 6 (3), 2016. doi:10.3390/app6030068.

Chapter 2

Methodology

This thesis is comprised primarily of three papers: a literature review (*A Review of Cybersecurity Concerns for Transactive Energy Markets*); an energy usage data privacy proposal (*Cyclic Homomorphic Encryption Aggregation (CHEA) - a Novel Approach to Data Aggregation in the Smart Grid*; and a market privacy proposal (*Enhancing Trust in Transactive Energy with Individually Linkable Pseudonymous Trading Using Smart Contracts*).

In the literature review, we seek to answer the first research question (TRQ1) in detail and thus provide motivation for the second (TRQ2) and third (TRQ3) thesis research questions. TQR4 evaluates the developments presented in TQR3.

2.1 Search and Literature Review

The literature review is structured such that its main research questions 1.RQ1-1.RQ3 are analogous to the three main points of TRQ1: the discovery of vulnerabilities (1.RQ1), compilation of proposed solutions (1.RQ2), and an analysis of the gaps between the two (1.RQ3).

We approach TRQ1 first by searching the SCOPUS database for publications related to transactive energy markets. 104 papers were considered, with 28 ultimately selected for further analysis based on their relevance to security topics within the domain.

From these papers, and several discovered through forward and backward snowballing [7], we extracted 14 categories of security concerns that pertain specifically to TEMs. We then combed the material for proposed or existing solutions to these concerns (which were sometimes but not always mentioned in the same paper as the concern). Finally, we performed

our own analysis of the material to determine both: a) the adequacy of the proposed solutions, if any; and b) the severity of the concern. This enabled us to ultimately provide recommendations for future work in the area based on these two axes, prioritizing more serious concerns with fewer solutions over less severe concerns with more coverage.

It was the results of this analysis, i.e., these recommendations, that we used to select the concerns to address for the thesis research question TRQ2.

2.2 Threat Selection

The concerns we chose to attempt to address were energy usage data privacy and distributed market privacy.

Energy usage data is widely understood to present significant risks to the consumer if leaked [1, 4, 6, 3]. This data can be used to determine facts about a household and its residents such as: activities being performed in the home, location of residents, who is or is not home, and the profile of electronics and appliances within the home [1, 4, 6]. These inferences can be further exploited to facilitate serious attacks, such as burglary or kidnapping.

In terms of severity: energy usage data presents significant risks to the consumer, not only from a privacy perspective but also from a physical safety perspective. Indeed, many argue that privacy of such data must be enshrined in law [6, 3], and in some cases it already has been set as a precedent [9].

Thus, solving the problem of energy usage data privacy contributes to: the safety of consumers using a transactive energy market, the trust that consumers have in this new technology, and regulatory aherence by parties seeking to introduce such markets (such as utility companies).

From a coverage perspective, energy usage data privacy was largely ignored in the TE literature at the time of writing, making it an excellent candidate for novel solutions.

Market privacy is another serious concern, both for user privacy and consumer trust in TE.

Having a user's real identity on the transaction ledger presents the potential for similar attacks mentioned earlier by analyzing their energy purchasing behaviour [6, 1, 4]. Anonymity on the transaction ledger is even more important than in traditional market models, since the ledger is public and thus all transactions can freely be viewed and analyzed by any party in the system; this information can be used to perform statistical

attacks (described earlier), to uncover personal or identifying information about market participants, and to facilitate phishing attacks against participants. This is significant as industry reports indicate that phishing attacks are on the rise in the energy sector [8].

While some market models did make an attempt to address this issue, the solutions were often drawn from existing blockchain-based privacy solutions, such as coin mixing [2, 10], and ignored the specific needs introduced by the energy setting. In particular, systems either offered full anonymity to users, ignoring the reality that identification may be desirable or necessary under some circumstances, or trusted the DSO to maintain user privacy. From a coverage perspective, no proposal in the literature addressed the problem holistically.

Both security/privacy concerns were analyzed to determine the feature space that a solution would have to cover. In addition, we had a clear set of design goals related to increasing decentralization, reducing points of failure, and operating without implicit trust.

2.3 Solution Design

We studied the field of solutions relevant to each concern to understand the existing conventions. Many of these, particularly the aggregation solutions, were not present within the TE literature but existed outside it. The related work and background sections of each solution paper correspond to this component of TRQ2. Analysis of the threat was a thesis-level concern and would not have made thematic sense to include in a journal paper presenting the solution. However, the design goal section of the ILPTS paper contributes to TRQ2.

We also gathered an array of techniques both from within the field of solutions and from outside sources. These techniques varied considerably in terms of purpose and area of research; some of the ones that ended up being useful include: *homomorphic encryption* (HE), secret sharing schemes, and cryptographic list shuffling.

With the techniques, design goals, and existing customs established, we set out to design our own solutions to each concern, namely CHEA for energy usage data privacy and ILPTS for market privacy. This was a lengthy and iterative process, involving discussions, brainstorming, trial and error, and partial implementations for verification. The solutions evolved even as their papers were written, in some cases due to discovery of an improvement, but in most cases due to the formalization process causing the surfacing of a detail that had not been considered.

2.4 Solution Evaluation

Both solutions (CHEA and ILPTS) were evaluated for their adherence to their design goals. Formal descriptions and analyses of the information flows and mathematical and cryptographic operations involved were used to justify their resilience to certain attacks and general security properties. These analyses comprise TRQ3, and are represented in the corresponding papers under the results section (ILPTS) and security analysis section (CHEA), although a more direct analysis is performed in Chapter 6.

To answer TRQ4 and inform the performance results section of the CHEA journal paper, we designed a custom simulation to test the scheme under different configurations. To compare performance directly with other schemes, we employed a technique introduced by Liu et al. [5]. ILPTS was not tested for performance, but was contrasted with related schemes which contributed to TRQ4.

The CHEA simulation was written in C on a Unix-based system. It consists of two separate programs, one that represents the DSO and one that represents a smart meter. A population of smart meters can be run alongside one DSO and commands can be used to test different topologies and other properties. The simulation was primarily used for validation of the concept and testing how well the solution scaled with increasing population size. Speed comparisons to other schemes were performed by benchmarking intense cryptographic operations on the testing machine (2017 MacBook Pro with Intel i5 processor @ 3.1GHz and 16GB ram) and extrapolating these results for our scheme and the ones we compared it with.

The paper comprising the literature review was published in the *Energies* journal. The paper containing the first solution (CHEA) has been published in *Energies* as well; the paper containing the second solution (ILPTS) has been submitted and is awaiting review.

Bibliography

[1] Muhammad Rizwan Asghar, György Dán, Daniele Miorandi, and Imrich Chlamtac. Smart meter data privacy: A survey. *IEEE Communications Surveys & Tutorials*, 19 (4):2820–2835, 2017. doi:10.1109/COMST.2017.2720195.

[2] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[3] Astrid Kalkbrenner and Jason Unger. Energy consumption data and rights to privacy: Climate change mitigation policy, privacy and the "internet of things" in alberta. *Environmental Law Centre of Alberta*, 1(1), Jan 2018.

[4] Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker. Inferring personal information from demand-response systems. *IEEE Security & Privacy*, 8(1):11–20, 2010. doi:10.1109/MSP.2010.40.

[5] Shuanggen Liu, Yaowei Liu, Wandi Liu, and Yuchen Zhang. A certificateless multi-dimensional data aggregation scheme for smart grid. *Journal of Systems Architecture*, 140:102890, 2023. ISSN 1383-7621. doi:10.1016/j.sysarc.2023.102890.

[6] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[7] Erica Mourão, João Felipe Pimentel, Leonardo Murta, Marcos Kalinowski, Emilia Mendes, and Claes Wohlin. On the performance of hybrid search strategies for systematic literature reviews in software engineering. *Information and Software Technology*, 123:106294, 2020. doi:10.1016/j.infsof.2020.106294.

[8] Shilpa PM and Lipi Gandhi. Cybersecurity threats to critical energy infrastructure: Business continuity in a changing geopolitical environment. *ISS Insights*, Oct 2023. URL `https://bit.ly/iss-insights-2023`.

[9] Molly Reynolds, Caitlin Morin, and Amir Eftekharpour. Ontario court of appeal clarifies privacy obligations for utilities. *Energy Regulation Quarterly*, 5(1), 2017.

[10] Tim Ruffing, Pedro Moreno-Sanchez, and Aniket Kate. Coinshuffle: Practical decentralized coin mixing for bitcoin. In Mirosław Kutyłowski and Jaideep Vaidya, editors, *Computer Security - ESORICS 2014*, pages 345–364, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11212-1.

Chapter 3

Literature Review

Advances in energy generation and distribution technology have created the need for new power management paradigms. Transactive energy markets are integrated software and hardware systems that enable optimized energy management and trading directly between prosumers.

This literature review covers unresolved security and privacy vulnerabilities in proposed implementations of such markets.

We first performed a coarse search for such implementations. We then combed the resultant literature for references to privacy concerns, security vulnerabilities, and attacks that their system was either vulnerable to or sought to address. We did so with a particular focus on threats that were not mitigated by the use of blockchain technology, a commonly employed solution.

Based on evidence from 28 peer-reviewed papers, we synthesized 14 categories of concerns, with their proposed solutions. We found that there are some concerns that have been widely addressed, such as protecting trading history when using a public blockchain. Conversely, there were serious threats that are not sufficiently being considered.

While a lack of real-world deployment has limited information about which attacks are most likely or feasible, there are clear areas of priority that we recommend to address going forward, including market attacks, false data injection attacks, single points of failure, energy usage data leakage, and privacy.

3.1 Introduction

Transactive energy (TE) grids, where anyone connected can produce or consume energy, is a burgeoning paradigm in the field of electrical power distribution [44, 11]. The past two decades have seen significant advancements in energy distribution, generation, and storage technology, including home batteries, electric vehicles, and renewable energy sources. This has led to the rise of *prosumers* who are capable of producing, storing and selling, but also consuming energy. While this is a positive development, as it can have benefits for efficiency while showing an increased demand for clean energy, TE also creates some logistical difficulties. As an increasing number of consumers transition to becoming prosumers, the inefficiencies of centralized energy markets compound. Additionally, new risks of local voltage instability are introduced.

As a result, TE market platforms are being researched as a potential solution to these effects. The new platforms promise to enable more localized trading of energy, better optimization of power flow, and altogether improved economic performance.

In parallel, developments in distributed computing have been spurred by the release of Bitcoin and the explosive growth in cryptocurrency that followed. The formation of Ethereum and subsequent advancement of smart contracts as a computing platform have enabled novel applications of distributed ledger technology such as blockchain, including transactive energy [25].

Researchers have discovered ways to use smart contracts to create distributed versions of many of the necessary mechanisms in the power grid, including demand response, optimal power flow, state estimation, and auction clearing [46, 37]. These distributed solutions have many significant benefits, including reducing single point of failure concerns and eliminating the need for trusted third parties in many cases [39].

The above security concerns and many others that arise when developing a TE market model can be addressed by implementing a solution using distributed ledger technology. For this reason, many of the models that are currently being discussed include blockchain as a platform pro forma [22, 19]. However, despite the perception of blockchains as implicitly secure due to their extensive use of encryption and redundancy, the systems that are built upon them must also be carefully considered to ensure reliability, security, trust, and privacy.

This literature review targets cybersecurity threats associated with transactive energy applications that are built on distributed ledger technology. The review then asks: *what privacy or data security concerns exist in a transactive energy context that are not addressed implicitly by the use of a blockchain?*

There are existing reviews that overlap with ours. For example, there are surveys of the use of smart contracts in energy systems [22]. There are also reviews of security threats to legacy power grids [29, 4]. However, as stated above, this survey focuses on cybersecurity threats associated with transactive energy applications built on distributed ledger technology. To our knowledge, there is no literature review that covers this exact juncture of research – perhaps in part due to the novelty of the field itself. This literature review provides researchers with an indication of gaps in the existing TE security literature, and thus areas that should be further explored. For practitioners, this review provides a comprehensive summary of existing solutions that should be considered for their architectures, as well as areas with security risks that should be addressed before deployment.

Note that as the focus of this literature review is on cybersecurity concerns; the purely physical components of grid infrastructures (power lines, transmission towers, utility poles, transformers, etc.) fall outside the scope of this study. The interested reader is invited to consult the work of Mar et al. [29] and Ardeshiri et al. [4] for a coverage of TE security concerns related to physical components. Note also that although several solutions addressing the surveyed cybersecurity concerns are also mentioned in this paper, it is beyond the scope of this literature review to evaluate the effectiveness of these solutions. Some of the surveyed papers further provide their own self-reported analyses and evaluations.

The rest of this paper is structured as follows. Section 3.2 presents the background and the motivation for this work. Section 3.3 explains the research questions that guided the data collection and analysis for this review, whereas Section 3.4 presents the review methodology. Section 3.5 presents our study's answer to RQ1, as well as threats discovered in the literature. Section 3.6 includes our response to RQ2 as well as solutions proposed in the literature, while Section 3.7 presents our answer to RQ3, as well as remaining threats that are yet to be satisfactorily addressed by existing research. Section 3.8 discusses limitations and threats to the validity of our review. Finally, in Section 3.9, we conclude the study with a reflection on the results and suggestions for further research.

3.2 Background and Motivation

In this section we describe some of the foundational concepts relevant to the study of TE markets and their associated cybersecurity properties. We will also seek to explain the relevance of each of these topics as they relate to TE.

3.2.1 Prosumer

A portmanteau of "producer" and "consumer" [19], the term *prosumer* refers to a member of the electrical grid that would traditionally have been a consumer (i.e., not a generation plant) but who also has the capability to produce or store electricity. This can be through renewable energy sources (RES), such as solar panels or wind turbines, or storage cells such as electric vehicles (EV) or battery energy storage systems (BESS).

In some cases, the term prosumer is used to refer to any (non-plant) member of the transactive energy market, i.e., households with or without power storage / production capabilities or, more simply, both prosumers and consumers. This can cause some confusion and should be noted by the reader. We will use prosumer to refer broadly to market participants, as in Section 3.5.11.

3.2.2 Transactive Energy

Transactive energy refers to a broad array of economic and control techniques related to electrical distribution, storage, trading, and generation [9, 39]. Transactive energy markets (TEM) are software and hardware systems that support these economic and control techniques. TE makes use of increased information flow in electricity markets, facilitated by advanced metering infrastructure (AMI) and smart meters (SMs) [28].

TE can improve many of the existing functions of the power grid, including demand response (DR), optimal power flow (OPF), and billing [23, 39]. Additionally, it presents opportunities to ease the adoption and integration of new technologies, such as distributed energy resources (DER), battery energy storage systems (BESS), and electric vehicles (EV) [28, 39, 24]. Finally, TE may enable new functionality within the power grid, for instance, energy trading between prosumers and supporting consumer energy preferences [9, 28].

TE also addresses some concerns within the traditional power grid implicitly. Since it is typically implemented in a decentralized fashion, it naturally improves data privacy, as well as addressing single point of failure concerns with existing grid infrastructure and reducing dependence on trusted third parties (TTP) [21].

Energy trading between prosumers is the most substantial paradigm shift promised by TE, and also where it gets the "transactive" in its name from. As DERs and renewable energy sources become more common among consumer households [24], traditional power distribution and billing paradigms present problems with scalability, communication overhead, and single point of failure concerns [48].

Enabling prosumers to trade between each other presents many benefits [39, 28]:

- Lower electrical transmission distances
- Improved market conditions
- More accurate, transparent, and fair billing
- Increased stability
- Greater efficiency (less loss of electricity, more efficient auctions)

The nature of electricity markets – that is, their high throughput and essential nature – means that TE applications must satisfy several requirements. These include data security, privacy, scalability, transaction speed, and low energy footprint [9].

It is apparent why TE and TEMs are being researched so thoroughly, as they are expected to confer many benefits to power grid operations, costs, and efficiency. However, their use does generate new security concerns that may not have existed in traditional grids or that may be exacerbated by these new paradigms.

For example, Kirli et al. [22] point out that the integration of DERs and RESs can increase the risk of operational failure due to system imbalance or local voltage excursion. The use of AMI and smart metering equipment captures much more data from consumers, resulting in greater privacy risks and thus greater need to protect such data [9].

Since TEMs are typically implemented as distributed applications, many of the proposed schemes and, indeed, surveys of those schemes take the use of blockchain or distributed ledger technology (DLT) for granted. As discussed, this review examines the security concerns that arise in the presence of such applications.

3.2.3 Blockchain

Blockchains, or DLTs, are peer-to-peer applications for distributed data storage and computation [37]. This technology is widely considered as a potential foundation for implementing transactive energy frameworks and other smart grid applications, due to its inherent ability to strongly assure integrity and non-repudiation for the data it manages on its distributed ledger [22, 20].

Some of the benefits of employing blockchain technology in these contexts include addressing some security concerns that exist within current legacy energy markets, as well as some that are created or exacerbated by a peer-to-peer or smart energy market.

An example of the former is that blockchain technology is recognized as solving the single point of failure issue that currently exists in the power grid.

Tampering with energy transactions is made vastly more difficult due to blockchain's high degree of redundancy, as noted by Lombardi et al. [28]. Blockchain can be used to improve privacy as well. Laszka et al. [23] propose PETra, an implementation built on DLT that is meant to provide privacy and, where possible, anonymity within necessary TE operations such as communication, bidding, offering, and trading.

There are reasons beyond just security that researchers have proposed the use of DLT. For example, Münsing et al. [37] suggest that the distributed and trustless nature of the blockchain could be used to address and mitigate the effects of monopolistic incentives, lack of cooperation from established utilities, and regulatory shortcomings with regards to peer-to-peer energy trading. DLT has also been touted as having the potential to reduce costs associated with TE by reducing or removing the need for trusted third parties, as noted by Mylrea and Gourisetti [39]. This could in turn increase the feasibility of DER integration, ultimately leading to a more resilient power grid [38]. Roaming electric vehicles (EV) present unique challenges for grid integration. Shuaib et al. [48] propose that DLT could be integral to supporting the integration of dynamic EVs into the electrical grid.

Blockchain technology is also necessary for enabling the use of smart contracts, another kind of software that underpins many TE implementations [39]. These DLT smart contracts differ from smart legal contracts, a related but ultimately distinct software engineering challenge. That smart contracts are built on top of DLT means that they inherit its security properties meaning, among other things, that they are immutable [22].

As noted, blockchains or DLT are distributed peer-to-peer applications that are most commonly used for data storage.

The prototypical example of a blockchain is the Bitcoin network. Bitcoin is a DLT application as well as its associated cryptocurrency [40]. In these implementations, the cryptocurrency is used both to facilitate transacting between members of the network, as well as to incentivize the process of network validation, commonly known as "mining". Every node in the network retains a full copy of the entire blockchain, comprising all of the transactions of cryptocurrency between blockchain addresses. Each new block is cryptographically linked to all of the past blocks via a hash function. This ensures that any attempt to edit the transaction history of the chain is easily detectable, lending the blockchain its non-repudiation.

The process of mining a block involves solving a difficult cryptographic puzzle that must be computed via brute force. This is known as "proof of work" consensus (POW). The complexity of this problem scales with the size of the network. This is meant to

ensure that no individual can have undue influence over the network, although this claim is regarded by some to be flawed. However, the common understanding is that in order to gain control of the ability to validate blocks, an attacker would have to compromise at least 51% of the network, leading to the term *51% attack* [48].

This is the traditional paradigm for DLT. Since its popularization, alternative strategies have been proposed to optimize network flow, increase transaction speed, improve safety, and decrease energy expenditure.

These strategies necessarily come with trade-offs. Private or permissioned blockchains, for example, have much better performance but are more susceptible to majority attacks, as noted by Münsing et al. [37]. These networks' integrity and availability become only as good as those in traditional decentralized applications .

Finally, we note that blockchain/DLT does not unilaterally solve security. Indeed, this fact provides the motivation for our review.

Mylrea and Gourisetti [39] point out that blockchain does not provide 100% security or prevention of attacks. Rather, it only improves security via authentication, encryption, and strong assurances for the integrity of the data. At the same time, it still leaves airgaps; for example, access to behind-the-meter systems [39]. Kirli et al. [22] point out that the immutability of smart contracts is a double-edged sword, and can leave the system vulnerable to re-entrancy attacks if proper checks and balances are not in place. As another example, Shuaib et al. [48] note that the supposed anonymity supported by the use of blockchain can be defeated in some circumstances, thereby introducing significant privacy concerns.

These examples provide an illustration of the kind of security gaps we hope to explore in this review. It is suspected that some TE research has a reduced focus on system security due to assumptions made about the guarantees associated with DLT. While DLT is a good foundation on which to facilitate TE, it is clearly not an encompassing security solution, and this review provides evidence of that as well as suggestions on how to improve research going forward.

3.2.4 Smart Contracts

Smart contracts (SCs) were introduced by Szabo in the mid-90s [52]. His vision for the technology was that of an automated legal contract, a computational system that would enable the execution of typical legal functions – such as sales - without human intervention. This would be done while respecting conditions like payment terms, liens, and confidentiality.

The modern understanding of a smart contract is slightly different. While Szabo did refer to digital cash protocols in his original description, "smart contracts" as implemented by Bitcoin strayed somewhat from his original smart "legal contract" definition. Smart contracts are best understood now as automated scripts that run on a distributed computing environment, enabled by DLT. This kind of smart contract was first introduced using Bitcoin's Script, which allowed basic scripting functionality via a stack-based programming language. This language enabled more complex transactions that could have built-in requirements for the recipient, such as requiring a set of private keys in order to retrieve the funds [6].

Ethereum and other platforms have since extended the idea with Turing-complete scripting languages, such as Solidity, which enable much more complex smart contracts to be implemented. Many of the platforms studied for this survey make use of such scripting languages, as they enable the complex functionality required to implement the various grid operations in a distributed manner that would simply not be possible with something like Bitcoin Script.

SCs are recognized by much of the research on distributed TE implementations as a necessary component of establishing TE on DLT. Münsing et al. [37] recognize SCs as a "key technology" for enabling distributed optimization at all scales of operation – optimization and scalability both being key requirements of a TE application. SCs facilitate the exchange of energy between prosumers [39]. They also remove the need for trusted third parties [38]. Finally, they are considered capable of providing a certain layer of guarantee with regards to money and energy transfer [48].

SCs also have benefits beyond just facilitating distributed TE implementations. They are seen as making energy auctions fairer by making the rules and their execution visible to all parties [28, 39]. Moreover, they enable additional security measures, such as the security enhancement layer described by Lombardi et al. [28].

However, there are some risks associated with the use of SCs. Chandra et al. [9] point out that privacy protection in TEMs has not been thoroughly studied. It is also noted that protecting the privacy of prosumers can have negative impacts on the safety of the market [20], a tension that will be seen as a running theme in this review. Finally, their immutability is another double-edged sword; meaning that they cannot be modified maliciously but neither can they be patched if a bug or vulnerability is discovered [22].

3.3 Research Questions

This review aims to answer three research questions:

RQ1. What are the security concerns that exist in blockchain-based transactive energy systems?

- This question will enable us to determine the (cyber-)security landscape in the transactive energy system marketplace.

RQ2. What security solutions currently exist in this space and which of the concerns discovered do they address?

- This will allow us to determine the current techniques that are currently being used to address security in this space. This, in turn, will enable us to find gaps in current security research which could potentially give rise to new techniques, as well as the efficacy of current techniques that may have room for improvement.

RQ3. What are the remaining security concerns that require attention?

- This question is particularly significant for researchers, as it can inform future research directions in the space of TEM security. By examining the threats found in RQ1 and the proposed solutions found in RQ2, we can identify threats for which solutions must be investigated further – either because no solutions have been proposed or because the ones that have been proposed expose significant limitations.

3.4 Methodology

This literature review is using a methodology inspired by Okoli [43], which involves an automated search on reliable and curated databases. The selection of relevant papers is based on repeatable exclusion criteria.

In addition, as suggested by Mourão et al. [36], the automated search results are supplemented by a snowballing phase on the selected papers (based on their references) to find additional relevant papers, satisfying the same criteria.

3.4.1 Source Database

We limited our search to the Scopus database (https://www.elsevier.com/solutions/scopus/why-choose-scopus), as it provides a comprehensive overview of the publications relevant to our topic, combined with an expressive and reliable search engine. Among its 80 million records, Scopus indexes all peer-reviewed journal and conference papers from IEEE, ACM, Elsevier, Springer-Nature, SAGE, Wiley, MDPI, Taylor & Francis, and many others. As this database is partially curated, the risk of including predatory publications is minimal (unlike with the use of Google Scholar).

3.4.2 Search Query

The following search query was used on Scopus:

```
TITLE-ABS-KEY (
   ( blockchain  OR  Ethereum  OR Hyperledger OR "smart contract" )
   AND
   ( security OR attack )
   AND ("transactive energy" OR DER OR "distributed energy resource" )
   AND  ( LIMIT-TO ( LANGUAGE ,  "English" ) )
)
AND  ( EXCLUDE ( DOCTYPE ,  "cr" ) )
```

As a first concept, we included the terms "blockchain" and "Ethereum" as synonymously indicating that the paper in question makes use of blockchain or smart contract technology. We also added "Hyperledger", a popular technology that supports permissioned blockchain development and that would likely be mentioned in a relevant paper. Finally, we included the possibility of "smart contract" in case the paper mentions smart contracts alone rather than the technology they exist upon. Note that Scopus also matches the plural forms of such terms.

As a second concept, we included the term "security" (which also covers the term cybersecurity) to ensure that the results returned present or review security solutions in particular. We include "attack" as an alternative in case a paper mentions a particular cyberattack – which is actually preferable to just mentioning security in general.

Finally, as a third concept, we included the term "transactive energy" to ensure that this research is related to the field of transactive energy markets (including the full "TEM" term

served to limit the search results beyond what was useful). We additionally included "DER" and "distributed energy resource" as synonyms, since these technologies are fundamental to transactive energy markets and are sometimes mentioned in lieu of the full TEM term.

We limited the search to papers written in English since this is the primary language spoken by the authors. Conference reviews (i.e., introduction to proceedings written by editors and conference organizers) were also excluded automatically.

This query, which focuses on the intersection between the three concepts, returned, in March 2023, 78 unique papers before exclusion.

3.4.3 Exclusion Criteria

Many of the results returned only mention security in passing or as a general benefit of their solution (i.e., "our solution promotes privacy, security, etc."). Others only talk of the security benefits of blockchain technology in general, or of specific concerns that are mitigated by the use of blockchain (i.e., the solution provides cryptographic security or consensus guarantees). Papers in either of these categories, or any other paper which does not discuss specific security concerns or cyberattacks, are excluded as irrelevant to the review.

3.4.4 Selected Papers

The application of the exclusion criteria led to 56 out of the initial 78 papers to be excluded. The backward snowballing phase, based on the older papers cited by the 22 papers selected from the databases, resulted in the addition of 6 papers, for a total of 28 selected papers, listed in Table 3.1. These papers are used in the next section to answer our three research questions.

3.5 RQ1 Results: Threats

This section answers RQ1 on the security concerns and threats about blockchain-based transactive energy systems discovered in the literature. Figure 3.1 summarizes the threats by their respective position in the TE infrastructure (application, network, or meter layer).

Table 3.1: List of selected papers, with their year and authors.

Year	Title	Authors
2023	Privacy protected product differentiation through smart contracts based on bilateral negotiations in peer-to-peer transactive energy markets	Chandra et al. [9]
2023	Blockchain and machine learning for future smart grids: a review	Mololoth et al. [35]
2022	Smart contracts in energy systems: A systematic review of fundamental approaches and implementations	Kirli et al. [22]
2022	Operational concerns and solutions in smart electricity distribution systems	Jayachandran et al. [19]
2022	Impact of blockchain technology on smart grids	Khan and Masood [20]
2022	Survey on blockchain for smart grid management, control, and operation	Aklilu and Ding [1]
2021	Security and privacy smart contract architecture for energy trading based on blockchains	Nazari et al. [41]
2021	Secure data access control with fair accountability in smart grid data sharing: An edge blockchain approach	Yang et al. [58]
2021	A multilayered semi-permissioned blockchain based platform for peer to peer energy trading	Zaman and He [62]
2021	A secure distributed ledger for transactive energy: The Electron Volt Exchange (EVE) blockchain	Saha et al. [46]
2021	Safe and private forward-trading platform for transactive microgrids	Eisele et al. [13]
2021	Blockchain for Future Smart Grid: A Comprehensive Survey	Mollah et al. [34]
2021	A blockchain-supported framework for charging management of electric vehicles	Dorokhova et al. [12]
2020	Cyber-attacks and mitigation in blockchain based transactive energy systems	Barreto et al. [5]
2020	An enhanced blockchain-based data management scheme for microgrids	Mbarek et al. [30]
2019	Blockchain for decentralized transactive energy management system in networked microgrids	Li et al. [25]
2019	Cyber-physical simulation platform for security assessment of transactive energy systems	Zhang et al. [63]
2019	Research on the application of blockchain in the energy power industry in China	Song et al. [51]
2019	Towards a semantic modelling for threat analysis of IoT applications: A case study on transactive energy	Fadhel et al. [14]
2019	A decentralised bilateral energy trading system for peer-to-peer electricity markets	Khorasany et al. [21]
2019	Secure blockchain-enabled DyMonDS design	Lauer et al. [24]
2019	Using blockchains to secure distributed energy exchange	Shuaib et al. [48]
2018	A blockchain-based infrastructure for reliable and cost-effective IoT-aided smart grids	Lombardi et al. [28]
2018	Review of cyber-physical attacks and counter defense mechanisms for advanced metering infrastructure in smart grid	Wei et al. [56]
2017	Blockchain: A path to grid modernization and cyber resiliency	Mylrea and Gourisetti [38]
2017	Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security	Mylrea and Gourisetti [39]
2017	Providing privacy, safety, and security in IoT-based transactive energy systems using distributed ledgers	Laszka et al. [23]
2017	Blockchains for decentralized optimization of energy resources in microgrid networks	Münsing et al. [37]

Figure 3.1: Simple taxonomy of the threats discovered. There are three overarching categories; the application layer, the network layer, and the hardware layer (including the smart meter itself). Each threat is relevant to at least one of these categories.

Table 3.2 presents a short summary of the threats discovered in the literature, as well as their security properties and sources. Please note from Table 2 that some threats are discussed by many papers, and that some papers discuss many different threats.

3.5.1 False Data Injection

A prominent cybersecurity issue is false data injection. From a data security principles perspective, the blockchain handles data integrity and non-repudiation to an extremely high degree of confidence, so those are of limited concern. False data injection attacks, however, concern data validity. While blockchain software ensures that data entered into

Table 3.2: Summary of security threats found in the literature and their respective sources.

Name	**Security Property**	**Found In**
False Data Injection (FDI)	Validity	[19, 46, 30, 54, 44, 18, 5]
Denial of Service (DoS)	Availability	[22, 62, 63, 5, 56, 35, 25]
Energy Usage Data	Confidentiality	[41, 13, 23, 30, 28]
51% Attack	Availability, Integrity	[22, 62, 48, 35]
Privacy	Confidentiality	[22, 41, 58, 46, 13, 23, 37, 9, 21]
Market Attacks	Integrity	[63, 14, 25, 46, 23, 22, 30, 28, 62, 41, 38, 39]
Single Point of Failure (SPOF)	Availability	[22, 58, 62, 46, 13, 30, 9, 24]
Edge Nodes	Confidentiality, Integrity	[63, 58, 19, 13]
Regulation & Standardization	Integrity, Availability	[51, 19, 28, 24, 25, 37, 38, 39, 14, 22, 20]
Smart Meter Firmware	Availability	[25, 28, 24, 14, 41]
Authenticating New Prosumers	Authentication	[62, 25, 19, 13]
Smart Contracts	Integrity, Availability, Confidentiality	[12, 22, 51, 20, 48]
Electric Vehicles	Confidentiality, Authentication	[5, 19, 10, 34, 46, 48, 20]
Communication	Confidentiality, Integrity	[30, 13, 23, 32, 61]

the chain is not changed, it cannot guarantee that the data entered is correct in the first place [46].

A false data injection attack, in general, involves modifying measurements before a system receives them or inserting fake measurements into a system in the place of real measurements. In a transactive energy setting, a false data injection attack may be carried out to observe the response it generates, to maliciously affect market operations by causing bids based on bad data, or to physically damage grid infrastructure [30, 27, 5].

State estimation is a process in which mathematical modeling is used to infer the power grid's total state based on real-world measurements collected by the Supervisory

Control and Data Acquisition (SCADA), such as the power injection of buses. This state is comprised of any number of variables that might assist with modeling or otherwise concern the control center, such as bus voltage angles or magnitudes. Such control center then makes decisions about grid operations based on the state estimation. It has been shown that an attacker could select values for their attack that thwart current algorithms for detecting bad measurements, ultimately allowing them to undermine the state estimation process [27, 53].

False data injection attacks in a traditional power grid would involve physically altering meter readings in some way. Additionally, the attacker would require knowledge of the configuration of the power system they are attempting to compromise [27]. Access to an energy management system controller would also provide an opportunity for such an attack, although it would be much more difficult to obtain than access to a measurement device [59, 3].

Smart grids and grids operating TE would also be vulnerable to these kinds of physical attacks, but may also present new vectors for the false data injection attack, an issue that would require further study.

In the Electron Volt Exchange transactive energy model, for example, aggregators are uniquely positioned to inject false readings into the system. In this case, the false data would be injected into the blockchain rather than the grid's control center. Since aggregators combine the measurements of many prosumers, they are then relied upon to accurately report on those measurements [46].

As stated, this attack can have several motivations, such as observing grid response or causing faulty bids. It can also be used to facilitate a denial-of-service attack by causing "algorithm divergence". In these cases, injected data is carefully computed to prevent the state estimation algorithm from converging on a result, thus preventing the process from moving forward by trapping it at a particular iteration [27, 46].

3.5.2 Denial of Service

A denial of service (DoS) attack is a type of cyberattack that involves overwhelming a network with traffic in order to render its services inaccessible. In the context of a transactive energy market a DoS attack can have several negative consequences. For example, a DoS attack can cause bids to be randomly discarded throughout the market [5].

This kind of attack is problematic because it is simple to implement, requiring minimal knowledge of the grid or network configuration (unlike the false data injection attack).

As noted by Barreto et al. [5], even an attacker who cannot necessarily perform a more complex or targeted attack can still launch a DoS attack to great effect.

As noted by Zaman and He [62], the proof of stake system of consensus that is employed by some blockchains in order to reduce the computational load – a necessity in the context of transactive energy, where there will be many transactions occurring rapidly due to its nature – can also make the system more vulnerable to this kind of attack.

Measures can be employed to detect and mitigate DoS attacks, such as the deep learning model described by Barreto et al. [5], which is trained to detect malicious nodes. The system uses this knowledge to dynamically modulate the cost of network access for offending nodes to make such attacks infeasible.

While DoS attacks are possible in many domains, in part due to their broad definition, advanced metering infrastructure can provide unique opportunities for DoS attack vectors. One of the most commonly referenced DoS attack in this domain is the puppet attack. In this attack, a malicious node receives a route request packet and returns a route with a nonexistent node at its end. This causes the penultimate node to enter the route discovery process when it cannot find a valid route to this nonexistent node. This process eventually results in a domino effect in which the last legitimate node sends route request packets to all of its neighbours, who then do the same, and so on. In this way, the network is flooded with route requests, causing a denial of service [60].

3.5.3 Energy Usage Data

A commonly referenced concern is that of the potential leakage of energy usage data. This data is of course accessible by prosumers who are providing energy to consumers. Due to the nature of the market and the fact that the system must make predictions about future energy usage patterns, they are privy to a large amount of information about these energy usage patterns of other members of the market.

In Figure 3.2, we can see the different ways this information might flow in a TE environment.

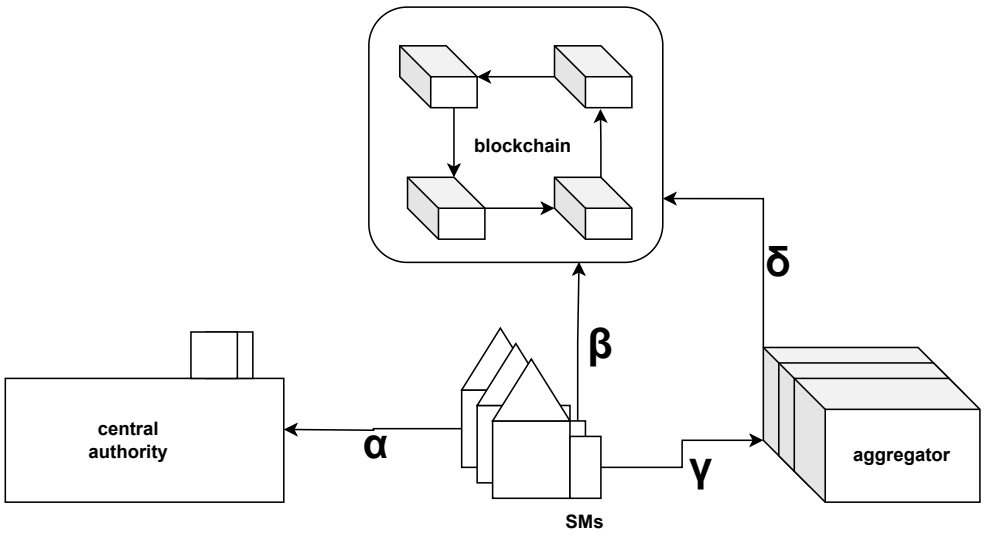

Figure 3.2: This diagram shows the flow of energy usage data in TE environments. Each letter corresponds to a different source of infrastructural risk with regards to the information's distribution.

- α - this represents the transmission of energy usage data from a household to a central authority. Although many TE implementations have distributed elements, most still have some operations that must be performed in a centralized manner, necessitating this flow of information.
- β - this represents data that goes straight from the smart meter to the blockchain. This is in some sense the safest path for the data, as it is never owned by a single party who might exploit it, however there is still the risk of a man-in-the-middle attack on β.
- γ - some TE models aggregate energy usage data before disseminating it in order to mitigate malicious analysis. This is a good solution, but introduces the risk of a malicious aggregator.
- δ - sending aggregated data to the blockchain is extremely secure since even a man-in-the-middle attack would glean very little useful information. However, it suffers from the fact that the data has to be aggregated by a trusted party.

A typical TEM model might include any or all of these vectors for energy usage data to flow according to the structure of the architecture. For example, most implementations do

not have a distributed solution for state estimation, which means it would be performed by a central utility, necessitating the inclusion of α. However, the Electron Volt Exchange [46] does include such a solution (distributed state verification), thus eliminating this vector.

As noted by Lisovich et al. [26], this data can be used to glean an undesirable amount of information about the households purchasing said power. This can include their appliance profiles, work schedules, and other behavioral data, such as when one is watching television [26, 31]. The extent to which this information can be exploited is not known, but even these limited examples are troubling.

The privacy of this data is therefore a commonly referenced concern. It is considered relevant that this data should be obfuscated in some way so that this privacy is not violated by energy producing members of the market.

3.5.4 51% Attack

The 51% attack is a scenario in which an attacker gains control of over half of the computing power in a distributed ledger system. In doing so, they would undermine the consensus algorithm that is intended to maintain the integrity of the information [62]. A successful 51% attack would enable the attacker to insert fake transactions into the blockchain, compromising the integrity of the market and auction in a TE setting [48]. While this attack can have outcomes similar to those of a false data injection attack (FDIA), its method of execution relies on an attack on the blockchain network as a whole – something that is generally taken for granted as secure by designers. This distinguishes the 51% attack from other methods of FDI.

This kind of attack is unlikely in a global blockchain such as Bitcoin or Ethereum [22]. This is due to the fact that these blockchains are comprised of many millions of nodes which, altogether, consist of enough computational resources that they consume roughly the same amount of power as the entire country of Sweden. This makes it extremely impractical for any singular entity to take over the network, as it is infeasible for one to generate or purchase enough new computing power to take over the network, and a hack taking over existing nodes would be at a scale that is unprecedented.

However, transactive energy markets would operate at a significantly smaller scale than the blockchains being discussed here. In some implementations, distributed ledgers could operate at a scale as small as a portion of a neighbourhood [46, 13]. These ledgers would then contribute to a larger ledger that would consist of the local power distribution grid. In both of these cases, the scale of computing resources is such that a 51% attack is much more feasible than in global blockchains.

This is an important note, since this kind of attack is generally ignored by larger DLTs due to its impracticality. Thus, there is a gap in security research on this topic – one that could have negative consequences for these smaller networks that may, in fact, be vulnerable to this kind of attack.

3.5.5 Privacy

Privacy is an important consideration in all cyber-physical systems, and especially so in the TE setting. TE systems necessarily have considerable information about their users, and even information that is not explicitly personal can be used to violate privacy or worse, as we discuss in Section 3.5.3. These peer-to-peer (P2P) energy markets in particular are cited as having significant privacy concerns, especially in those employing a public blockchain. Smart contracts also present unique privacy concerns [22].

Many studies and reviews acknowledge an overall insufficient focus on privacy within the TE literature [41, 9, 13]. As an example, [28] explicitly state that they do not address privacy concerns in their proposal. [9] note that only a small subset of TE studies focus on securing SM data, and furthermore claim that privacy with regards to smart contracts in P2P market settings has not been studied at all. [13] reference the literature review of [2], who found that among 47 DLT TE applications, privacy preservation was a key area that has yet to be addressed.

There are two major areas of concern when it comes to privacy in TEMs, market privacy and data privacy, which we will use to organize our discussion here.

3.5.5.1 Market Privacy

Market privacy involves protecting consumers' identities and market activities within the TE system, typically in the context of the application layer. Solutions in this space involve protocols that enable trading energy while protecting participants' identities.

Indeed, transaction history can be used to glean a great deal of information about prosumers, and thus most researchers agree that their identity should be obfuscated in some way in the market application [9, 13, 63].

Some systems can operate with extreme levels of privacy. Bitcoin, for example, provides total anonymity to its users, at the expense of safety mechanisms that might otherwise be desirable in a financial setting, such as legal enforcement, repudiation, and fraud protection. On the other hand, some systems can operate with limited privacy in favor of integrity

and trust. In any case, there is a tension between privacy, trust, and safety that cannot be reconciled. Since TE cannot sacrifice either quality totally, it must make its tradeoffs with extreme consideration.

This tension means that attempts to address privacy can often have unintended consequences. For example, in the TRANSAX platform [13], prosumers are totally anonymous but can be associated with groups to allow for safety constraints. While this addresses the grid safety problem, an inability to associate a trading profile with a particular prosumer under any circumstances still harms dispute resolution and other legal considerations. Conversely, attempts to improve safety without consideration can accidentally create privacy risks. Mbarek et al. [30] introduce a solution in which malicious trading activity is sent to endorsing peers. It is clear that the goal is to identify and track malicious behaviour, which is important, but doing so in this manner presents privacy concerns that might reduce consumer trust in the TEM.

As a final example of such a tradeoff, Laszka et al. [23] present the issue of a prosumer trading maliciously and causing grid instability. In order to mitigate this, the DSO must be able to enforce safety restrictions, but this requires access to the prosumer's transaction history, which can be used to estimate their future state.

Another significant issue is that privacy presents a barrier to dispute resolution, as noted by Khan and Masood [20]. They point out that most works provide total anonymity on the premise that the market application and associated smart contracts will execute flawlessly, something which should not be relied upon.

3.5.5.2 Data Privacy

Data privacy primarily concerns data flow in the network, although it can also involve information used by the application layer. This kind of data is typically less identity-oriented than market privacy data. Solutions in this space are more likely to involve noising, encryption, and aggregation. Impersonal data that might be subject to privacy concerns can include commercial and operational information, user preferences, consumption and production constraints, energy bids, and energy usage data [9, 62, 46, 63, 49].

This data must be communicated and stored in ways that prevent unauthorized access [63]. Examples like personal information are obvious and thus handled by most researchers, but there are less apparent sources of privacy risk that must also be considered. Metering data can also be used maliciously, again, as we discuss in Section 3.5.3.

Many operations in a TE system require some amount of potentially compromising information to function, exacerbating the problem. Processes like market clearing require

private information such as participants' utility and cost curves and preferences [21]. Shuaib et al. [48] note the difficulty of performing auction operations, like negotiating prices, contracts, and payments, while also preserving user privacy. The challenge of negotiating prices in a P2P setting while maintaining privacy is also discussed by Kirli et al. [22] and Son et al. [50]. Implementing safety measures also often requires a significant degree of detail about a prosumer's trading history, consumption and production history, and production and storage capacities. The desire to protect this information presents challenges to implementing safety measures, and vice versa [13, 23]. Finally, Li et al. [25] mention the necessity of certain operational information, such as production cost coefficients or operating states, as well as the preference to keep this information secret.

Applications such as home energy management (HEM), which operate as extensions of TE systems, can present additional privacy concerns due to the sensitive positioning of the hardware involved and the information collected [22].

3.5.6 Market Attacks

Market attacks are a broad class of cyberattacks within the TE framework that can be performed either by market participants or by external parties. We differentiate these attacks by their positioning within the TE framework because they are specified this way in the literature, rather than by their technical execution. Unfortunately, details about such technical execution were lacking in the literature, which instead focused on the kinds of outcomes that could potentially be achieved by an adversary, assuming they had some particular ability within the system (i.e., manipulating bids). The attacks aggregated here have particular relevance to TE markets for one of three reasons:

- The kind of data involved (i.e., energy usage);
- The possible outcomes of an attack (i.e., grid instability);
- How TE structures might affect an adversary's approach (i.e., manipulating a smart meter to falsify consumption metrics).

A term that comes up frequently is *malicious trading*. This refers to the actions of malicious market participants, who might perform a number of adversarial actions within the market [22, 41, 25], including:

- Falsely reporting consumption / production;

- Not following through on payment / energy production;
- Claiming falsely that energy / payment was not received;
- Posting fake bids / offers.

Market participants are understood to have financial incentives to cheat [46], so creating a robust TE system that is resilient to attacks that prosumers might try to carry out is critical. Any of these effects can also be produced by error rather than malice. The intent does not necessarily change the impact, though, so we will not differentiate between erroneous vs. malicious trades.

Malicious trading is considered a serious threat to TEMs. It is understood that such practices can cause outcomes as serious as physical grid instability [25, 23]. Malicious trading can also threaten the integrity of market operations and produce instability in the market application [25, 23]. In addition, it can be used to steal energy or money [46]. For example, a consumer might falsely report that no energy was received after a transaction, causing the transaction to fail and preventing their account from being charged (as well as the producer from receiving their compensation). This demonstrates another risk of FDI (Section 3.5.1) [30].

The problem of fraudulent energy transactions is somewhat novel to the TE paradigm, as observed by Mylrea and Gourisetti [39]. This new framework also introduces other concerns, such as bid manipulation [63]. The generation of fake bids and offers is also discussed in [22].

Attacks can include maliciously altering pricing and quantity of energy bids and offers. These in particular can cause financial harm, TEM malfunctioning, and even widespread power outages [25]. Li et al. [25] also point out that the amount of power transferred might deviate from the contractual obligation due to various physical uncertainties, another issue that must be accounted for when designing TE.

Adversaries (market participants or otherwise) can also use market information to invade the privacy of participants. These situations can sometimes be intractable, a problem we discuss in Section 3.5.5. For example, Laszka et al. [23] point out that the TE system must have access to a prosumer's transactions for safety, but that this information can be used inappropriately to infer future energy consumption.

3.5.7 Single Point of Failure

Decentralization is an overall operational goal of TE systems, often supported by DLT. As Lauer et al. [24] point out, a centralized entity should not be trusted with full control of a transactive grid, since it would represent a single point of failure (SPOF).

SPOF is a universal infrastructural concern, whether in the domains of software, physical, or process architectures. The essential concern is that a system reliant on the presence, reliability, or trustworthiness of a single entity is consequently vulnerable to the failure or compromise of this entity. SPOFs should be avoided both for data security and operation. As Eisele et al. [13] note, the market should remain available regardless of whether some nodes or the DSO are offline.

Single points of failure are present in the existing power delivery system. Generation is typically handled by a single or few plants, meaning that access to electricity is dependent on the reliable operation of these plants. Traditional attempts to mitigate this dependence might include backup generators or, more recently, battery storage systems.

TE attempts to ameliorate energy accessibility by enabling prosumers to trade directly between each other, meaning that in some cases energy delivery can still be successfully performed even if central generation is compromised. Distributed TE further enhances these qualities by enabling the operation of the energy market completely independently of a centralized server.

Distributed TEMs, however, still contain single points of failure. Because TEMs are complex, interconnected systems that contain many software and hardware components, SPOFs still arise within them, even if their primary functionality, such as market clearing, is handled in a decentralized manner [46]. In some cases, these SPOFs are simply an oversight, while in others there may be technical reasons that a distributed solution has not yet been proposed.

Even if a TE implementation avoids total centralization, it often will have single points of failure within its architecture. For example, Chandra et al. [9] mention that cloud-based aggregation creates a SPOF concern at the aggregator, since hackers now have a single target that they can compromise to access all of the data being aggregated. Beyond this, the aggregator is required to be operational for the system to function, an assumption that would be preferable to avoid as the system would be more reliable without it.

Zaman and He [62] reference a SPOF concern created in the NRGcoin [55] solution, wherein the DSO is required to maintain private data of all market participants. This, again, represents a SPOF both in terms of operation and data security. They similarly discuss a solution by Mihaylov et al. [33], who attempt to address scalability by using

a Redundant Byzantine Fault Tolerance consensus mechanism, but in doing so result in needing a central authority (i.e., a SPOF). In many cases, attempts to address scalability and security concerns of public blockchains require some compromise of decentralization, increasing SPOF concerns.

Another subproblem that can create SPOFs is the authentication of new prosumer nodes [62, 13]. Having a central authority that performs this task can lead to operational risks as well as service delays. Additionally, it presents trust implications that are undesirable in a distributed application context.

Cryptographic key issuance is another area in which SPOFs can be unwittingly created. As Yang et al. [58] point out, many existing schemes that employ asymmetric key distribution require centralized key generation, an undesirable trait in a distributed TE setting.

Finally, Saha et al. [46] observe that most TE blockchain solutions still require centralized computation to achieve market convergence, among other necessary computations.

From these examples we can see that even highly decentralized, DLT-based systems can suffer from single points of failure within specific processes. With this in mind, it is important that researchers carefully consider each process that their system will be required to perform, and whether the design of that process creates any SPOFs.

3.5.8 Edge Nodes

Edge nodes, sometimes also referred to as "fog nodes", are auxiliary computing devices that provide additional computer power to client devices on a distributed network application, such as a blockchain.

Edge nodes are often employed in TEM proposals for services such as data aggregation, performing cryptographic functions or other computations that are too intense for IoT devices like smart meters, or providing additional communication infrastructure [63, 58, 19].

The use of edge computing devices presents obvious risks regarding data ownership and access, since sensitive data relating to prosumer behaviour and TE operations may be communicated to compromised or faulty nodes. Yang et al. [58] mention these risks, as well as some that are specific to their access control implementation. For one, nodes might collude to reconstruct secret information (such as cryptographic keys) that was intentionally discarded. Another such concern is incorrect calculations returned by faulty or malicious nodes, which may result in bad data. TE operations acting on bad data can affect both market stability, as well as physical grid stability.

3.5.9 Regulation & Standardization

While regulation and standardization are not cybersecurity threat surfaces in the traditional sense, they are referenced frequently enough as general concerns that affect TE cybersecurity overall that it is appropriate to enumerate them here.

3.5.9.1 Regulation

The lack of clear regulation surrounding the operation of TE markets is frequently cited as a concern in the literature. This kind of regulation falls into two categories: that of Internet of Things (IoT) device standardization and interoperability, and regulation of the TE market operations. TE market operations include market clearing, trading practices, and reconciliation when there is a trade dispute.

In the work of Münsing et al. [37], regulatory expenses are presented as potentially not worthwhile in a bid to demonstrate why a distributed market, despite possibly introducing market inefficiency, would be preferable to a highly regulated centralized market.

The issue of IoT device regulation and interoperability is introduced by Mylrea and Gourisetti [39]. In fact, the simulation testbed they introduce is explicitly positioned to be capable of exploring this area of concern to ensure the health and adoption of the proposed decentralized TE systems. They also mention questions of legality with regards to DLT-based platforms, and while they do not provide suggestions, they draw attention to the problem of regulation and reconciliation in these environments. Fadhel et al. [14] also point out that adoption of decentralized TE solutions is limited by the absence of a clear regulatory body.

Finally, the issue of an underdeveloped legal framework for decentralized energy applications with regards to their safety is referenced by Kirli et al. [22] as well as Khan and Masood [20].

3.5.9.2 Software Standards

DLT software, and particularly smart contracts, are still relatively new software paradigms. This novelty on its own brings forth concerns that it may not yet be ready for widespread adoption [30].

The recency of the development of these technologies also means that there are no established software standards yet. Song et al. [51] point out that the lack of a standardized

security architecture for blockchain and smart contracts has led to an exceptional number of vulnerable smart contracts being deployed [42]. These concerns are extremely relevant to DLT-based TE, as the latter will inherit the security posture of the architecture it is built upon, and this posture will be facilitated by standardized security and development practices. Additionally, standard protocols and software architecture of TE itself have yet to be established, as noted by Jayachandran et al. [19].

3.5.9.3 Hardware Standards

In addition to inheriting the security posture of the software systems they are built on, TEMs will inherit the security of the hardware they are run on. Of particular concern are IoT devices, which are heavily involved in TE operations, often limited by processing power, and a relatively recent technological development. While we discuss these limitations in more detail elsewhere, the specific concern of a lack of standardization in TE and IoT hardware is mentioned in several studies [15].

Lombardi et al. [28] mention the security concerns raised by the lack of security standards for SM hardware. Additionally though, they note the difficulty raised in designing TE software specifications when the hardware choices are inconsistent and difficult to predict. Li et al. [25] note that the variety of vendors producing communication and control devices, and the lack of security standardization between them, produce unpredictable cybersecurity threats. This concern is also noted by Lauer et al. [24], with an emphasis on the fact that integrating vulnerable technology into a critical service like the power grid presents additional risks.

3.5.10 Smart Meter Firmware

Smart meters are considered a significant weak point in TEM security stature. In particular, firmware vulnerabilities are often cited as an area of concern. This is a highly relevant threat point for our purposes, since SM firmware security is not inherently improved by a blockchain-based architecture. In fact, data integrity on the blockchain can be harmed by vulnerable smart meters.

Lombardi et al. [28] point out that smart meters are often poorly designed, which can lead to firmware bugs and vulnerabilities. These present challenges that have limited their deployment thus far despite efforts to do so by large utility companies.

Adversaries who control vulnerable smart meters can cause a variety of unwanted effects, including severe outcomes such as a blackout [14].

Due to the positioning within the TE architecture as monitoring devices supporting trading, load balancing, state estimation, and many more critical grid operations, keeping SM firmware up to date with security patches is widely recognized as an area of great importance and research [28, 25, 41, 24, 14].

3.5.11 Authenticating New Prosumers

New prosumers in a TEM will need to be authenticated by the network [62]. In most cases, it is expected that they will be registered by a regulatory authority such as the DSO before they can be activated on the market [25, 13].

This presents several security concerns. Often, a TTP is required to perform the authentication, which is an inherent risk [41]. In most cases, this reliance also results in a single point of failure. Additionally, requiring market participants to register with a centralized authority presents privacy concerns.

3.5.12 Smart Contracts

Smart contracts are present in most proposed distributed TEM architectures to enable auction trading on DLT, as well as facilitating distributed versions of other critical functionality in some cases. While smart contracts enable distribution of many TE applications, addressing threats of SPOF and TTPs, they also present new security concerns.

Smart contracts are immutable by nature once they are deployed on the blockchain [48]. This is a double-edged sword, as this prevents tampering, but also limits the ability to patch contracts when vulnerabilities are discovered [22]. Such vulnerabilities are a real threat; Nikolić et al. [42] found that 34,200 smart contracts on the Ethereum blockchain contain dangerous code. Smart contracts also face implementation limitations that affect their ability to enable some security functionality Khan and Masood [20].

In addition to vulnerabilities, Kirli et al. [22] note that there is a risk of misuse of smart contracts by malicious actors with profit incentives. These parties can use backdoors in the SC code to perform re-entrancy attacks. Such attacks have already seen effective use in the decentralized finance sector, with a particular variant known as a rug-pull thrifting users of their crypto assets [22].

Smart contracts can have impacts beyond just financial assets. Song et al. [51] note that vulnerable or malicious SCs could cause auction malfunction resulting in extreme deviations in power trading demand, an effect that threatens security as well as physical stability of the grid.

3.5.13 Electric Vehicles

Electric vehicles (EVs) present some unique challenges in the TE domain. They form a new attack surface for adversaries, who can use EVs as a vulnerability to affect grid operations, as noted by Barreto et al. [5]. Malicious EV users could also misuse their own vehicle for such purposes [19]. Additionally, the ability for hackers to exploit vehicles with autonomous capabilities can lead to extremely harmful outcomes, ranging from damaging the vehicle to serious injury or even death [10]. The concern is that connecting to a TE when roaming charging could present a novel attack vector for hackers to target autonomous EVs. This concern is only going to become more relevant as vehicles continue to adopt autonomous driving features [57].

EVs also suffer from familiar challenges. Mollah et al. [34] note that EVs are subject to many of the privacy concerns of smart meters, including leakage of identifying information [46] and energy usage data, as well as some novel concerns such as location data. These concerns can also be exacerbated by the fact that EVs may have to connect to different TE networks depending on where they are charging.

In fact, this leads to probably the most unique concern associated with EVs. Unlike smart meters connected to homes, EVs may wind up charging in a TE network operated by a different DSO. This presents a two-sided challenge: the DSO does not have the required information to authenticate the EV, and sharing personal information with a foreign DSO represents a privacy risk for the EV [48, 20].

3.5.14 Communication

Although it is not explicitly mentioned often, networking protocols and communication devices provide a substantial attack surface for TEM cybersecurity.

Mengelkamp et al. [32] note that even a secure smart meter and TE system can be undermined by an insecure communication network. In a similar vein, the authors of TRANSAX, Eisele et al. [13], point out that communication privacy is a baseline foundation that must be present in order to provide privacy and anonymity in a distributed application. Finally, Mbarek et al. [30] discuss the particular risk of the operating environment that TE components are deployed in. They note that smart meters and other sensors collect data on the customer premises, an open environment, and communicate via wireless protocols. These factors, as well as the sensitivity of the data in question, make them especially likely cyberattack targets, both in terms of appeal to hackers and vulnerability.

Another consideration is the security of the blockchain itself. While this is often taken for granted, the blockchain's security posture is based on the difficulty of cryptographic operations, as we discussed in Section 3.2.3. A looming concern is that the development of increasingly powerful quantum computers will render cryptography based on prime factorization – currently the most popular cryptographic technique – virtually insecure [61]. This is a problem for systems built on top of blockchains that employ this kind of cryptography, as the integrity of the system's records will be nullified if difficult hash puzzles can be solved at will.

3.6 RQ2 Results: Solutions

This section answers RQ2 on the security solutions that currently exist in this space, together with the concerns they address.

Table 3.3 presents a summary of the solutions discovered in the literature organized by the threats they are associated with, and the papers they were found in. Solutions without titles or that appear as components of larger systems are denoted as "various".

3.6.1 False Data Injection

As noted in the discussion on false data injection attacks in RQ1, power grid state estimation is the process of estimating various state variables about the grid using mathematical models and real-world measurements. However, real-world measurements are unreliable – whether due to tampering or simple equipment failure. For this reason, methods were developed for detecting bad measurements. These methods presume that good measurements should produce state estimations closer to the actual state, while bad measurements should do the opposite, meaning that there should be detectable inconsistencies between the good and bad measurements [27].

The robust state verification of the Electron Volt Exchange (EVE) is strongly inspired by these methods and assumptions [46]. It approaches the new challenges imposed by decentralization by introducing distributed measurement verification. This method shifts the focus from full state estimation to real power injection within an aggregator region and power flow between aggregator regions. It does so by solving a regression problem – similar to state estimation, but a continuous function – one time-step after a given market iteration [46].

Table 3.3: Summary of security threats found in the literature, with their respective solutions and sources.

Threat	**Solutions**	**Found In**
False Data Injection (FDI)	RSV, ETSE, various	[46, 30, 54, 44, 18]
Denial of Service (DoS)	Smart contracts, various	[62, 25]
Energy Usage Data	HE, Aggregation, various	[46, 9]
51% Attack	Various	[62]
Privacy	Pseudonymity, anonymity, various	[13, 23, 62, 41, 48, 46, 21]
Market Attacks	Reputation, access blocking, various	[41, 62, 13, 25, 28, 8]
Single Point of Failure (SPOF)	Various	[62, 13, 58, 46]
Edge Nodes	Incentives, distributed authority, BFT, various	[58, 19, 13]
Regulation & Standardization	ETSE	[28]
Smart Meter Firmware	BBPS, various	[28, 14, 41, 25]
Authenticating New Prosumers	Smart contracts, distributed verification, various	[25, 62, 13]
Smart Contracts	Third-party verification, various	[22, 48]
Electric Vehicles	Various	[46, 20, 7, 47]
Communication	Quantum-safe cryptography, quantum key distribution, PETra	[17, 61, 23]

Similar to the older methods for bad measurement detection in state estimation described by Liu et al. [27], the robust state verification of EVE is itself vulnerable to certain attacks. An attacker can cause a circumstance known as algorithm divergence, a type of denial of service attack in which the market operations are unable to iterate since the verification process will become stuck in an infinite loop. Additionally, malicious aggregators may be able to inject false data that is passed along to neighboring aggregators [46].

A novel system for remedial (rather than preventative) action is mentioned by Onu-

manyi et al. [44]. They describe a scheme based on thyristor-controlled series capacitors that was found to successfully ameliorate the effects of some FDI-based cyberattacks after they occur.

There are other solutions proposed in articles that did not show up in the primary results. Extended distributed state estimation is another mechanism for attempting to detect false data injection attacks, specifically those that are thought to be tolerable to usual state estimation algorithms as discussed earlier [54]. Introduced by Wang et al. [54], the technique predates EVE's solution by several years. Finally, He et al. [18] use a machine learning-based method to detect FDI attacks by comparing incoming measurements to models of historical data.

3.6.2 Denial of Service

Smart contracts are positioned by Li et al. [25] as potentially mitigating the effects of DoS attacks on edge devices by monitoring and validating their solution process while it occurs.

Zaman and He [62] explicitly address DoS attacks in their proposal by continuously modifying which blockchain nodes are performing prosumer verification, as well as which are validating each transaction. Additionally, they use different sets of nodes for each of these tasks. Finally, their Q-score mechanism, which we discuss in more detail in Section 3.6.6, ensures that only well behaved nodes should induce transactions. A misbehaving node will quickly have its Q-score decreased, preventing it from launching transactions en masse and reducing its ability to perform a DoS attack.

3.6.3 Energy Usage Data

A common solution to protecting energy usage data is aggregation. In some cases, temporally aggregated data is enough information to perform a necessary function, such as billing a customer every month. In others, aggregated data from multiple prosumers can be used effectively [46]. Methods of aggregating data typically employ intermediate edge nodes (Section 3.5.8) as aggregators. This data is often encrypted using a partially homomorphic cryptosystem such as the Paillier cryptosystem, in order to enable the aggregators to aggregate the data without compromising it [9].

Homomorphic encryption (HE) refers to encryption schemes that have a specific property. This property is that cyphertext encrypted by an HE scheme can have some mathematical operations performed on it, and decrypting the result of these operations will provide the same answer as if the operation had been performed on the plain text.

There are a variety of such schemes. RSA, which was not created with this in mind, is consistent under multiplication, making it homomorphic. Fully homomorphic encryption schemes, which support arbitrary operations on cyphertext consistent with plain text, have been devised as well. However, these are typically computationally demanding. As such, most real-world solutions, especially those in the IoT space (where computational resources are constrained and time is often a factor), limit their homomorphism to just the necessary operations. For data aggregation, the required operation is usually addition.

3.6.4 51% Attack

51% attacks are not given a great deal of consideration in the TE literature, most likely due to their difficult execution in most real-world blockchain applications. However, Zaman and He [62] note that their solution presents a hurdle to potential 51% attacks: rather than simply obtaining more compute power to gain control over the network, an attacker would have to take over 51% of the smart meters, a much more challenging prospect.

3.6.5 Privacy

Coin mixing is introduced as a privacy solution in the TRANSAX protocol [13]. It addresses the problem that all transactions in the blockchain are publicly available, meaning that prosumers' energy trades could be tracked.

In the traditional public blockchain model, such as Bitcoin, users are anonymous since they are not connected publicly to their account addresses. However, this anonymity is limited: since the transaction are public, if the account address is ever deanonymized then the users' transactions can be linked to their actual identity [45]. This problem is even more apparent in the transactive energy setting since, in that case, the main account addresses are in fact connected to the prosumers' identity.

To address this, TRANSAX allows prosumers to create anonymous accounts from which to perform trades. However, if the transfer of resources into these accounts can be easily tracked, then the additional security they offer is negligible, since attaching them to an identity would be as simple as looking up the transaction that added funds to their account and tracing back the source account.

Thus, TRANSAX employs CoinShuffle, a coin mixing technology. This platform is fully distributed and requires no trusted third party. It allows multiple prosumers to engage in a coin mixing protocol, allowing them to create several anonymized accounts which cannot

be linked back to any individual prosumer. In this way, they are now able to perform trades anonymously.

Other implementations also seek to decouple a prosumer's identity from its trading activity, such as those introduced by Laszka et al. [23], Zaman and He [62], Nazari et al. [41], and Shuaib et al. [48]. While these approaches improve user privacy, they can harm system security and trust by impeding legal enforcement and dispute resolution, as discussed in Section 3.5.5.

The Electron Volt Exchange [46] addresses bid and constraint privacy, ensuring that energy scheduling is performed without revealing this information. Finally, Khorasany et al. [21] discuss a solution which involves decomposing the market optimization problem. In doing so, the solution limits the information required about the market participants while still accounting for their utility curves and preferences.

3.6.6 Market Attacks

The fundamental problem of forging energy transactions on the market database is handled by the blockchain [28, 25]. However, as we discussed in Section 3.5.6, many threat vectors remain beyond this basic attack pattern.

Market integrity is also vulnerable to FDI attacks, for which solutions are outlined in Section 3.6.1. These attacks can facilitate outcomes such as falsely claiming failed energy delivery and falsifying consumption data.

Eisele et al. [13] do not specifically address market attacks, but do bring forward suggestions such as employing reputation mechanisms or requiring security deposits to enable fines for malicious actors. However, it is questionable whether security deposits would be a desirable mechanism for TEM membership, because they could be introducing consumer wariness and/or social inequality.

The ability to detect attacks, and especially to differentiate malicious data from erroneous data, is as important as a system's response to bad data. This is even more true when considering a critical service such as energy delivery; it is essential that punishment is meted out with restraint and accuracy. Methods of detecting and responding to attacks have been extensively researched in the domain of control systems [8]. This research includes consideration of automatic response mechanisms leading the system to an unsafe state, a concern we reference later in this section. Techniques from this field, such as sequential detection and change detection, could be employed directly or tailored for TEMs by designers to further enhance market security beyond the blockchain.

Reputation mechanisms are systems or protocols designed to keep track of a node / prosumer's behaviour over its lifetime on the network. These systems intend to promote positive and honest behaviour and discourage malicious or dishonest behaviour. They must be carefully designed such that they promote desired behaviours and maintain an accurate estimate of a node's value to the network.

These reputation mechanisms can address many of the discovered security concerns, including general structural threats like DoS attacks, in addition to market attacks such as price manipulation, fake sale, fake purchase, and others. Most attacks that require repeated malicious interactions with the network can be mitigated to some extent by a reputation mechanism, although it is not recommended as a sole security solution – rather it could work as a deterrent in concert with other cybersecurity techniques.

One such implementation we found was Q-score, a reputation mechanism developed by Zaman and He [62]. It attempts to track the reliability of a node by assigning a numerical value to each node. This score is applied to all parties; including prosumers, consumers, verifying nodes, and validation nodes. Q-score uses a simple mechanism that rewards successful transactions and punishes unsuccessful transactions. In a successful transaction, all parties receive a positive boost to their Q-score. This includes the buyer, the seller, and the verifier and validation peers and leaders. If a transaction is unsuccessful, both the verifier and validation peers receive a negative impact on their Q-score. Depending on the reason for the transaction failure either the seller or the buyer may receive a negative impact on their Q-Score: if the seller fails to provide energy, they will be punished; alternatively, if the purchaser fails to pay then they will receive a score punishment.

Access blocking is a blunt approach to addressing misbehaving nodes, proposed explicitly by Nazari et al. [41]. Unlike the more refined reputation mechanisms discussed earlier, access blocking involves simply banning a particular node from the transactive energy network altogether. This unrefined approach has the benefit of being extremely secure but has the potential drawback of unfairness. On the one hand, preventing access altogether eliminates undesirable behaviour entirely. On the other, it may unreasonably punish mistaken behaviour. With a service as essential as electricity, it is important that social fairness be maintained [19].

3.6.7 Single Point of Failure

As discussed in Section 3.5.7, single points of failure can appear in a variety of circumstances within a TEM architecture. As such, there is no single solution, but rather methods of decentralizing different parts of the process.

Zaman and He [62] present a solution for decentralizing the authentication of new prosumers, which we discuss in more detail in Section 3.6.11. TRANSAX offers a robust solution for market operations that is tolerant to disruptions of any critical entities [13]. However, both approaches still require the DSO for prosumer registration. Yang et al. [58] employ a secret sharing scheme to enable distributed key generation, solving the SPOF of centralized key distribution in their application.

The Electron Volt Exchange [46] takes significant strides in the pursuit of comprehensive decentralization, proposing a TE-tailored distributed state estimation algorithm RSV, which we discussed in detail in Section 3.6.1. Additionally, their solution uses a decentralized price optimization algorithm to fully avoid SPOF in the market clearing process, unlike many similar works.

3.6.8 Edge Nodes

Edge nodes themselves are a solution to some of the threats presented by IoT device limitations (Section 3.5.8) by enabling stronger security operations than would otherwise be possible. However, the edge nodes themselves then present potential points of failure.

In the scheme proposed by Yang et al. [58], edge nodes are used extensively to assist with cryptographic computations. As such, it was critical that they address the concerns presented by their involvement. They used a decentralized key generation scheme to distribute authority. They also mention that other schemes employ user-performed verification in order to reduce the risk of malicious or erroneous edge nodes.

Jayachandran et al. [19] propose that standards for IoT blockchain applications should be expanded upon. The lack of software design standards in this space is noted as a threat in Section 3.5.9.

In TRANSAX [13], prosumers embody edge node responsibility by acting as solvers for market operations. In this scheme the prosumers are implicitly incentivised to perform this task, since it enables them to create beneficial trades. Eisele et al. [13] do note, however, that this is safe since the overall application will only accept new solutions if they are strictly better than the current solution, creating a mutually beneficial incentive structure.

Another solution that impacts edge nodes is reputation, which we discuss in Section 3.6.6. These reputation mechanisms can be applied similarly to edge nodes as they are to prosumer nodes in order to mitigate adversarial behaviour by reducing exposure to bad actors.

Byzantine fault tolerance is a mathematical technique for reducing the impact of incorrect or malicious results when consensus is required. It is used extensively in blockchain implementations to reduce the impact of failure for any given node, making malicious behavior and collusion much more difficult as it means that bad actors must control much larger portions of the network to have an impact on its consensus mechanisms. It can also be used in validation applications to ensure that data is being reported and calculated accurately. For example, in the scheme of Yang et al. [58], multiple fog nodes may perform the same calculation to prevent a single node from unilaterally affecting a client node's reputation or ability to purchase power.

3.6.9 Regulation & Standardization

While many papers highlight a lack of standardization and regulation of various processes and components in TEMs, few present solutions to such issues. This is sensible in some cases as the issue would be out of scope, such as IoT hardware standards, and focus should be on mitigating the effects of vulnerable or inconsistent hardware. One example of such a technique is presented by Lombardi et al. [28] with their Energy Trading and Security Enhancement (ETSE) module, intended to make their solution agnostic to SM hardware. However, the authors do not propose an actual implementation of such a module, noting the significant difficulty of doing so.

In cases where proposals would be more expected, such as architectural standards, we suppose that the field is still in its innovation stage and is not mature enough to yet support the creation of such standards. With regards to regulation, work will need to be coordinated with government bodies and other stakeholders. Such work is out of scope of our review.

3.6.10 Smart Meter Firmware

SM firmware security is critically important due to TEM's high dependence on accurate readings from SMs.

Lombardi et al. [28] introduce a framework that includes a Security Enhancement Layer (SEL), which attempts to mitigate the threat of vulnerable smart meters. Before a transaction is executed, the smart meter on each side of the transaction is checked for known firmware vulnerabilities. If any are found, the transaction is not executed and the SM is temporarily quarantined from the network until a patch becomes available or is installed.

This concept of device block-listing is also mentioned by Fadhel et al. [14] as well as by Nazari et al. [41], who reference the SEL study.

Zero-day exploits are of particular concern in such highly sensitive and vital industries as power distribution. The lack of electrical power can cause economic losses or even loss of life in some circumstances (see the 2021 Texas power crisis [16]). As such, keeping relevant software and firmware up to date with the latest patches and bug fixes is of even greater importance than it is in most cybersecurity settings.

To this end, Li et al. [25] have proposed a blockchain-based patch system, shown in Figure 3.3. Such a system would keep security patches in a public cryptographic ledger, making them universally accessible and unmodifiable. The first benefit ensures that all users will be able to access the patches at all times and removes single point of failure – there is no reliance on a particular server or supplier to be available for them to download the new binary. The second benefit prevents malicious modifications to patch binaries, hopefully addressing some phishing and Trojan attacks that would otherwise be possible.

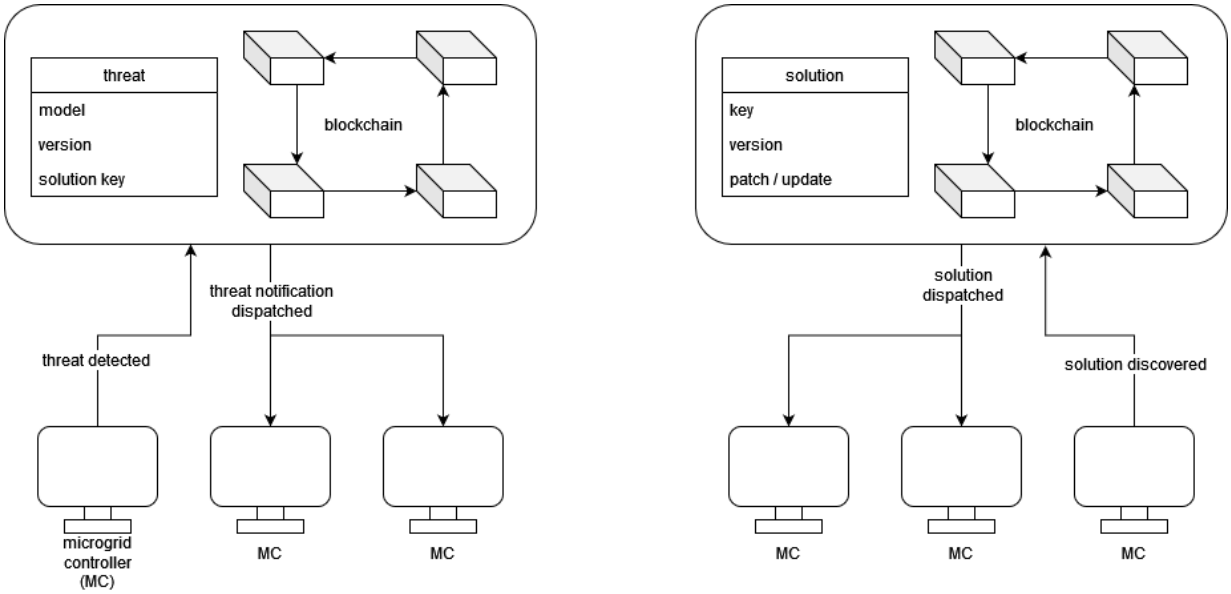

Figure 3.3: Visualization of the DLT-based patch system proposed by Li et al. [25], with the threat detection and notification (left) followed by the solution discovery and dispatching (right).

There may be drawbacks to such a system, such as the potential for exploit by adversarial microgrid controllers. These parties would suddenly be in a position to report

fake threats or post fake solutions that may be automatically downloaded by other microgrid controllers. These cases demonstrate a need for further security consideration of DLT-based software or firmware update solutions.

These kinds of systems are still being explored and would likely benefit any transactive energy implementation, although they could also be conceived as separate systems to handle primarily smart meter firmware updates, for example.

3.6.11 Authenticating New Prosumers

There are three solutions proposed in our selected papers. Li et al. [25] suggest that smart contracts might be used to automate the authentication process, eliminating the SPOF concern (3.5.7) while maintaining market integrity. Zaman and He [62] propose a distributed solution in which prosumers already on the market can verify new participants via their SMs. Finally, TRANSAX also has its own protocol for authentication [13].

3.6.12 Smart Contracts

The effects of vulnerable smart contracts are somewhat outside of the domain of TEM designers, however, some ideas to address these are brought forward.

Kirli et al. [22] mention the possibility of employing third parties to verify smart contracts. They note that this technique is common practice in the decentralized finance industry. Shuaib et al. [48] discuss the lack of standard security architecture. This could be considered invocation of such standardization as a solution for SC security and fraud prevention, although the authors do not make specific suggestions towards producing such a standard.

3.6.13 Electric Vehicles

Regarding charging outside of an EV's home network, Khan and Masood [20] reference a mutual authentication scheme that disguises the EV's information from the supplier of energy, while still maintaining safety requirements.

In terms of privacy, many schemes have been proposed to facilitate privacy-preserving energy trading between EVs, some of which are found in [46, 20, 7, 47].

From what we observed, attention has not been given to the particular problem of EVs disrupting stability intentionally. This makes sense since this concern is present in many

DERs, and as such research to prevent load imbalances as a result of malicious behavior would be device-agnostic.

3.6.14 Communication

As stated in RQ1, network security is not frequently referenced in TEM proposals. Solutions referenced typically exist outside the TE domain, such as onion routing, garlic routing, or the Matrix protocol [13]. Laszka et al. [23] also claim to provide communication anonymity with their protocol, PETra.

In order to address the threat of quantum computers, efforts have begun to produce the next generation of cryptography which will be impenetrable even to quantum algorithms, with these techniques being generally referred to as *quantum-safe cryptography*. Yin et al. [61] go further by developing a quantum-secure network, laying the foundation for future quantum-safe blockchain technology. This technology will require continued development before it can be integrated into TEMs.

3.7 RQ3 Results: Remaining Security Concerns

This section answers RQ3 on the remaining security concerns that require attention.

3.7.1 False Data Injection

While some solutions exist to address false data injection attacks, such as EVE's robust state verification (Section 3.6.1), they themselves present certain threats. Due to the risk of algorithm divergence present in state estimation algorithms, it would be interesting to investigate whether there is a way to mitigate this threat or perhaps another method of FDIA detection that avoids this threat entirely.

3.7.2 Denial of Service

While DoS attacks are not considered extensively in the TE literature, we do not consider them a top remaining threat. DoS attacks within the market application, which are the kind that TEM designers should be most concerned with, are well handled by existing concepts like reputation mechanisms.

DoS attacks on communication infrastructure are out-of-scope of TEM design, but that does not mean that researchers should not attempt to mitigate their effects. However, since solutions addressing SPOF concerns implicitly address DoS concerns, we reason that DoS attacks and their outcomes will be mitigated by the general trend of decentralization within TE, and are thus not a priority in their own right.

3.7.3 Energy Usage Data

While some solutions for energy usage data privacy exist, primarily data aggregation, these are not currently sufficient. The variety of use cases for energy usage data means that aggregation is not always optimal or even possible while maintaining the usefulness of the data. Additionally, the threat has received insufficient attention in general in the TE literature, indicating a need for further research.

In traditional energy management models, this information would only be accessible to the system operator, thus presenting significantly less risk [27]. However, in distributed models, this information is often required for market operations and is therefore sent to verifying or aggregating nodes.

The leakage of energy usage data presents very acute, real-world threats to safety and privacy [26, 31]. For these reasons, we consider this to be a serious gap in the security posture of current TE models.

3.7.4 51% Attack

As we note in Section 3.5.4, 51% attacks are rare in global blockchain networks due to the scale of computational resources required to launch such an attack. However, many TEM implementations have sub-blockchains that service very small areas, such as a neighbourhood. This can make the prospect of a 51% attack much more likely, something that we did not see discussed in the literature. We thus consider 51% attacks to be a remaining threat, and of particular concern to TE research. Either research would have to demonstrate that the attack is not a concern despite the small network sizes or solutions should be proposed, but at the moment neither is true.

3.7.5 Privacy

As we noted in Section 3.5.5, there is an insufficient level of privacy consideration in the existing literature. The two main areas of privacy we discuss are market privacy and data

privacy.

Solutions that address market privacy currently do so at the expense of traceability, and thus inhibit dispute resolution and tracking of criminal behavior. For example, TRANSAX [13] and other proposals present privacy solutions that disconnect a prosumer's identity from their trading history. This prevents attacks that rely on using trading history to predict features about the consumer, as well as discovering the identity of a trading partner. However, these proposed systems do not enable this connection under any circumstances; a desirable attribute for the consumer but one that prevents recourse altogether, even by the DSO.

Data privacy solutions are being proposed, such as those within the Electron Volt Exchange [46] that protect user constraints and bids. However, given the sprawling, complex nature of TEMs, there still remain gaps in data privacy (notably energy usage data, which we discuss in Section 3.5.3). Furthermore, there is a lack of consensus about which approaches best suit which applications, as data must be obfuscated while remaining useful to the requesting party. Some approaches employ noising techniques, others use algebraic transformation, just to name a few. Some of these techniques will be more appropriate for certain kinds of data than others, and until an optimum is realized protocols that treat data differently with respect to its operational function will need to continue to be developed.

Privacy is a broad and multifaceted issue. Combined with the fact that TEMs are large, complex systems means that it is difficult to address privacy holistically. Given the broad scope of the problem, the intricacies in the systems involved, and the lack of coverage of the solutions proposed, we consider privacy to continue to be a threat that requires further research.

3.7.6 Market Attacks

While reputation mechanisms are a promising method of encouraging good behaviour in the TEM, they do not guarantee it. Furthermore, they are not a proactive solution, something that was largely missing from the literature.

Attack detection needs to be improved. At present, it is impossible to distinguish between malicious and erroneous input, instead relying on repeated behaviour to influence eventual response. Furthermore, dispute resolution needs to be considered by TEM designers if the effects of malicious behaviour cannot be mitigated. Prosumers will not want to join a market where occasional theft is considered a cost of doing business, and DSOs will not want to subsidize fraud.

Overall, we consider market attacks to be a medium-level priority. Assessing such a broad category introduces some limitations, but the presence of some solutions, such as reputation, mitigate the potentially disastrous effects of unregulated markets (examples of which have been seen in the Non-Fungible Token space [22], an unregulated market with vulnerabilities similar to TE's that we discussed briefly in Section 3.5.12). However, it would be preferable to have more proactive and reliable solutions going forward.

3.7.7 Single Point of Failure

As we discussed in Sections 3.5.7 and 3.6.7, single point of failure is a broad threat that can present in different ways. As such, it would not be reasonable to present a single solution that achieved full coverage of the problem. Rather, we would hope to find that it has been eliminated in all processes involved with TEM operation, from onboarding to market clearing to billing. Since this is not the case, we consider SPOF to remain a high-priority threat, and work should continue to decentralize TEM architectures holistically.

3.7.8 Edge Nodes

Edge nodes present a variable security threat depending on their positioning within a given architecture.

Edge nodes often perform tasks that require sensitive data. Combined with the fact that they are typically independently operated nodes, they can be seen as a significant security threat. Indeed, edge nodes are only as secure as the systems that use them, and there are many examples in the literature of edge nodes being deployed insecurely in TEM architectures.

However, despite often being employed to augment the TEM computing environment, edge nodes are *not* a mandatory component of a TEM infrastructure. Additionally, existing solutions to edge node concerns, such as distributing authority, reputation mechanisms, and incentives, are fairly comprehensive (Section 3.6.8). For these reasons, we do not consider edge nodes to be a top security priority for TEM designers, and we hence classify them as a medium-level threat.

Future solutions should decrease reliance on edge nodes when possible, whether through hardware improvements or more efficient algorithms. Where they cannot be avoided, researchers should incorporate them with care, and consider inventing more secure protocols or employing existing mechanisms to mitigate the threat of malicious or error-prone nodes.

3.7.9 Regulation & Standardization

As we discussed in Section 3.6.9, much of the solution space to this issue is outside of the scope of technical DLT-based TE research, which we surveyed for this review. However, work can and should be done towards mitigating the potential side effects of inconsistent and vulnerable hardware, as well as those of vulnerable smart contracts.

3.7.10 Smart Meter Firmware

While vulnerable SM firmware presents a significant security risk in a TEM framework, proposed solutions look promising towards mitigating the effects of compromised SMs. Research that more directly addresses the problem is out of scope of TEM designers. We believe that the techniques present in the proposed solutions, such as device block-listing [28] or blockchain-based patching [25], provide a promising foundation for addressing this threat. That said, it is a serious threat, and so must be handled comprehensively. Additionally, some existing solutions, such as the Blockchain-Based Patch System (BBPS), present their own concerns and must be studied further. For these reasons we consider SM firmware to present a medium-priority threat going forward.

3.7.11 Authenticating New Prosumers

While a de facto standard protocol for authenticating new prosumers has yet to arise, we do not consider this area of TEM operation a significant security risk. Many papers that raise the concern of authentication focus on availability [41]. The trend is towards decentralization of this process, which inherently improves availability. Overall the presence of several proposed solutions compared to the relatively minor security risk presented by this stage of operations lead us to rank this threat as a low-priority one.

3.7.12 Smart Contracts

Smart contracts clearly present a significant source of vulnerability in TE systems that employ them. This is mostly due to the fact that popular smart contract implementations themselves are relatively insecure [13], and have caused problems in many of the industries that have attempted to use them [22].

While much of the work to improve SC security posture happens outside the TE domain, in the meantime TEM researchers should strongly consider this factor. They should seek

to employ SCs only where absolutely necessary, and write the chaincode with utmost consideration when doing so.

3.7.13 Electric Vehicles

Although EVs present a unique challenge in terms of integration into the energy market, their cybersecurity implications are not as significant. In terms of presenting a novel attack surface within the market, they share this property with other innovations such as RESs, BESSs, and smart meters. For this reason, EVs cannot be singled out as presenting a uniquely concerning safety threat to the stability of the system.

Solutions were not discovered for preventing TE-to-EV hacking, however, this is considered external to the TE domain, and vehicle cybersecurity should be handled by automakers.

The other main concern is user data management when roaming. While this does present some design challenges, the nature of roaming charging also makes this data less threatening than, say, SM energy usage data, which gives adversaries insights into a user's home. Additionally, this problem has received considerable research attention. For these reasons, we classify EVs as a low-level threat overall.

3.7.14 Communication

Although we noted in Section 3.5.14 and Section 3.6.14 that network security has not been a strong consideration for TE researchers thus far, we do not consider it a priority remaining issue for TE research. This is due to the fact that network security is already a richly researched field outside of the context of TE; as Eisele et al. [13] note, communication security research is orthogonal to TE security research. Similarly, cryptography research is a field that supports TE development, rather than being a subset of it. TE resources would be better spent enhancing the security of TE protocols and applications to reduce the impact of potentially insecure communication devices.

3.8 Limitations and Threats to Validity

There are common validity threats in literature reviews that also apply here [64]. From an internal validity perspective:

- We use only one database and our query was limited to English papers and only a few synonyms per concept. This means that we likely missed relevant papers in our selection. This was partly mitigated by the wide coverage of that database (Scopus) and its constant updates. Further mitigation was provided by relying on a complementary search technique, namely backward snowballing.
- Another decision was to exclude non-peer-reviewed documents, including white papers and patents. This means that the selected papers, and our conclusions, are biased towards academic contributions, at the expense of purely industrial concerns and solutions that might have brought complementary views and information.
- As most of the paper selection and data extraction was performed by one person (the first author), the process might have been subject to various unconscious biases. This was partly mitigated by involving some of the other co-authors for borderline decisions. We have also made the raw data extracted from the selected papers available online for documenting, reproducing, and extending our literature review.

From external validity and conclusion validity perspectives, there are also important threats and limitations that can be identified:

- The relatively narrow subject of this review, namely cybersecurity concerns that affect TEMs that use blockchains, presents a limitation in terms of available research. TE and blockchains are both relatively new fields. Among papers that fall into this category, many include little to no security analysis. Often, security concerns that were discussed were those that were addressed by including a blockchain, i.e., concerns that were not relevant to our review.
- Perhaps just as impactful is the limited real-world deployment of TE systems. Most of the systems investigated for this review are in the proof-of-concept stage, meaning that they have not been tested in the wild, leading to a paucity of empirical data. It is difficult for researchers to predict which attacks will be most feasible or rewarding for hackers, but this is the analysis which we must rely on until TEMs are implemented broadly. As such, it is likely that some weaknesses have gone unnoticed, leaving gaps in security solutions.

3.9 Conclusion and Future Work

While TE promises to revolutionize power management operations, it introduces cybersecurity threats that need to be managed before it can be widely implemented. DLT is a

widely employed solution for maintaining market integrity, but it only solves a subset of the issues brought on by TE paradigms.

In rigorously studying our selection of 28 peer-reviewed contributions to the DLT-based TE literature, we found 14 important cybersecurity concerns that remain beyond the scope of the blockchain. Among those, we identified five that we consider as presenting the highest risks: market attacks, FDI, SPOF, energy usage data leakage, and privacy. Although other security threats remained, they were either of a lower priority or outside the scope of purely TE-centric research.

For market attacks, we believe proactive attack detection and mitigation measures within the TE market / application layer should be researched. Existing literature on network security could likely be mined for techniques that identify malicious nodes and handle them in a more sophisticated manner than with access blocking. It is also possible that new techniques will have to be designed due to the ethical considerations related to power access, which could present an interesting research opportunity.

Considering FDI, existing solutions are themselves vulnerable to additional threats, such as algorithm divergence. Efforts should be made to design FDIA countermeasures that are robust to such attacks.

SPOF is a critical structural threat and should be avoided with great effort, both to protect data as well as the availability of a vital utility. Improving this factor will require researchers to find ways around operations that have traditionally relied on centralized computation or trusted entities. This might include designing new power management protocols that are tailored to the distributed nature of TE, or discovering algorithms to decentralize existing operations (as was done with OPF [37]), both of which present exciting research opportunities.

Energy usage data is often ignored as a privacy element within the TE literature. However, we found it to be one of the most dangerous potential data leaks and feel that more robust solutions need to be crafted. Data anonymization techniques are insufficient due to the network topology; as such, future research should focus on methods of obfuscating or encrypting the data while maintaining its usefulness to relevant parties. This could be via improved aggregation techniques, which we discuss in Section 3.6.3, noising methods that preserve statistical qualities of the data, or some novel technique that suits the limited computing power and sensitive nature of the domain.

Finally, privacy research is found to be lacking overall in the TE literature. Additionally, many solutions that address privacy ignore practical outcomes of total anonymity and introduce their own problems with security and trust. Balance must be sought between anonymity and trust in order to establish a legitimate and fair market. Future work should

focus on creating systems within TE markets that maximize consumer privacy without sacrificing traceability, an ideal that has not yet been met in our view.

Bibliography

[1] Yohannes T. Aklilu and Jianguo Ding. Survey on blockchain for smart grid management, control, and operation. *Energies*, 15(1), 2022. ISSN 1996-1073. doi:10.3390/en15010193.

[2] Merlinda Andoni, Valentin Robu, David Flynn, Simone Abram, Dale Geach, David Jenkins, Peter McCallum, and Andrew Peacock. Blockchain technology in the energy sector: A systematic review of challenges and opportunities. *Renewable and Sustainable Energy Reviews*, 100:143–174, 2019. ISSN 1364-0321. doi:10.1016/j.rser.2018.10.014.

[3] Ioannis Andrea, Chrysostomos Chrysostomou, and George Hadjichristofi. Internet of things: Security vulnerabilities and challenges. In *2015 IEEE Symposium on Computers and Communication (ISCC)*, pages 180–187, 2015. doi:10.1109/ISCC.2015.7405513.

[4] Ali Ardeshiri, Amir Lotfi, Reza Behkam, Arash Moradzadeh, and Ashkan Barzkar. *Introduction and Literature Review of Power System Challenges and Issues*, pages 19–43. Springer International Publishing, Cham, 2021. doi:10.1007/978-3-030-77696-1_2.

[5] Carlos Barreto, Taha Eghtesad, Scott Eisele, Aron Laszka, Abhishek Dubey, and Xenofon Koutsoukos. Cyber-attacks and mitigation in blockchain based transactive energy systems. In *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, volume 1, pages 129–136, 2020. doi:10.1109/ICPS48405.2020.9274708.

[6] Bitcoin Wiki. Script, 2021. `https://en.bitcoin.it/wiki/Script`.

[7] Sergio Cantillo-Luna, Ricardo Moreno-Chuquen, Harold R. Chamorro, Vijay K. Sood, Shahriar Badsha, and Charalambos Konstantinou. Blockchain for distributed energy resources management and integration. *IEEE Access*, 10:68598–68617, 2022. doi:10.1109/ACCESS.2022.3184704.

[8] Alvaro A. Cárdenas, Saurabh Amin, Zong-Syun Lin, Yu-Lun Huang, Chi-Yen Huang, and Shankar Sastry. Attacks against process control systems: Risk assessment,

detection, and response. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '11, page 355–366, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450305648. doi:10.1145/1966913.1966959.

[9] Rohit Chandra, Krishnanand Kaippilly Radhakrishnan, and Sanjib Kumar Panda. Privacy protected product differentiation through smart contracts based on bilateral negotiations in peer-to-peer transactive energy markets. *Sustainable Energy, Grids and Networks*, 34:100997, 2023. ISSN 2352-4677. doi:10.1016/j.segan.2023.100997.

[10] Matthew Channon and James Marson. THE liability for cybersecurity breaches of connected and autonomous vehicles. *Computer Law & Security Review*, 43:105628, 2021. ISSN 0267–3649. doi:10.1016/j.clsr.2021.105628.

[11] William Cox and Toby Considine. Structured energy: Microgrids and autonomous transactive operation. In *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6, Feb 2013. doi:10.1109/ISGT.2013.6497919.

[12] Marina Dorokhova, Jérémie Vianin, Jean-Marie Alder, Christophe Ballif, Nicolas Wyrsch, and David Wannier. A blockchain-supported framework for charging management of electric vehicles. *Energies*, 14(21), 2021. ISSN 1996-1073. doi:10.3390/en14217144.

[13] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[14] Nawfal Fadhel, Federico Lombardi, Leonardo Aniello, Andrea Margheri, and Vladimiro Sassone. Towards a semantic modelling for threat analysis of iot applications: A case study on transactive energy. In *Living in the Internet of Things (IoT 2019)*, pages 1–6, 2019. doi:10.1049/cp.2019.0147.

[15] Earlence Fernandes, Jaeyeon Jung, and Atul Prakash. Security analysis of emerging smart home applications. In *2016 IEEE Symposium on Security and Privacy (SP)*, pages 636–654, 2016. doi:10.1109/SP.2016.44.

[16] Nina Flores, Heather McBrien, Vivian Do, Mathew Kiang, Jeffrey Schlegelmilch, and Joan Casey. The 2021 texas power crisis: distribution, duration, and disparities. *Journal of Exposure Science & Environmental Epidemiology*, 33:1–11, 08 2022. doi:10.1038/s41370-022-00462-5.

[17] Jie Gu, Xiao-Yu Cao, Yao Fu, Zong-Wu He, Ze-Jie Yin, Hua-Lei Yin, and Zeng-Bing Chen. Experimental measurement-device-independent type quantum key distribution with flawed and correlated sources. *Science Bulletin*, 67(21):2167–2175, 2022. ISSN 2095-9273. doi:10.1016/j.scib.2022.10.010.

[18] Youbiao He, Gihan J. Mendis, and Jin Wei. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid*, 8(5):2505–2516, Sep. 2017. ISSN 1949-3061. doi:10.1109/TSG.2017.2703842.

[19] M. Jayachandran, K. Prasada Rao, Ranjith Kumar Gatla, C. Kalaivani, C. Kalaiarasy, and C. Logasabarirajan. Operational concerns and solutions in smart electricity distribution systems. *Utilities Policy*, 74:101329, 2022. ISSN 0957-1787. doi:10.1016/j.jup.2021.101329.

[20] Hamzah Khan and Tariq Masood. Impact of blockchain technology on smart grids. *Energies*, 15(19), 2022. ISSN 1996-1073. doi:10.3390/en15197189.

[21] Mohsen Khorasany, Yateendra Mishra, and Gerard Ledwich. A decentralised bilateral energy trading system for peer-to-peer electricity markets. *IEEE Transactions on Industrial Electronics*, 06 2019. doi:10.1109/TIE.2019.2931229.

[22] Desen Kirli, Benoit Couraud, Valentin Robu, Marcelo Salgado-Bravo, Sonam Norbu, Merlinda Andoni, Ioannis Antonopoulos, Matias Negrete-Pincetic, David Flynn, and Aristides Kiprakis. Smart contracts in energy systems: A systematic review of fundamental approaches and implementations. *Renewable and Sustainable Energy Reviews*, 158:112013, 2022. ISSN 1364-0321. doi:10.1016/j.rser.2021.112013.

[23] Aron Laszka, Abhishek Dubey, Michael Walker, and Douglas Schmidt. Providing privacy, safety, and security in iot-based transactive energy systems using distributed ledgers. In *Proceedings of the Seventh International Conference on the Internet of Things*, 10 2017. doi:10.1145/3131542.3131562.

[24] Michelle Lauer, Rupamathi Jaddivada, and Marija Ilić. Secure blockchain-enabled dymonds design. In *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, COINS '19, page 191–198, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366403. doi:10.1145/3312614.3312654.

[25] Zhiyi Li, Shay Bahramirad, Aleksi Paaso, Mingyu Yan, and Mohammad Shahidehpour. Blockchain for decentralized transactive energy management system in networked microgrids. *The Electricity Journal*, 32(4):58–72, 2019. ISSN 1040-6190.

doi:10.1016/j.tej.2019.03.008. Special Issue on Strategies for a sustainable, reliable and resilient grid.

[26] Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker. Inferring personal information from demand-response systems. *IEEE Security & Privacy*, 8(1):11–20, 2010. doi:10.1109/MSP.2010.40.

[27] Yao Liu, Peng Ning, and Michael K. Reiter. False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur.*, 14(1), jun 2011. ISSN 1094-9224. doi:10.1145/1952982.1952995.

[28] Federico Lombardi, Leonardo Aniello, Stefano De Angelis, Andrea Margheri, and V. Sassone. A blockchain-based infrastructure for reliable and cost-effective iot-aided smart grids. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, 01 2018. doi:10.1049/cp.2018.0042.

[29] Adriana Mar, Pedro Pereira, and JoÃ£o F. Martins. A survey on power grid faults and their origins: A contribution to improving power grid resilience. *Energies*, 12(24), 2019. ISSN 1996-1073. doi:10.3390/en12244667.

[30] Bacem Mbarek, Stanislav Chren, Bruno Rossi, and Tomáš Pitner. *An Enhanced Blockchain-Based Data Management Scheme for Microgrids*, pages 766–775. Springer, 03 2020. ISBN 978-3-030-44037-4. doi:10.1007/978-3-030-44038-1_70.

[31] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[32] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Designing microgrid energy markets: A case study: The brooklyn microgrid. *Applied Energy*, 210:870–880, 2018. ISSN 0306-2619. doi:10.1016/j.apenergy.2017.06.054.

[33] Mihail Mihaylov, Sergio Jurado, Narcís Avellana, Kristof Van Moffaert, Ildefons Magrans de Abril, and Ann Nowé. Nrgcoin: Virtual currency for trading of renewable energy in smart grids. In *11th International Conference on the European Energy Market (EEM14)*, pages 1–6, May 2014. doi:10.1109/EEM.2014.6861213.

[34] Muhammad Baqer Mollah, Jun Zhao, Dusit Niyato, Kwok-Yan Lam, Xin Zhang, Amer M. Y. M. Ghias, Leong Hai Koh, and Lei Yang. Blockchain for future smart grid: A comprehensive survey. *IEEE Internet of Things Journal*, 8(1):18–43, Jan 2021. ISSN 2327-4662. doi:10.1109/JIOT.2020.2993601.

[35] Vidya Krishnan Mololoth, Saguna Saguna, and Christer Åhlund. Blockchain and machine learning for future smart grids: A review. *Energies*, 16(1), 2023. ISSN 1996-1073. doi:10.3390/en16010528.

[36] Erica Mourão, João Felipe Pimentel, Leonardo Murta, Marcos Kalinowski, Emilia Mendes, and Claes Wohlin. On the performance of hybrid search strategies for systematic literature reviews in software engineering. *Information and Software Technology*, 123:106294, 2020. doi:10.1016/j.infsof.2020.106294.

[37] Eric Münsing, Jonathan Mather, and Scott Moura. Blockchains for decentralized optimization of energy resources in microgrid networks. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 2164–2171, 08 2017. doi:10.1109/CCTA.2017.8062773.

[38] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain: A path to grid modernization and cyber resiliency. In *2017 North American Power Symposium (NAPS)*, pages 1–5, 09 2017. doi:10.1109/NAPS.2017.8107313.

[39] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23, 09 2017. doi:10.1109/RWEEK.2017.8088642.

[40] Satoshi Nakamoto. A peer-to-peer electronic cash system, Oct 2008. URL `https://bitcoin.org/en/bitcoin-paper`.

[41] Masoumeh Nazari, Siavash Khorsandi, and Jaber Babaki. Security and privacy smart contract architecture for energy trading based on blockchains. In *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, pages 596–600, 2021. doi:10.1109/ICEE52715.2021.9544155.

[42] Ivica Nikolić, Aashish Kolluri, Ilya Sergey, Prateek Saxena, and Aquinas Hobor. Finding the greedy, prodigal, and suicidal contracts at scale. In *34th Annual Computer Security Applications Conference*, ACSAC '18, page 653–663, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450365697. doi:10.1145/3274694.3274743.

[43] Chitu Okoli. A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37, 2015. doi:10.17705/1CAIS.03743.

[44] Adeiza J. Onumanyi, Sherrin J. Isaac, Carel P. Kruger, and Adnan M. Abu-Mahfouz. Transactive energy: State-of-the-art in control strategies, architectures, and simulators. *IEEE Access*, 9:131552–131573, 2021. ISSN 2169-3536. doi:10.1109/ACCESS.2021.3115154.

[45] Tim Ruffing, Pedro Moreno-Sanchez, and Aniket Kate. Coinshuffle: Practical decentralized coin mixing for bitcoin. In Mirosław Kutyłowski and Jaideep Vaidya, editors, *Computer Security - ESORICS 2014*, pages 345–364, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11212-1.

[46] Shammya Saha, Nikhil Ravi, Kári Hreinsson, Jaejong Baek, Anna Scaglione, and Nathan G. Johnson. A secure distributed ledger for transactive energy: The electron volt exchange (eve) blockchain. *Applied Energy*, 282:116208, 2021. ISSN 0306-2619. doi:10.1016/j.apenergy.2020.116208.

[47] Shammya Shananda Saha, Christopher Gorog, Adam Moser, Anna Scaglione, and Nathan G. Johnson. Integrating hardware security into a blockchain-based transactive energy platform. In *2020 52nd North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2021. doi:10.1109/NAPS50074.2021.9449802.

[48] Khaled Shuaib, Juhar Abdella, Farag Sallabi, and Mohammed Abdel Hafez. Using blockchains to secure distributed energy exchange. In *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 622–627, 04 2018. doi:10.1109/CoDIT.2018.8394815.

[49] F. Siddiqui, S. Zeadally, C. Alcaraz, and S. Galvao. Smart grid privacy: Issues and solutions. In *2012 21st International Conference on Computer Communications and Networks (ICCCN)*, pages 1–5, July 2012. doi:10.1109/ICCCN.2012.6289304.

[50] Ye-Byoul Son, Jong-Hyuk Im, Hee-Yong Kwon, Seong-Yun Jeon, and Mun-Kyu Lee. Privacy-preserving peer-to-peer energy trading in blockchain-enabled smart grids using functional encryption. *Energies*, 13(6), 2020. ISSN 1996-1073. doi:10.3390/en13061321.

[51] Wenjie Song, Yun Li, and Dongmei Yang. Research on the application of blockchain in the energy power industry in china. *Journal of Physics: Conference Series*, 1176 (4):042079, mar 2019. doi:10.1088/1742-6596/1176/4/042079.

[52] Nick Szabo. Smart contracts, 1994. https://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html.

[53] Ognjen Vuković and György Dán. Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks. *IEEE Journal on Selected Areas in Communications*, 32(7):1500–1508, 2014. doi:10.1109/JSAC.2014.2332106.

[54] Dai Wang, Xiaohong Guan, Ting Liu, Yun Gu, Chao Shen, and Zhanbo Xu. Extended distributed state estimation: A detection method against tolerable false data injection attacks in smart grids. *Energies*, 7(3):1517–1538, 2014. ISSN 1996-1073. doi:10.3390/en7031517.

[55] Shen Wang, Ahmad F. Taha, Jianhui Wang, Karla Kvaternik, and Adam Hahn. Energy crowdsourcing and peer-to-peer energy trading in blockchain-enabled smart grids. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(8):1612–1623, Aug 2019. ISSN 2168-2232. doi:10.1109/TSMC.2019.2916565.

[56] Longfei Wei, Luis Puche Rondon, Amir Moghadasi, and Arif I. Sarwat. Review of cyber-physical attacks and counter defense mechanisms for advanced metering infrastructure in smart grid. In *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–9, 2018. doi:10.1109/TDC.2018.8440552.

[57] Yair Wiseman. Autonomous vehicles. In *Research Anthology on Cross-Disciplinary Designs and Applications of Automation*, pages 878–889. IGI Global, 2022. doi:10.4018/978-1-6684-3694-3.ch043.

[58] Wenti Yang, Zhitao Guan, Longfei Wu, Xiaojiang Du, and Mohsen Guizani. Secure data access control with fair accountability in smart grid data sharing: An edge blockchain approach. *IEEE Internet of Things Journal*, 8(10):8632–8643, 2021. doi:10.1109/JIOT.2020.3047640.

[59] Charithri Yapa, Chamitha de Alwis, Madhusanka Liyanage, and Janaka Ekanayake. Survey on blockchain for future smart grids: Technical aspects, applications, integration challenges and future research. *Energy Reports*, 7:6530–6564, 2021. ISSN 2352-4847. doi:10.1016/j.egyr.2021.09.112.

[60] Ping Yi, Ting Zhu, Qingquan Zhang, Yue Wu, and Jianhua Li. A denial of service attack in advanced metering infrastructure network. In *2014 IEEE International Conference on Communications (ICC)*, pages 1029–1034, 2014. doi:10.1109/ICC.2014.6883456.

[61] Hua-Lei Yin, Yao Fu, Chen-Long Li, Chen-Xun Weng, Bing-Hong Li, Jie Gu, Yu-Shuo Lu, Shan Huang, and Zeng-Bing Chen. Experimental quantum secure network with

digital signatures and encryption. *National Science Review*, 10(4), 10 2022. ISSN 2095-5138. doi:10.1093/nsr/nwac228.

[62] Ishtiaque Zaman and Miao He. A multilayered semi-permissioned blockchain based platform for peer to peer energy trading. In *2021 IEEE Green Technologies Conference (GreenTech)*, pages 279–285, 2021. doi:10.1109/GreenTech48523.2021.00052.

[63] Yue Zhang, Scott Eisele, Abhishek Dubey, Aron Laszka, and Anurag K. Srivastava. Cyber-physical simulation platform for security assessment of transactive energy systems. In *2019 7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, pages 1–6, 2019. doi:10.1109/MSCPES.2019.8738802.

[64] Xin Zhou, Yuqin Jin, He Zhang, Shanshan Li, and Xin Huang. A map of threats to validity of systematic literature reviews in software engineering. In *2016 23rd Asia-Pacific Software Engineering Conference (APSEC)*, pages 153–160, 2016. doi:10.1109/APSEC.2016.031.

Chapter 4

Energy Usage Data Privacy (CHEA)

The transactive energy market is an emerging development in energy economics built on advanced metering infrastructure. Data generated in this context is often required for market operations, while also being privacy sensitive. This dual concern has necessitated the development of various methods of obfuscation in order to maintain privacy while still facilitating operations. While data aggregation is a common approach in this context, many of the existing aggregation methods rely on additional network components or lack flexibility. In this paper, we introduce Cyclic Homomorphic Encryption Aggregation (CHEA), a secure aggregation protocol that eliminates the need for additional network components or complicated key distribution schemes, while providing additional capabilities compared to similar protocols. We validate our scheme with formal security analysis as well as a software simulation of a transactive energy network running the scheme. Results indicate that CHEA performs well in comparison to similar works, with minimal communication overhead. Additionally, CHEA retains all standard security properties held by other aggregation schemes, while improving flexibility and reducing infrastructural requirements. Our scheme operates on similar assumptions as other works, but current smart metering hardware lags in terms of processing power, making the scheme infeasible on the current generation of hardware. However, these capabilities should quickly advance to an accommodating state. With this in mind, and given the results, we believe CHEA is a strong candidate for aggregating transactive energy data.

4.1 Introduction

Advanced metering infrastructures (AMI) have enabled new functionalities in the power delivery sector. Smart meters enable frequent collection of detailed energy profile data from the homes of consumers [39]. Additionally, smart energy grids facilitate the transmission, storage, and analytic computations on consumers' energy profile data by means of Information and Communication Technology (ICT) networks that inter-operate with energy grids. Novel developments of intelligent demand response by means of smart thermostats, smart plugs, smart lighting, and smart appliances, and optimal power flow algorithms, AI-based electric and thermal/cooling loads predictions are examples of technological progresses in the smart energy grid sector [16]. These progressions will enable greater energy economic efficiency through models such as transactive energy (TE) markets, further enabling high frequency peer-to-peer energy trading between prosumers (consumers who also produce energy) and power utilities [10]. These models will also contribute to greater energy efficiency, due to shorter electrical transmission distances as well as improved load balancing, demand response, and AI-assisted predictions.

Although a standardized transactive energy market (TEM) platform has yet to arise, innovative research works and proposals of TEM platforms are being published regularly, as recently surveyed by Garcia et al. [14]. These platforms often employ blockchain technology to secure private information such as consumer trading history, identifying information, and account balances, among others [1]. However, an often-overlooked privacy concern is that of energy usage data. Usage data is regularly required for critical operations such as automated demand response, power flow optimization, and billing. While it is not inherently identifying, many researchers believe it should fall under the category of protected private information [28]. This is due to the fact that this data can facilitate many malicious attacks on consumers, including profiling, trading pattern recognition, and malicious trading [38].

Because energy usage data is required for many critical functions, it cannot be obfuscated entirely, and instead must be accessible in some form to the distribution systems operator (DSO). Proposals for adapting this information have included adding noise to the data, performing algebraic transformations that preserve large scale statistical properties, and using battery banks to disguise a household's true energy needs [3, 17, 40]. However, the approach we believe to be the most effective is aggregation [24]. Data aggregation is successfully used in many other industries, including in healthcare [6]. It has the benefit of not changing the data (like noising and algebraic transformation would), as well as not requiring additional equipment to be installed at the household (like battery-based techniques would).

Many protocols have been proposed to facilitate data aggregation in smart grids. However, as will be seen in the related work, these protocols suffer from drawbacks in terms of hardware requirements and/or flexibility. Additionally, these protocols were not designed with transactive energy in mind.

Our proposed protocol, named *Cyclic Homomorphic Encryption Aggregation (CHEA)* was designed from the ground up to be integrated into a transactive energy market environment. It is designed to be both flexible and distributed, which are features that make it a natural complement to blockchain-based TEM models.

In this paper, we outline the design of the CHEA protocol, investigate its strengths and weaknesses from a security standpoint, examine the simulation results of the protocol, and compare its performance with other energy data aggregation schemes.

The **key contributions** of CHEA, introduced in this paper, include:

1. Network infrastructure: while blockchain-enabled TE networks and smart grids often have auxiliary nodes such as fog nodes (FNs) or edge nodes, CHEA does not rely on them. This enables networks that do not have auxiliary nodes to use our scheme without installing additional hardware. Networks that already have such nodes will still benefit, as these nodes will have more resources available to them to perform other tasks (e.g., blockchain validation) than if they were involved with aggregation.

2. Key distribution: many proposed aggregation techniques in this space rely on complicated cryptographic key distribution schemes in order to function. This is because data in this context must be aggregated regularly, requiring synchronization between data generating, requesting, and aggregating parties. In contrast, other cases – such as medical data – might only need to be aggregated once for a study. Our scheme sidesteps the issue by performing the aggregation process entirely locally to the group of data owners.

3. Single points of failure: trusted authorities (TA) and FNs can be single points of failure from both a security and operational standpoint. If a FN goes offline, then its region of smart meters will no longer be able to report to the DSO. If a TA goes offline then the entire network will not be able to report data. If either are compromised then an adversary would be able to acquire individual readings from consumers, enabling attacks discussed in Section 4.2. Our scheme does not rely on either of these mechanisms.

4. Scalability: it is noted by Ming et al. [29] and others that the resources of TAs could become overwhelmed if the area covered by a DSO became too large. Since key

generation happens locally in CHEA, this concern does not exist. Additionally, as a neighbourhood grows, new aggregating nodes would need to be installed to support additional meters (in schemes that require them). Our scheme is not subject to this scaling factor. The distributed nature of the protocol means that scaling is limited only by the communication capacity of the DSO.

5. Decentralization: as discussed in [38] the trend in TEM development is towards decentralized applications. A decentralized solution for data aggregation supports this trend and confers the same benefits seen by turning other power management mechanisms into distributed applications.

6. Flexibility: CHEA supports flexible group membership, group size (which corresponds to aggregated data resolution), and locality, meaning that groups can be sparse or dense. These variables allow the scheme to support diverse applications that might have different data requirements.

7. Fault tolerance: on-the-fly group generation means that member smart meters are guaranteed to be active, provided that they do not malfunction during the ∼500ms aggregation cycle.

This paper will be of interest to practitioners who are designing Transactive Energy Markets (TEMs), smart meters, and smart grid (SG) infrastructure. It provides a novel solution to a common problem in these domains, making it directly useful to SG and TEM operators. Smart meter hardware designers should consider these solutions to inform the technical specifications of the meters. In addition, researchers can use our results as a comparison basis for novel schemes, or they may choose to build on our scheme or adapt it for other application fields.

The rest of this paper is structured as follows. Section 4.2 presents the background and motivation behind the design and implementation of an innovative cybersecurity protocol for smart grids and transactive energy markets' stakeholders. Section 4.3 outlines similar research works in the field and discusses the various approaches to smart grid data aggregation that have been considered. In Section 4.4, the CHEA protocol is described in detail. In Section 4.5, a formal security analysis is performed to demonstrate the privacy-related benefits of the proposed approach. In Section 4.6, the implemented CHEA protocol is demonstrated in a simulated smart grid network environment. We also discuss the simulation results, as well as how they compare to simulations of similar schemes. Section 3.8 discusses limitations of our experiments. Finally, in Section 4.8, we conclude the study with a reflection on the results and suggestions for further research.

4.2 Background and Motivation

Transactive energy markets (TEMs) are meant to represent autonomous distributed platforms for trading energy and reserve among prosumers of energy (with numerous and different types of distributed energy resources – DERs) and distribution system operators (DSOs). Built on top of smart grid infrastructure, a TEM consists of software that facilitates the trading of energy directly between consumers, prosumers, and DSOs. This capability has the potential to increase efficiency, lower costs, and reduce environmental impacts [7, 27].

A typical TEM architecture consists of power consumers equipped with smart meters (SMs), prosumers with distributed energy resources, battery energy storage systems (BESS), electric vehicles (EVs), and a traditional power generation station [27, 30, 23]. Additionally, there will be some form of distributed system operator, sometimes referred to as control center (CC) or cloud control center (CCC), who handles billing, data management, demand response, demand prediction, and other critical maintenance and operational tasks [22, 30].

There is no shortage of data generated within smart grids and transactive energy markets. While some of the mentioned data is personal, such as a user's transaction history, identity, or credit standing, other sources are more opaque.

CHEA is concerned with the privacy protection of energy usage data. This data is required for many critical functions:

- physical operation of the power grid, for example, when performing state estimation;
- novel functionality offered by the smart grid, such as automated demand response;
- operating a transactive energy market, for billing, transaction verification, etc.

A recent literature review that examined privacy concerns in various transactive energy market implementations found that, while steps were generally taken in TEM proposals to protect identifying information, energy usage data is often mishandled [38].

Several studies, including those from McDaniel and McLaughlin [28] and from Lisovich et al. [25], note that energy usage data should be considered private data, and that its leakage can lead to a number of undesirable outcomes. These can include revealing appliance profiles in consumer households, spying, facilitating theft, and allowing hackers to understand activity within the home [1, 25, 28].

While proposed TE models have neglected this area of concern, smart grid research has produced a number of novel methods for making such data available for important functions, while maintaining user privacy [1]. Additionally, existing methods of aggregating data have been adapted for smart grid environments.

Some proposed methods include:

- *Algebraic Transformation* [3]: Algebraic transformation refers to a family of mathematical techniques that enable modifying a set of data so that the individual values are altered, but results of certain computations remain consistent.

- *Battery Filtering*: Battery filtering is a method of disguising usage data proposed by Kalogridis et al. [17]. This method suggests using a battery energy storage system in the home (this could be an EV, a Tesla Powerwall, or another BESS product) as a buffer between the home and the power grid. The BESS would be discharged to meet the short term-term energy demands of the home and charged regularly from the grid. The use of such an energy buffer has the effect of obfuscating the real-time energy usage patterns in the home while maintaining power availability.

- *Data Aggregation* [24]: Commonly used in the medical sciences [6], data aggregation is the process of summarizing data for analysis. Generally, this will consist of performing a summation of each dimension or feature of the data before transmitting it to the data receiver. This has the benefit of allowing analysis of real data without compromising the privacy of individuals – since only the aggregate is analyzed, the data receiver cannot tie any individual measurement to a specific person.

- *Random Noise* [40]: The introduction of random noise is a common method of providing access to data while preserving privacy and preventing statistical attacks. This involves generating noise, or random datapoints, using a random function, such as Perlin noise [21]. Noise can have different statistical properties, such as smooth transitions between points. The generated noise is then used to modify the real measured values (for example, by adding the absolute value of the noise at a particular point) enough to disguise them from attackers, but subtly enough that the result is still useful to the party analyzing the data. In circumstances where approximations suffice, this can be an appropriate solution.

Table 4.1 displays the attributes of the privacy-preserving data collection methods discussed. **MA1** is hardware independence; **MA2** is preservation of exact measurements; **MA3** is discrete values reporting; and **MA4** is timely data reporting (for use in demand response applications).

Our scheme (CHEA) performs data aggregation in order to preserve privacy. We find this to be most appropriate for the TE environment for several reasons. Unlike introducing random noise and algebraic transformation, aggregation allows the DSO to work with real values. While some operations – such as demand response – might function adequately given approximate values, others – for example, transaction verification – are not so forgiving. Battery filtering is a compelling solution, but suffers from requiring prosumers to own expensive hardware. User privacy should not be predicated on purchasing additional equipment, and for this reason, a software solution is preferable. Of course, users with EVs or BESSs can still employ battery filtering if they desire.

Homomorphic encryption (HE) is a mechanism commonly employed to facilitate privacy-preserving data aggregation [44]. HE refers to encryption in which mathematical operations can be performed on the cyphertext, and the result will be equivalent to having encrypted the result of the same operation performed on the plaintext. For example, assuming an HE-based encryption function $E()$ and the use of the sum operator $(+)$ as the aggregation function, then $E(x) + E(y) = E(x + y)$. Hence, the aggregated result can be decrypted without knowing the nature of the individual operands.

While there are many aggregation schemes based on HE, simulation results (presented later in Section 4.5) indicate that CHEA excels uniquely in a few key areas (on top of being a "truer" distributed application), including privacy, collusion-resistance, and flexibility, which make it a valuable contribution to the field.

4.3 Related Work

There is extensive literature on the topic of data privacy. Even if we restrict ourselves only to a smart grid context, there exist many solutions, as hinted by Section 4.2. In this section, we focus our analysis of similar works on privacy solutions that employ data aggregation,

Table 4.1: Comparison of privacy-preserving data collection methods

Method	**MA1 (HW)**	**MA2 (exact)**	**MA3 (discr.)**	**MA4 (timely)**
Algebraic Transformation	yes	no	yes	yes
Battery Filtering	no	no	yes	no
Data Aggregation	yes	yes	no	yes
Random Noise	yes	no	yes	yes

and ideally those that are also based on homomorphic encryption. A search on Scopus (the most comprehensive index for such literature) with the query privacy AND ("smart grid" OR "transactive energy") AND "data aggregation" AND "homomorphic encryption" executed on December 1, 2023, returned 100 journal and conference papers. The related work closest to ours is discussed below.

There are many privacy-preserving aggregation schemes that rely on homomorphic encryption [44, 45, 34, 9, 11, 46, 29]. Of those, a majority [45, 34, 9, 11] use the Paillier cryptosystem [31, 32], which we also chose to use when implementing CHEA due to its relative computational efficiency. Other HE cryptosystems, such as Elliptic Curve ElGamal [13, 20], are occasionally used and some researchers compared the performance of multiple cryptosystems [42].

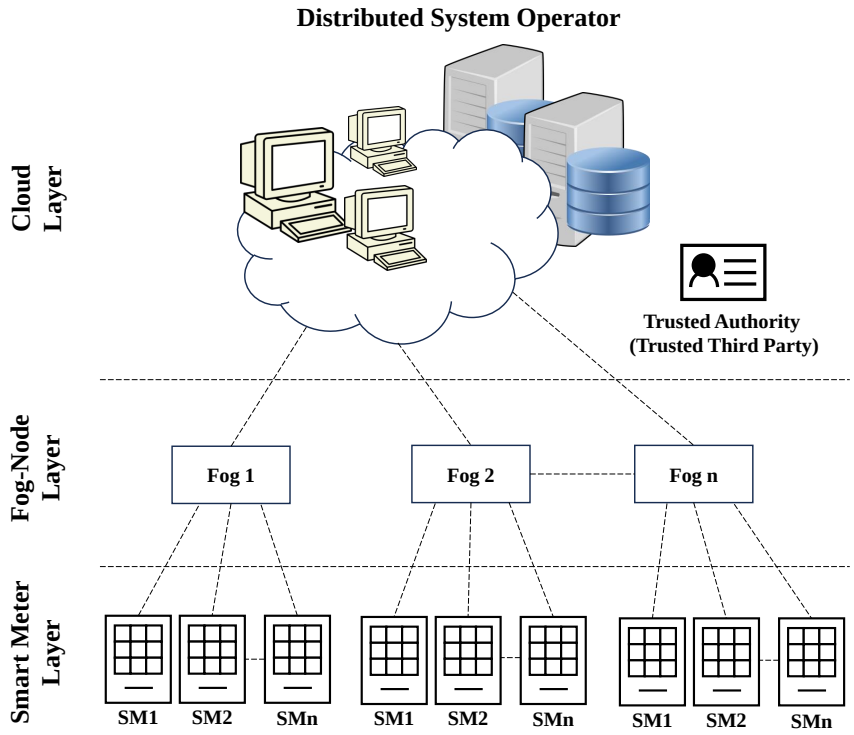

Figure 4.1: Typical architecture for an aggregation protocol employing fog nodes. (Figure adapted from Khan et al. [19])

The vast majority of aggregation protocols [19, 35, 18, 8, 29, 41, 46, 26] use a system

in which smart meters connect to an intermediate node who aggregates the data before sending it to the distributed system operator. This architecture is outlined in Figure 4.1.

There are several drawbacks to this architecture. One is that each aggregating node represents a single point of failure, whose outage would prevent a large region of smart meters (the ones aggregated by the failing node) from reporting data. Another is that collusion between the DSO (who owns the HE private key) and an aggregating node is trivial and hence may compromise many households. Finally, there are concerns of scalability given the need to add physical nodes to support neighbourhood expansion.

Another common theme is the use of a trusted authority (TA) to generate and distribute cryptographic keys. This can leave the system vulnerable to a number of attacks, including man-in-the-middle, false data injection, and data deletion if the TA is compromised [44, 45, 29]. Key distribution is a central problem for traditional schemes, such that it has become its own area of research. For example, Cheng et al. [12] propose a key distribution scheme that reduces communication overhead and improves security. One can thus conclude that avoiding key distribution altogether would be a valuable proposition.

Some solutions do attempt to address these concerns. For example, the scheme by Chen et al. [11] attempts to improve resilience and flexibility by enabling variable subsets of meters, but still relies on a TA to generate and distribute keys.

Instead of relying on external fog nodes, several of the aggregation schemes use in-network aggregation, or aggregation that happens on the smart meters as the data is sent along. This means that intermediate meters will add their encrypted measurements to those sent by previous meters, a strategy we also use in CHEA. In fact, one of the earliest examples of a smart-grid aggregation scheme for data privacy used this technique. Li et al. [24] describe a technique in which an aggregation "tree" is created to describe the path the data will take. In their case they are employing in-network aggregation for the sake of computation and communication efficiency, so the aggregation tree remains static based on the network topology (as it would not make sense to send the data down a less efficient path). Despite making use of this technique for a different purpose than CHEA, it is encouraging to see that the assumption we make about smart meters being capable of performing such aggregation is not unprecedented in the literature.

Table 4.2 presents the attributes of each of the relevant schemes.

Table 4.2: Comparison of related schemes (SA1: fog node independence; SA2: multi-dimensionality; SA3: collusion resistance; SA4: forward/backward secrecy; SA5: fault tolerance; SA6: dynamic membership; SA7: dynamic/variable group size; SA8: trusted third party/trusted authority independence)

Scheme	SA1	SA2	SA3	SA4	SA5	SA6	SA7	SA8
Liu et al. [26]	no	yes	yes	yes	yes	yes	no	no
Song et al. [36]	no	no	no	yes	no	yes	no	yes
Zuo et al. [47]	no	yes	yes	no	no	no	no	yes
Zhang and Liu [45]	no	yes	yes	yes	no	yes	yes	no
CHEA	yes	no	yes	yes	yes	yes	yes	yes

Another example which shows conceptual similarities to our protocol is the proposed solution by Gomez Marmol et al. [15]. In their scheme, bilinear homomorphism is employed because it allows the aggregate of the data to have a separate decryption key from the components that make up the aggregate. This way, the meters can send their encrypted data to the DSO directly (avoiding intermediate nodes), followed by an aggregate key, and the DSO can only decrypt the data once it has summed the ciphertexts (with some exceptions).

4.4 Scheme

CHEA is built around homomorphic encryption (HE), with algorithmic parameters described in Table 4.3. Although a specific HE implementation is not required to employ the protocol, for the purposes of implementation we have used the Paillier cryptosystem [31, 32] due to its computational efficiency [42]. This last property is essential due to the relatively weak processing power of AMI equipment. In general, the only requirement is that the cryptosystem is additively homomorphic and supports unlimited additions (unless the group size parameter α is fixed at 3, in which case it only needs to support 2 additions – in general, the HE must support $\alpha_{max} - 1$ additions, where α_{max} is some maximum acceptable value for α).

The parameters α and β enable the DSO to dynamically adjust the level of aggregation as well as the precision of the aggregation regions. Zhang and Liu [45] note that variable subset sizes enable improved data analysis by the DSO. The capacity to target different regions and/or degrees of aggregation allows the scheme to support multiple privacy-preserving applications making use of aggregated data.

Table 4.3: Nomenclature

N	number of aggregation groups
G_n	n^{th} aggregation group
α	group size parameter (must be 3 or above)
β	locality parameter
T	total SM population
SM_i	i^{th} smart meter
$SM_j^{G_n}$	smart meter at position j in group G_n
$SM_{leader}^{G_n}$	leader of group G_n ($= SM_0^{G_n}$)
SM_x	currently selected smart meter
P_i	i^{th} selection pool
L_i	leadership status of i^{th} smart meter
M_j	measurement of $SM_j^{G_n}$
A_j	aggregate at position j in group (note: $A_0 = M_0$)
Pk	public key of $SM_{leader}^{G_n}$
Vk	private key of $SM_{leader}^{G_n}$
$E(X)Pk$	cyphertext of X encrypted using Pk
p, q	large prime numbers
m, γ, μ	cryptographic intermediaries

4.4.1 Overview

CHEA works by dynamically splitting the region of SMs controlled by the DSO into distinct sets, referred to as groups. Each of these groups is a set of SMs that will have their data aggregated together. It is noted that different applications might require different densities of aggregation, or more or less tightly localized aggregation. In order to support these different applications, CHEA offers flexibility along these dimensions, expressed using the parameters α and β. α represents the number of SMs that will be in each group; a higher number equates to larger groups and less precise information, and vice versa. β controls the localization of aggregation; a high β value corresponds to a wide search region when generating groups, meaning that meters within groups may be physically distant. Conversely, a low β value corresponds to a tight search region, meaning that meters within groups will be physically close (a necessity for some applications, such as demand response).

Once these parameters have been set, the DSO will generate a "plan" consisting of a set of groups and cycles within those groups. An example of this can be seen in Figure 4.2.

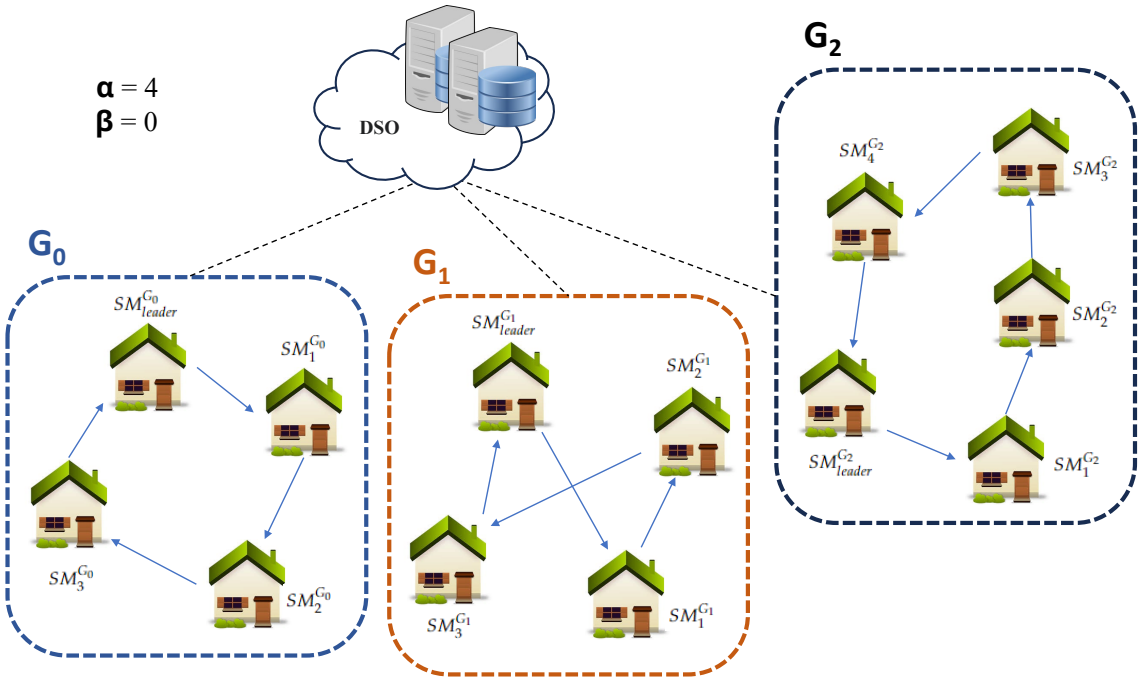

Figure 4.2: Sample CHEA plan construction at DSO is demonstrated with $T = 13, \alpha = 4, \beta = 0$.

Once the plan has been generated, the DSO sends a signal to each meter SM_i to take a measurement along with a packet specifying their position in the group and the ID or address of the meter to which they will be sending data (Figure 4.3).

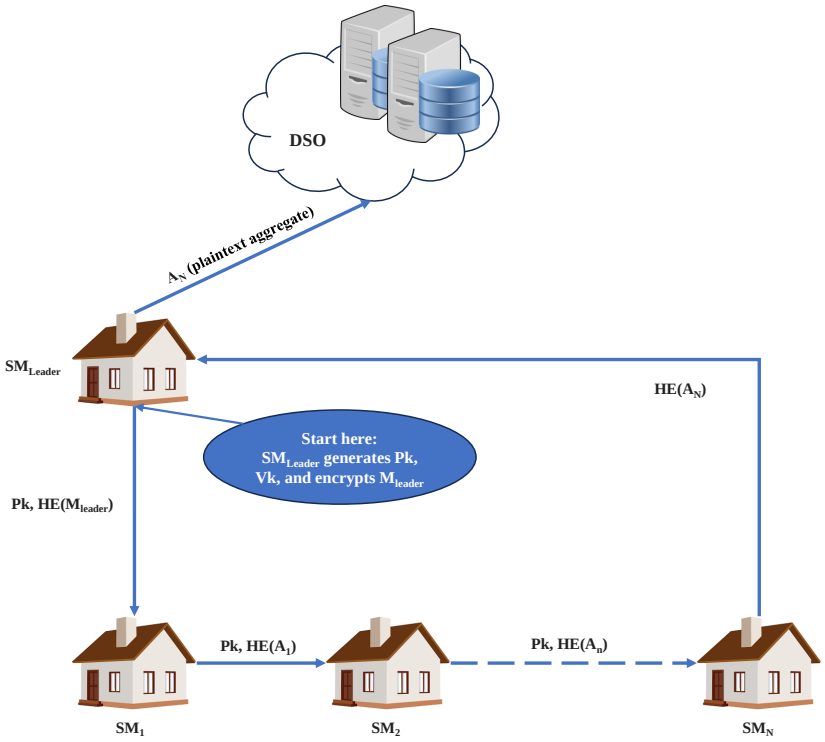

Figure 4.3: CHEA protocol aggregation phase is demonstrated; the dotted line represents an unknown number of intermediate nodes.

4.4.2 Formal Description

The aggregation process begins with the DSO initiating an aggregation request. The DSO will select values for the parameters α and β, which respectively determine the group size (or data resolution) and locality. These parameters are set based on the application that the aggregated data will be used for.

At this stage, the DSO must generate aggregation groups of smart meters from within its population. It begins by sending a ping signal to all of the registered meters to ensure that only active meters are added to the plan. These active meters are split into regions of size β^2 using Algorithm 4.1. These pools are then used to generate groups of size α using Algorithm 4.2. While it cannot be guaranteed that a given pool will have a number of meters divisible by α, the pools are constructed such that each subsequent pool will be geographically close to the previous pool. This means that during Algorithm 4.2 the DSO

can simply iterate over the pools to find the next closest meter if needed. This guarantees minimum deviation from the desired plan while maintaining computational efficiency.

Algorithm 4.1 Beta Pool Creation

Input: β ▷ size of aggregation regions
Input: $S = [SM_0, SM_1, ..., SM_n]$ ▷ list of smart meters
Input: $(Lat_{min}, Lon_{min}), (Lat_{max}, Lon_{max})$ ▷ DSO region boundaries (latitude, longitude)
Output: $[P_0, ..., P_{max}]$ ▷ list of pools
$i_{max} \leftarrow \lfloor Lat_{max}/\beta \rfloor$
$j_{max} \leftarrow \lfloor Lon_{max}/\beta \rfloor$
for $i = 0..i_{max}$ **do**
  for $j = 0..j_{max}$ **do**
    ▷ ensures that IDs are assigned in a snake-like pattern, i.e., subsequent pools will
    ▷ also be physically close
    if $i\%2 == 0$ **then**
      $pool.id \leftarrow i * j_{max} + j$
    else
      $pool.id \leftarrow i * j_{max} + j_{max} - j - 1$
    end if
    Initialize empty list $P_{pool.id}$
    Define $P_{pool.id}$ region as:
      $((Lat_{min} + \beta * i, Lon_{min} + \beta * j), (Lat_{min} + \beta * (i+1), Lon_{min} + \beta * (j+1))$
    Add all SM_x to $P_{pool.id}$ where:
      $(Lat_{min} + \beta * i) \leq SM_x.Lat \leq (Lat_{min} + \beta * (i+1))$ **and**
      $(Lon_{min} + \beta * j) \leq SM_x.Lon \leq (Lon_{min} + \beta * (j+1))$
  end for
end for

Each group that is generated can be stored as a list of tuples, one for each smart meter, where the first half of the tuple is the SM identifier, and the second half of the tuple is the identifier points to the next SM in the group. For example, group G_0 would look like the following circular list: [($SM_0^{G_0}$,$SM_1^{G_0}$), ($SM_1^{G_0}$,$SM_2^{G_0}$), ..., ($SM_{\alpha-1}^{G_0}$,$SM_0^{G_0}$)].

Additionally, a list containing the identifiers of the SM leader for each group is stored: [(G_0,$SM_{leader}^{G_0}$), ..., (G_N,$SM_{leader}^{G_N}$)]

The DSO will now communicate this plan to the meters. Each meter receives a packet containing:

Algorithm 4.2 Dynamic Group Generation

```
Input: α                                                    ▷ group size
Input: T                                                    ▷ total population
Input: P = [P_0, P_1, ..., P_max]                           ▷ list of pools
Output: [G_0, ..., G_N]                                     ▷ list of groups
N ← ⌊(T/α)⌋                                                 ▷ number of groups
j = 0                                                       ▷ pool index
for i = 0..N − 1 do
    Initialize empty list G_i
    while P_j is empty do                                   ▷ find the earliest pool with ungrouped meters
        j = j + 1
    end while
    Select random SM_x from P_j                             ▷ select the group leader
    Add SM_x to G_i as SM_leader                            ▷ SM_0 is the leader
    Remove SM_x from P_j
    Prev = SM_x
    if i < N − 1 then size ← α − 1                          ▷ regular group size = α
    else  size ← T − 1 − i ∗ α                              ▷ last group size ≥ α
    end if
    for j = 1..size do                                      ▷ add the other group members
        while P_j is empty do                               ▷ find the earliest pool with ungrouped meters
            j = j + 1
        end while
        Select random SM_x from P_j
        Add SM_x to G_i as SM_j
        Remove SM_x from P_j
        Prev.next = SM_x
        Prev = SM_x
    end for
    SM_x.next = SM_leader
end for
```

- The aggregation request.
- Its leadership status.
- The identifier/address of the next meter in the group (i.e., the one it will be sending data to).

Algorithm 4.3 is the algorithm used to generate public and private keys in the Paillier homomorphic cryptosystem. This algorithm will be run by the SM_{leader} of each group in order to generate the public key, which will be sent to each meter in the group and used to encrypt each reading, as well as the private key, which will be used to decrypt the resulting aggregate before sending it to the DSO.

Algorithm 4.3 Paillier Key Generation (from Paillier [31])

Output: Pk ▷ public key
Output: Vk ▷ private key
Choose large primes $p, q \ni \gcd(p \times q, (p-1) \times (q-1)) = 1$ ▷ greatest common divisor
$m \leftarrow p \times q$
$\lambda \leftarrow \text{lcm}(p-1, q-1)$ ▷ least common multiple
Choose integer $g \in \mathbb{Z}^*_{m^2}$ ▷ sush that g is relatively prime to m^2
$\mu = (\frac{m}{(g^\lambda mod(m^2))-1})mod(m)$ ▷ modular multiplicative inverse
$Pk \leftarrow (m, g)$
$Vk \leftarrow (\lambda, \mu)$

Once the groups have been generated and the DSO has communicated these plans to the meters, aggregation can begin (see Algorithm 4.4). The full sequence of events, including the plan distribution, can be observed in terms of message exchanges in Figure 4.4.

4.4.3 Initialization

The DSO first selects values for α and β based on the application that data is being aggregated for (e.g., demand response, state estimation, etc.). The parameter α indicates the number of members in each aggregation group, and is thus inversely proportional to aggregation resolution. The parameter β, used in Algorithm 4.1, determines the maximum physical distance between smart meters in the same group.

The DSO begins by generating $N = T/\alpha$ arbitrary groups of size α (except the last group, which can be larger), which is a tunable parameter based on the immediate needs of the DSO specified at the time of request. More specifically, Algorithm 4.2 generates a number of groups equal to the floor of the population divided by the group size. We ensure that groups contain at least three members each (as implied by the last comment in Algorithm 4.2). This is because we do not want to risk having a group of size two or one, which would remove the security benefit for those meters. In the case of one, it is obvious that security cannot be attained via aggregation, and they would be relying purely

Algorithm 4.4 Aggregation

Input Group $G_n[SM_0..SM_{z-1}]$ ▷ group of size $z \geq \alpha$
Output A_n ▷ aggregated value for G_n
for $j = 0..z-1$ **do** ▷ simultaneous
    $M_j \leftarrow SM_j^{G_n}$ measurement ▷ taking measurement
end for
$i \leftarrow 1$
$SM_{leader} \leftarrow SM_0$
SM_{leader} generates public key P_k and private key V_k using Algorithm 4.3
SM_{leader} encrypt(M_0) as $E(A_0)$ ▷ $A_0 = M_0$ in this case
$SM_{leader} \rightarrow (SM_{leader}.next) : [E(A_0), Pk]$ ▷ send encrypted measure & key to next node
$SM_i \leftarrow SM_{leader}.next$
while $SM_i \neq SM_{leader}$ **do**
    SM_i encrypt(M_i) as $E(M_i)$
    SM_i add($E(M_i) + E(A_{i-1})$) as $E(A_i)$
    $SM_i \rightarrow (SM_i.next) : [E(A_i), P_k]$ ▷ send encrypted aggregated value & key to next
    $SM_i \leftarrow SM_i.next$
    $i \leftarrow i + 1$
end while
SM_{leader} decrypt($E(A_{z-1})$) as A_n
$SM_{leader} \rightarrow DSO : A_n$

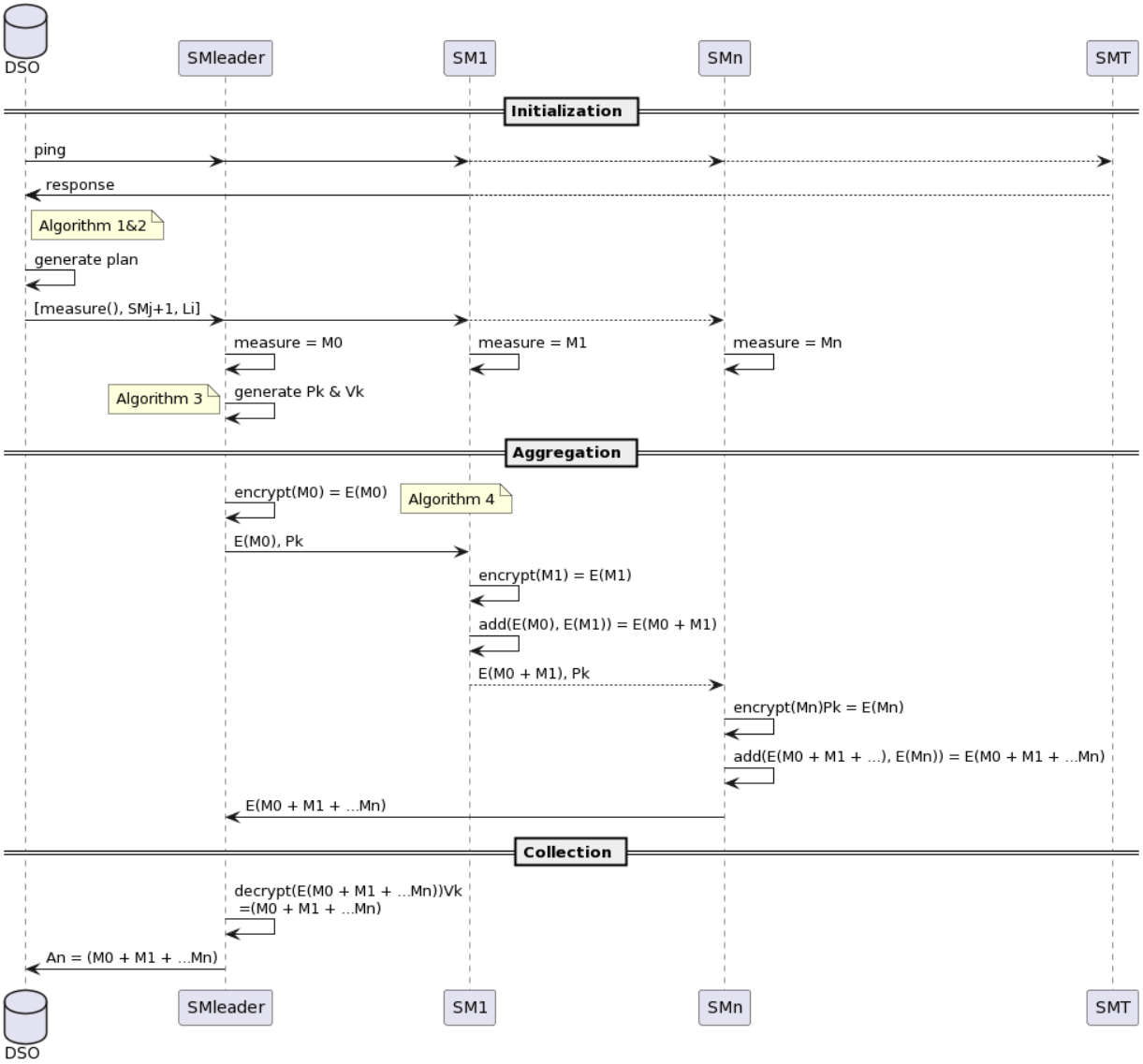

Figure 4.4: Execution cycle of the CHEA protocol with a group of n smart meters from a population of T smart meters illustrated as a sequence diagram. $SM_{leader}..SM_n$ are part of the group under focus but the others (SM_T) are part of other groups.

on security of communication with the DSO, eliminating the benefit CHEA provides. In the case of a group size of two, the leader of the group would be able to determine the other meter's consumption data by subtracting their own data after the aggregation is received

and decrypted.

Instead, we guarantee sufficient group size using the method described; as a result, all but one group are guaranteed to have size alpha, with the final group having $\alpha \leq size \leq 2 * \alpha - 1$.

In each of these groups, we identify one smart meter as the "leader", who begins the aggregation process, generates the required cryptographic keys, and reports the aggregated result to the DSO . Although all the meters will take their measurement simultaneously, the aggregation process necessarily takes place sequentially within each group.

4.4.4 Aggregation

The lead smart meter ($SM_{leader}^{G_n}$) will begin by generating a public and private key for the HE cryptosystem (Algorithm 4.3). It will then use this public key to encrypt its own measurement (Algorithm 4.4). The leader then sends its encrypted measurement along with the public key it generated to the next smart meter in the group (each meter will be sent the address of the next meter in the chain, the order of which is generated by the DSO at the same time the group is generated).

The next meter in the chain then encrypts its own measurement using the public key sent by the leader. By the property of additive homomorphism, it can add its encrypted measurement to the one sent by the leader, resulting in a cyphertext that contains the total usage by both the leader and the current meter.

As α is minimally 3 (to avoid privacy issues in small groups), the second meter sends the public key and current aggregate to a third meter, who then encrypts their own data using the private key and adds it to the current aggregate, creating a new encrypted aggregate that contains the sum of all three meters' usage. This goes on until the last meter of the group sends the cyphertext to the leader.

4.4.5 Decryption

The lead smart meter will have received the cyphertext containing the aggregate of the group's usage data from the last meter in the group. At this point, the leader will decrypt the usage data using the private key it generated at the beginning of the process. There is then no way for the leader to discover the usage data of any of the previous meters. This is because even if a leader maliciously decrypts the aggregate and subtracts its own usage

data, the result will simply be the aggregate of the other $\alpha - 1$ meters' data, from which no precise information can be gained.

This aggregate data can now be sent to the DSO (end of Algorithm 4.4). Because the data is pre-aggregated, it can be sent in plaintext. However, the cost of encrypting with a public key from the DSO is insignificant and likely worth the small security improvement.

Note also that the groups are created dynamically and randomly at each request from the DSO to prevent statistical attacks that could be performed over multiple iterations.

4.4.6 Faults

Smart meters are sent a ping immediately prior to the generation of aggregation groups. This ensures that they are online and active before they are added to a group for a particular aggregation cycle (input S of Algorithm 4.2).

If groups are large enough (α), this may not be a sufficient guarantee. In these cases, some communication efficiency can be sacrificed by including the full plan in each intermediate step, i.e., the full list of addresses in the group. Each meter in a given group would wait for a response from the subsequent meter before discarding the aggregate. If no response is received, they would defer transmission to the next address in the list.

4.5 Security Analysis

This section analyses how CHEA protects privacy under common attacks. This analysis goes beyond the points related to security, privacy, and fault tolerance made in the previous section.

4.5.1 Man-in-the-Middle Attack

A man-in-the-middle attack is a privacy attack where two parties are communicating and a third party, the attacker, intercepts this communication. The attacker may then use the intercepted information to facilitate burglary or blackmail, or may modify the information before it reaches its intended recipient.

A smart grid with TE running CHEA has several distinct interception points:

(a) Communication from the DSO to the SMs;

(b) Communication from the SM_{leader} to the next SM;

(c) Communication from a non-leader SM to the next SM;

(d) Communication from the SM_{leader} to the DSO.

In (a), the greatest risk would be altering the plan. The attacker may be able to change the address for the meter $SM_j^{G_n}$ to send information to. This would not enable false data injection (FDI), since they would need to intercept multiple transmissions in order to change the plan coherently, however, it could disrupt the cycle path and cause an aggregation failure for the group in question.

Communication from SM_{leader} to the first intermediate meter in the group (b) poses a slight risk because it is the only transmission that contains non-aggregated data. However, it is encrypted using SM_{leader}'s public key, so they should be the only one capable of decrypting it. If they were to collude, they would only be putting their own data at risk.

Communication between two intermediate meters (c) poses little risk, since the data would be both encrypted and aggregated. Even in the unlikely scenario that the attacker can decrypt the message, it will not expose any private information.

Considering the final transmission (d), there is no privacy risk to the users in the aggregation region since the data has already been aggregated before this step occurs. The greatest risk would be FDI, which is the main reason it should be protected using the public key of the DSO before sending.

4.5.2 Collusion Resistance

CHEA is extremely collusion resistant. Because group membership is allocated on the fly, it is impractical to form an adversarial coalition in advance. Additionally, an attacker would need the aggregate total that is sent *by* the victim (which may not be sent to them, depending on the plan), as well as the aggregate total that was sent *to* the victim. In the best case scenario, where the group size is 3, this would require cooperation with one other user. Where $\alpha > 3$, it would require cooperation with two other users, and specifically the two users interacting with the victim (which is determined by the flow of the plan). Even if this contrivance was successfully achieved during one iteration, the plan would be extremely unlikely to succeed in the next iteration as group assignments and data flow would be reassigned.

Requirements (SM_v is the victim):

- SM_{leader} colludes
- SM_{v-1} colludes
- SM_{v+1} colludes
- Attacker colludes or has (by chance) one of these three roles.

The probability of the DSO generating a plan that enables attackers to target a particular meter during a particular aggregation cycle is:

$$\frac{(\frac{\alpha}{N})^3 * (\prod_{i=1}^{3} \frac{i}{\alpha})}{(3!)^2}$$

where N is the total smart meter population (or selection pool) and α is the group size during that cycle. The value 3 comes from the three parties required to collude with the attacker.

This expression can be simplified algebraically to:

$$\frac{1}{6N^3}$$

This function shrinks extremely quickly, demonstrating that the probability of successful collusion is negligible. With only 100 smart meters in a selection pool, the odds of success are already exceeding low (1 in 6 million). In this case, even if aggregation was performed every second (an improbably high rate), a plan enabling a collusion attack would only be generated once every 70 days; this frequency is not useful for any class of cyberattack.

4.5.3 Quantum Attacks

Although our implementation and simulation employed the Paillier encryption scheme, the core CHEA protocol is cryptosystem agnostic. Therefore, it would be trivial to implement the protocol with a quantum resistant cryptosystem, such as lattice-based cryptography [4, 5].

4.6 Performance Analysis

4.6.1 Overview

We simulated the operations carried out by the smart meters and the DSO respectively using two custom C programs that communicate using sockets, as they would in a real networking application. Simulations were performed on a 2017 MacBook Pro with a Intel Core i5 CPU at 3.1GHz and 8 GB RAM. The code is available online in a replication package [37].

Each smart meter process simulates the operations carried out by a smart meter, as well as containing a unique identity that comprises physical variables associated with the smart meter being simulated to improve the accuracy of the simulation. The DSO application acts as a server for the smart meters and generates and distributes the plan as described in Section 4.4.

In addition to employing a realistic network implementation, we further enhanced the simulation's accuracy by employing current IEEE smart meter communication protocols as outlined by Zaraket et al. [43] and Shanmukesh et al. [33].

4.6.2 Results

We compared the performance of our scheme against schemes by Liu et al. [26], Song et al. [36], and Zuo et al. [47]. Table 4.4 summarizes the results in terms of net communication overhead for each scheme. Table 4.5 summarizes the results of processing time impacts for each phase of aggregation for each scheme. Processing time for competing schemes was produced by simulating several tests of the most impactful operations on our hardware and extrapolating the assessment method employed by Liu et al. [26].

Processing time for competing schemes was produced by simulating several tests of the most impactful operations on our hardware and extrapolating the assessment method employed by Liu et al. [26]. The Stanford Pairing-Based Cryptography library was employed to perform the tests, which were run on the aforementioned 2017 MacBook Pro. The testing parameters were $population = 30$ and $dimensionality = 10$. CHEA was extended theoretically to support multidimensionality to ensure a fair comparison with the other schemes.

We found that our scheme has communication overhead comparable to the other schemes. Specifically, our initialization phase and end phase are as good as or better than the schemes

Table 4.4: Comparison of communication overhead

Scheme	Initialization phase	Aggregation phase	To DSO
CHEA	1024 bits	2072 bits	1024 bits
Liu et al. [26]	3456 bits	2368 bits	2368 bits
Song et al. [36]	4192 bits	1056 bits	1024 bits
Zuo et al. [47]	1088 bits	1600 bits	1600 bits

Table 4.5: Comparison of performance (times averaged over 10 runs)

Scheme	Initialization phase	Aggregation phase	Decryption phase
CHEA	125 ms	497 ms	29 ms
Liu et al. [26]	79 ms	340 ms	27 ms
Song et al. [36]	122 ms	1036 ms	34 ms
Zuo et al. [47]	119 ms	731 ms	1337 ms

compared. In the aggregation phase, our scheme has the second highest communication overhead. This result was anticipated, as sending the public key (P_k) along with the intermediate aggregate ($E(A_i)$) necessarily incurs additional communication cost.

With regards to processing time, we found that our scheme is comparable to the others in the decryption phase, except for the much slower approach from Zuo et al. [47]. This result is not surprising since there is no additional processing required. CHEA's initialization phase is the slowest of the group (but not by much); this can be attributed to the extra processing required to generate the plan each round. In the aggregation phase, CHEA is the second fastest. It is slower than Liu et al. [26]'s scheme because encryptions must be performed sequentially (other schemes have a parallel reporting phase). However, the structure of the scheme requires fewer cryptographic operations overall, resulting in performance gains compared to the rest of the schemes. Additionally, scaling analyses indicate that our scheme could see comparably superior performance as neighbourhood sizes increase, as we note and explain in the next section.

Overall, our scheme incurs a slight communication disadvantage compared to *some* similar schemes, as well as a slower initialization time. However, aggregation time is competitive and scaling (discussed in Section 4.6.3) is improved. Additionally, CHEA improves flexibility, fault tolerance, and security. For example, although Liu et al. [26] remains the fastest scheme, it lacks the flexibility of CHEA's dynamic group membership, which offers versatile locality and aggregation resolution.

4.6.3 Scaling

Registration

Since there is no centralized key distribution, the registration/initialization phase is essentially of complexity $\mathcal{O}(1)$, incurring no extra time cost as the population of smart meters grows. As we discuss in Section 4.1, local in-group key generation, the distributed nature of the protocol, and avoidance of complex key distribution make our scheme significantly more scalable than comparable protocols [36, 2].

Aggregation

As is the case with registration, the aggregation happens locally and concurrently within each group, thus there is no time scaling associated with having larger neighbourhoods (i.e., more groups). Larger groups (corresponding to a larger α value), on the other hand, *will* incur an increased time penalty during this phase, since the aggregation is generated sequentially withing a group.

Scaling tests were performed using the custom simulation, the results of which are presented in Figure 4.5. The simulation used libhcs instead of PBC for HE, hence the faster times relative to the earlier test. For these tests, β was set to zero (which is interpreted as nearest neighbour instead of pooling) and the neighbourhood size was $10km^2$. Population and α were varied as seen in the graph.

It can be seen that the results support the theoretical scaling properties, that is, increasing the population has a limited effect on performance while increasing α has a small but noticeable effect. Unfortunately, simulating all devices (and, in particular, DSO operations) on the same computer somewhat dampens the visibility of these effects, and future tests should include a distributed hardware simulation to better isolate them.

4.7 Limitations

Due to the lack of available hardware for testing, the protocol was tested using a software simulation of a smart grid environment. Ideally this would have been tested in a more accurate setup using real, physical smart meters to demonstrate feasibility and gather data for prospective users. However, given that most current-generation AMIs are not capable of performing the necessary operations, it is difficult to test HE-based aggregation

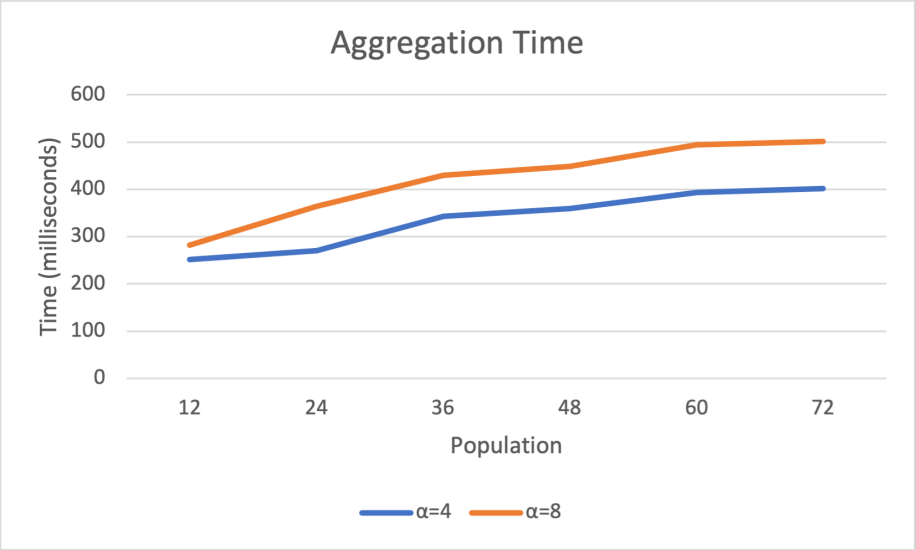

Figure 4.5: CHEA aggregation times for different neighbourhood populations. The sub-linear increase can largely be explained by the computer being under greater load as it simulates more smart meters.

schemes in practice at this time. As a result, related literature largely relies on software simulations, as we did.

Another limitation lies in the assumptions made about capabilities of smart meters. Although we did not make any unprecedented assumptions – that is, assumptions that have not already been made in similar literature – we do presume smart meters to be capable of certain tasks that are not currently among their commonly-found functionalities. This may call into question the feasibility of this architecture as a real-world solution. In discussions with industry experts, it was noted that manufacturers prefer to avoid unnecessary upgrades to grid hardware, which may limit short-term adoption. However, in spite of the current state of smart meter hardware, it seems reasonable to assume that they will follow the trend of computational ability in virtually every other domain, and that they will receive increased computing power at decreased costs as time goes on. As this happens, it will become cost effective for utility companies to provide increased processing capabilities to metering equipment in order to support more comprehensive security measures, as this will increase consumer trust, provide marketing opportunities, and reduce expensive breaches, all at limited cost. For these reasons, we believe security research regarding the smart grid should not be limited to current hardware capabilities, and continue the example set by other researchers with CHEA.

4.8 Conclusion and Future Work

This paper introduced Cyclic Homomorphic Encryption Aggregation (CHEA) as a new scheme for protecting privacy. We found CHEA to be effective at providing privacy-protecting aggregation of energy usage data for distributed smart grid energy trading settings, including TEMs. Formal analysis and software simulation confirm that the protocol provides significant security benefits without sacrificing performance.

Although the scheme performs similarly to other smart grid aggregation schemes, current smart meters are likely incapable of supporting the requisite cryptographic operations, so real-world deployment will depend on hardware improvements in AMI.

The CHEA protocol presents a novel, distributed HE-based aggregation solution for TE that could potentially be generalized to other environments with similar infrastructures, e.g., environments consisting of networked devices that generate data to be consolidated, and operate on distributed applications. One potential candidate may be smart electric vehicles, which generate driving quality and accident data to inform insurance providers for different demographics. Other metered utilities, such as water and gas, may also potentially benefit from our solution, although extending the scheme to these areas may present novel domain-specific challenges.

Some other potential areas for future research include:

- Improving upon CHEA by making it even more robust to communication or meter failures during the aggregation phase. Future iterations could include plans that are dynamically adjusted based on where communication drops off, but this will require making the meters even more autonomous (thus increasing their computational load).
- Investigating different methods of group generation. It may be possible to forgo the requirement of creating a plan centrally by the DSO if smart meters create the plan dynamically in a more procedural manner, perhaps using cellular automata, for example. While there is no guarantee that this would be more efficient, it could be an interesting research topic and would at least confer the benefit of increasing distribution, reducing reliance on centralized computing even further.
- Adding finer control to the locality parameter to enable the DSO to be more specific about how regions are divided. For example, it may want to consider network topology, neighbourhoods, or other currently unsupported factors when requesting an aggregate reading.

- Looking into other applications of HE to grid management operations. This technology may enable distributed applications for functionality such as state estimation, transaction verification, or power flow optimization. Distributing these tasks could present further privacy, security, and reliability benefits to TEM users and operators, similar to those seen by CHEA.

Bibliography

[1] Muhammad Rizwan Asghar, György Dán, Daniele Miorandi, and Imrich Chlamtac. Smart meter data privacy: A survey. *IEEE Communications Surveys & Tutorials*, 19 (4):2820–2835, 2017. doi:10.1109/COMST.2017.2720195.

[2] Mohamad Badra and Sherali Zeadally. Lightweight and efficient privacy-preserving data aggregation approach for the smart grid. *Ad Hoc Networks*, 64:32–40, 2017. ISSN 1570-8705. doi:10.1016/j.adhoc.2017.05.011.

[3] Alex R. Borden, Daniel K. Molzahn, Parmeswaran Ramanathan, and Bernard C. Lesieutre. Confidentiality-preserving optimal power flow for cloud computing. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1300–1307, 2012. doi:10.1109/Allerton.2012.6483368.

[4] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) lwe. Cryptology ePrint Archive, Paper 2011/344, 2011. URL https://eprint.iacr.org/2011/344. https://eprint.iacr.org/2011/344.

[5] Zvika Brakerski and Vinod Vaikuntanathan. Lattice-based fhe as secure as pke. Cryptology ePrint Archive, Paper 2013/541, 2013. URL https://eprint.iacr.org/2013/541. https://eprint.iacr.org/2013/541.

[6] Suryadip Chakraborty. *Data Aggregation in Healthcare Applications and BIGDATA set in a FOG based Cloud System*. PhD thesis, University of Cincinnati, 2016. URL http://rave.ohiolink.edu/etdc/view?acc_num=ucin1471346052.

[7] Rohit Chandra, Krishnanand Kaippilly Radhakrishnan, and Sanjib Kumar Panda. Privacy protected product differentiation through smart contracts based on bilateral negotiations in peer-to-peer transactive energy markets. *Sustainable Energy, Grids and Networks*, 34:100997, 2023. ISSN 2352-4677. doi:10.1016/j.segan.2023.100997.

[8] Siguang Chen, Li Yang, Chuanxin Zhao, Vijayakumar Varadarajan, and Kun Wang. Double-blockchain assisted secure and anonymous data aggregation for fog-enabled smart grid. *Engineering*, 8:159–169, 2022. ISSN 2095-8099. doi:10.1016/j.eng.2020.06.018.

[9] Siguang Chen, Li Yang, Yanhang Shi, and Qian Wang. Blockchain-enabled secure and privacy-preserving data aggregation for fog-based its. *Computers, Materials & Continua*, 75(2):3781–3796, 2023. ISSN 1546-2226. doi:10.32604/cmc.2023.036437.

[10] Tao Chen, Qais Alsafasfeh, Hajir Pourbabak, and Wencong Su. The next-generation U.S. retail electricity market with customers and prosumers–a bibliographical survey. *Energies*, 11(1), 2018. doi:10.3390/en11010008.

[11] Yuwen Chen, Shisong Yang, José-Fernán Martínez-Ortega, Lourdes López, and Zhen Yang. A resilient group-based multisubset data aggregation scheme for smart grid. *IEEE Internet of Things Journal*, 10(15):13649–13661, 2023. doi:10.1109/JIOT.2023.3262731.

[12] Yuan Cheng, Yanan Liu, Zheng Zhang, and Yanxiu Li. An asymmetric encryption-based key distribution method for wireless sensor networks. *Sensors*, 23(14), 2023. ISSN 1424-8220. doi:10.3390/s23146460.

[13] Taher Elgamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985. doi:10.1109/TIT.1985.1057074.

[14] Yuly V. Garcia, Oscar Garzon, Carlos J. Delgado, Jan L. Diaz, Cesar A. Vega Penagos, Fabio Andrade, Adriana C. Luna, and J. C. Hernandez. Overview on transactive energy–advantages and challenges for weak power grids. *Energies*, 16(12), 2023. doi:10.3390/en16124607.

[15] Felix Gomez Marmol, Christoph Sorge, Osman Ugus, and Gregorio Martinez Perez. Do not snoop my habits: Preserving privacy in the smart grid. *IEEE Communications Magazine - IEEE Commun. Mag.*, 50:166–172, 05 2012. doi:10.1109/MCOM.2012.6194398.

[16] Ateeb Hassan, Hadi Nabipour Afrouzi, Chua Hong Siang, Jubaer Ahmed, Kamyar Mehranzamir, and Chin-Leong Wooi. A survey and bibliometric analysis of different communication technologies available for smart meters. *Cleaner Engineering and Technology*, 7:100424, 2022. ISSN 2666-7908. doi:10.1016/j.clet.2022.100424.

[17] Georgios Kalogridis, Rafael Cepeda, Stojan Z. Denic, Tim Lewis, and Costas Efthymiou. Elecprivacy: Evaluating the privacy protection of electricity management algorithms. *IEEE Transactions on Smart Grid*, 2(4):750–758, 2011. doi:10.1109/TSG.2011.2160975.

[18] Hayat Mohammad Khan, Abid Khan, Farhana Jabeen, Adeel Anjum, and Gwanggil Jeon. Fog-enabled secure multiparty computation based aggregation scheme in smart grid. *Computers & Electrical Engineering*, 94:107358, 2021. ISSN 0045-7906. doi:10.1016/j.compeleceng.2021.107358.

[19] Hayat Mohammad Khan, Abid Khan, Farhana Jabeen, and Arif Ur Rahman. Privacy preserving data aggregation with fault tolerance in fog-enabled smart grids. *Sustainable Cities and Society*, 64:102522, 2021. ISSN 2210-6707. doi:10.1016/j.scs.2020.102522.

[20] Neal Koblitz. *A course in number theory and cryptography*, volume 114. Springer Science & Business Media, 1994.

[21] Ares Lagae, Sylvain Lefebvre, Rob Cook, Tony DeRose, George Drettakis, David S Ebert, John P Lewis, Ken Perlin, and Matthias Zwicker. A survey of procedural noise functions. *Computer Graphics Forum*, 29(8):2579–2600, 2010. doi:10.1111/j.1467-8659.2010.01827.x.

[22] Aron Laszka, Abhishek Dubey, Michael Walker, and Douglas Schmidt. Providing privacy, safety, and security in iot-based transactive energy systems using distributed ledgers. In *Proceedings of the Seventh International Conference on the Internet of Things*, 10 2017. doi:10.1145/3131542.3131562.

[23] Michelle Lauer, Rupamathi Jaddivada, and Marija Ilić. Secure blockchain-enabled dymonds design. In *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, COINS '19, page 191–198, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366403. doi:10.1145/3312614.3312654.

[24] Fengjun Li, Bo Luo, and Peng Liu. Secure and privacy-preserving information aggregation for smart grids. *IJSN*, 6:28–39, 04 2011. doi:10.1504/IJSN.2011.039631.

[25] Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker. Inferring personal information from demand-response systems. *IEEE Security & Privacy*, 8(1):11–20, 2010. doi:10.1109/MSP.2010.40.

[26] Shuanggen Liu, Yaowei Liu, Wandi Liu, and Yuchen Zhang. A certificateless multi-dimensional data aggregation scheme for smart grid. *Journal of Systems Architecture*, 140:102890, 2023. ISSN 1383-7621. doi:10.1016/j.sysarc.2023.102890.

[27] Federico Lombardi, Leonardo Aniello, Stefano De Angelis, Andrea Margheri, and V. Sassone. A blockchain-based infrastructure for reliable and cost-effective iot-aided smart grids. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, 01 2018. doi:10.1049/cp.2018.0042.

[28] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[29] Yang Ming, Yabin Li, Yi Zhao, Pengfei Yang, and Yu Yao. Efficient privacy-preserving data aggregation scheme with fault tolerance in smart grid. *Sec. and Commun. Netw.*, 2022, jan 2022. ISSN 1939-0114. doi:10.1155/2022/5895176.

[30] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23, 09 2017. doi:10.1109/RWEEK.2017.8088642.

[31] Pascal Paillier. Public-key cryptosystems based on composite degree residuosity classes. In Jacques Stern, editor, *Advances in Cryptology — EUROCRYPT '99*, pages 223–238, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg. ISBN 978-3-540-48910-8.

[32] Pascal Paillier and David Pointcheval. Efficient public-key cryptosystems provably secure against active adversaries. In Kwok-Yan Lam, Eiji Okamoto, and Chaoping Xing, editors, *Advances in Cryptology - ASIACRYPT'99*, pages 165–179, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg. ISBN 978-3-540-48000-6.

[33] Pudi Shanmukesh, Lagineni Mahendra, Katta JaganMohan, R.K.Senthil Kumar, and B.S. Bindhumadhava. Secure dlms/cosem communication for next generation advanced metering infrastructure. *Asian Journal For Convergence In Technology (AJCT) ISSN -2350-1146*, 7(1):92–98, Apr. 2021. doi:10.33130/AJCT.2021v07i01.020.

[34] Ashutosh Kumar Singh and Jatinder Kumar. A privacy-preserving multidimensional data aggregation scheme with secure query processing for smart grid. *The Journal of Supercomputing*, 79(4):3750–3770, Mar 2023. ISSN 1573-0484. doi:10.1007/s11227-022-04794-9.

[35] Parminder Singh, Mehedi Masud, M. Shamim Hossain, and Avinash Kaur. Blockchain and homomorphic encryption-based privacy-preserving data aggregation model in smart grid. *Computers & Electrical Engineering*, 93:107209, 2021. ISSN 0045-7906. doi:10.1016/j.compeleceng.2021.107209.

[36] Jingcheng Song, Yining Liu, Jun Shao, and Chunming Tang. A dynamic membership data aggregation (dmda) protocol for smart grid. *IEEE Systems Journal*, 14(1):900–908, March 2020. ISSN 1937-9234. doi:10.1109/JSYST.2019.2912415.

[37] Daniel Sousa-Dias. Cyclic homomorphic encryption aggregation – simulation replication package, 2023. https://github.com/Smart-Contract-Modelling-uOttawa/CHEA-Simulation.

[38] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, and John Mylopoulos. A review of cybersecurity concerns for transactive energy markets. *Energies*, 16(13), 2023. ISSN 1996-1073. doi:10.3390/en16134838.

[39] Noelia Uribe-Pérez, Luis Hernández, David De la Vega, and Itziar Angulo. State of the art and trends review of smart metering in electricity grids. *Applied Sciences*, 6 (3), 2016. doi:10.3390/app6030068.

[40] Ognjen Vuković, György Dán, and Rakesh B. Bobba. Confidentiality-preserving obfuscation for cloud-based power system contingency analysis. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 432–437, 2013. doi:10.1109/SmartGridComm.2013.6687996.

[41] Junhua Wu, Zhuqing Xu, Guangshun Li, Cang Fan, Zhenyu Jin, Yuanwang Zheng, and Jinbo Xiong. E-lpdae: An edge-assisted lightweight power data aggregation and encryption scheme. *Sec. and Commun. Netw.*, 2022, jan 2022. ISSN 1939-0114. doi:10.1155/2022/6218094.

[42] Shiying Yao, Jian Zeng, Shuangxing Wang, Xiaolong Yang, Jingtang Luo, and Ziqi Wang. A secure data aggregation scheme enabling abnormal smart meters traceback for smart grid. In *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*, EITCE '22, page 911–916, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450397148. doi:10.1145/3573428.3573780.

[43] Carine Zaraket, Ioannis Dogas, Dimitrios Kalyvas, Panagiotis Papageorgas, Michel Aillerie, and Kyriakos Agavanakis. Open source LoRaWAN telemetry test bench

for smart grid - A DLMS/COSEM implementation case study. In *Technologies and Materials for Renewable Energy, Environment and Sustainability: TMREES21Gr*, page 020196, Athens, Greece, Sep 2021. doi:10.1063/5.0095471.

[44] Jianhong Zhang and Chenghe Dong. Privacy-preserving data aggregation scheme against deletion and tampering attacks from aggregators. *Journal of King Saud University - Computer and Information Sciences*, 35(4):100–111, 2023. ISSN 1319-1578. doi:10.1016/j.jksuci.2023.03.002.

[45] Liying Zhang and Yining Liu. Fsda: Flexible subset data aggregation for smart grid. *IEEE Systems Journal*, 17(1):569–578, 2023. doi:10.1109/JSYST.2022.3199386.

[46] Xiaojun Zhang, Wei Tang, Dawu Gu, Yuan Zhang, Jingting Xue, and Xin Wang. Lightweight multidimensional encrypted data aggregation scheme with fault tolerance for fog-assisted smart grids. *IEEE Systems Journal*, 16(4):6647–6657, Dec 2022. ISSN 1937-9234. doi:10.1109/JSYST.2022.3146504.

[47] Xiangjian Zuo, Lixiang Li, Haipeng Peng, Shoushan Luo, and Yixian Yang. Privacy-preserving multidimensional data aggregation scheme without trusted authority in smart grid. *IEEE Systems Journal*, 15(1):395–406, March 2021. ISSN 1937-9234. doi:10.1109/JSYST.2020.2994363.

Chapter 5

Market Privacy (ILPTS)

The transactive energy market (TEM) is a recent development in energy economics used for trading energy and ancillary services among power distribution companies (e.g., Distribution System Operator – DSO) and prosumers of energy (e.g., virtual power plants and microgrids). TEMs use advanced metering infrastructures and smart meters data for market monitoring and periodical energy bill settlements. In such markets, energy prosumers can also trade amongst themselves directly, as long as the DSO knows about those transactions and checks their feasibility in terms of grid safety, reliability, and resiliency. This raises a number of privacy concerns, as data must be shared among prosumers to facilitate operations. To address these concerns, many solutions anonymize market participation to prevent leakage of trading history or other personal information. However, this leads to issues with potential abuse or malfunction. This paper proposes the *Individually Linkable Pseudonymous Trading Scheme* (ILPTS), a new solution based on smart contracts that enables an automated system to detect fraudulent or malfunctioning behaviour and temporarily reveal the identity of the offending trader to the DSO for intervention. This solution offers benefits over unethical alternatives such as disconnecting power.

5.1 Introduction

Advanced metering infrastructure (AMI) has led to progress in functionality in the power delivery space [17]. Smart meters enable the collection of much more frequent and detailed data from virtual power plants (VPPs), smart homes/buildings, and smart grid facilitates for improved power delivery services and better analytics for distribution system planning and optimal operations [26]. Developments like transactive energy (TE), improved

demand response and power flow optimization, consumption prediction profiling, and machine learning-supported surge prediction, are just a few examples of technological progress and innovations in power systems [10]. These innovations will enable greater economic efficiency through models like transactive energy, enabling high frequency peer-to-peer energy trading. These models will also contribute to greater energy efficiency, due to improved generation and load balancing, intelligent demand response programs, and AI-assisted predictions [7].

Blockchain technology has emerged as a *de facto* foundation for transactive energy implementations due to its strong cryptographic security and resistance to false data injection [24, 13, 12]. However, publishing electricity transactions on a public ledger poses significant privacy risks to the energy prosumers, consumers, and distribution system operators (DSO) [1, 18].

Current solutions make a tradeoff between privacy and regulation, either choosing to enforce greater privacy measures at the expense of being able to regulate market activity [8], or reducing prosumers' privacy to facilitate improved regulatory ability.

This paper contributes a new privacy scheme named *Individually Linkable Pseudonymous Trading* (ILPTS), which aims to alleviate the need for this tradeoff. In the proposed solution, prosumers are able to trade energy under fully anonymous pseudonyms, the true identities of whom are not known to any party including the DSO itself. Meanwhile, a recovery database is stored on a blockchain that enables re-association of a user's true identity with their pseudonym if they consistently fail transactions. This process requires multiple parties to cooperate and can only be initiated by a self-governing smart contract, thus no prosumer can be manually exposed.

The contributions of this paper are as follows:

- **New functionality**: previous systems had to decide between total anonymity and capacity to respond to errors. By enabling total but revocable privacy, ILPTS enables functionality not seen in other transactive energy market proposals. This would enable a DSO to be more responsive to network errors and to provide novel services such as in-home maintenance to users.

- **Enhanced privacy**: ILPTS would enhance the privacy of the basic blockchain-based TEM model by providing anonymity to energy clients. When implemented in a model that already provides this anonymity, ILPTS improves its functionality while retaining the existing anonymity. When implemented in a model with partial trust – i.e., anonymity in the public market but identities registered with the DSO –

ILPTS improves privacy by hiding the user's identity more comprehensively, creating an effectively trustless system.

- **Equity**: ILPTS improves the fairness of TEM models by enabling repudiation and mediation, obviating the need for coarse punitive measures such as disconnecting power, and reducing false positives in fraud and malware detection.
- **Improved security**: while enabling the previously mentioned benefits, ILPTs's novel registration process is secure and fully distributed. Information is concealed and encrypted at every stage. In addition, the deanonymization process can only be initiated by the monitoring smart contract, greatly reducing the potential for abuse.

ILPTS will be useful to purveyors of smart grid technology as a proof-of-concept of the privacy and security stature of this new paradigm. It will also be useful to TEM designers to improve the equity, privacy, and legality of their distributed energy market systems. Finally, ILPTS provides a novel technique to achieve goals that are seemingly at odds in a secure, distributed, and validated manner, which will be of interest to researchers in the fields of TE, cryptography, and distributed computing.

The paper is structured as follows. Section 5.2 examines the state of privacy solutions in the field and works that motivated the development of this scheme. In Section 5.3, we describe the technologies involved, the architecture of TEMs, and the fundamental problem we seek to address. We describe the design goals of our solution in Section 5.4. Section 5.5 introduces our ILPTS solution, providing both a high-level overview as well as a formal description. In Section 5.6 we evaluate the scheme's adherence to the design goals, as well as the security properties of the scheme and potential attacks that could be carried out. Section 5.7 concludes the study by evaluating the success of the scheme in achieving its design goals and discussing future research directions.

5.2 Related Work

Efforts to improve privacy in TEMs have focused on a few key areas, including: privacy-preserving data aggregation, market privacy and anonymization, and privacy-preserving billing mechanisms. These advancements are critical for prosumer's safety and trust; adoption of smart grids (SG) and TEM technology – which stand to present environmental and economic benefits – will be limited by prosumers' trust in their privacy being maintained by such technology.

Much privacy research in the TE field has originated in electric vehicle-specific research [23, 2, 3, 16, 9, 15]. This is due to the unique position of electric vehicles (EVs) in transactive energy and smart grid architecture. Their ability to move freely and the identifying information they contain present novel challenges to TE and SG designers [9]. EVs can reveal different kinds of information to attackers compared to other advanced metering infrastructure (AMI) which can be more damaging in some cases, such as facilitating stalking [16].

Li et al. [15] propose a privacy-preserving double auction scheme that employs homomorphic encryption (HE), a cryptographic technique our scheme also uses. Another double auction scheme is proposed by Li et al. [16], who note that their scheme achieves positive results with social welfare – a key component of our scheme – and ensures locational privacy. Baza et al. [3] propose another EV-centric scheme that focuses on providing trading that cannot be linked to the location of the user, while still handling the risks associated with anonymity. The authors extend this proposal for privacy-preserving trading in charging station-to-vehicle trading to vehicle-to-vehicle trading in [2]. Sharma et al. [23] acknowledge the importance of decentralization to avoid single point of failure and trust concerns, and propose a decentralized scheme to facilitate EV energy trading with minimal mediation.

Bergquist et al. [4] propose a solution meant to extend the existing PETra protocol proposed by Laszka et al. [14] with additional communication security measures and transactional anonymity. It is similar to our solution in the sense that it aims to enhance security within an existing architecture.

This scheme is meant to enhance schemes like TRANSAX, a TEM architecture proposed by Eisele et al. [8]. TRANSAX introduced the idea of employing coin mixing to support market anonymity in a TEM. In their solution, users are anonymous within groups, which are used to ensure maintenance of safety restrictions with regards to energy transmission. Although TRANSAX provides the greatest anonymity among proposed TEM architectures (to the best of our knowledge), it suffers from the fact that the anonymity is irreversible, a guiding motivator for our solution.

5.3 Background and Motivation

This section provides important background information that motivates the need for developing a new solution such as ILPTS.

5.3.1 Network Architecture

The network architecture is presumed to consist of:

- Distribution system operator (DSO): a centralized authority that authenticates users and manages the distribution (or retail) energy market.
- Smart meters: users connect to the grid and distributed market via smart meters.
- Fog nodes: auxiliary network computing nodes with elevated status and trust.
- Blockchain: a distributed ledger on which transactions and other information will be recorded.
- Smart contracts: scripts that execute on the blockchain and that may facilitate additional functionality or public information transfer.

5.3.2 Blockchain Markets

Blockchain, or distributed ledger technology, is a distributed database of transactions whose state is maintained by members of the *blockchain network*. Users on the network contribute computing resources to verify the state of the blockchain and facilitate transactions between members by way of *cryptocurrency*, the primary unit of value on the chain.

Blockchain technology offers several unique security benefits, including non-repudiation, an immutable record of transactions, and extreme cryptographic resistance against false data injection [19].

Modern implementations, such as Ethereum, have introduced the notion of scripting, employing the ledger as a state machine to facilitate distributed virtual computation [5]. The scripts that run on these blockchains are known as *smart contracts*.

Blockchain-based transactive energy markets are automated energy auctions where the market clearing functionality is handled by smart contracts, the blockchain ledger is used to maintain the record of energy exchange, and often an intermediary cryptocurrency token is used to manage energy transactions.

5.3.3 Reputation Mechanisms

An issue that arises with the use of highly anonymous distributed computing methods to facilitate real-world auctions is that of malicious users and associated cyberattacks or fraud.

In an earlier review [24], we discovered a common method for handling such concerns: reputation mechanisms. In some markets, users would be associated with a reputation value. An example of this mechanism, Q-Score, is introduced by Zaman and He [27]. This Q-Score reputation value would be penalized for participating in failed transactions. This provides a simple heuristic that identifies users who may be gaming the system in some way – for example, by forcing their meter to repeatedly claim that energy was not delivered, thus preventing an outbound transfer of funds.

Common methods of handling low-reputation users include access blocking [20] and separate trading pools. The failed transaction metric illustrates the issue at hand regarding punishment, since users may also experience failed transactions due to hardware error, or other circumstances that are no fault of their own. Regardless of the reason, cutting off a user's access to energy is unethical due to the vitality of electricity in modern life [11].

5.3.4 Coin Mixing

Coin mixing is a protocol proposed by Ruffing et al. [21] intended to provide users on a blockchain a heightened level of anonymity. Blockchain technology is regarded as anonymous because a users' real identity (i.e., name, address, social security number, etc.) is not associated with their account, unlike with a credit card or bank account. However, because all transactions are public, if one's identity *is* ever connected to a particular account, their activity would be completely compromised.

A consideration for how to manage this might be to set up puppet accounts and trade using their funds. Again, though, the public nature of the blockchain makes this ineffective; it would be trivial to trace funds from the original, compromised account to its puppets.

Coin mixing aims to solve this problem by enabling a batch of users to untraceably send coins to a batch of proxy addresses.

5.3.5 Problem Statement

The technologies described all seek to solve problems that arise when designing systems to facilitate peer-to-peer energy trading, but their confluence creates novel problems that

have not been encountered elsewhere.

Blockchain technology enables the creation of truly distributed markets, but necessarily (by design) makes transaction information public. In an energy distribution setting, this can lead to dangerous attacks on user privacy.

One solution to this problem is to disconnect users' real identities from their trading accounts, as is proposed by Eisele et al. [8] in their TRANSAX protocol. This has the benefit of foiling privacy attacks based on trading or energy usage patterns, but introduces safety, fraud, and repudiation concerns.

Reputation mechanisms such as those described above constitute a blunt instrument that can be used to identify users who are repeatedly participating in failed transactions, as either buyer or seller of energy. This can be the result of fraud but, notably, can also occur as a result of equipment malfunction or user error.

Reputation mechanisms only detect misuse; they do not address it. Most proposed schemes employ these mechanisms by either: disconnecting users whose reputation falls below a certain threshold, or placing such users in a lower priority auction pool. These options are not acceptable for a vital utility such as electricity [11].

5.4 Design Requirements

This section introduces important requirements that must be met by a privacy-preserving scheme for transactive energy markets.

1. **Trustless**: The system architecture supporting the scheme shall not rely on any participant being 100% honest.
2. **Distributed**: The scheme shall rely on as few central authorities as possible in order to function.
3. **Individual Identification**: The scheme shall allow an *individual* to be identified from their pseudonym; not just a batch of users.
4. **Total Anonymity**: When trading, users shall be anonymous to all parties, including the DSO.
5. **Restorable Privacy**: After a user has been identified, they shall be able to resume trading under a new pseudonym with total anonymity.

6. **Automated**: The process of registering, validating, deanonymizing, and reanonymizing users shall be automated and require no manual intervention.

7. **Snoop-Resistant**: A curious party shall be prevented from revealing the true identity of a particular pseudonym, regardless of their level of privilege in the system. This shall be true at every stage of the process, including registration and deanonymization.

5.5 ILPTS Scheme

The purpose of our ILPTS scheme is to enable users to maintain anonymity on the TEM blockchain from all parties (including the DSO and other users) while still being able to be identified in edge cases. Additionally, it is important that this anonymity can be restored after they have been identified so that they may resume trading safely and privately.

5.5.1 Overview

Transactive energy market applications run using smart contracts to enable functions like trading, automatic auctions, market clearing, and transaction verification, among others.

Total anonymity in this case refers to the market participant being anonymous from all parties, including trading partners, other market participants, and the market operator (or DSO). This distinction is important because it is trivial to make a user anonymous to market participants while being visible to the market operator, but this sacrifices privacy. Similarly, making a user anonymous to all parties is possible at the expense of governance. Our ILPTS scheme seeks to achieve both goals.

First, anonymity is achieved using an existing cryptographic protocol known as *coin mixing* [21]. This process enables a batch of users on a blockchain to transfer their funds into new accounts that are not associated in any way with their original accounts. The process, which will be described in greater detail in the following sections, is impenetrable even to the batch of users involved in the mix. In the context of a TEM, batches of market participants associated with particular smart meters can transfer their funds (in whatever form they may appear) into anonymous accounts that can trade on the market with no association – digital or physical – to the original identity, account, or meter.

This creates a problem in situations where identification may be desirable. To amend this, a secondary blockchain (or separate storage on the main chain) is employed to store

recovery information. Since information on the chain is public, it is mandatory that this recovery information be protected. Fog nodes are employed to encrypt and, when necessary, decrypt this recovery information. However, if the recovery information consisted of an encrypted ID and pseudonym pair, then an individual fog node would be able to deanonymize users. This reduces the privacy benefit that is being aimed to achieve.

To combat this, the real identity of the users is split using another existing protocol, *Shamir's secret sharing scheme* [22], which allows information to be split in a secure but recoverable way, where a threshold number of parts are required to reconstruct the original information.

The users' real identities are split into n parts, each of which are then encrypted – by the user – using a different fog node's public key for each part.

The user then sends a tuple containing each of these encrypted parts along with their pseudonym to the smart contract, which activates the pseudonym and stores the parts and pseudonym on the recovery blockchain. This way, no individual fog node can deanonymize the user. However, if deanonymization becomes necessary, the smart contract can make a request to the fog nodes to decrypt their parts and send them to the DSO. Fog nodes are semi-trusted, as in not trusted to not snoop, but trusted to cooperate (send the correct information) lest they lose their privileged status.

This description is somewhat simplified. In reality, a batch of users must send their ID parts and pseudonyms together; otherwise it is trivial to see who sent which pseudonym to be activated. To facilitate this, a batch of users creates their ID part, pseudonym tuples. They then shuffle these among themselves using *cryptographic shuffling* [6], which enables them to create a master list of tuples where ID parts are correctly associated with pseudonyms, but no member of the batch knows which other member of the batch corresponds to which pseudonym.

It is this master list that is sent to the smart contract, who then sends each tuple to the recovery blockchain. However, before doing so, it performs a validation step. This step is to check that the ID parts being sent are indeed correlated with the real IDs of the users in the batch. It does so by homomorphically adding the ID parts associated with each fog node, then requesting these sums be decrypted by the fog nodes. These sums of parts can then be used to reconstruct the purported sum of the IDs using Shamir's secret sharing scheme. If this reconstructed sum matches the sum of the IDs in the batch of users, then validation is considered successful. At this point, the smart contract will activate the pseudonyms of all members of the batch and transmit their recovery data to the recovery blockchain and the registration process is complete.

This is a full description of onboarding. All data transmission happens on-chain via smart contract to support transparency and verifiability.

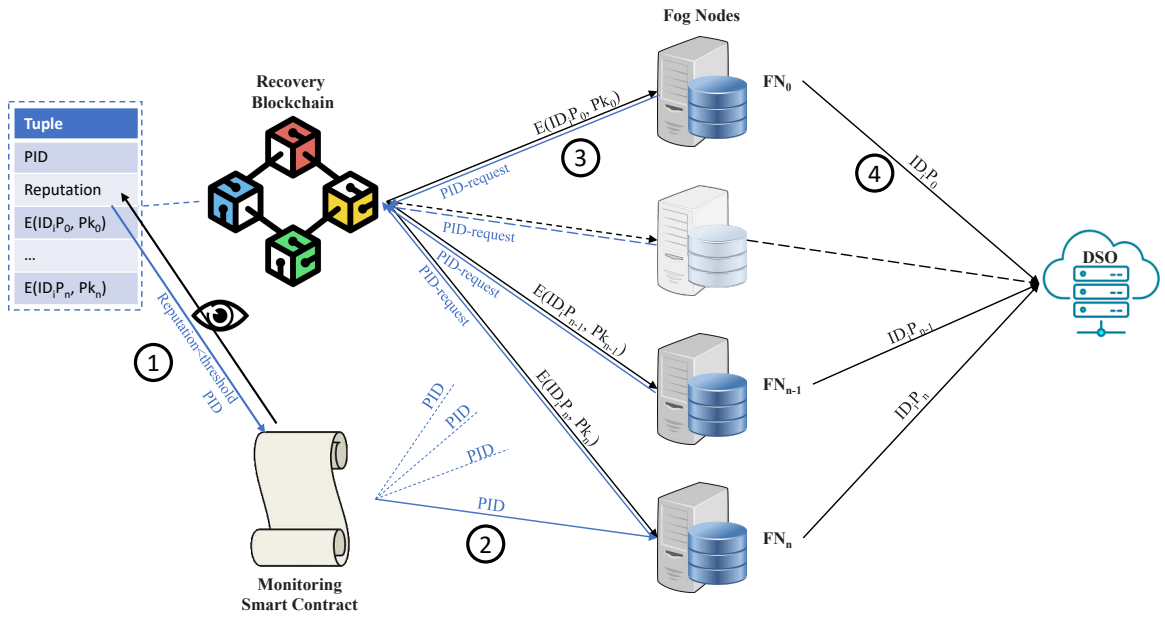

Figure 5.1: Deanonymization of a user under the ILPTS scheme.

The deanonymization process is demonstrated schematically in Figure 5.1 as well. A smart contract monitors the reputation of all market participants (who are all using pseudonyms). If the reputation of a pseudonymous market participant falls below a threshold, the contract locates the pseudonym ① and associated ID partitions in the recovery blockchain. It then sends a signal ② to each fog node containing the pseudonym. The fog node requests its ID partition ③, decrypts it and sends it to the DSO ④, which can then use the decrypted ID partitions to reconstruct the pseudonymous user's real identity.

From here, the DSO can investigate the cause of the user's reputation decline, which may be user error, hardware failure, or intentional hacking. It can then take appropriate actions. After a predetermined duration, the user can re-register on the market with a new pseudonym, recovering the ability to trade anonymously (until and unless their reputation declines again).

5.5.2 Formal Description

Here we will describe the processes and algorithms in ILPTS formally for supporting implementations.

Table 5.1: Nomenclature

ID_i	real ID of i^{th} user
ID_ip_j	part j of ID_i
FN_i	i^{th} fog node
Pk_i	Pk of fog node FN_i
Pk_i^u	Pk of i^{th} user ID_i
PK^B	list of Pks in batch B
PID_i	pseudonym of user ID_i
$\oplus$	homomorphic addition
SC	smart contract
$E(X, Pk)$	cyphertext of X encrypted using Pk
n_0	divisions of data
n	max FN ID when registering $(n_0 - 1)$
k_0	threshold to reconstruct data
k	max FN ID when deanonymizing $(k_0 - 1)$
B	batch of users registering
N_0	number of users in batch
N	max user ID when registering $(N_0 - 1)$
$X \rightarrow Y : W$	X sends W to Y
$X \leftarrow Y : Z$	X receives Z from Y
ω_i	registration data of i^{th} user
Ω	list containing registration data
Ω_i	state of list Ω at i^{th} user
$Shuffle(Y)$	randomize order of elements in list Y
$Partition(X)$	use [22] to generate $\gamma = [Xp_0, ..., Xp_n]$

5.5.2.1 Initialization

The initialization process is as follows:

1. A batch of N_0 users with identifiers ID_0 through ID_N join.
2. A subset of n_0 fog nodes is requested, denoted FN_0 through FN_n.

5.5.2.2 Key Exchange

The key exchange process is as follows:

1. Fog nodes FN_0 through FN_n generate public keys Pk_0 through Pk_n respectively.
2. Each fog node FN_i sends its public key Pk_i to the smart contract SC.
3. The smart contract SC sends all n_0 public keys $Pk_0 - Pk_n$ to each of the users $ID_0 - ID_n$.

5.5.2.3 Registration

To begin the registration phase, users $ID_i \in B$ split their IDs and encrypt each part with a corresponding fog node's Pk. The user's pseudonym PID_i is appended to this list of encrypted parts. These tuples are then shuffled among the users so that no one user in the batch knows which ID parts are associated with which real ID (which would break security). The tuples are ultimately sent to the smart contract.

The registration process is detailed in Algorithm 5.1.

Algorithm 5.1 Registration

Input: n_0 ▷ number of FNs and ID parts
Input: $B = [ID_0..ID_N]$ ▷ batch of user IDs
Output: $\Omega = [\omega_0..\omega_N]$ ▷ list of recovery tuples
for each $ID_i \in B$ **do** ▷ IDs in batch
  $\omega_i := Partition(ID_i) = [ID_ip_0, ..., ID_ip_n]$ ▷ list of ID parts
  for each $ID_ip_j \in \omega_i$ **do** ▷ ID parts in list
    $\omega_i[ID_ip_j] = E(ID_ip_j, Pk_j)$
  end for ▷ replace ID parts with encrypted ID parts
  $\omega_i := append(\omega_i, PID_i)$ ▷ append pseudo ID
  ID_i: generate Pk_i^u, Vk_i^u ▷ user keys
  $ID_i \rightarrow B$: Pk_i^u ▷ broadcast Pk_i^u to batch B
  $ID_i \leftarrow B$: $PK^B \ni Pk_{i-1}^u, ..., Pk_{i-i}^u$ ▷ retrieve all previously generated Pk^u from batch B
  for each $Pk_j^u \in PK^B$ **do** ▷ Pks currently available in batch
    $\omega_i = E(\omega_i, Pk_j^u)$ ▷ sequentially encrypt ω_i with each $Pk^u \in PK^B$
  end for
  $ID_i \leftarrow B$: Ω_{i+1} ▷ ID_i will wait until ID_{i+1} makes this available
  for each $E(\omega_j, Pk_i^u) \in \Omega_{i+1}$ **do** ▷ encrypted lists in batch list
    decrypt$(E(\omega_j, Pk_i^u)) \rightarrow \omega_j$ using Vk_i^u
    $\Omega_{i+1}[E(\omega_j, Pk_i^u)] = \omega_j$ ▷ replace encrypted list with list
  end for
  $\Omega_i = \Omega_{i+1}$
  $\Omega_i.append(\omega_i)$ ▷ append current user tuple to batch list
  Shuffle(Ω_i)
  $ID_i \rightarrow B$: Ω_i ▷ send updated batch list to batch
end for

5.5.2.4 Validation

The validation phase must ensure that the encrypted data being sent does indeed contain the real IDs of the users without revealing that data. To initiate this process, the contract homomorphically sums the parts of each tuple corresponding to each fog node and sends each of these sums to its respective fog node. Each fog node decrypts its sum and sends this back to the smart contract. The smart contract then reverses the splitting process and reconstructs a number. This number is checked for equality against the sum of the real IDs of all of the users in the batch. If the equality check succeeds, the information is

considered accurate and the pseudonyms can be activated.

The validation process is detailed in Algorithm 5.2.

Algorithm 5.2 Validation

Input: n_0 ▷ number of FNs and ID parts
Input: Ω ▷ batch list of tuples
Output: Ω ▷ validated batch list of tuples
Initialize: $\Lambda = [\lambda_0 = E(0, Pk_0), ..., \lambda_n = E(0, Pk_n)]$ ▷ homomorphically encrypted 0s
for each $\lambda_j \in \Lambda$ **do** ▷ one for each FN
  for each $\omega_i \in \Omega$ **do** ▷ user tuples in batch list
    $\lambda_j = \oplus(\lambda_j, E(ID_i p_j, Pk_j))$ ▷ homomorphically sum ID parts encrypted by FN_j
  end for
  $\lambda_j = \sum_{i=0}^{n} E(ID_i p_j, Pk_j)$ ▷ sums of ID parts encrypted by each FN
  $SC \rightarrow FN_j$: λ_j ▷ send each FN their sum
  FN_j: decrypt$(\lambda_j) \rightarrow \sum_{i=0}^{n} ID_i p_j$
  $FN_j \rightarrow SC$: $\sum_{i=0}^{n} ID_i p_j$ ▷ send decrypted sum to SC
  $\Lambda[\lambda_j] = \sum_{i=0}^{n} ID_i p_j$ ▷ replace encrypted sums with sums
end for
$\phi \leftarrow Partition^{-1}(\Lambda)$ ▷ reconstruct sum of IDs by reversing secret scheme [22]
if $\phi = \sum_{i=0}^{n} ID_i$ **then** ▷ if reconstructed sum equals sum of known IDs
  Validation Successful
else
  Validation Failed
end if

5.5.2.5 Finalization

Once validation is successful, the identity recovery data will be sent onto the recovery blockchain for storage and retrieval. The pseudonyms of the users in the batch will be activated on the market blockchain.

Storage of recovery data: $SC \rightarrow BC : \Omega$

Activation of pseudonymous users: Activate$(PID_0, ..., PID_N)$

5.5.2.6 Full Sequence Diagrams

Figure 5.2 shows, using a sequence diagram, the process of deanonymizing a user when their reputation falls below the threshold. While Figure 5.1 illustrates the operation well conceptually, Figure 5.2 better visualizes the flow of information between participants. Figure 5.3 shows the secure registration process involving the users in the registration batch, the verifying fog nodes, and the smart contract coordinating the process. Included in this as well is the validation process that ensures users are submitting their real IDs for recovery in the future.

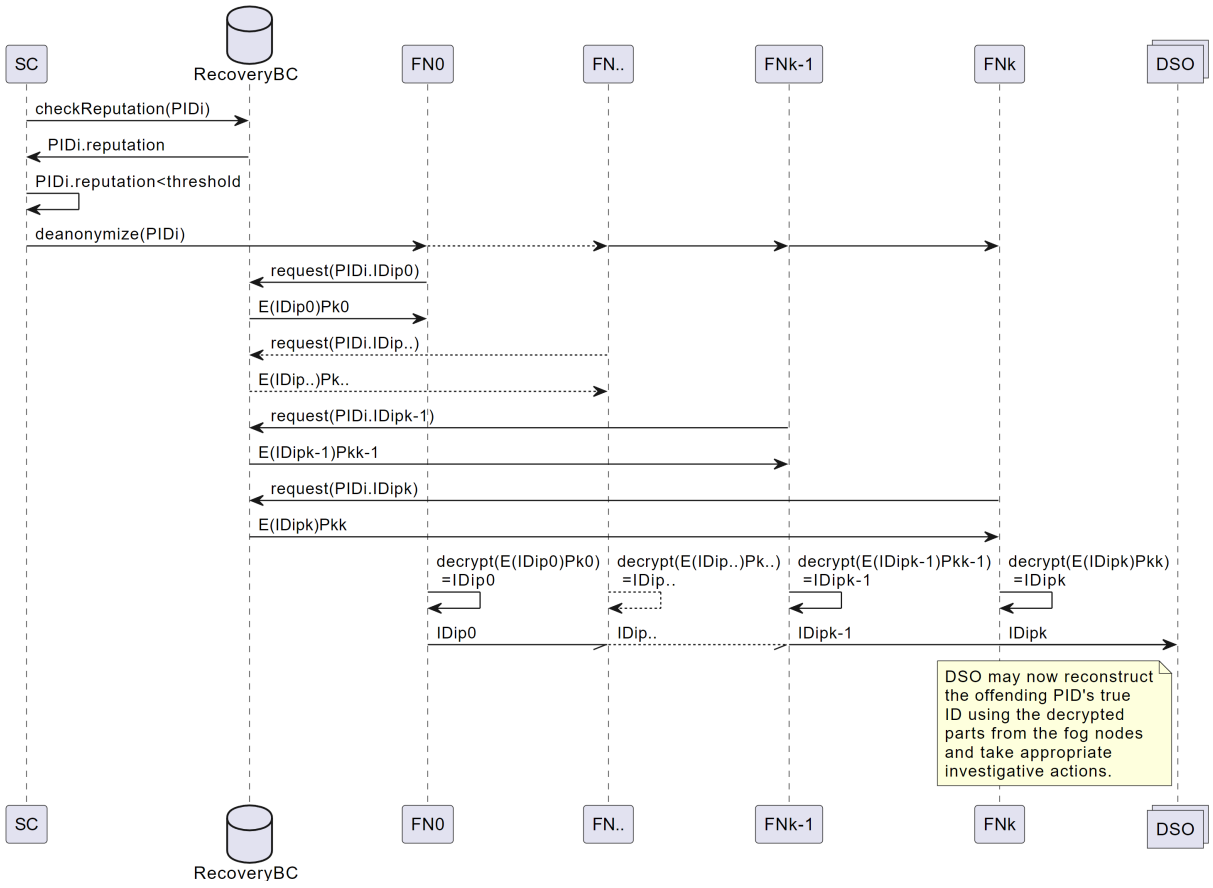

Figure 5.2: Full sequence diagram of the deanonymization phase.

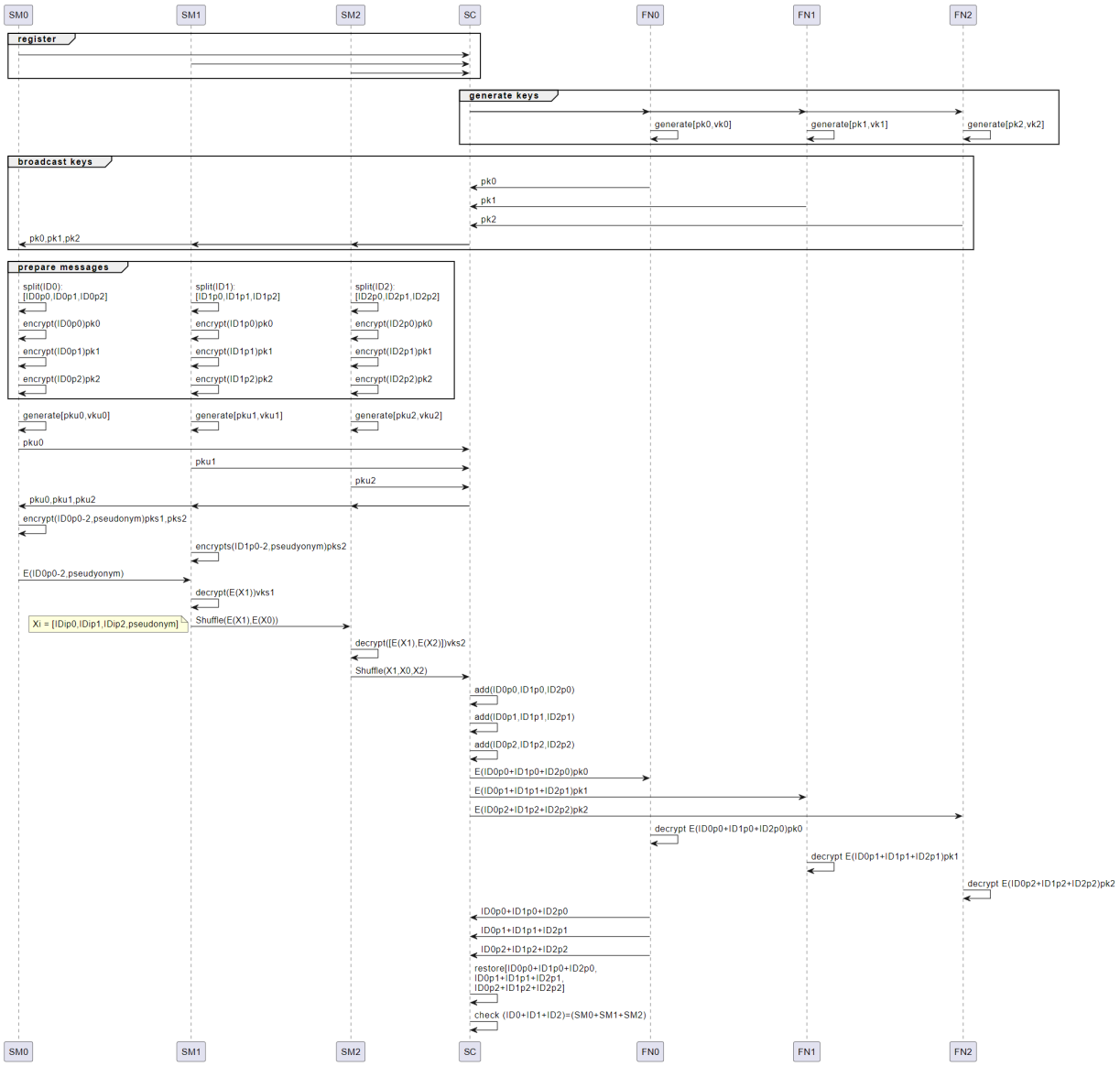

Figure 5.3: Full sequence diagram of the registration phase, operating on a batch of 3 users and 3 fog nodes.

5.6 Discussion and Analysis

This section first assess how well the requirements from Section 5.4 are met by ILTPS, and then reports on an informal security analysis of this proposed scheme.

5.6.1 Design Evaluation against Requirements

1. **Trustless**: The system successfully achieves a nearly trustless architecture. Registration and deanonymization are both automated and distributed, meaning no central authority or trusted third party is required. The only trust involved is that the fog nodes will correctly decrypt and forward the appropriate information to the DSO during deanonymization, but they are incentivized to do so as we discuss in Section 5.6.2. Fog nodes could theoretically collude to deanonymize individual users, but due to the properties of the secret sharing scheme employed such an attack would require k or more malicious nodes [22].

2. **Decentralized**: The registration process requires a blockchain, which is decentralized by nature, smart contracts, and fog nodes – no central authority. From this perspective, the scheme succeeds at this design goal. The fog nodes are semi-trusted; however, only a subset are required and they are interchangeable to a point. The monitoring and deanonymization processes are handled entirely by a smart contract and fog nodes. The information is of course transmitted to the DSO, but the process of deanonymizing the user for them is entirely automated and distributed. Moreover, the information sent to the DSO is not pre-reconstructed, making it less vulnerable to attacks (which we discuss in Section 5.6.2).

3. **Individual Identification**: ILTPS successfully allows users to sign up in batch, with validation that does not compromise the anonymity of their pseudonym, while still enabling individual identification when necessary.

4. **Total Anonymity**: Total anonymity is achieved since the true identities of the pseudonyms are split into many pieces which are each encrypted by different fog nodes. This property is important because previous schemes have had to decide between sacrificing a level of anonymity (by registering with the DSO) or traceability (having no mechanism to identify a pseudonym).

5. **Restorable Privacy**: A user can be reanonymized by simply undergoing the registration process again. If no new users are being added, a subset of existing users can be re-registered automatically to facilitate the batch registration process.

6. **Automated**: All information exchange and validation are handled automatically by an on-chain smart contract, fulfilling this design goal. Additionally, reputation monitoring and requests to fog nodes to deanonymize a target user are also handled automatically by a smart contract.

7. **Snoop-Resistant**: This design goal is achieved because the deanonymization process is initiated by the smart contract. The contract is written in such a way that manually initiating this process is not possible; it only occurs when a user's reputation is detected to have fallen below a certain threshold.

Overall, ILPTS achieves the primary objective of providing total anonymity while maintaining safety. The system successfully enables total anonymity while simultaneously retaining a link to the pseudonymous user's original identity. Users in markets employing this scheme will have their identities protected from all parties, including the DSO. The lack of centralized storage of links means that there is no single target for attackers, nor an internal leakage threat. Using a smart contract to monitor reputations and activate the deanonymization policy prevents the policy from creating a potential attack surface; as long as the blockchain remains secure (a reasonable assumption) the policy will be protected.

5.6.2 Security Analysis

This section evaluates how ILPTS handles certain attacks and security properties, as well as discusses attacks methods that may be unique to the scheme. Additionally, it evaluates how the scheme performs when used in combination with other cybersecurity solutions designed for TEMs. Finally, this section considers the security properties associated with theoretical modifications to the schemes, and whether these may serve as overall improvements that could warrant future research and development.

5.6.2.1 Trust

It can be seen that our ILPTS scheme successfully pulls off private *and* regulatable market activities with validation, all without trusted parties.

Fog nodes are not trusted not to snoop (hence the splitting of IDs), but are trusted to send information when requested. While this is a form of trust, it is reasonable given:

- Failure to oblige will cost their privileged status in the network, creating an implicit incentive to comply.

- Only k of n fog nodes need to comply (and not fault) in order for deanonymization to be successfully performed, meaning that some number of nodes could, in fact, maliciously not comply without affecting the integrity of the scheme. This ratio can be set arbitrarily by the DSO to their desired security level.

While fog nodes do have privileged status in the scheme, they do not need to be fully trusted nor do they all need to be trustworthy even simply with compliance. In addition, as stated, while there is no incentive for them not to snoop, doing so would not reveal any information.

5.6.2.2 Reputation Attack

In a reputation attack, a user could potentially configure their smart meter settings to ensure trading is done with a particular prosumer, and then cause those transactions to fail consistently leading to a decline in reputation for the victim. Luckily, in our system the consequences of such an action is significantly reduced, since it would only cause an investigation rather than a wholesale ban from the market. As can be seen in Section 5.5, the anonymity of the user can easily be restored by re-registering with a new batch. Newly acquired anonymity is not damaged by past activation of the protocol; it is totally restored to its original state and even an individual at the DSO would not be able to trace a user between market identities.

5.6.2.3 Man-in-the-Middle Attack

There are several locations at which communication could be intercepted during each phase of the scheme. These include:

A $SM \rightarrow SC$ during registration

B $SC \rightarrow FN$ during registration

C $FN \rightarrow BC$ during registration

D $BC \rightarrow FN$ during deanonymization

E $FN \rightarrow DSO$ during deanonymization

Concerning locations A through D, there is little risk even if communication is intercepted. This is due to the fact that all information at these stages is both split and encrypted with multiple different keys, making the information inaccessible unless k fog nodes were to collude.

The greatest risk lies in location E; when the fog nodes send the ID parts to the DSO. This can be mitigated by encrypting the information with a public key from the DSO, but this is less secure than the multiple stages of encryption during the registration process. A positive confounding factor is that this does require multiple points of interception (specifically between k fog nodes and the DSO).

An *inside job* at the DSO is not a risk since the DSO cannot manually initiate the deanonymization protocol for an arbitrary user.

5.6.2.4 Synergies

ILPTS can be used independently to provide anonymity on the transaction ledger in a safe, distributed, and ethical manner. However, it can also be used cohesively with other security solutions designed for transactive energy markets.

For example, TRANSAX [8], which we discussed earlier, would benefit directly from being used in conjunction with ILPTS. This combination would allow TRANSAX to continue offering its boundary safe anonymity, while mitigating the negative effects associated with group-level identification.

Another example is a solution we previously developed called cyclic homomorphic encryption aggregation (CHEA) [25]. CHEA contributes to the security provided by ILPTS by protecting the dual of the information at risk on the transaction ledger: the energy consumption data produced and sent by the smart meters themselves. This information exfiltrated from either source can be cross referenced to identify a prosumer or used independently to invade their privacy. Thus, the combined effect of these schemes is to enhance this privacy from both directions.

Additionally, the unique properties of each scheme reinforce each other. For example, CHEA is an aggregation scheme that does not require fog nodes. Limiting the role of fog nodes in this way enhances overall security by reducing the trust requirements of each node; in particular, CHEA combined with ILPTS involves fog nodes *only* in the protection of identity recovery information. ILPTS used with traditional aggregation schemes would mean that fog nodes are involved in processing both channels of usage information, increasing the risk of collusion or malicious behaviour.

5.6.2.5 Potential Improvement

ILPTS was initially designed with the intention of employing fog nodes to provide computational and storage resources while maintaining a distributed environment to reduce trust requirements and enhance reliability.

A goal throughout was to reduce the role of the fog nodes by structuring the scheme such that most if not all of the data could be kept on the blockchain. This was to decrease reliance on fog nodes, lower the chances of critical failure, and also to reduce the computational demands of the nodes. By the final stage of the ILPTS' development the role of the fog nodes, from a processing and storage standpoint, had become limited enough that allowing the role to be handled by smart meters themselves became a consideration.

Replacing the fog nodes would require additional considerations beyond the scope of this work, but the basic idea would be to require each smart meter to generate a set of keys to be used when they are required to fulfill the *fog node* role in either of the policies. Registration and deanonymization would look largely similar, with a separate subset of smart meters replacing the fog nodes.

This concept confers benefits including increased reliability due to greater redundancy (as there are more smart meters to pull from than fog nodes), reduced collusion potential, and decreased infrastructural requirements. On the other hand, such a modification would need to take into account the reduced computational capacity of smart meters, mechanisms for ensuring a viable subset is available, fault tolerance, and the fact that smart meters often may be busy with other tasks.

5.7 Conclusion and Future Work

Our market privacy solution (ILPTS) aims to provide privacy, safety, and security to prosumers participating in a blockchain-based transactive energy market. These goals have been achieved and enhance prosumer safety, regulatory compliance, and market adoption of TE solutions.

A limitation of the work is that the scheme was not tested experimentally, either via implementation or software simulation, so security results are based solely on formal analysis. In particular, it was found by formal analysis that the ILPTS privacy solution successfully creates an environment where the design specifications are met, namely:

- Users are totally anonymous to all parties;

- This anonymity can be uncovered only by an autonomous system under specific circumstances;
- Anonymity can be restored when required;
- The solution does not contain centralized vulnerabilities.

This solution improves equity and ability to regulate in a TE environment, reducing the risk of fraud and the need for coarse punitive measures like priority pooling or disconnection.

Future work should include (ideally *in situ*) simulations of the solution and cyberattacks that may be performed, to further verify the robustness of the scheme proposed. This would necessitate creating a functional version of the scheme, with appropriate blockchain environment, smart contract code, and client code, another future challenge. Finally, exploration of a further decentralized scheme, as discussed in Section 5.6.2.5, could prove an interesting research direction.

Bibliography

[1] Muhammad Rizwan Asghar, György Dán, Daniele Miorandi, and Imrich Chlamtac. Smart meter data privacy: A survey. *IEEE Communications Surveys & Tutorials*, 19 (4):2820–2835, 2017. doi:10.1109/COMST.2017.2720195.

[2] Mohamed Baza, Ramy Amer, Amar Rasheed, Gautam Srivastava, Mohamed Mahmoud, and Waleed Alasmary. A blockchain-based energy trading scheme for electric vehicles. In *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–7, 2021. doi:10.1109/CCNC49032.2021.9369517.

[3] Mohamed Baza, Ahmed Sherif, Mohamed M. E. A. Mahmoud, Spiridon Bakiras, Waleed Alasmary, Mohamed Abdallah, and Xiaodong Lin. Privacy-preserving blockchain-based energy trading schemes for electric vehicles. *IEEE Transactions on Vehicular Technology*, 70(9):9369–9384, 2021. doi:10.1109/TVT.2021.3098188.

[4] Jonatan Bergquist, Aron Laszka, Monika Sturm, and Abhishek Dubey. On the design of communication and transaction anonymity in blockchain-based transactive microgrids. In *Proceedings of the 1st Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers*, SERIAL '17, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450351737. doi:10.1145/3152824.3152827.

[5] Vitalik Buterin. Ethereum white paper: A next generation smart contract & decentralized application platform. *ethereum.org*, 2013. URL https://ethereum.org/en/whitepaper/.

[6] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM*, 24(2):84–90, feb 1981. ISSN 0001-0782. doi:10.1145/358549.358563.

[7] Tao Chen, Qais Alsafasfeh, Hajir Pourbabak, and Wencong Su. The next-generation U.S. retail electricity market with customers and prosumers–a bibliographical survey. *Energies*, 11(1), 2018. doi:10.3390/en11010008.

[8] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[9] Sahil Garg, Kuljeet Kaur, Georges Kaddoum, Francois Gagnon, and Joel J. P. C. Rodrigues. An efficient blockchain-based hierarchical authentication mechanism for energy trading in v2g environment. In *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, 2019. doi:10.1109/ICCW.2019.8756952.

[10] Ateeb Hassan, Hadi Nabipour Afrouzi, Chua Hong Siang, Jubaer Ahmed, Kamyar Mehranzamir, and Chin-Leong Wooi. A survey and bibliometric analysis of different communication technologies available for smart meters. *Cleaner Engineering and Technology*, 7:100424, 2022. ISSN 2666-7908. doi:10.1016/j.clet.2022.100424.

[11] M. Jayachandran, K. Prasada Rao, Ranjith Kumar Gatla, C. Kalaivani, C. Kalaiarasy, and C. Logasabarirajan. Operational concerns and solutions in smart electricity distribution systems. *Utilities Policy*, 74:101329, 2022. ISSN 0957-1787. doi:10.1016/j.jup.2021.101329.

[12] Hamzah Khan and Tariq Masood. Impact of blockchain technology on smart grids. *Energies*, 15(19), 2022. ISSN 1996-1073. doi:10.3390/en15197189.

[13] Desen Kirli, Benoit Couraud, Valentin Robu, Marcelo Salgado-Bravo, Sonam Norbu, Merlinda Andoni, Ioannis Antonopoulos, Matias Negrete-Pincetic, David Flynn, and Aristides Kiprakis. Smart contracts in energy systems: A systematic review of fundamental approaches and implementations. *Renewable and Sustainable Energy Reviews*, 158:112013, 2022. ISSN 1364-0321. doi:10.1016/j.rser.2021.112013.

[14] Aron Laszka, Abhishek Dubey, Michael Walker, and Douglas Schmidt. Providing privacy, safety, and security in iot-based transactive energy systems using distributed ledgers. In *Proceedings of the Seventh International Conference on the Internet of Things*, 10 2017. doi:10.1145/3131542.3131562.

[15] Donghe Li, Qingyu Yang, Wei Yu, Dou An, Xinyu Yang, and Wei Zhao. A strategy-proof privacy-preserving double auction mechanism for electrical vehicles demand response in microgrids. In *2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC)*, pages 1–8, 2017. doi:10.1109/PCCC.2017.8280481.

[16] Donghe Li, Qingyu Yang, Dou An, Wei Yu, Xinyu Yang, and Xinwen Fu. On location privacy-preserving online double auction for electric vehicles in microgrids. *IEEE Internet of Things Journal*, 6(4):5902–5915, 2019. doi:10.1109/JIOT.2018.2872444.

[17] Federico Lombardi, Leonardo Aniello, Stefano De Angelis, Andrea Margheri, and V. Sassone. A blockchain-based infrastructure for reliable and cost-effective iot-aided smart grids. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, 01 2018. doi:10.1049/cp.2018.0042.

[18] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[19] Eric Münsing, Jonathan Mather, and Scott Moura. Blockchains for decentralized optimization of energy resources in microgrid networks. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 2164–2171, 08 2017. doi:10.1109/CCTA.2017.8062773.

[20] Masoumeh Nazari, Siavash Khorsandi, and Jaber Babaki. Security and privacy smart contract architecture for energy trading based on blockchains. In *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, pages 596–600, 2021. doi:10.1109/ICEE52715.2021.9544155.

[21] Tim Ruffing, Pedro Moreno-Sanchez, and Aniket Kate. Coinshuffle: Practical decentralized coin mixing for bitcoin. In Mirosław Kutyłowski and Jaideep Vaidya, editors, *Computer Security - ESORICS 2014*, pages 345–364, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11212-1.

[22] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, nov 1979. ISSN 0001-0782. doi:10.1145/359168.359176.

[23] Giriraj Sharma, Amit M. Joshi, and Saraju P. Mohanty. strade: Blockchain based secure energy trading using vehicle-to-grid mutual authentication in smart transportation. *Sustainable Energy Technologies and Assessments*, 57:103296, 2023. ISSN 2213-1388. doi:10.1016/j.seta.2023.103296.

[24] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, and John Mylopoulos. A review of cybersecurity concerns for transactive energy markets. *Energies*, 16(13), 2023. ISSN 1996-1073. doi:10.3390/en16134838.

[25] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, Masoud Bashari, and John Mylopoulos. Cyclic homomorphic encryption aggregation (chea)—a novel approach to data aggregation in the smart grid. *Energies*, 17(4), 2024. ISSN 1996-1073. doi:10.3390/en17040878.

[26] Noelia Uribe-Pérez, Luis Hernández, David De la Vega, and Itziar Angulo. State of the art and trends review of smart metering in electricity grids. *Applied Sciences*, 6 (3), 2016. doi:10.3390/app6030068.

[27] Ishtiaque Zaman and Miao He. A multilayered semi-permissioned blockchain based platform for peer to peer energy trading. In *2021 IEEE Green Technologies Conference (GreenTech)*, pages 279–285, 2021. doi:10.1109/GreenTech48523.2021.00052.

Chapter 6

Discussion

This chapter discusses the results of the three papers (Chapters 3 to 5) along the four thesis research questions (TRQ1-TRQ4) identified in the introduction, Section 1.2. We also discuss some interesting features of the proposed schemes, as well as the limitations of the studies. Finally, we discuss suggestions, requirements, and potential challenges surrounding implementation of the schemes in real-world scenarios.

6.1 TRQ1

The first thesis research question asked: *What privacy and security concerns exist in transactive energy implementations that are not currently being adequately addressed?* In order to answer this, we first had to comb the literature on the topic for areas of concern, specific attacks, and security flaws, as well as proposed and existing solutions to these threats.

We discovered 15 threat vectors (although we combined market and data privacy for clarity) that can be summarized into 5 overarching security categories: cyberattacks, trusted entities, infrastructure, privacy, and novel attack surfaces. Some threats fit into multiple categories. In the following sections, we will describe each of these categories and the threats they describe.

6.1.1 Cyberattacks

Threats in this category consist of existing cyberattacks that would be effective, or could be extended to be effective, in the transactive energy environment. These include:

False Data Injection (FDI)

FDI refers to maliciously reporting incorrect information to a system. In the context of TE, this will generally require modifying or compromising a smart meter in order to cause it to report incorrect or arbitrary information. This can be done with the intention of fuzzing the grid by observing the effect of specific values, damaging physical infrastructure by falsely reporting load, manipulating energy prices, or tampering with bid verification.

FDI existed as a concern in traditional power grids, so state estimation algorithms have already been developed to minimize the impact of false or incorrect measurements. For TE, the Electron Volt Exchange introduced Robust State Verification, a distributed method of verifying measurements tailored for TE that builds off of past approaches [27]. Other solutions attempt to reduce the impact of FDI attacks after they occur, such as that introduced by Onumanyi et al. [24].

Due to the existence of such solutions, we considered FDI to be relatively low priority. Although the solutions have some weaknesses, exploiting these requires an unrealistic degree of effort and coordination relative to other threats.

Denial of Service

DOS, or in some cases distributed DOS (DDOS), attacks are cyberattacks that attempt to overwhelm a network or server with traffic in order to make the network or application unresponsive and thus inaccessible to legitimate users.

DOS attacks can be carried out with little sophistication, and can be performed by unique means in a transactive energy market [34, 1]. They can result in damaging effects such as causing bids to be discarded sporadically [1].

Despite their effectiveness and ease of execution, we did not classify DOS attacks as a significant concern for TE researchers. Existing measures such as reputation mechanisms easily detect and manage the patterns of behaviour associated with such an attack on the market layer [35]. DOS attacks on infrastructure outside the application are beyond the scope of TE designers.

51% Attack

The 51% attack is a type of attack on a blockchain in which an adversary takes control of over half of the computing resources on a blockchain, enabling them to unilaterally alter

on-chain data. Traditional global blockchains are simply too large for this to be a realistic concern, so solutions are rare. However, in a TEM a blockchain would likely oversee a much smaller area, such as a city or neighbourhood, making this threat more concerning. Some models, such as that introduced by Zaman and He [35], present obstacles but, overall, we consider this threat to be of concern to TE researchers.

6.1.2 Infrastructure & Points of Failure

These concerns represent problems with either: the foundation upon which a TEM system is built, or pressure points within the system that have the potential to compromise reliability or privacy.

Single Point of Failure

Single points of failure (SPOFs) were one of the most frequently mentioned security concerns in the TEM literature. It is a general threat that can present in a variety of ways. A single point of failure can affect information privacy, operations, and access. For example, sensitive customer data all stored in a single location is a security point of failure that threatens customer privacy if it is compromised. It also presents a point of failure for access to said information if the server goes down for any reason, whether malicious or incidental. Both are similarly true with communication that must always pass through a particular node. On the other hand, points of failure can be entities like key generation centers who are necessary for continued operations.

It makes sense that this concern would be brought up frequently, as points of failure are a large part of what TEM and microgrids hope to address. By making energy access less centralized, these models aim to reduce the impact of generator outages or failure to meet demand. However, SPOF is not only mentioned in terms of cases where TE solves it. Indeed, all TEM models we examined retained points of failure somewhere in their architecture.

Solutions to SPOFs are particular to the process in which the SPOF arises. For example, Zaman and He [35] introduce a method of authenticating new prosumers in a distributed manner. The Electron Volt Exchange [27] introduce a decentralized state estimation algorithm, solving a SPOF that exists in traditional grid management.

A complete solution to SPOF would involve a fully decentralized TEM, where no processes rely on any single entity. Since such a model has not been proposed (and partial

solutions, as described above, still leave gaps in coverage), we considered SPOF to be unsolved. SPOFs are also a serious concern, as they can affect access to energy. It was for these reasons that we determined SPOFs to be a high-priority threat.

Edge Nodes

Edge nodes (also called fog nodes) are auxiliary nodes in the TEM network. That is, they are not direct participants in the market application. These nodes are often used to facilitate additional functionality, as in the Electron Volt Exchange (EVE) [27] where they may serve as regional aggregators.

Edge nodes pose an infrastructural risk by presenting a quasi-single point of failure. Many schemes require some percentage of edge nodes to be active or in agreement to maintain operations, hinging reliability on their uptime. Others, like the EVE mentioned earlier, may employ edge nodes to service regions within a grid – meaning that an edge node represents an SPOF for its particular region.

Communication

Network security underpins the security of a TEM. However, it is a research area that intersects with TE research, rather than being a component of it [8]. As in other similar cases, TEM designers are recommended to build around potentially insecure networks, rather than concern themselves with network security itself.

Regulation & Standardization

Regulation of TE will be an important factor in its development. Governments are wary of blockchain-based solutions (and cryptography in general) due to their difficulty to regulate [3]. However, regulation and standardization can also be important security considerations. Regulations can be employed to ensure that Internet of Things (IoT) devices meet certain security and interoperability standards [20]. Regulation can also help to define market operations and reconciliation practices, which would influence TE design [9]. Finally, a lack of standardization for DLT software best practices was found to be a source of security weakness, leading to inconsistent experiences between scripts and applications.

While these topics are relevant, the development of such standards and regulations were determined to be out-of-scope for TE designers, who should instead focus on mitigating the effects of their absence.

Smart Contracts

Smart contracts are a critical underpinning of most blockchain-based transactive energy implementations. They facilitate complicated applications like price and preference matching between prosumers, auction clearing, and registration, to name a few examples [2, 21]. These functions cannot be achieved with a pure blockchain; the additional scripting a smart contract provides is necessary [19]. Unfortunately, popular implementations of smart contracts were found to be insecure and have had negative effects in industries that have built on them [13, 23].

Similar to regulation & standardization, improving smart contract security is outside of the domain of TE design research. Instead, researchers should seek to limit smart contract use where possible and ensure schemes are carefully designed around them.

6.1.3 Trusted Entities

Trusted entities are parties within a TEM architecture who have an elevated permission level, handle sensitive information, or are otherwise necessarily trusted to act in a cooperative, non-destructive, and non-invasive manner. They present an implicit security threat because good faith is not a robust security solution; any of these trusted entities could compromise the system should they become malicious.

Single Point of Failure

Single points of failure in a trust context arise whenever a single party is privy to or responsible for the safe storage of some sensitive information. In these cases, either compromise by an external adversary or a malicious insider can result in data leakage, loss of customer privacy, or operational failure. These trust points of failure can often be identified in the literature when a scheme or model refers to a trusted third party or trusted authority.

Edge Nodes

Edge nodes often serve as trusted entities where they are employed. In many cases edge nodes are counted on not to collude to reconstruct sensitive information. Some researchers note that they are also trusted to provide reliable responses to requests, and are often in a unique position to inject false information – as is the case with aggregating nodes in the scheme proposed by Saha et al. [28], for example.

Authentication

Authentication of users almost always relies on a trusted authority. The DSO is an implicitly trusted entity, but as we will demonstrate in TRQ2, even it can be treated as adversarial.

6.1.4 Privacy

The threats in this section concern the concealment of private information and the risks, especially to the consumer, that its leakage could bring forth.

Energy Usage Data

Patterns of energy usage are considered a significant privacy and safety concern in the literature. This data becomes significantly more appealing to hackers in a smart grid TE environment due to the increased frequency of reporting, which can facilitate more sophisticated attacks, as well as the novel attack surfaces.

Energy consumption data is a serious threat because it can enable more grave attacks than others. Adversaries can use this data to find out what kind of appliances are in a home (and thus if they are valuable), when people are home, and even what activities inhabitants might be engaged in [15]. This knowledge can assist with burglary and even kidnapping [17].

In spite of this risk, consumption data is required for many critical operations. As such, solutions exist outside of the TEM literature, but they were not invoked in any of the models we investigated. Additionally, these solutions are not tailored for the distributed TEM environment. Due to the high risk associated with this data and the lack of consideration it received in the literature, we found it to be a high priority for TE research.

Market Privacy

Market privacy concerns consumer identities and trading activity in the TEM. The goal is to decouple a user's real identity from their trades to prevent energy usage-style attacks against them from the market layer. However, it is challenging to anonymize users while also maintaining reasonable authority and safety in the grid environment. This issue lead us to consider market privacy a high priority research direction.

Data Privacy

Data privacy is used to describe the privacy of impersonal data such as operational information, consumption and production constraints, trading preferences, and any regularly reported statistics. These kinds of data would typically be flowing from the prosumer to a central analyzer like the DSO. Techniques for enhancing privacy in this area are in line with traditional information privacy techniques: encryption, random noise, aggregation, buffering, etc. As noted in regards to energy usage data, these techniques are not necessarily employed in TEMs, and we considered research into adapting them for the distributed environment valuable.

6.1.5 Novel Attack Surfaces

Each of the threats in this category represents a relatively new point of attack to consider within the TEM architecture. These can create risk simply by virtue of less existing research or improperly assuming that existing methods can be ported to this new environment.

Market Attacks and Privacy

The market in a transactive energy market is a unique combination of software, hardware, and utility. It creates opportunities for new kinds of attacks, as we discuss earlier, which are as yet unmanaged by existing implementations. Privacy management also presents a unique challenge in this domain. New privacy violations are possible when the utility being traded on a blockchain is necessarily inventoried in multiple locations. The distinctive properties of the transactive energy environment create obstacles for traditional privacy-preserving mechanisms like pseudonymity; a fact that is critically relevant to and expanded upon in our discussions of TRQ2 and TRQ3.

Electric Vehicles

Electric vehicles (EVs) present a special challenge to TE designers. They interact with the grid in novel ways, regularly connecting at new locations unlike most other equipment which is stationary. While they do not present a distinct security risk to the TEM itself, protecting EV user privacy involves original challenges. Their ability to roam means a security compromise could enable stalking.

Smart Meter Firmware

IoT firmware is notoriously insecure, often deployed by manufacturing companies with little to no cybersecurity experience. Smart meters being fundamentally insecure is a serious risk to TEMs, as it makes the cyberattacks and privacy attacks discussed earlier much more viable and likely.

While firmware is outside of the scope of TE research, designers can and should work towards detecting and limiting the impact of compromised hardware.

Smart Contracts

Smart contracts remain a relatively new application type. As we note in the discussion on regulations, standardized best practices have yet to arise [30]. Even scripting languages with which to write contracts have not settled into a reliable standard. The status quo, Solidity, is notoriously prone to erroneous and insecure scripts [23].

6.2 TRQ2

The second thesis research questions asked: *Can systems be created to improve a selection of these concerns and what properties would they need to have?*

6.2.1 Threat 1: Energy Usage Data Privacy

As noted earlier, energy usage data can present a significant privacy and security risk to the consumer [17, 15]. Said data can be used to deanonymize users on the market platform, as well as discover details about a consumers' whereabouts, activities, and possessions. These kinds of information can be used to invade the user's privacy, but also can facilitate physical attacks on the user such as burglary or kidnapping [17].

Nevertheless, energy usage data is required for many purposes, including billing, trade verification, state estimation, load profiling, and demand response. Thus, the problem of preserving the privacy of this information while making it usable by the relevant parties (in the case of TE this is typically the DSO) has been considered. Some broad techniques that have been proposed include battery filtering, algebraic transformation, adding noise to data, and data aggregation [4, 11, 33]. Of these, we considered data aggregation to be

the most appropriate for the transactive energy domain. This was due to the fact that it retains exact values for measured data, does not require additional consumer hardware to work, and reports demand as it occurs.

Data aggregation techniques exist in other domains, including closely related techniques designed for the smart grid [18, 38]. While the motivating concern was energy usage data privacy, the process of aggregating said data introduces other security concern categories beyond privacy – specifically, trusted entities and points of failure.

As such, the threats associated with a solution of this sort are:

- Energy usage data privacy
  - The leakage of consumption data is the primary threat being addressed.
- Market privacy
  - Market privacy can be directly compromised by the leakage of consumption data. This information can be compromising at any point in the process, not just directly from the meter (where it would be most easily associated with the consumer). Indeed, even aggregated data can pose a risk in some cases where statistical attacks are viable.
- Trust
  - Edge nodes
    - Edge nodes present an implicit trust risk in most schemes as they are provided with unaggregated data (in these contexts they are often referred to as aggregating nodes).
  - Authentication
    - Authentication ties into market privacy. As noted earlier, the DSO is often implicitly trusted with identifying information. This can increase the risk of data reconstruction and subsequent privacy violations if aggregation regions are identifiable.
  - Single points of failure
    - Other trusted entities that can singularly violate trust include key generation centers (KGC) that are third parties tasked with generating cryptographic keys to be distributed among the network to facilitate privacy-preserving aggregation.

- Infrastructure & points of failure
  - Edge nodes
    - The failure of an edge node that is being used to aggregate data will make reporting impossible for the region it supports, presenting a serious infrastructure risk.
  - Single point of failure
    - While edge nodes are an obvious SPOF, others can exist. As noted earlier, KGCs are necessarily responsive for some schemes. In schemes with intranetwork aggregation and static topology, any root node becomes a single point of failure. Others may fail if a particular SM fails due to the nature of the cryptosystem employed
  - Communication
    - Dependence on externally secure communication channels to address man-in-the-middle attacks is undesirable; a scheme should be robust to such infiltrations.

Accordingly, it can be stated that the primary design goal of such a scheme is to address these security concerns.

6.2.2 Threat 2: Market Privacy

Market privacy specifically concerns the protection of consumers within the market application. This includes the blockchain on which transactions are recorded, the smart contracts used for market clearing, and generally the information that must be shared to facilitate energy trading.

The mission to protect consumers within the market layer faces a seemingly intractable dilemma: full anonymity inhibits regulatability, while safer solutions necessarily compromise on privacy. The former means that a fully anonymous solution prevents an authoritative body, such as the DSO, from being able to handle misuse or malfunction with nuance. Existing solutions to address malicious market activity such as alternative trading pools or altogether disconnecting users from the grid [35, 22] are not compliant with legal or ethical guidelines [10].

Addressing market privacy will encounter the following threats:

- Market privacy
  - Market privacy is the primary threat being addressed.
- Trust
  - Authentication
    - Authentication is central to market privacy. The security stature of the parties and processes involved will directly impact the quality of market privacy achieved. Which is to say, market privacy *cannot* be attained without secure, privacy-preserving authentication.
  - Single points of failure
    - Trust points of failure can arise during the authentication process, but also when storing user identities. In fact, this relates to the dilemma described earlier; a trusted authority overseeing the relationship between anonymous and real identities is a potential compromise, but a compromise nonetheless as it introduces the threat in question.
- Infrastructure & points of failure
  - Single point of failure
    - Storing identities with a trusted third party can also introduce operational SPOF; if the server is not live the DSO will have no recourse for malicious or malfunctional market activity.
  - Communication
    - Assuming that a solution involves transmitting a user's real identity at some point (when it becomes necessary), the process for doing so should be designed under the assumption that some kind of network attack may occur.
  - Regulation & standardization
    - Interoperability between disparate softwares will be necessary to create any complex process in a TEM. Also, as noted, standard practices in DLT would ease the necessary risk of using smart contracts.
  - Smart contracts
    - Due to the complexities involved with constructing a market and anonymization process, smart contracts are a necessary evil. However, as stated in TRQ1, the aim should be to use them only when called for.

6.2.3 Themes

It can be seen that addressing these two concerns involves consideration of many similar security concerns. These include:

- Trust
- Points of Failure
- Privacy
- Authentication
- Communication Infrastructure

Security concerns in these areas share a unified origin: centralization.

Transactive energy and blockchain are both responses to the novel capacity of information technology to support decentralized applications. Decentralization of processes can enhance security, privacy, and reliability by circumventing the need for trusted authorities, removing single points of failure, enabling distributed authentication, and enabling anonymous interactivity.

TEMs built on blockchains seek to capitalize on the benefits of decentralized economics to enhance the benefits of decentralized energy management via microgrids [14, 19]. It is reasonable, then, that seeking to enhance these features in the other processes in the TEM would naturally lead one to this convergent solution: increased decentralization.

6.3 TRQ3

The third thesis research question asked: *How well do the systems proposed align with their design and security goals?*

In order to address the selected threats, we designed a privacy-preserving process to manage each threat. Solving each threat drew up similar security concerns, which resulted in highly aligned design requirements and conceptual similarities. Most notably, both schemes support the trend towards decentralization in TEM architecture.

6.3.1 Cyclic Homomorphic Encryption Aggregation (CHEA)

CHEA was designed to address the concern of energy usage data privacy by facilitating its use in critical grid management functionality without compromising consumer privacy.

The main process of CHEA is:

1. The DSO specifies the parameters that adjust aggregation locality and resolution.
2. The DSO sends a signal to determine which smart meters are active and available for aggregation.
3. The DSO generates an aggregation plan that splits the active meters into regions and then randomly assigns groups of meters within each region.
4. Each group has a "leader" smart meter who generates a public and private homomorphic cryptographic key.
5. The leader encrypts their measurement with the public key.
6. The leader then sends their encrypted measurement and public key to the next group member.
7. The recipient encrypts their measurement using the leader's public key and homomorphically adds their encrypted measurement to the leader's.
8. The recipient then forwards this encrypted aggregate to the next member of the group.
9. The next member repeats this process, encrypting their measurement, homomorphically adding it to the running total, and forwarding the aggregate and leader's public key to the next member.
10. Eventually the next member will be the leader, who receives the total encrypted aggregate.
11. The leader decrypts the aggregate and sends it to the DSO.

The key insight of CHEA is that homomorphic encryption enables privacy-preserving aggregation without a centralized aggregating node. As long as a cycle is above a critical threshold of three members, no party at any point in the process will be able to infer the

measurements of any other. Since aggregated data is implicitly safe it can be sent secured or unsecured to the DSO, although encryption is still recommended to thwart statistical attacks. Though they are unlikely, encryption is a cheap countermeasure.

- Energy usage data privacy
  - Energy usage data is encrypted within a group of smart meters using a new public key each round. The data is aggregated by passing the encrypted measurements between members of the group who sequentially homomorphically add their measurement to the total. This total is returned to the key generating meter who decrypts the total and sends it to the DSO. This way, no party ever receives a plaintext, unaggregated reading and privacy is preserved throughout the process.
- Market privacy
  - Market privacy is enhanced by protecting energy usage data. The scheme prevents deanonymization attacks by keeping the necessary data hidden at every stage.
- Trust
  - Edge nodes
    - The concern of trusted edge nodes is avoided entirely by structuring the scheme such that all processes are handled by the smart meters.
  - Authentication
    - Creating an environment where it is impractical for collusion to occur even if an insider at the DSO is involved reduces the risk of privacy attacks. This reduces the load on authentication security.
  - Single points of failure
    - No single points of trust failure are present in the scheme. Key generation is handled in-group, removing the need for a KGC. Group topology is randomized each round, preventing any particular meter from being in a uniquely informed or trusted position. As noted, even the DSO would not have a single collusion target due to the randomization.
- Infrastructure & points of failure

- Edge nodes
  - The non-presence of edge nodes in the scheme means that there is no risk of a region going dark due to a single component failure.
- Single point of failure
  - No single points of access failure are present in the scheme. A KGC is not required, so failure to generate and distribute keys for the entire process is not a threat. While a meter may fail during a cycle and cause its group to fail to report, dynamic group generation guarantees that each meter is responsive at the beginning of a cycle. Thus, a meter that goes offline during one cycle will not be included in the next, meaning it cannot become a SPOF.
- Communication
  - Communication security is not taken for granted; data is encrypted and/or aggregated at every stage. Since there is no single point of trust, there is no risk of key leakage whether through hacking or collusion.

6.3.2 Individually Linkable Pseudonymous Trading Scheme (ILPTS)

ILPTS is designed to enable total anonymity in the market application (i.e. not even the DSO has a database of real identities), while enabling an automated deanonymization policy that cannot be manually triggered by any party. This was accomplished by leveraging distributed systems, secret sharing, homomorphic encryption, and smart contracts.

ILPTS consists of two main processes: registration and deanonymization.

Registration

1. A batch of users request to register.
2. The registration smart contract requests a set of fog nodes.
3. The fog nodes send their public keys to the users via the contract.
4. Each user splits their ID into parts using a secret sharing scheme and encrypts each part with a different public key.
5. Each user creates a list containing the encrypted parts and their pseudonym.

6. These lists are cryptographically shuffled among the batch before being sent in plaintext form (the ID parts inside are still encrypted) to the contract.

7. The contract validates the encrypted data by homomorphically summing the parts associated with each fog node, sending each sum to the appropriate fog node for decryption, and verifying equality between the sum of decrypted sums and the sum of the batch's real IDs.

8. If validation is successful, the contract stores the lists on the recovery blockchain and activates the batch's pseudonyms on the market.

Deanonymization

1. A pseudonymous user's reputation falls below the threshold.

2. The monitoring smart contract sends a deanonymization request containing the pseudonym to a valid subset of fog nodes.

3. Each fog node requests its associated ID part from the pseudonym's record on the recovery blockchain.

4. Each fog node decrypts its associated ID part and transmits the result to the DSO.

5. The DSO can reconstruct the real ID by reversing the secret sharing scheme.

ILPTS addresses each of the security goals as follows:

- Market privacy
  - Market privacy is achieved by creating a system in which no party, authoritative or not, has access to the true identities of the pseudonymous market users. Prosumers are able to trade energy under the assumption of full anonymity with regards to trades and preferences. The deanonymization policy cannot be manually activated, preventing the creation of novel attack vectors.
- Trust
  - Authentication

    * ILPTS proposes a novel registration policy that operates in a distributed manner and is fully automatically coordinated with no trusted authorities (including the DSO). The validation process is also completely decentralized, automated, privacy-preserving, and performed on-chain. This contributes to the market anonymity being total, since no central authority has to know the information at any point.
  - Single points of failure
    * The use of cryptographic shuffling enables batch registration to occur without any information leakage, even to members of the batch. Homomorphic encryption is utilized to enable public validation while preserving privacy. Validation would otherwise be required to be performed in secret, which would necessarily create a trusted authority and subsequent trust point of failure. This is avoided through our method for on-chain batch validation.

- Infrastructure & points of failure
  - Single point of failure
    * ILPTS avoids single points of failure by storing the pseudonym, identity pairs in a blockchain, ensuring that the deanonymization policy will be able to be executed when needed. Additionally, registration and retrieval of identities is performed with redundant edge nodes, such that failure of a particular edge node does not disable either function.
  - Communication
    * When a user's real identity is requested to be sent to the DSO, it is first transmitted in several pieces with each encrypted by a different edge node. The edge nodes receive and decrypt these parts before sending them to the DSO (still encrypted with the DSO's public key). Intercepting the ID would require an unrealistic amount of coordination, including knowing which edge nodes will receive the request (which is determined at the time of request), the capacity to capture packets along these paths, and the private key of the DSO. The degree of distribution ensures that the scheme remains secure even if communication channels are compromised.
  - Regulation & standardization
    * This solution does not directly address interoperability concerns. In terms of regulation as an external concern, ILPTS enables stronger adherence to existing legal standards surrounding utilities.

- Smart contracts
  - Smart contracts were necessary for two parts of the process. First, onboarding of market participants. A contract was required so that the hidden true identities could be validated prior to storage and user activation, to ensure that all parties were indeed sending the appropriate information at each stage, and ultimately to coordinate the registration process. Second, a smart contract is used to monitor pseudonymous user reputation and activate the deanonymization policy. This was necessary to ensure that the policy could not be activated manually; any other solution would have introduced undesirable backdoors. The use of a smart contract ensures that the policy is protected with the same level of security as the blockchain itself.

6.3.3 Synergies

This section discusses the various synergies between the two schemes and DLT-based transactive energy markets as a whole. Note that although CHEA and ILPTS can be used *individually* (as they do not depend on one another), they can also be used *together* as they do not conflict with each other, and their benefits can be combined to enhance privacy even further.

Elimination of Trust

Both schemes are constructed such that they can operate in as trustless an environment as possible. This enhances security by reducing reliance on assumptions regarding the intentions of parties in the system. It also enhances privacy because in most cases this requires that private information is never accessible to a third party in intelligible form.

Homomorphic Encryption

Homomorphic encryption is a kind of cryptosystem where there are homomorphisms between operations on the plaintext and cyphertext [25]. This means that one could, for example, perform an operation on two encrypted numbers, and the output would be an encrypted message containing the result of the homomorphic operation on the plaintext.

This concept is more central to CHEA, but it appears in both schemes as it is a useful tool for operating on data in a low trust environment.

Increased Decentralization

Both schemes perform their functions in a more decentralized manner than preceding schemes. This strategy emerged naturally as a method of reducing points of failure (both in terms of trust and access), inhibiting collusion, and increasing reliability.

Privacy Enhancing

Both schemes contribute to enhanced privacy in the transactive energy environment, particularly where there had been a lack of attention prior. Although schemes for market privacy existed, they were either insufficient or not feasible to be implemented due to the trade off mentioned in TRQ2.

Customers will be more likely to adopt a system with heightened safety guarantees, so these schemes could contribute to a wider spread of TEM usage if implemented.

Inhibition of Coordinated Attacks Between Layers

Each scheme, by enhancing its target layer, naturally contributes to the inhibition of attacks that require information from both layers, such as statistical deanonymization.

Robustness to Interception

Both schemes maintain message encryption and data privacy throughout all processes. This is a natural consequence of the design goal of maximal trustlessness, requiring that no party receive undue access to privileged information. These facts mean that communication security is not a requirement for security of the scheme, directly addressing a concern from TRQ1. Additionally, the stochastic nature of both schemes results in complex information flow, making man-in-the-middle and other listening-based attacks exceptionally difficult to pull off.

Reduced Infrastructure Requirements

Existing schemes in both areas typically employ network and computing infrastructure beyond the minimally essential DSO and smart meters. For example, many aggregation schemes rely on fog nodes [28, 16, 32, 40, 38]. Aggregation schemes and authentication

protocols relating to market security both often rely on trusted authorities for tasks like key generation and validation [16, 8].

In contrast, ILPTS and CHEA both significantly reduce reliance on additional architecture. Thus, the advancements proposed in this thesis contribute to lower network requirements, which can reduce costs for TE developers and lower the barrier to adoption for TEMs.

Transactive Energy Principles

Both schemes contribute to the TE principles we have referred to throughout by capturing the benefits of decentralization and distributed applications. They are uniquely suited to the TE environment, having been designed with the consideration of a public blockchain of transactions. Although smart grid research occasionally intersects with the concepts contained in this thesis, solutions designed for a smart grid environment do not have to contend with this foundational assumption that all transaction data is public.

Each scheme creates a process that is more closely aligned with the functionality of the blockchain network underlying the market mechanism. In doing so, they contribute to the ultimate goal of bringing the security level of the surrounding processes up to that of the blockchain.

6.4 TRQ4

The third thesis research question asked: *Are these proposed systems satisfactory in terms of performance and compared to competing solutions?*

6.4.1 ILPTS

The closest conceptual comparison to ILPTS is the market privacy protocol introduced in TRANSAX by Eisele et al. [8]. In this protocol, users are totally anonymous as in ILPTS. However, their original identities cannot be recovered; each user belongs to a group within which they are irrevocably anonymous. For this reason, we consider ILPTS to be the superior solution. While TRANSAX provides similar privacy guarantees, it does not have a protocol for handling malfunctioning users. Instead, it treats any reputation drop as malicious and punishes the user. As for performance, ILPTS could not be tested empirically at this time.

6.4.2 CHEA

6.4.2.1 Performance

We found CHEA's performance to be slightly better than similar schemes. It performed worse (but not by much) in the initialization phase, where it has to do the most processing. In the aggregation phase, CHEA was the second fastest, and considerably faster than its laggards. The decryption phase was similar for all but one particularly slow scheme.

These performance estimates were created using a process borrowed from Liu et al. [16]. The first step involved timing computationally intensive cryptographic operations, in this case on a 2017 MacBook Pro with a Intel Core i5 CPU at 3.1 GHz and 8 GB RAM using the Stanford Pairing-Based Cryptography library. These times were then extrapolated to simulate the time that would be taken to perform the major operations (namely, initialization, aggregation, and decryption) using each of the compared protocols. This process was done for CHEA and the schemes we compared it to, ensuring accurate comparison.

6.4.2.2 Features & Assumptions

More importantly than raw performance (which is similar among most schemes), CHEA manages to make many fewer assumptions about the system architecture than related schemes.

Table 6.1 presents the assumptions made by each of the relevant schemes (no is better).

Table 6.1: Comparison of related schemes (A1: fixed aggregation regions; A2: fog nodes; A3: trusted authority; A4: SM perform encryption; A5: SM perform key generation; A6: static network topology; A7: authentication certificate/signature required; A8: static membership)

Scheme	A1	A2	A3	A4	A5	A6	A7	A8
CHEA	no	no	no	yes (Paillier)	yes	no	no	no
Liu et al. [16]	yes	yes	partial	yes (ECC)	no	yes	no	no
Song et al. [32]	yes	yes	no	yes (Paillier)	partial	yes	no	no
Khan et al. [12]	no	yes	yes	yes (Paillier)	partial	yes	yes	yes
Zuo et al. [40]	yes	yes	no	yes (ElGamal)	partial	yes	yes	yes
Zhang and Liu [38]	yes	yes	yes	yes (Paillier)	no	yes	no	no

In addition, it offers a novel feature set not rivaled by existing literature.

Notably, it offers the following features:

- Flexible aggregation subsets
- Variable data density
- Costless scaling
- Improved security (discussed in Section 6.4.2.3)
- Scalability

CHEA manages to accomplish these all while maintaining performance in the top quartile of related schemes.

While some schemes address individual concerns, such as flexible subset sizing [37] or collusion [16], none that we found managed to capture all of these features. CHEA provides the greatest flexibility by enabling variable subset sizes and, because it does not rely on aggregating nodes, it can also vary the clustering of aggregation groups. Finally, CHEA was designed for blockchain-based TEMs specifically, ensuring that intermediate data never passes through a centralized node that could otherwise cross reference the public transaction information to compromise customers.

Another feature of CHEA is its performance as network parameters change; one-shot performance does not tell the whole story. Because of CHEA's decentralized nature, it has both temporal and architectural advantages when it comes to scaling in terms of adding new users, adjusting aggregation parameters, and handling large neighborhoods.

CHEA was simulated in a variety of settings using a custom program. These results confirmed that it scales well as expected, experiencing minimal processing time increase as populations get larger. Part of this increase can also be attributed to having performed the simulation on a single machine (thus having to simulate more entities). A more robust test would involve a distributed environment with auxiliary machines to represent smart meters.

6.4.2.3 Security

One of the features of CHEA is that it allows smart meters to be grouped arbitrarily; this allows for flexible subset sizing, data density, and data resolution, but it also functions as a

security feature. The β parameter allows the DSO application to specify how distant smart meters within the same group can be. The higher the DSO application sets this parameter at aggregation time, the further apart smart meters may be grouped together. This heightens security by increasing the variety of group participants, thus making interception of data more difficult and statistical attacks even more unlikely to succeed (or be carried out in the first place). By enabling this parameter, CHEA allows the DSO application to enable the maximum level of security that is possible for a given application – as opposed to being locked into a particular arrangement.

CHEA also improves security by using a distributed model which does not make use of fog nodes or centralized key generation. These avoidances mean that there are fewer opportunities for collusion under CHEA, and fewer "insiders" who could seek to leak information. In addition, the more complicated information transfer associated with CHEA (compared to topologically static schemes) makes eavesdropping on communication significantly less feasible. All of these factors also contribute to the reduction in practicality of probablistic or statistical attacks against a TEM using CHEA.

Once again, CHEA achieves these security gains while maintaining competitive performance.

6.5 Limitations

6.5.1 Review

- Only one source database (SCOPUS) was used to discover initial literature, which can introduce bias into the dataset.
- Single author performed data collection and analysis; this causes inherent limitations, as well as potential author bias in analysis.

6.5.2 CHEA

- Unable to test / validate on real hardware.
- Performed simulations on single processor (may try to address this in the future using virtualization).
  - Calls into question the source of some data trends in the results.

In practice, one of the main limitations holding CHEA back from widespread implementation is the computational demands it makes of smart meters. Although most (if not all) HE-driven aggregation schemes require smart meters to perform encryption operations [36, 38, 31, 6, 7, 39, 18], and many require them to participate in key or certificate assignment [16], it remains a limitation of the scheme. Consumers and utility companies alike will resist expensive upgrades to metering equipment (which has thus far impacted the deployment of transactive energy in general). While CHEA is an improvement to smart grid aggregation schemes with respect to decentralized transactive energy (and while it costs less to implement than other ideas, like battery filtering [11]), it retains this weakness. Either improvements to HE efficiency will be required, or supplemental schemes will have to be used in the interim until metering technology catches up to more advanced aggregation protocols.

6.5.3 ILPTS

- Did not consider implications of open membership (how disputes will be resolved, how tax burdens will be calculated / determined).
- Time and resource limitations making implementation or even a mock implementation (or simulation) infeasible to develop, and thus validate.
- Assumptions were made about the architecture of the parent TEM. These assumptions were based on the literature, but they make the scheme less universal regardless.

Limitations that may affect ILPTS in practice may include the requirement that smart meters are required to perform many calculations, similar to CHEA. In this case these calculations primarily arise during the batch registration process, particularly when splitting their IDs into parts using Shamir's secret sharing scheme [29], encrypting those parts homomorphically [25, 26], and when performing the cryptographic shuffling to disguise their pseudonymous IDs from other batch members [5].

Other challenges presented to implementing ILPTS concern the role of fog nodes in the scheme. Fog nodes must be trusted to be "honest but curious" [32], meaning that while they are not "trusted parties" (as they may maliciously try to gather information), they are trusted to provide honest results of calculations and store data correctly. Additionally, fog nodes must be reliably accessible above a certain threshold, and must be capable of maintaining storage of important data long-term.

Another complication is consumer trust in a system in which their personally identifying information is stored in an ostensibly public manner, although protected via encryption. Most consumers are used to systems in which their data is stored on a private server, although as cryptocurrency and related technologies become more mainstream this sentiment may change. Finally, utilities and distributed system operators may not be comfortable giving up centralized control over storage of identifying information. While TRANSAX [8] and other schemes propose this option, it is not clear whether TEM operators would be interested in such solutions.

6.6 Implementation

6.6.1 CHEA

In order to implement CHEA, smart meters would have to be sufficiently powerful to:

- Generate and store cryptographic keys.
- Perform encryption and decryption of data.
- Send and receive encrypted data & keys.

The network architecture would have to support communication between smart meters, would need to be high enough bandwidth to support transmission of encrypted data (which is inherently larger than plaintext data), and would need to be fast enough to enable rapid aggregation in order to support all aggregation purposes. Realistically, these factors will not be a bottleneck; it is more likely that smart meter capabilities would be a performance bottleneck.

Blueprint:

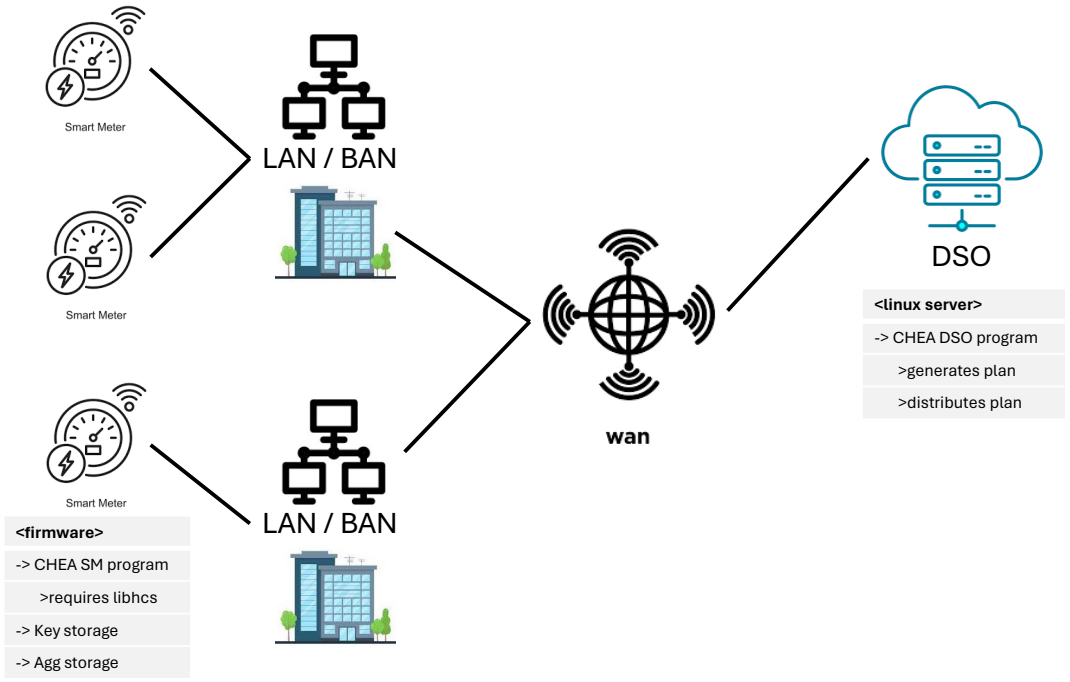

Figure 6.1: Software and network blueprint for implementing CHEA protocol within a TEM. Programs are specified in the algorithms in Chapter 4.

Figure 6.1 shows that under this typical network architecture all components have paths for communication; that is, any smart meter can communicate with any other, and the DSO can delegate to all other components. Note that although the diagram appears hierarchical, this does not represent the aggregation scheme, only the network topology. This is an example (and a typical one) but, as long as the network is a spanning tree, any topology will be compatible with CHEA.

Smart meters are shown with their associated software and storage components; this

will assist with hardware specification when real-world implementation is desired. Exact values for storage and data transfer requirements can be found in Chapter 4.

6.6.2 ILPTS

This section outlines the network components and suggested software requirements for a real-world implementation of ILPTS. First, the required components for the scheme:

- Recovery Blockchain
  - Supported by DSO and fog nodes.
- DSO
- Fog nodes
- Smart contracts
  - SC for coordinating batch registration.
  - SC to monitor reputations and initiate deanonymization protocol when appropriate.
- Smart meters

Blueprint:

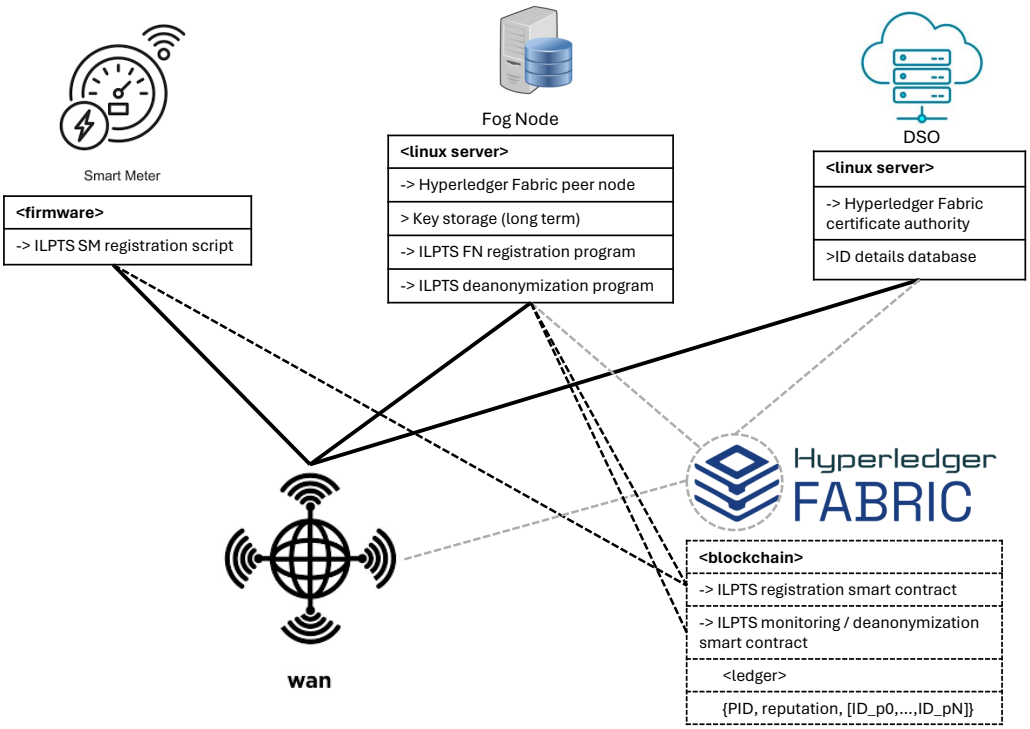

Figure 6.2: Software and network blueprint of ILPTS that is intended to assist with implementation. Scripts and programs are specified in the algorithms in Chapter 5.

Figure 6.2 displays the network architecture in ILPTS. The DSO and fog nodes collaborate to support the distributed Hyperledger Fabric (in this case) blockchain over a wide-area network. This blockchain in turn supports the two smart contracts that coordinate the two main phases of ILPTS - registration and deanonymization.

Smart meters are required to have the ILPTS registration script which enables them to perform their duties during registration; that is, splitting their real IDs using Shamir's secret sharing scheme [29], encrypting these parts homomorphically using the keys provided by fog nodes, generating their own cryptographic keys to support the cryptographic

shuffling process [5], as well as encrypting the lists of user submissions to further support the shuffling.

Fog nodes, similarly, are required to have the fog node-specific registration program, to facilitate their role in the process. As well, they are required to have the deanonymization program: this enables them to execute the deanonymization process when called upon by the monitoring smart contract. Finally, they are required to store their homomorphic public and private keys long term, to ensure that deanonymization remains possible (as well as registration and validation of future users).

Bibliography

[1] Carlos Barreto, Taha Eghtesad, Scott Eisele, Aron Laszka, Abhishek Dubey, and Xenofon Koutsoukos. Cyber-attacks and mitigation in blockchain based transactive energy systems. In *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, volume 1, pages 129–136, 2020. doi:10.1109/ICPS48405.2020.9274708.

[2] Bitcoin Wiki. Script, 2021. `https://en.bitcoin.it/wiki/Script`.

[3] Adam C. Bonin. Protecting protection: First and fifth amendment challenges to cryptography regulation. *University of Chicago Legal Forum*, 1996(1):495–517, 1996.

[4] Alex R. Borden, Daniel K. Molzahn, Parmeswaran Ramanathan, and Bernard C. Lesieutre. Confidentiality-preserving optimal power flow for cloud computing. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1300–1307, 2012. doi:10.1109/Allerton.2012.6483368.

[5] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM*, 24(2):84–90, feb 1981. ISSN 0001-0782. doi:10.1145/358549.358563.

[6] Siguang Chen, Li Yang, Yanhang Shi, and Qian Wang. Blockchain-enabled secure and privacy-preserving data aggregation for fog-based its. *Computers, Materials & Continua*, 75(2):3781–3796, 2023. ISSN 1546-2226. doi:10.32604/cmc.2023.036437.

[7] Yuwen Chen, Shisong Yang, José-Fernán Martínez-Ortega, Lourdes López, and Zhen Yang. A resilient group-based multisubset data aggregation scheme for smart grid. *IEEE Internet of Things Journal*, 10(15):13649–13661, 2023. doi:10.1109/JIOT.2023.3262731.

[8] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[9] Nawfal Fadhel, Federico Lombardi, Leonardo Aniello, Andrea Margheri, and Vladimiro Sassone. Towards a semantic modelling for threat analysis of iot applications: A case study on transactive energy. In *Living in the Internet of Things (IoT 2019)*, pages 1–6, 2019. doi:10.1049/cp.2019.0147.

[10] M. Jayachandran, K. Prasada Rao, Ranjith Kumar Gatla, C. Kalaivani, C. Kalaiarasy, and C. Logasabarirajan. Operational concerns and solutions in smart electricity distribution systems. *Utilities Policy*, 74:101329, 2022. ISSN 0957-1787. doi:10.1016/j.jup.2021.101329.

[11] Georgios Kalogridis, Rafael Cepeda, Stojan Z. Denic, Tim Lewis, and Costas Efthymiou. Elecprivacy: Evaluating the privacy protection of electricity management algorithms. *IEEE Transactions on Smart Grid*, 2(4):750–758, 2011. doi:10.1109/TSG.2011.2160975.

[12] Hayat Mohammad Khan, Abid Khan, Farhana Jabeen, Adeel Anjum, and Gwanggil Jeon. Fog-enabled secure multiparty computation based aggregation scheme in smart grid. *Computers & Electrical Engineering*, 94:107358, 2021. ISSN 0045-7906. doi:10.1016/j.compeleceng.2021.107358.

[13] Desen Kirli, Benoit Couraud, Valentin Robu, Marcelo Salgado-Bravo, Sonam Norbu, Merlinda Andoni, Ioannis Antonopoulos, Matias Negrete-Pincetic, David Flynn, and Aristides Kiprakis. Smart contracts in energy systems: A systematic review of fundamental approaches and implementations. *Renewable and Sustainable Energy Reviews*, 158:112013, 2022. ISSN 1364-0321. doi:10.1016/j.rser.2021.112013.

[14] Donghe Li, Qingyu Yang, Dou An, Wei Yu, Xinyu Yang, and Xinwen Fu. On location privacy-preserving online double auction for electric vehicles in microgrids. *IEEE Internet of Things Journal*, 6(4):5902–5915, 2019. doi:10.1109/JIOT.2018.2872444.

[15] Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker. Inferring personal information from demand-response systems. *IEEE Security & Privacy*, 8(1):11–20, 2010. doi:10.1109/MSP.2010.40.

[16] Shuanggen Liu, Yaowei Liu, Wandi Liu, and Yuchen Zhang. A certificateless multidimensional data aggregation scheme for smart grid. *Journal of Systems Architecture*, 140:102890, 2023. ISSN 1383-7621. doi:10.1016/j.sysarc.2023.102890.

[17] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security & Privacy*, 7(3):75–77, 2009. doi:10.1109/MSP.2009.76.

[18] Yang Ming, Yabin Li, Yi Zhao, Pengfei Yang, and Yu Yao. Efficient privacy-preserving data aggregation scheme with fault tolerance in smart grid. *Sec. and Commun. Netw.*, 2022, jan 2022. ISSN 1939-0114. doi:10.1155/2022/5895176.

[19] Eric Münsing, Jonathan Mather, and Scott Moura. Blockchains for decentralized optimization of energy resources in microgrid networks. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 2164–2171, 08 2017. doi:10.1109/CCTA.2017.8062773.

[20] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain: A path to grid modernization and cyber resiliency. In *2017 North American Power Symposium (NAPS)*, pages 1–5, 09 2017. doi:10.1109/NAPS.2017.8107313.

[21] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23, 09 2017. doi:10.1109/RWEEK.2017.8088642.

[22] Masoumeh Nazari, Siavash Khorsandi, and Jaber Babaki. Security and privacy smart contract architecture for energy trading based on blockchains. In *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, pages 596–600, 2021. doi:10.1109/ICEE52715.2021.9544155.

[23] Ivica Nikolić, Aashish Kolluri, Ilya Sergey, Prateek Saxena, and Aquinas Hobor. Finding the greedy, prodigal, and suicidal contracts at scale. In *34th Annual Computer Security Applications Conference*, ACSAC '18, page 653–663, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450365697. doi:10.1145/3274694.3274743.

[24] Adeiza J. Onumanyi, Sherrin J. Isaac, Carel P. Kruger, and Adnan M. Abu-Mahfouz. Transactive energy: State-of-the-art in control strategies, architectures, and simulators. *IEEE Access*, 9:131552–131573, 2021. ISSN 2169-3536. doi:10.1109/ACCESS.2021.3115154.

[25] Pascal Paillier. Public-key cryptosystems based on composite degree residuosity classes. In Jacques Stern, editor, *Advances in Cryptology — EUROCRYPT '99*, pages 223–238, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg. ISBN 978-3-540-48910-8.

[26] Pascal Paillier and David Pointcheval. Efficient public-key cryptosystems provably secure against active adversaries. In Kwok-Yan Lam, Eiji Okamoto, and Chaoping Xing, editors, *Advances in Cryptology - ASIACRYPT'99*, pages 165–179, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg. ISBN 978-3-540-48000-6.

[27] Shammya Saha, Nikhil Ravi, Kári Hreinsson, Jaejong Baek, Anna Scaglione, and Nathan G. Johnson. A secure distributed ledger for transactive energy: The electron volt exchange (eve) blockchain. *Applied Energy*, 282:116208, 2021. ISSN 0306-2619. doi:10.1016/j.apenergy.2020.116208.

[28] Shammya Shananda Saha, Christopher Gorog, Adam Moser, Anna Scaglione, and Nathan G. Johnson. Integrating hardware security into a blockchain-based transactive energy platform. In *2020 52nd North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2021. doi:10.1109/NAPS50074.2021.9449802.

[29] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, nov 1979. ISSN 0001-0782. doi:10.1145/359168.359176.

[30] Khaled Shuaib, Juhar Abdella, Farag Sallabi, and Mohammed Abdel Hafez. Using blockchains to secure distributed energy exchange. In *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 622–627, 04 2018. doi:10.1109/CoDIT.2018.8394815.

[31] Ashutosh Kumar Singh and Jatinder Kumar. A privacy-preserving multidimensional data aggregation scheme with secure query processing for smart grid. *The Journal of Supercomputing*, 79(4):3750–3770, Mar 2023. ISSN 1573-0484. doi:10.1007/s11227-022-04794-9.

[32] Jingcheng Song, Yining Liu, Jun Shao, and Chunming Tang. A dynamic membership data aggregation (dmda) protocol for smart grid. *IEEE Systems Journal*, 14(1):900–908, March 2020. ISSN 1937-9234. doi:10.1109/JSYST.2019.2912415.

[33] Ognjen Vuković, György Dán, and Rakesh B. Bobba. Confidentiality-preserving obfuscation for cloud-based power system contingency analysis. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 432–437, 2013. doi:10.1109/SmartGridComm.2013.6687996.

[34] Ping Yi, Ting Zhu, Qingquan Zhang, Yue Wu, and Jianhua Li. A denial of service attack in advanced metering infrastructure network. In *2014 IEEE International Conference on Communications (ICC)*, pages 1029–1034, 2014. doi:10.1109/ICC.2014.6883456.

[35] Ishtiaque Zaman and Miao He. A multilayered semi-permissioned blockchain based platform for peer to peer energy trading. In *2021 IEEE Green Technologies Conference (GreenTech)*, pages 279–285, 2021. doi:10.1109/GreenTech48523.2021.00052.

[36] Jianhong Zhang and Chenghe Dong. Privacy-preserving data aggregation scheme against deletion and tampering attacks from aggregators. *Journal of King Saud University - Computer and Information Sciences*, 35(4):100–111, 2023. ISSN 1319-1578. doi:10.1016/j.jksuci.2023.03.002.

[37] Jianhong Zhang and Chenghe Dong. Privacy-preserving data aggregation scheme against deletion and tampering attacks from aggregators. *Journal of King Saud University - Computer and Information Sciences*, 35(4):100–111, 2023. ISSN 1319-1578. doi:10.1016/j.jksuci.2023.03.002.

[38] Liying Zhang and Yining Liu. Fsda: Flexible subset data aggregation for smart grid. *IEEE Systems Journal*, 17(1):569–578, 2023. doi:10.1109/JSYST.2022.3199386.

[39] Xiaojun Zhang, Wei Tang, Dawu Gu, Yuan Zhang, Jingting Xue, and Xin Wang. Lightweight multidimensional encrypted data aggregation scheme with fault tolerance for fog-assisted smart grids. *IEEE Systems Journal*, 16(4):6647–6657, Dec 2022. ISSN 1937-9234. doi:10.1109/JSYST.2022.3146504.

[40] Xiangjian Zuo, Lixiang Li, Haipeng Peng, Shoushan Luo, and Yixian Yang. Privacy-preserving multidimensional data aggregation scheme without trusted authority in smart grid. *IEEE Systems Journal*, 15(1):395–406, March 2021. ISSN 1937-9234. doi:10.1109/JSYST.2020.2994363.

Chapter 7

Conclusion

Transactive energy markets promise to revolutionize power delivery and management by enabling consumers to trade energy directly, a feature that will only continue to become more useful as people invest in renewable energy sources and home batteries [4]. The infrastructure that makes this possible, however, comes with the cost of increased attack surface for hackers, higher bandwidth information that enables novel attacks, and more personal data requirements that present privacy risks.

Blockchain technology has become the de facto standard on which to build the market applications due to its high security guarantees and distributed nature, which aligns with transactive energy [2, 3]. While blockchain guarantees market security, though, the rest of the system remains vulnerable.

The goal of this thesis was to improve the security and privacy stature of transactive energy markets built on blockchain by identifying and proposing solutions to high priority vulnerabilities in the domain.

We successfully captured the field of security concerns in the target domain, as discussed in research question TRQ1. We graded them on priority based on the severity of consequences of an attack, the ease of execution, and the presence or lack of solutions to the concern. From there, we isolated and analyzed two concerns as discussed in TRQ2. We used these analyses to create two processes designed to address the web of security and privacy concerns associated with them; these are described in TRQ3. Finally, we analyzed the processes from a performance and comparative perspective, as outlined in TRQ4.

Overall, we succeeded at creating innovative and functional solutions to relevant security gaps in the transactive energy literature, thus improving the security and privacy stature of

such markets. The publication of these solutions, as well as a literature review identifying and analyzing security concerns in the field, are our contributions to this field of scientific study.

7.1 Thesis Research Question TRQ1

This research question is answered in the first paper of this thesis, *A Review of Cybersecurity Concerns for Transactive Energy Markets* [5] (Chapter 3). Note that the research questions in the paper directly correlate to the subheadings motivated by TRQ1: discovery of TE vulnerabilities, assembly of proposed solutions to said vulnerabilities, and an analysis of the resulting gap in security coverage.

In Chapter 3, we analyzed 28 papers and discovered 14 major categories of security concerns. Within these categories we explored the specific kinds of attacks raised by the TEM designers, and solutions that have been proposed to address them.

We found that there was a gradient of severity among the threat categories discovered, as well as significant differences in the quantity and quality of proposed solutions.

In brief, some classes of attacks, such as false data injection (Section 3.5.1) or denial-of-service (Section 3.5.2) were found to be well-handled by existing measures in proposed designs. Others, such as direct attacks on communication infrastructure (Section 3.5.14) or smart meter firmware (Section 3.5.10), were determined to be outside of the scope of consideration by transactive energy designers.

The remaining threats were considered severe, insufficiently addressed in TEM designs, and relevant to TEM design. These threats were smart contract vulnerabilities, edge or fog nodes, single point of failure, personal data privacy, and energy usage data privacy.

There is a common theme among these threats: many of them can be addressed or mitigated by increased decentralization. This is an unsurprising result, as transactive energy, microgrids, and blockchain technology were all developed with the intent of harnessing the benefits of decentralized operations.

7.2 Thesis Research Question TRQ2

For TRQ2, we selected the following threats:

- Energy usage data privacy
- Market privacy

In order to analyze the threats, they had to be considered in context:

Energy Usage Data Privacy

Energy usage data is typically required by the DSO to perform a variety of grid management functions, including demand response, load balancing, state estimation, billing, etc. Leakage of this data can compromise consumer privacy and facilitate physical threats, and thus protecting it is critical.

There are several methods for protecting data while enabling its analysis. These include data aggregation, adding random noise to data, and algebraic transformation. Battery filtering is a method of obfuscating demand data that is specific to this domain. A more detailed analysis of these candidate techniques can be found in Section 4.2.

Data aggregation was selected strategy due to its properties aligning well with the use case, namely: timely data reporting, preservation of exact values, and hardware independence.

Upon examination of existing data aggregation policies in related domains (Section 6.2.1), we found that they encounter threats associated with: edge nodes, trust, single points of failure, and communication security.

Generally, their architecture relied on edge nodes, which pose an access risk and, in many cases, a trust risk as well. Many schemes used trusted authorities such as key generating centers.

Our design goal, then, was to design an aggregation scheme that could provide timely updates to the DSO while mitigating these factors to the highest degree possible.

Market Privacy

Market privacy concerns the concealment of a consumer's identity in the trading application. This concern has been considered by most schemes, but existing solutions present problems of their own. Critically, a balance between anonymity and the ability to regulate consumer behavior must be struck.

In examining existing ideas and potential paths for improvement (Section 6.2.2), we found that addressing this issue would require consideration of: authentication, trust, single points of failure, communication security, and regulations.

Every market privacy solution we discovered fell short in some regard, either providing full anonymity with no path for recourse or sacrificing the security of consumer identities to facilitate corrective action.

The challenge for this solution was to allow consumers to trade under pseudonyms, which were totally anonymous, while somehow maintaining a link between said pseudonym and their true identity in case of emergency. The solution must address this challenge while avoiding pitfalls that would lead to the threats outlined earlier.

7.3 Thesis Research Question TRQ3

Answering this research question entailed designing processes for transactive energy markets that met the high security standards set out in TRQ2, as well as adhering to the design goals as closely as possible. The results, CHEA and ILPTS, successfully enhance privacy in each of their respective areas and achieve the heightened security expectations. The schemes can be used independently, as they have no dependencies, but they can also be used *in tandem*, as they do not conflict. Furthermore, security and privacy are enhanced to a greater degree when they are deployed together, as their security properties reinforce each other while targeting different system layers.

CHEA for Energy Usage Data Privacy

Our proposal for energy usage data aggregation is *Cyclic Homomorphic Encryption Aggregation (CHEA)* [6], introduced in Chapter 4.

CHEA creates groups of smart meters from which data is to be aggregated. Within each of these groups a cyclic aggregation occurs, in which each smart meter from the group sequentially adds its measurement to the running total. A full description of the process can be found in Section 4.4. This process is facilitated by homomorphic encryption, which enables the process to occur while the data is encrypted, preserving privacy along the way.

Notably, this process does not involve edge nodes. Cryptographic keys are generated by the leader of each cycle, obviating the need for key generating centers as well. Each cycle of CHEA involves the dynamic generation of groups and new keys, thwarting collusion and ensuring backward and forward secrecy.

CHEA circumvents issues of trust, edge nodes, and single points of failure by avoiding centralized aggregation nodes entirely. The sporadic flow of information also makes it more robust to compromised communication infrastructure. Overall, CHEA successfully meets the design requirements and even introduces novel functionality in terms of subset and region flexibility.

ILTPS for Market Privacy

Our proposal for market privacy is the Individually Linkable Pseudonymous Trading Scheme (ILPTS), introduced in Chapter 5.

ILPTS consists of two main processes: registration and deanonymization.

During the registration process, a batch of users break their real identities into parts and encrypt each part with a different edge node's public key. These encrypted parts are then associated with their pseudonyms and shuffled amongst the batch using a specialized protocol, before being stored on a blockchain adjacent to the market chain. This process is coordinated by a smart contract that also validates the data on chain.

The deanonymization policy is activated by a different smart contract that monitors each pseudonymous user's reputation. Once triggered, a request is sent to a critical subset of edge nodes, which request the ID part associated with their node and the pseudonym. These nodes decrypt the partial ID before sending it to the DSO, which is now free to reconstruct the real identity of the offending user. Both policies are fully specified in Section 5.5.

ILPTS successfully creates a system within which users are totally anonymous to every member of the overall transactive energy market architecture, but which retains an association between their market identity and real identity. It avoids issues of trust and single points of failure by employing distributed storage and specially coordinated protocols to circumvent the need for trusted authorities. The deanonymization policy does not become a threat vector because it is not possible to trigger manually and, being on chain, is secured by the same cryptographic assurance as the blockchain itself.

7.4 Thesis Research Question TRQ4

CHEA was tested against similar schemes using mathematical analysis; the results of cryptographic benchmarks were extrapolated based on the number of operations used in each

scheme. Using this method, CHEA was found to perform well compared to similar schemes. In the aggregation phase, it was second fastest. It was slowest by a small margin in the initialization phase, and all schemes were similar in the decryption phase. As the aggregation phase is the most time consuming, we consider this a positive result.

Regarding ILPTS, the scheme has not yet been implemented, so its performance was not tested.

As far as we know, ILPTS is the only scheme of its kind to provide the combination of features it promises. The most similar scheme is the market privacy solution in TRANSAX [1], which does not enable revocable privacy. For this reason, we consider ILPTS superior.

Similarly, CHEA promises a unique cross section of features. While other schemes have approached subsets of the problem it sets out to solve, such as operating without fog nodes, inhibiting collusion, and flexible subset sizing, none that we found shared all of its benefits. CHEA was also simulated in a custom testbed to assess its ability to scale. It was found to perform well, showing a sublinear time increase (small coefficient) relative to population size. Other schemes had higher coefficients, indicating that increases in population size will have a greater effect.

Both schemes make strides in enhancing transactive energy market privacy. They were each created with DLT-based markets as a target, and thus have unique features to handle certain attack surfaces that do not exist in other areas, such as the public transaction ledger. The schemes presented in this thesis help to further the trend toward decentralization that was discovered during the literature review. They contribute to the overall goal of bringing the security stature of the non-market TEM processes up to that of the blockchain.

7.5 Closing Thoughts and Future Work

Overall, this thesis successfully achieves the goals set for it. It presents significant enhancements to the security and privacy of transactive energy markets. It also provides a comprehensive analysis and compendium of threats and solutions in the domain, a useful reference for researchers and practitioners alike.

Regardless, there are many remaining research opportunities that emerge from this work. For one, only two threats and associated processes of concern were addressed. Many more continue to exist that either lack solutions altogether or lack quality solutions. These threats, contained in the published literature review, provide a concise jumping off point for future security researchers in the transactive energy field.

Of particular note is the single point of failure threat. This threat is commonly addressed via decentralization and distribution. While these concepts are central to the philosophies of transactive energy and microgrids, many processes surrounding them remain centralized. Researchers can develop unconscious biases that cause them to overlook certain possibilities; there remain many opportunities to further decentralize and improve process design in these systems, as we discovered.

There are also possibilities to improve upon the design of the schemes introduced in this thesis. While we believe CHEA is a large step forward in decentralized management of aggregation, some components of the scheme remain underdeveloped. For example, the method of segmenting an area into aggregation regions could be more sophisticated, perhaps taking into account network topology, user preferences, energy sources, or functional requirements.

An even more transformative idea could be to eliminate the DSO from an active role altogether. Instead of generating the plan, the DSO would send a request to the smart meters which would then self-organize into aggregation groups and proceed with the operation. This would have the benefit of removing the DSO from any knowledge of the plan, making the already difficult task of colluding under CHEA practically impossible and eliminating an entire threat vector. This approach could also improve fault tolerance and reliability because smart meters would have to be active much closer to the time of aggregation to be included.

There are opportunities for improvements in ILPTS as well. While the scheme was designed with fog nodes in mind, the desire to reduce their storage load led to a scheme that could potentially be adapted to a fully fog node-free architecture like CHEA. In this setup, a cluster of smart meters would be drawn from during the registration and deanonymization phases to stand in for the role currently held by fog nodes. This setup would enhance decentralization, reduce trust requirements, and eliminate the need for auxiliary nodes. Such a system would also require a high amount of redundancy and would need the meters to store their ILPTS keys indefinitely. Working out the details of such a scheme could present an interesting and concrete research direction.

Transactive energy will be an important industry going forward, especially as climate change becomes more urgent. We hope to contribute to the adoption of these systems through improved privacy and security, and to inspire continued research in the field.

Bibliography

[1] Scott Eisele, Taha Eghtesad, Keegan Campanelli, Prakhar Agrawal, Aron Laszka, and Abhishek Dubey. Safe and private forward-trading platform for transactive microgrids. *ACM Trans. Cyber-Phys. Syst.*, 5(1), dec 2021. ISSN 2378-962X. doi:10.1145/3403711.

[2] Mohsen Khorasany, Yateendra Mishra, and Gerard Ledwich. A decentralised bilateral energy trading system for peer-to-peer electricity markets. *IEEE Transactions on Industrial Electronics*, 06 2019. doi:10.1109/TIE.2019.2931229.

[3] Desen Kirli, Benoit Couraud, Valentin Robu, Marcelo Salgado-Bravo, Sonam Norbu, Merlinda Andoni, Ioannis Antonopoulos, Matias Negrete-Pincetic, David Flynn, and Aristides Kiprakis. Smart contracts in energy systems: A systematic review of fundamental approaches and implementations. *Renewable and Sustainable Energy Reviews*, 158:112013, 2022. ISSN 1364-0321. doi:10.1016/j.rser.2021.112013.

[4] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23, 09 2017. doi:10.1109/RWEEK.2017.8088642.

[5] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, and John Mylopoulos. A review of cybersecurity concerns for transactive energy markets. *Energies*, 16(13), 2023. ISSN 1996-1073. doi:10.3390/en16134838.

[6] Daniel Sousa-Dias, Daniel Amyot, Ashkan Rahimi-Kian, Masoud Bashari, and John Mylopoulos. Cyclic homomorphic encryption aggregation (chea)—a novel approach to data aggregation in the smart grid. *Energies*, 17(4), 2024. ISSN 1996-1073. doi:10.3390/en17040878.