JOSIP JURAJ STROSSMAYER UNIVERSITY OF OSIJEK

FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY OSIJEK

Miljenko Švarcmajer

Optimization Strategies for Enhancing Robustness and Security in Distributed Systems

Doctoral Dissertation

Osijek, 2026.

This doctoral thesis was prepared at the Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek.

Supervisor: Assoc. Prof. Mirko Köhler, Ph.D., Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek

Co-supervisor: Prof. Ivica Lukić, Ph.D., Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek

Doctoral thesis number: [To be assigned].

Committee for the assessment of the doctoral thesis:

1. Prof. Krešimir Nenadić, Ph.D., FERIT, J. J. Strossmayer University of Osijek
2. Assoc. Prof. Zdravko Krpić, Ph.D., FERIT, J. J. Strossmayer University of Osijek
3. Prof. Kristijan Lenac, Ph.D., RITEH , University of Rijeka

Date of the defence of the doctoral thesis: [To be assigned].

Contents

Acknowledgements

No dissertation is written alone. Although my name appears on the cover, many people helped me reach this point and deserve my gratitude. Therefore, before turning to the science itself, I would like to dedicate a few words to them.

First and foremost, I would like to thank my supervisor, Associate Professor Mirko Köhler, Ph.D., and my co-supervisor, Professor Ivica Lukić, Ph.D. You introduced me to scientific research and blockchain and helped me realize that I was capable of contributing to this field. Looking back, I believe we achieved much together over the past several years.

This path would never have begun without Professor Željko Hederić, Ph.D., who encouraged me to pursue a scientific career, and Professor Krešimir Nenadić, Ph.D., who helped me find my place within the broad field of computer science.

Before thanking all of my colleagues, I would like to single out one person who has been much more than a colleague. Robert, thank you for twenty years of friendship, countless memes, reels, inside jokes understood only by the two of us, as well as for your advice and motivational speeches whenever things were not exactly smooth. Also, while listening about your research, there were moments when it sounded neither like Croatian nor as if it had been written in the Latin alphabet. Yet those conversations always reminded me how much there still is to learn. More importantly, thank you for being a constant presence throughout this journey.

To all of my colleagues, thank you for sharing your knowledge, encouragement, and occasionally a healthy dose of impostor syndrome acquired through discussions about machine learning. Nevertheless, thanks to all of you, I managed to figure out the basics, I think.

I would also like to thank the Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, under whose auspices this dissertation was conducted and which provided me with the opportunity to pursue an academic career.

A doctorate is more than science. It is also a journey filled with unexpected challenges, successes, setbacks, and experiences that leave a lasting mark.

To my mother, Anica, thank you for looking after Jakov, especially in these final months, when the demands of finishing this work were greatest, and for celebrating each of my successes, at times more loudly than I did myself. That help, arriving when I needed it most, gave me the room to see this work through.

Ana and Darko, thank you for your friendship, support, and for always being there when

it mattered most. You never understood my research, nor what a Q1 paper meant, yet you rejoiced in every milestone as though you did, and that meant more than any technical understanding could. There is a saying that friends are the family you choose for yourself. You are undoubtedly part of mine. That is why I chose you to be the godparents of my child.

Finally, I leave the last words to those who have been part of this story from the very beginning and every hour of every day.

Marina and Jakov, although you always had a spouse and a father, this doctorate occasionally took the husband and dad from you. Marina, I apologize for all the conversations in which I was physically present but mentally occupied by research problems. Jakov, I apologize for confiscating the computer in the name of science and at the expense of Minecraft.

Please know that I could not have wished for a better wife or a better son. Your names and your contribution are woven into this dissertation, into every chapter and every page, written in invisible ink.

Abstract

Optimization Strategies for Enhancing Robustness and Security in Distributed Systems

Modern distributed networks, increasingly exemplified by blockchain systems, have evolved from largely homogeneous nodes into heterogeneous environments encompassing devices with markedly different computational, energy, and connectivity profiles, ranging from wearable sensors to high-performance computing infrastructure. This growing heterogeneity introduces structural and security challenges that existing approaches address only partially. This thesis develops three original scientific contributions addressing distinct aspects of this problem, unified by a shared commitment to computational efficiency suitable for resource-constrained nodes. The first contribution (TC1) is a model for assessing the stability of distributed network topology based on node color distribution, applying greedy graph coloring algorithms (Greedy Ascending, Greedy Descending, and DSATUR) to simulated 100-node networks, where color class evenness serves as a practical, computationally inexpensive indicator of structural balance. The second contribution (TC2) repurposes the same coloring procedure for node load balancing, identifying an edge density of approximately 2000 edges as a practical operating point that balances load distribution evenness against network resource overhead. The third contribution (TC3) is a method for offline entropy harvesting from wearable device sensor data, implemented on a commercial smartwatch under two acquisition modes; motion-triggered acquisition achieved a Shannon entropy of 0.997 and a min-entropy of 0.918, outperforming passive acquisition (0.991 and 0.851, respectively), demonstrating that wearables can serve as self-contained cryptographic entropy sources without dedicated hardware. These contributions are examined primarily through the lens of blockchain networks, a representative example of heterogeneous distributed systems. Four supporting publications, addressing quantum computing threat timelines, consensus mechanism selection, smart city applications, and post-quantum migration strategies, situate the three contributions within a broader security context relevant to the continued evolution of distributed systems toward greater heterogeneity and post-quantum readiness.

Keywords: distributed systems, graph coloring, greedy algorithms, load balancing, entropy harvesting, blockchain security, robustness, optimization

Sažetak

Strategije optimizacije za povećanje robusnosti i sigurnosti u distribuiranim sustavima

Suvremene distribuirane mreže, čiji su sve češći primjer mreže ulančanih blokova, razvile su se od pretežno homogenih čvorova prema heterogenim okruženjima koja obuhvaćaju uređaje s bitno različitim računalnim, energetskim i komunikacijskim karakteristikama, od nosivih senzora do visokoučinkovite računalne infrastrukture. Ova rastuća heterogenost donosi strukturne i sigurnosne izazove koje postojeći pristupi rješavaju samo djelomično. Ova disertacija razvija tri izvorna znanstvena doprinosa koja se bave različitim aspektima ovog problema, objedinjena zajedničkom usmjerenošću na računalnu učinkovitost prikladnu za čvorove s ograničenim resursima. Prvi doprinos (TC1) je model za procjenu stabilnosti topologije distribuirane mreže temeljen na raspodjeli boja čvorova, koji primjenjuje pohlepne algoritme bojanja grafova (Greedy Ascending, Greedy Descending i DSATUR) na simuliranim mrežama od 100 čvorova, pri čemu ujednačenost klasa boja služi kao praktičan, računalno nezahtjevan pokazatelj strukturne ravnoteže. Drugi doprinos (TC2) primjenjuje isti postupak bojanja za ujednačavanje opterećenja čvorova, utvrđujući gustoću veza od približno 2000 bridova kao praktičnu radnu točku koja uravnotežuje ujednačenost raspodjele opterećenja s mrežnim troškovima resursa. Treći doprinos (TC3) je metoda za izvanmrežno prikupljanje entropije iz senzorskih podataka nosivih uređaja, implementirana na komercijalnom pametnom satu pod dva načina akvizicije; akvizicija pokrenuta pokretom postigla je Shannonovu entropiju od 0,997 i min-entropiju od 0,918, nadmašujući pasivnu akviziciju (0,991 odnosno 0,851), čime se pokazuje da nosivi uređaji mogu poslužiti kao samostalni izvori kriptografske entropije bez namjenskog hardvera. Ovi doprinosi razmatraju se prvenstveno kroz prizmu mreža ulančanih blokova mreža, reprezentativnog primjera heterogenih distribuiranih sustava. Četiri prateća rada, koja se bave vremenskim okvirom kvantne prijetnje, odabirom mehanizma konsenzusa, primjenama u pametnim gradovima i strategijama migracije prema postkvantnoj kriptografiji, smještaju tri doprinosa unutar šireg sigurnosnog konteksta relevantnog za daljnju evoluciju distribuiranih sustava prema većoj heterogenosti i postkvantnoj spremnosti.

Ključne riječi: distribuirani sustavi, bojanje grafova, pohlepni algoritmi, ujednačavanje opterećenja, prikupljanje entropije, sigurnost mreža ulančanih blokova, robusnost, optimizacija

1 Introduction

The growing complexity of distributed computing systems has made their design and management increasingly demanding. Early distributed networks were built around uniform nodes with comparable hardware and software characteristics. Modern networks, however, bring together a wide range of heterogeneous devices, from low-power wearables and embedded sensors to high-performance computing infrastructure, which introduces a new class of challenges that existing approaches only partially address.

This thesis focuses on three such challenges within heterogeneous distributed networks: assessing topological stability, balancing node load, and harvesting entropy from wearable devices for cryptographic use. The first two challenges are addressed through greedy graph coloring algorithms, chosen for their computational efficiency and suitability for resource-constrained environments. The third challenge is addressed through an offline, sensor-based entropy acquisition method, motivated by the same underlying concern for computational efficiency but methodologically independent of graph theory.

This chapter presents the research background and motivation in Section 1.1, identifies the research gap in Section 1.2, defines the research goals and hypotheses in Section 1.3, lists the thesis contributions in Section 1.4, and outlines the thesis structure in Section 1.5.

1.1 Background and Motivation

Distributed computing systems emerged in the latter half of the twentieth century as a means of connecting multiple independent computers into a cooperative network capable of sharing resources and processing workloads collectively. In their earliest form, these networks

were built around nodes that were largely uniform in nature. Participating devices shared comparable hardware architectures, operated under identical or closely related operating systems, and communicated through standardized protocols designed with homogeneity as an implicit assumption, as shown in Figure 1.1. Under such conditions, distributed algorithms could rely on predictable node behavior, consistent processing capacities, and stable communication patterns. Network management was therefore relatively straightforward, as the system designer could treat all nodes as functionally equivalent participants [15, 38, 47].

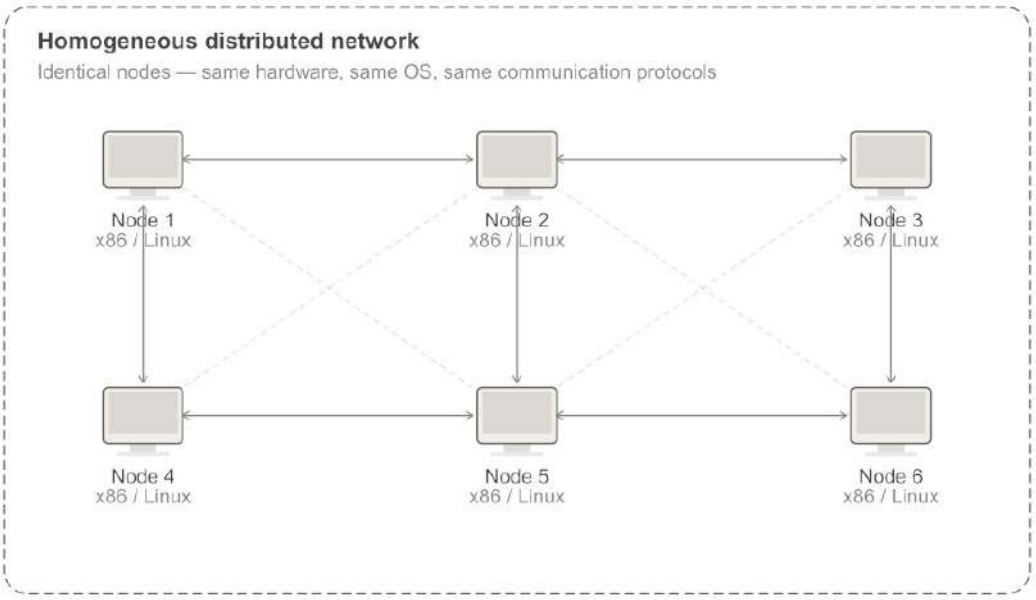

Figure 1.1: Example of a homogeneous distributed computer network.

The introduction of blockchain technology in 2008 marked a significant turning point in the evolution of distributed systems [31]. Originally conceived as a peer-to-peer electronic cash system operating on commodity personal computers, blockchain demonstrated that large-scale decentralized consensus could be achieved without relying on a central authority. Bitcoin's earliest network consisted of largely homogeneous nodes, predominantly desktop computers running identical mining and validation software. In the years that followed, blockchain networks expanded far beyond this original scope. The participating nodes diversified rapidly, encompassing smartphones and handheld devices, embedded IoT sensors, dedicated ASIC and FPGA mining hardware, cloud computing infrastructure, and high-performance computing clusters operating simultaneously within a single decentralized network [13, 24, 25, 39, 65]. This observed shift, documented within the Bitcoin network itself, illustrates the broader trend toward node heterogeneity that motivates the problem addressed in this thesis.

The consequences of this shift are visible across a broad range of contemporary application domains. Smart city platforms, for instance, integrate heterogeneous sensing infrastructure,

digital twin representations, machine learning pipelines, and private blockchain networks into unified urban management ecosystems [26]. In such environments, a single distributed system may simultaneously involve low-power street sensors with severely constrained computational resources, mid-tier edge computing nodes performing local data aggregation, and powerful cloud servers responsible for global coordination and analytics. The coexistence of these fundamentally different device classes within a shared network introduces structural and operational challenges that uniform-node models were never designed to handle. An example of such a network is shown in Figure 1.2.

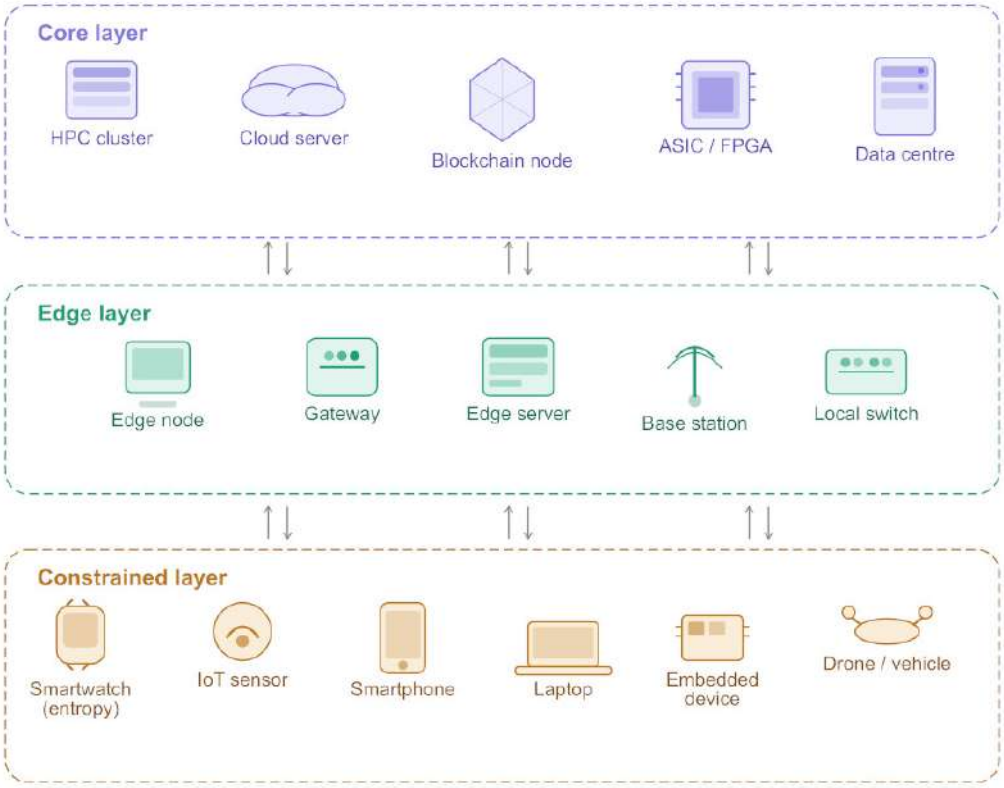

Figure 1.2: Example of a heterogeneous distributed computer network.

Three such challenges are of particular relevance to this thesis. The first concerns topological stability. In a heterogeneous network, the uneven distribution of node capabilities produces structural imbalances within the network graph. Certain nodes accumulate a disproportionate number of connections relative to their capacity, leading to saturation and reduced fault tolerance across the network. The second challenge concerns load balancing. Differences in computational power across nodes result in unequal processing loads, whereby some nodes become operational bottlenecks while others remain underutilized. The third challenge concerns cryptographic entropy generation. Resource-constrained nodes, particularly wearable and handheld devices, lack the dedicated hardware required to produce randomness meeting established security standards [3], which is essential for secure cryptographic key generation in distributed security protocols [10].

These three challenges motivate the research presented in this thesis and define the scope of its contributions.

1.2 Research Gap and Problem Statement

The evolution of distributed networks toward heterogeneous node compositions has been well documented in the literature. Studies on blockchain scalability, Internet of Things architectures, and smart city platforms consistently highlight the increasing diversity of participating devices and the operational challenges this diversity introduces [25, 26, 41, 50]. However, despite the growing recognition of these challenges, the existing body of research addresses them only in isolation, and several critical gaps remain.

With respect to topological stability, existing approaches to network analysis in distributed systems predominantly focus on connectivity metrics such as degree centrality, clustering coefficients, and path length [37]. While these metrics provide useful descriptors of network structure, they do not directly capture the relationship between node saturation and fault tolerance in heterogeneous environments. Graph coloring has been applied extensively in scheduling, frequency assignment, and register allocation problems [6, 9, 29, 34, 40], yet its application as a topology stability assessment tool in distributed peer-to-peer networks remains largely unexplored.

With respect to load balancing, the predominant approaches in distributed computing rely on heuristic and metaheuristic optimization methods, including genetic algorithms, simulated annealing, and ant colony optimization [59]. These methods are capable of producing near-optimal solutions, but they impose significant computational overhead, rendering them impractical for deployment on resource-constrained nodes. Greedy algorithm-based approaches have been theoretically studied in the context of graph partitioning [14, 56], but their direct application to load balancing in heterogeneous distributed networks, particularly through the lens of chromatic number optimization, has not been systematically investigated.

With respect to entropy generation, the security of distributed systems relies fundamentally on the availability of high-quality randomness for cryptographic key generation [18, 46]. Dedicated hardware random number generators address this requirement on capable devices, but resource-constrained nodes such as wearables and embedded sensors lack such hardware. Prior work has explored sensor-based entropy sources on smartphones [27] and IoT devices [41], demonstrating that raw sensor data contains exploitable randomness. However, the specific challenge of offline entropy harvesting from wearable devices, where network connectivity cannot be assumed and motion-triggered acquisition modes have not been systematically compared, has not been addressed in the context of blockchain and distributed system security.

These three gaps collectively define the research space of this thesis. Each gap corresponds

directly to one of the research goals defined in the following section.

1.3 Research Goals and Hypotheses

The overall research goal of this thesis is formulated as follows:

> *To develop optimization strategies for enhancing robustness and security in distributed systems through graph-theoretic topology modeling, greedy load balancing methods, and offline entropy harvesting from wearable sensor data.*

The overall research goal is divided into three Research Goals (RGs).

Research Goal 1

RG1: Develop a model for assessing the stability of distributed network topology based on node color distribution, aimed at reducing congestion and increasing system resilience.

The hypothesis associated with RG1 is formulated as follows:

H1: Graph coloring of distributed network topologies using greedy-based algorithms (Greedy Ascending, Greedy Descending, DSATUR) enables quantitative assessment of network stability, whereby an optimal chromatic number combined with uniform color class distribution indicates reduced congestion and improved fault tolerance.

Research Goal 2

RG2: Develop a method for node load balancing in heterogeneous distributed networks based on greedy algorithms, identifying an edge density that balances load distribution evenness against network resource overhead, thereby enhancing system robustness.

The hypothesis associated with RG2 is formulated as follows:

H2: Greedy graph coloring applied to distributed network topologies with varying edge densities produces measurably more even load distribution across nodes as edge density increases up to a certain threshold, beyond which further increases yield diminishing improvements relative to the additional resource overhead incurred, while remaining computationally more efficient than heuristic and metaheuristic alternatives.

Research Goal 3

RG3: Develop a method for offline entropy harvesting from sensor data of wearable devices for application in distributed systems, evaluating motion-triggered and passive acquisition modes against established cryptographic randomness standards.

The hypothesis associated with RG3 is formulated as follows:

H3: Sensor data collected from wearable devices via offline sampling produces entropy of sufficient quality, as validated by NIST SP 800-90B statistical testing, to serve as a viable randomness source for cryptographic key generation in resource-constrained distributed system nodes, with motion-triggered acquisition yielding measurably higher entropy quality than passive, time-triggered acquisition.

1.4 Original Scientific Contributions

This thesis makes the following original scientific contributions, each corresponding to one of the research goals defined in Section 1.3:

TC1: A model for assessing the stability of distributed network topology based on node colour distribution, aimed at reducing congestion and increasing system resilience.

TC2: A method for balancing node load to enhance the robustness of distributed networks, based on greedy algorithms.

TC3: A method for offline entropy harvesting from sensor data of wearable devices for application in distributed systems.

Table 1.1 presents the mapping between these contributions, the research goals defined in Section 1.3, and the publications included in this thesis. Papers A and B represent the main contributing publications, directly establishing the methods and validation underlying TC1, TC2, and TC3. Papers C, D, E, and F serve as supporting publications, providing broader security context, illustrative application scenarios, and complementary distributed systems analysis relevant to the thesis contributions.

Table 1.1: Mapping of publications to research goals and original scientific contributions.

Paper	**Title (short)**	**Type**	**Mapped contribution(s)**
Paper A	Greedy graph coloring for distributed networks	Main	TC1, TC2
Paper B	Entropy extraction from wearable sensors	Main	TC3
Paper C	Quantum computing threats to distributed systems	Supporting	Security context for TC1–TC3
Paper D	CD-7 consensus framework	Supporting	Distributed consensus context for TC1, TC2
Paper E	Advancing Smart City through heterogeneous networks	Supporting	Illustrative context for TC1, TC2
Paper F	Post-quantum business intelligence	Supporting	Security context for TC3

The relationships between these challenges, contributions, and targeted system properties

are summarised in Figure 1.3.

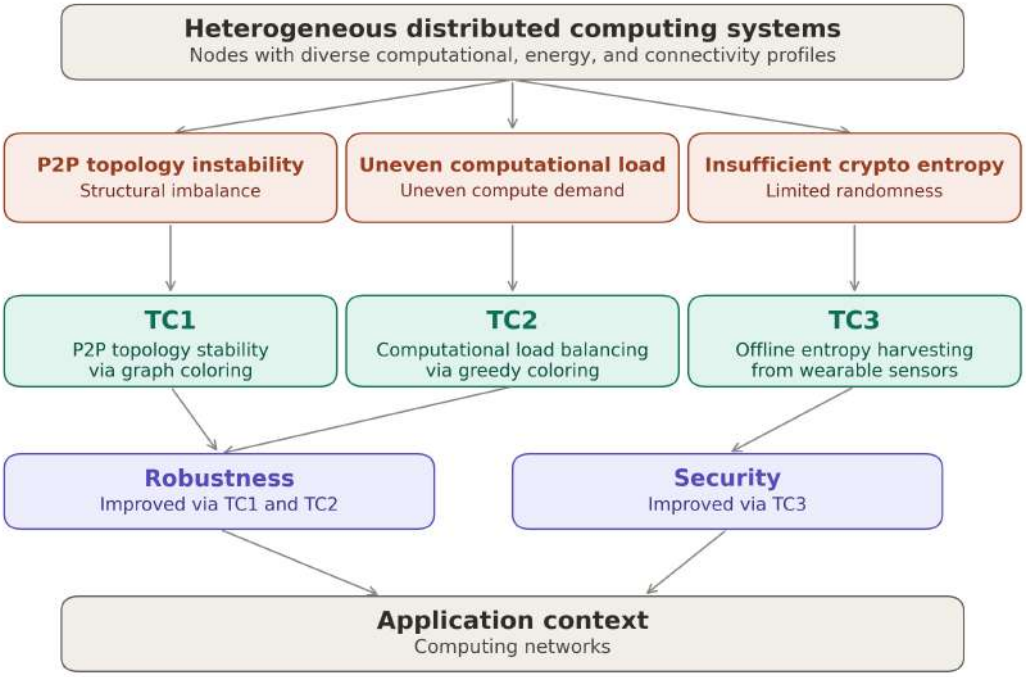

Figure 1.3: Conceptual framework of the thesis, illustrating how the heterogeneity of distributed computing systems motivates three key challenges addressed by TC1, TC2, and TC3, targeting robustness and security in computing networks.

1.5 Thesis Structure

This thesis consists of two main parts. Part I provides an overview of the research conducted and is organized into five chapters. Part II contains the full text of the six publications included in this thesis.

Chapter 1 (Introduction) presents the research background and motivation, identifies the research gap, defines the research goals and hypotheses, lists the original scientific contributions, and outlines the structure of the thesis.

Chapter 2 (Background and Related Works) provides the technical background required to understand the contributions of this thesis, covering graph theory and graph coloring, load balancing in distributed systems, and entropy generation for cryptographic applications. The chapter also reviews the relevant state of the art for each of the three research goals.

Chapter 3 (Research Contributions) presents the three original scientific contributions in detail, describes the methodology and validation approach underlying each contribution, and summarises the six publications included in this thesis together with their mapping to the research goals.

Chapter 4 (Discussion) synthesises the findings presented across the included publications, discusses the extent to which the research goals and hypotheses have been fulfilled, and addresses the limitations of the proposed methods.

Chapter 5 (Conclusions and Future Work) summarises the main findings of the thesis and outlines directions for future research.

Part II contains the full text of the following publications, presented in the order in which they are referenced throughout Part I: Paper A, Paper B, Paper C, Paper D, Paper E, and Paper F.

2

Background and Related Works

This chapter establishes the technical and theoretical foundation required to understand the contributions presented in this thesis. It begins with an overview of distributed systems and the structural implications of node heterogeneity, followed by the theoretical background on graph theory and graph coloring as applied to topology assessment and load balancing. The chapter then reviews entropy generation methods relevant to cryptographic applications, and concludes with a discussion of security implications specific to heterogeneous distributed environments. Throughout the chapter, the relevant state of the art is reviewed for each of the three research goals defined in Chapter 1.

2.1 Distributed Systems: Robustness and Security Challenges

A distributed system is commonly defined as a collection of independent computing nodes that communicate and coordinate their actions through message passing in order to achieve a common goal [53]. The defining characteristic of such systems is the absence of shared memory and global clock synchronization, which introduces fundamental challenges in coordination, consistency, and fault tolerance that do not arise in centralized computing architectures [16].

Early distributed systems research largely assumed a degree of node homogeneity. Nodes were treated as functionally interchangeable units with comparable processing capacity and reliable network connections, an assumption that simplified the design of consensus protocols, scheduling algorithms, and fault tolerance mechanisms [2]. This assumption underpinned much of the foundational work on distributed consensus, including the formulation of practical Byzantine fault tolerance [15] and early peer-to-peer file-sharing architectures.

The emergence of blockchain technology introduced a qualitatively different model of distributed computation, characterized by open participation, adversarial conditions, and the absence of a central coordinating authority [31]. As blockchain networks matured, the population of participating nodes diversified substantially. Contemporary blockchain networks comprise full nodes running on dedicated servers, lightweight clients on mobile devices, specialized mining hardware, and increasingly, embedded and wearable devices interacting with the network through gateway intermediaries [5, 13, 25, 44]. This diversification is mirrored in related domains. Internet of Things deployments combine sensors, actuators and gateway devices with markedly different computational and energy profiles [7], while smart city platforms integrate sensing infrastructure, edge computing nodes, and centralized analytics services into cohesive but structurally heterogeneous systems [26].

This heterogeneity has direct consequences for the design of distributed algorithms. Node heterogeneity affects three aspects of system behavior that are central to this thesis: the structural stability of the network topology, the distribution of computational load across nodes, and the availability of resources required for security-critical operations such as cryptographic key generation. Each of these aspects is examined in the following sections. The heterogeneity of participating nodes has a direct implication for the choice of analytical and optimization methods. In a network where devices range from resource-constrained wearables and embedded sensors to high-performance computing infrastructure, any method intended for deployment on network nodes must remain computationally tractable even on the least capable devices. This requirement rules out optimization approaches that depend on iterative search, global state knowledge, or substantial memory overhead — characteristics common to metaheuristic methods such as genetic algorithms, simulated annealing, and ant colony optimization [59]. Graph theory provides a computationally efficient alternative framework for analyzing and optimizing distributed network structure. By representing nodes as vertices and communication links as edges, the structural properties of a heterogeneous network — including its connectivity patterns, degree distributions, and load characteristics — can be formally captured and analyzed. Within this framework, graph coloring algorithms, and greedy variants in particular, offer a favorable trade-off: a single pass over the network graph yields a partition of nodes into non-conflicting groups, from which both structural stability indicators and load balancing assignments can be derived, at a computational cost that scales linearly with the number of nodes and edges [14, 56]. The following section develops this framework in detail.

2.2 Graph Theory and Network Topology Optimization

Graph theory provides the mathematical framework through which the topology and operational behavior of distributed networks can be formally analyzed [23, 29, 30, 58]. A distributed

network is naturally represented as a graph, where nodes correspond to vertices, and communication links correspond to edges. Within this representation, two classes of problems are of particular relevance to this thesis: the assessment of structural stability through graph coloring and the balancing of computational load across network nodes. Both problems are addressed in this thesis using greedy algorithmic approaches, motivated by their favorable trade-off between solution quality and computational cost.

The graph coloring problem consists of assigning a color to each vertex of a graph such that no two adjacent vertices share the same color, while minimizing the total number of colors used. The minimum number of colors required to color a graph is known as its chromatic number [6]. Graph coloring is classified as an NP-hard problem, meaning that no known algorithm can solve it optimally in polynomial time for arbitrary graphs [4]. As a consequence, exact algorithms become computationally infeasible for large graphs, motivating the widespread use of heuristic and approximation methods.

Several heuristic algorithms have been proposed in the literature for solving the graph coloring problem, including First Fit, Largest Degree Ordering, Welsh-Powell, Incidence Degree Ordering, Recursive Largest First, and the Degree of Saturation (DSATUR) algorithm [6, 22, 40, 43]. Among these, greedy-based methods, including DSATUR as a greedy variant that orders vertices dynamically by saturation degree, consistently achieve the most favorable trade-off between computational time and solution quality. In contrast, methods relying on metaheuristic search, such as Tabu Search [33] and Ant Colony Optimization [1], achieve marginally better colorings at substantially higher computational cost, particularly on large graphs.

Greedy algorithms construct a solution incrementally by making the locally optimal choice at each step, without reconsidering previous decisions [56]. While this approach does not guarantee a globally optimal solution, it offers substantially lower computational complexity than exhaustive or metaheuristic search methods, making it particularly suitable for large-scale or resource-constrained applications. In the context of graph partitioning, greedy coloring algorithms partition the vertex set of a graph into color classes, each of which forms an independent set within the graph [14]. The relative sizes of these color classes provide a structural indicator of how evenly the graph's vertices are distributed across non-conflicting groups. Computational complexity analyses of greedy partitioning confirm that, although determining the exact number of colors used by the greedy algorithm under a given vertex ordering is itself a hard problem in the general case, the algorithm remains efficient in practice and produces partitions whose properties can be analyzed systematically.

In this thesis, vertices are ordered by their degree, that is, by the number of neighboring nodes, prior to coloring. Two ordering directions are considered: Greedy Descending, in which vertices are colored in order of decreasing degree, and Greedy Ascending, in which the order is reversed. The Greedy Descending variant corresponds to the Welsh-Powell algorithm [35, 36], which has been shown to produce favorable colorings on a range

of graph structures despite its computational simplicity. The Greedy Ascending variant, which processes lower-degree vertices first, is included to evaluate whether reversing this ordering produces measurable differences in the resulting color class distribution, a question of particular relevance when coloring is subsequently used as the basis for load balancing.

This thesis builds on these theoretical foundations to apply greedy graph coloring, implemented through Greedy Ascending, Greedy Descending, and DSATUR variants, as the basis for both topology stability assessment (TC1) and node load balancing (TC2) in heterogeneous distributed networks. The rationale for this choice, and its specific application to the two contributions, is detailed in Chapter 3.

2.3 Entropy and Randomness for Cryptographic Applications

Cryptographic systems rely fundamentally on the availability of high-quality randomness [18]. Secure key generation, nonce creation, and the initialization of cryptographic protocols all require unpredictable bit sequences, since any structure or bias in the underlying randomness source can be exploited by an adversary to compromise the security of the resulting cryptographic material. The quality of an entropy source is therefore not a peripheral implementation detail but a foundational requirement for the security guarantees of any cryptographic system.

Random number generators used in cryptographic applications are generally classified into two categories [57]. True random number generators (TRNGs) derive randomness from physical processes, such as electronic noise or environmental variation, which are inherently unpredictable [51]. Pseudo-random number generators (PRNGs), by contrast, are deterministic algorithms that expand a relatively short random seed into a longer sequence that is computationally indistinguishable from true randomness [52]. PRNGs depend critically on the quality of their seed material, and weaknesses in the underlying entropy source have repeatedly been shown to compromise systems that otherwise rely on cryptographically sound algorithms. Documented vulnerabilities in widely deployed PRNGs, including the Linux kernel random number generator, have demonstrated that insufficient entropy accumulation or flawed mixing functions can allow an adversary to reconstruct internal states and predict future outputs [17, 20].

Conventional approaches to generating high-quality entropy typically rely on dedicated hardware random number generators or operating system entropy pools that aggregate noise from multiple system events [52]. While effective, these mechanisms presuppose the availability of suitable hardware components or a sufficiently rich set of system events from which entropy can be harvested. This assumption does not hold uniformly across heterogeneous distributed networks. Resource-constrained nodes, including wearable devices, embedded sensors, and other low-power components, are frequently not equipped with dedicated entropy-generating hardware, and their limited interaction with users and external

systems restricts the availability of conventional operating system entropy sources.

This limitation has motivated growing interest in sensor-based entropy generation as an alternative source of randomness for resource-constrained devices. Sensors embedded in smartphones and wearable devices, originally intended for motion sensing, environmental monitoring, or biometric measurement, capture physical processes whose inherent noise can be extracted and processed into usable entropy [27]. Sensor-derived entropy generation has been demonstrated using bio-related signals captured by wearable devices, with the resulting random sequences validated against standard statistical test suites [60]. This body of work establishes that physical sensors embedded in commercially available devices can serve as viable entropy sources without requiring additional dedicated hardware, a property of particular relevance to heterogeneous distributed networks in which a substantial proportion of nodes possess limited computational and energy resources.

The statistical quality of entropy sources intended for cryptographic use is commonly assessed through standardized test suites. The NIST SP 800-22 statistical test suite provides a set of tests for evaluating whether a sequence of bits exhibits properties consistent with true randomness [42]. The NIST SP 800-90B recommendation specifically addresses the validation of entropy sources intended for use in random bit generation, providing methods for estimating min-entropy and assessing whether an entropy source provides sufficient unpredictability for cryptographic applications [28, 52]. These standards provide the methodological basis against which the entropy harvesting method proposed in this thesis is validated.

Building on this body of work, this thesis investigates offline entropy harvesting from smartwatch sensor data as a source of cryptographic randomness for resource-constrained nodes in heterogeneous distributed networks, extending prior findings on sensor-derived entropy to a multisensor acquisition model validated under both static and motion-triggered conditions [64]. The motivation, methodology, and validation results underlying this contribution (TC3) are presented in detail in Chapter 3.

2.4 Security Implications in Heterogeneous Distributed Systems

The structural and operational challenges addressed in the preceding sections, namely topology stability, load balancing, and entropy generation, were examined in this thesis primarily through the lens of blockchain networks, which serve as a representative and practically significant example of heterogeneous distributed systems. Studying blockchain networks as the motivating context revealed several security implications that extend beyond the three original scientific contributions and connect them to the broader research conducted as part of this thesis [11, 19].

One such implication concerns the relationship between node heterogeneity and the

choice of consensus mechanism. Proof-of-Work consensus, as originally proposed for Bitcoin, requires substantial and continuous computational effort, rendering it impractical for resource-constrained nodes such as wearable or embedded devices [31]. Alternative mechanisms, including Proof-of-Stake and its variants, reduce computational demands but introduce different trust and security assumptions [8, 32]. Systematic reviews of consensus mechanisms confirm that no single approach is universally optimal, and that the appropriate choice depends on the specific deployment context, including the heterogeneity of participating nodes and the security guarantees required by the application [21, 61]. This observation, namely that the increasing heterogeneity of blockchain network participants directly constrains which consensus mechanisms can realistically be deployed, motivated the development of a data-driven framework for consensus mechanism selection in distributed systems, undertaken as part of the supporting work of this thesis [66].

A further implication emerged from examining the long-term cryptographic assumptions underlying blockchain systems. Many of the cryptographic primitives currently used for digital signatures and key exchange in blockchain systems, including elliptic curve cryptography, rely on computational problems that are believed to be intractable for classical computers but that would be efficiently solvable by a sufficiently capable quantum computer running Shor's algorithm [48, 49]. Although large-scale, fault-tolerant quantum computers capable of executing such attacks do not yet exist, the development trajectory of quantum computing hardware has prompted growing concern regarding the long-term security of data and transactions that depend on cryptographic assumptions that may not hold indefinitely [12, 45]. This concern proved particularly relevant for blockchain systems, where transaction histories and public keys remain permanently accessible, creating the possibility of retrospective decryption once sufficiently powerful quantum computers become available. Investigating this development trajectory, and its projected impact on blockchain security in heterogeneous distributed networks, formed the basis of a further supporting contribution of this thesis [67].

These findings, in turn, connect directly to post-quantum cryptography, which addresses the quantum threat by developing cryptographic primitives believed to remain secure against adversaries equipped with quantum computers. The transition toward post-quantum cryptographic standards was found to have direct implications for entropy requirements, since post-quantum algorithms frequently require larger key sizes and correspondingly greater quantities of high-quality randomness during key generation [63]. This observation links back to the entropy harvesting contribution of this thesis (TC3): resource-constrained nodes adopting post-quantum cryptographic schemes face increased demand for locally generated entropy, a requirement that offline entropy harvesting from wearable sensors is positioned to address.

The relevance of entropy harvesting to post-quantum cryptographic readiness extends beyond a general motivational link. The National Institute of Standards and Technology finalized the first post-quantum cryptographic standards in 2024, specifying the Module-Lattice-Based Key-Encapsulation Mechanism (ML-KEM, derived from CRYSTALS-Kyber)

in FIPS 203 [54] and the Module-Lattice-Based Digital Signature Algorithm (ML-DSA, derived from CRYSTALS-Dilithium) in FIPS 204 [55]. Both standards rely on lattice-based hardness assumptions and require substantially larger key material than classical elliptic-curve schemes: public keys for ML-KEM range from approximately 800 to 1568 bytes depending on the selected security level, compared to 32 bytes for a typical elliptic-curve public key. Notably, FIPS 204 specifies that key generation for ML-DSA can be seeded from a 32-byte (256-bit) random value, a length directly compatible with the SHA-256 digest produced by the entropy harvesting method developed in this thesis. This indicates that the offline, sensor-derived entropy validated for TC3 is, in principle, of sufficient length to seed post-quantum key generation on resource-constrained wearable devices, extending the practical relevance of the proposed method beyond classical cryptographic applications and reinforcing the motivation for the supporting investigation of post-quantum migration strategies presented in Papers C and F.

Taken together, these findings, drawn from studying blockchain as a heterogeneous distributed network, illustrate how consensus mechanism selection, quantum computing threat timelines, and post-quantum cryptographic readiness are interconnected security considerations that reinforce the relevance of the three original scientific contributions of this thesis. The relationship between these supporting publications and the main contributions of this thesis is examined in detail in Chapter 3.

3

Research Contributions

This chapter presents the three original scientific contributions of this thesis in detail. For each contribution, the underlying methodology is described, the validation approach is explained, and the relevant publication in which the contribution was first presented and evaluated is identified. The chapter also discusses the role of the supporting publications included in this thesis in relation to these three contributions. Figure 3.1 provides an overview of the mapping between research goals, hypotheses, contributions, and publications.

3.1 TC1: Topology Stability Assessment Model

The first contribution of this thesis addresses the assessment of topological stability in heterogeneous distributed networks. As discussed in Section 2.2, the structural properties of a network graph have direct implications for its resilience to node saturation and its overall fault tolerance, yet existing approaches to network analysis do not directly capture this relationship through the lens of graph coloring.

To address this gap, this thesis proposes a topology stability assessment model based on node color distribution. The model treats each node in a distributed network as a vertex in a graph, with edges representing communication links between nodes. Graph coloring is then applied to the network using three algorithmic variants: Greedy Ascending, Greedy Descending, and DSATUR, as discussed in Section 2.2. The resulting coloring assigns each node to a color class, where nodes within the same color class share no direct connection.

The stability of a given network topology is assessed through two related indicators derived from the coloring outcome. The first indicator is the number of colors used, which

Figure 3.1: Mapping between research goals, hypotheses, thesis contributions, publications, and validation approach.

approximates the chromatic number of the network graph and reflects the degree of structural conflict among nodes [30]. The second indicator is the distribution of nodes across color classes. A topology in which nodes are distributed evenly across color classes is interpreted as more structurally balanced and, by extension, more resilient to node saturation than a topology in which color classes are highly uneven [23].

Color class evenness is quantified as the ratio between the smallest and largest color class in the resulting partition. A ratio approaching unity indicates that nodes are distributed approximately equally across color classes, while a ratio approaching zero indicates that one or more color classes are disproportionately large, corresponding to a structurally imbalanced topology in which a small number of nodes bear a disproportionate share of the network's connectivity burden. This ratio provides a single, interpretable scalar indicator of structural balance that can be computed directly from the coloring output without additional graph analysis.

The simulated networks used for evaluation were generated as random graphs under a fixed node count of 100, with edges assigned uniformly at random to achieve the target edge densities of 1500, 2000, and 2500 edges. This generation model was selected because it produces networks whose degree distributions reflect the uneven connectivity patterns characteristic of heterogeneous distributed networks, while remaining reproducible and amenable to systematic comparison across edge density levels [58]. The relationship between edge density and chromatic number across the three evaluated greedy variants is illustrated in Figure 3.2.

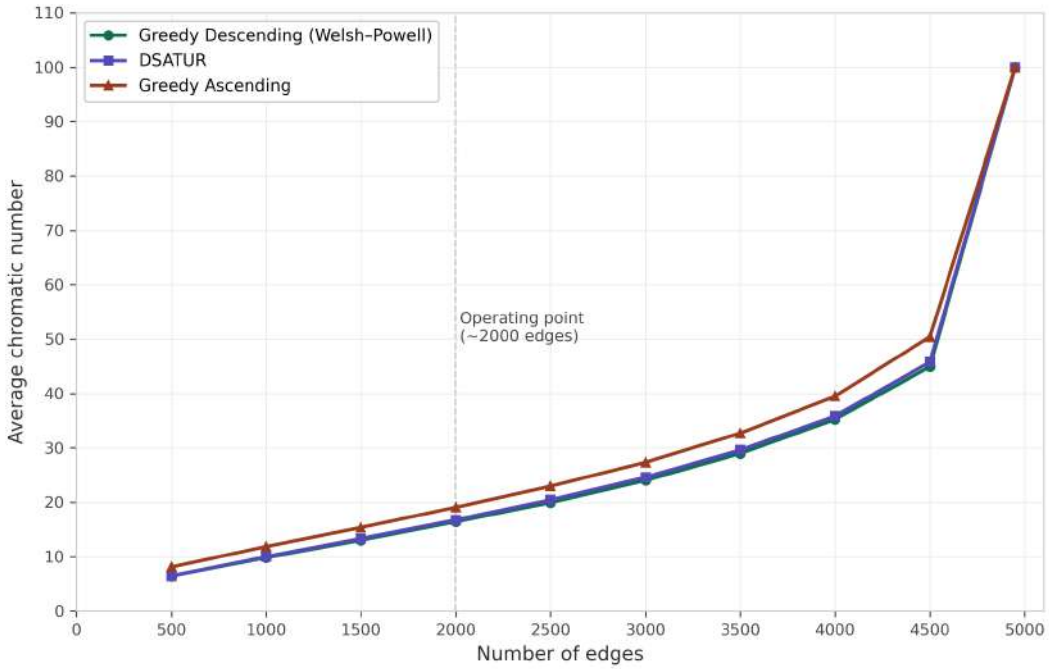

Figure 3.2: Chromatic number as a function of edge density for three greedy graph coloring variants applied to simulated 100-node distributed networks.

This range of edge densities was selected to represent a realistic operating envelope for heterogeneous distributed networks: networks at the lower bound of connectivity are structurally fragile, since the removal of any single edge can disconnect the network, while networks approaching the maximum possible number of edges impose computational and communication overheads that are impractical for resource-constrained nodes. The methodology, simulation setup, and complete results of this evaluation are presented in Paper A [62].

3.2 TC2: Node Load Balancing Method

The second contribution of this thesis addresses node load balancing in heterogeneous distributed networks, building directly on the topology stability assessment model presented in Section 3.1. As discussed in Section 2.2, conventional load balancing approaches in distributed systems frequently rely on metaheuristic optimization methods that are computationally expensive and therefore poorly suited to resource-constrained nodes [59].

This thesis proposes a node load balancing method based on the same greedy graph coloring procedure used for topology stability assessment. The underlying premise is that the color classes produced by graph coloring can be repurposed as load balancing groups. Since nodes within the same color class share no direct connection, they can be treated as candidates for parallel task assignment or load distribution without introducing direct communication conflicts between them. Throughout this section, "load" refers to structural load, that is, the distribution of communication links and concurrent task slots across the network topology, rather than to the computational capacity of individual nodes. Accordingly, an even color-class partition indicates that no group of mutually non-adjacent nodes is structurally overburdened relative to the others; it does not assume that nodes sharing a color are equivalent in processing power. The extension of this method to account for heterogeneous node capacities is discussed as a limitation in Section 4.5. An even distribution of nodes across color classes, the same indicator used to assess topological stability in Section 3.1, simultaneously indicates a balanced allocation of load across the network, since no single group of mutually connected nodes is disproportionately large relative to the others [14].

The number of colors used in the coloring and the evenness of their distribution across nodes together encode two complementary properties of the network. The number of colors reflects the degree of structural diversity in the network: a larger chromatic number indicates that the network contains more groups of mutually non-adjacent nodes, which corresponds to greater connectivity and higher fault tolerance, since no single node failure can isolate an entire group. The evenness of color distribution — that is, how uniformly nodes are spread across color classes — reflects the degree of load balance: when all color classes contain approximately the same number of nodes, no single group of parallel workers is

disproportionately large, and the computational load is distributed without bottlenecks. A network is considered optimized when both conditions are simultaneously satisfied: sufficiently many color classes to ensure structural resilience, and sufficiently even distribution across those classes to ensure balanced resource utilization. Crucially, this optimized state must be achieved without unnecessarily increasing the number of communication links, since each additional link beyond the operating point consumes network bandwidth and node processing capacity without producing proportional improvements in either stability or load balance. The method proposed in this thesis identifies this operating point empirically, providing a computationally efficient guideline for network configuration that achieves robustness while preserving resource efficiency.

Concretely, the load balancing procedure operates as follows. Once the greedy coloring is applied to the network graph, the resulting color classes define a partition of the node set into non-conflicting groups. Incoming computational tasks are distributed across these groups in a round-robin or demand-driven fashion, with all nodes within a given color class eligible to receive tasks in parallel. Because no two nodes within the same color class share a direct communication link, this assignment does not introduce direct message conflicts at the link level, avoiding the coordination overhead that arises when directly connected nodes are assigned concurrent tasks. The load distribution quality is then assessed through the same color class evenness indicator used for topology stability: an even partition ensures that no single group of nodes is systematically overloaded while others remain idle [56].

Beyond this structural interpretation, the same color-class evenness indicator was used to determine a practical operating point for network connectivity, expressed as the number of edges relative to the number of nodes. Across the simulated 100-node networks described in Section 3.1, increasing edge density from 1000 to 2000 produced a measurable improvement in color class evenness across all three greedy variants, indicating progressively better load distribution. Beyond approximately 2000 edges, further increases in density yielded only marginal improvements in evenness relative to the additional resource and communication overhead incurred, identifying an edge density in this range as a practical compromise between load balancing performance and network resource cost.

The numerical results obtained from Paper A quantify this trade-off concretely. Among the three evaluated variants, DSATUR achieved the most uniform color class distribution across all tested edge densities, reflecting its saturation-aware vertex selection strategy. The difference in chromatic efficiency between the Ascending and Descending variants was 20.56% at the sparsest tested configuration, narrowing to 10.91% at the densest configuration excluding the fully connected case, confirming that the performance gap between variants diminishes as edge density increases. At approximately 2000 edges, the network achieved high color class evenness and structural resilience without incurring the redundancy costs associated with denser configurations, while increasing edge count to 2500 produced only marginal further improvement [62].

Critically, this operating point was identified consistently across all three evaluated greedy variants — Greedy Ascending, Greedy Descending, and DSATUR — indicating that the threshold of approximately 2000 edges represents a property of the network structure itself rather than an artifact of any particular algorithmic choice. This cross-algorithm consistency strengthens the practical validity of the finding: regardless of which greedy variant is deployed on a resource-constrained node, the same connectivity guideline applies. The method thus provides a computationally efficient means of determining the optimal number of communication links in a heterogeneous distributed network — sufficient to achieve structural robustness and balanced load distribution, but without incurring the resource overhead of maintaining unnecessary connections [62].

Figure 3.3 illustrates the progressive evening of color class distribution across increasing edge densities for all three greedy variants applied to simulated 100-node networks.

This dual use of the same coloring procedure, applied for both topology stability assessment and the identification of a favorable edge density for load balancing, reflects the methodological efficiency that motivated the choice of greedy algorithms in the first place, as discussed in Section 2.2. A single, computationally inexpensive coloring operation yields insight into both the structural stability and the load distribution of the network, without requiring two separate optimization procedures. The complete evaluation of this method is presented in Paper A [62].

3.3 TC3: Offline Entropy Harvesting Method

The third contribution of this thesis addresses the generation of cryptographic entropy on resource-constrained nodes within heterogeneous distributed networks, specifically wearable devices. As discussed in Section 2.3, such nodes are frequently not equipped with dedicated hardware random number generators, while their continuous physical contact with the user provides access to a rich set of sensor data that can serve as an alternative entropy source.

This thesis proposes an offline entropy harvesting method based on multisensor data collected from a commercial smartwatch. A custom Android application, implemented using Jetpack Compose and Kotlin, collects data from a subset of the device's onboard sensors during measurement windows of 2.2 seconds in duration, a window selected to ensure that each active sensor reports at least one reading. The sensors activated during each measurement window included the accelerometer, gyroscope, gravity sensor, linear acceleration sensor, rotation vector sensor, barometric pressure sensor, heart rate sensor, and step counter, selected for their sensitivity to physical motion and environmental variation and for their availability on commercially deployed smartwatch hardware. Each sensor contributes a set of floating-point readings during the measurement window; these readings are concatenated and serialized into a binary representation, forming the input to the

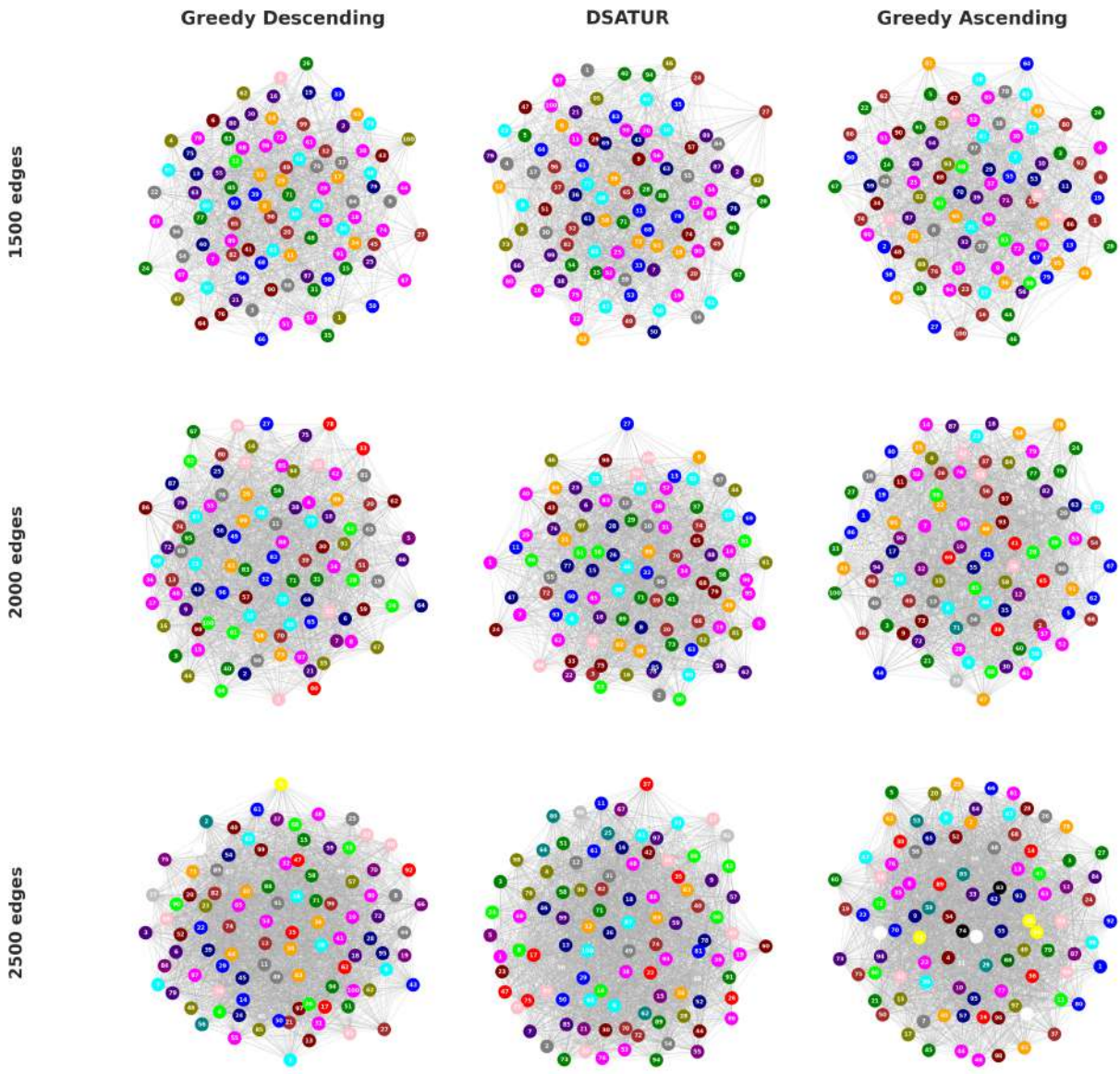

Figure 3.3: Color class distribution for three greedy graph coloring variants applied to simulated 100-node distributed networks at edge densities of 1500, 2000, and 2500 edges. The progressive evening of color distribution with increasing edge density is visible across all three variants, with the operating point at approximately 2000 edges providing a favorable balance between load distribution quality and network resource overhead.

entropy extraction pipeline. The resulting bit vector is then passed through a SHA-256 hash function, producing a 256-bit output that serves as the final entropy contribution for that measurement.

Two acquisition modes were implemented and compared. In the still mode, measurements are taken every 2.5 seconds without requiring user interaction, capturing environmental and physiological variability from sensors such as the barometer and heart rate monitor. In the shake mode, data collection is triggered only once device motion exceeds a predefined acceleration threshold, requiring the user to perform a deliberate movement before a measurement is recorded. This design choice was intended to introduce additional physical randomness into the sensor readings through user-induced motion, exploiting the stochastic nature of wrist movement as a source of unpredictability.

A total of 4800 measurements were collected under each acquisition mode. The resulting bit sequences were truncated to a fixed length of 11,400 bits and evaluated for cryptographic suitability using the NIST SP 800-90B entropy assessment methodology discussed in Section 2.3, including Shannon entropy, min-entropy, Markov dependency analysis, and compression-based redundancy estimation.

The validation results confirmed the effectiveness of the proposed method. In the shake acquisition mode, the generated bit sequences achieved a Shannon entropy of 0.997 bits per bit and a min-entropy of 0.918 bits per bit, approaching the theoretical maximum and demonstrating that motion-triggered sampling produces randomness of near-ideal quality. The still acquisition mode yielded a Shannon entropy of 0.991 bits per bit and a min-entropy of 0.851 bits per bit, confirming that passive acquisition, while viable, produces measurably lower entropy quality than motion-triggered acquisition, in direct support of the hypothesis formulated in H3. Both modes passed the NIST SP 800-90B dependency analysis, indicating that the generated sequences are free of detectable statistical structure [64].

Figure 3.4 presents a comparative evaluation of the proposed method against established entropy sources, including hardware-based true random number generators (TRNGs) and software-based cryptographically secure pseudorandom number generators (CSPRNGs).

The shake mode results are competitive with operating system RNGs such as `/dev/urandom`, showing only a minor deficit in min-entropy while exceeding its Shannon entropy value. Although the still mode does not reach the threshold required for standalone cryptographic use, it remains a viable supplementary entropy source when combined with other inputs. Neither mode matches the theoretical security of dedicated TRNGs or mature CSPRNGs; however, the proposed method offers unique advantages aligned with the requirements of heterogeneous distributed systems: it operates entirely offline, requires no dedicated cryptographic hardware, is implementable on widely available consumer devices, and produces entropy sufficient for cryptographic key generation when motion-triggered acquisition is used. The complete methodology, sensor configuration, and validation results are presented in Paper B [64].

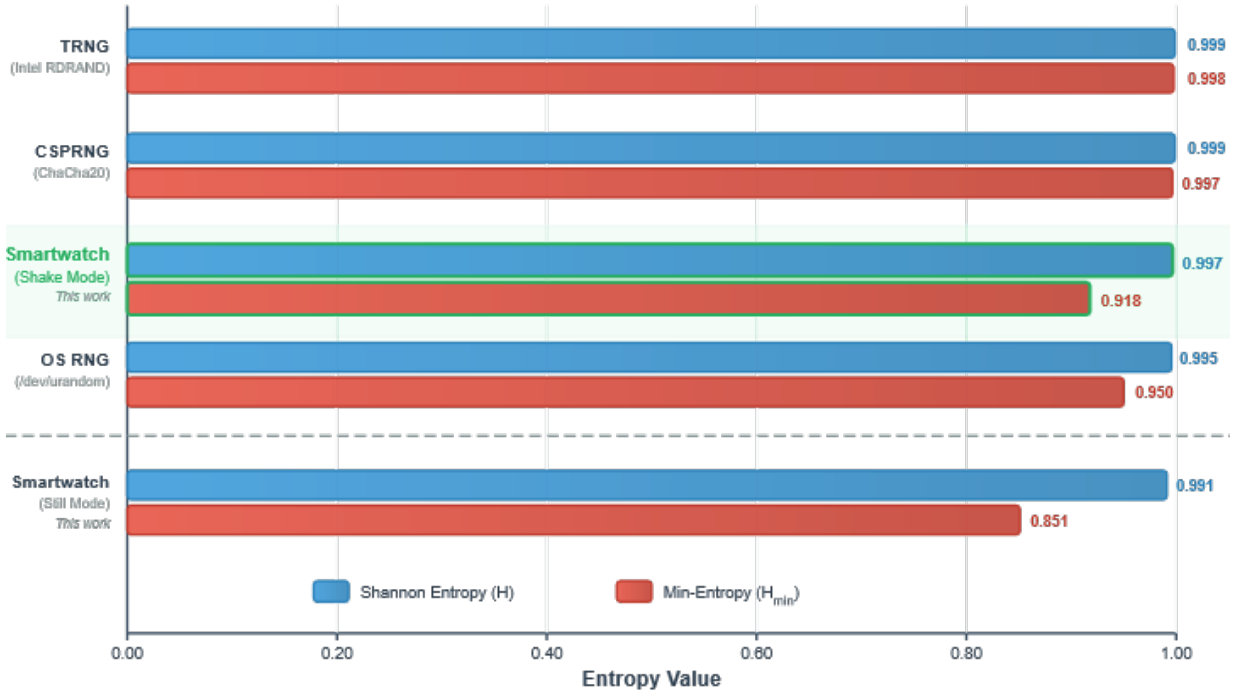

Figure 3.4: Comparative evaluation of Shannon entropy and min-entropy for the proposed smartwatch-based entropy harvesting method against established entropy sources. The shake acquisition mode achieves entropy levels competitive with operating system RNGs, while the still mode falls below the threshold for cryptographic suitability. Reference method data taken from [64].

3.4 Role of the Supporting Publications

While Papers A and B directly establish the methods and validation underlying TC1, TC2, and TC3, four further publications are included in this thesis as supporting contributions, situating the three original scientific contributions within a broader security and application context.

Paper C provides a predictive analysis of quantum computing development and its projected impact on blockchain security, motivating the long-term relevance of robust entropy generation as addressed by TC3. Paper D proposes a consensus-driven framework for the selection of blockchain consensus mechanisms, illustrating the broader dependence of distributed system design on node-level resource constraints relevant to TC1 and TC2. Paper E examines the integration of artificial intelligence, digital twin technology, and blockchain in a smart city context, providing a concrete illustration of the heterogeneous distributed network architecture motivating this thesis, as introduced in Section 1.1. Paper F addresses the design of post-quantum secure blockchain infrastructure for business intelligence systems, extending the security context established in Paper C to a specific application domain and reinforcing the entropy requirements addressed by TC3.

The relationship between these six publications, the three research goals, and the three original scientific contributions is summarized in Table 1.1.

4 Discussion

This chapter synthesizes the findings presented across the publications included in this thesis, evaluating the extent to which the research goals and hypotheses defined in Chapter 1 have been addressed. Each of the three research goals is discussed individually with reference to the corresponding contribution and its supporting evidence, followed by a discussion of cross-cutting themes that connect the three contributions, and a discussion of the limitations of the research presented in this thesis.

4.1 Addressing Research Goal RG1: Topology Stability

RG1 sought to develop a topology stability assessment model for heterogeneous distributed networks based on graph coloring, with the aim of reducing node saturation and increasing system resilience. This goal was addressed through the topology stability assessment model presented as TC1 and evaluated in Paper A.

The results obtained in Paper A support hypothesis H1, which proposed that graph coloring using greedy-based algorithms enables quantitative assessment of network stability through the number of colors used and the evenness of color class distribution. Across simulated 100-node networks with 1000, 1500, and 2000 edges, the coloring outcomes produced by the Greedy Descending, Greedy Ascending, and DSATUR variants consistently reflected the underlying connectivity structure of the network, with denser networks requiring fewer colors relative to their node count and exhibiting more even color class distributions than sparser networks. This finding is consistent with the theoretical relationship between edge density and chromatic number discussed in Section 2.2, and confirms that color distribution

evenness can serve as a practical, computationally inexpensive indicator of structural balance in a distributed network [62].

Compared to existing approaches to network stability assessment, which predominantly rely on connectivity metrics such as degree centrality and clustering coefficients [37], the coloring-based model proposed in this thesis offers a complementary perspective. Rather than describing the connectivity of individual nodes in isolation, the model provides a global, structural indicator derived from the coloring of the entire network, allowing the comparison of different network configurations using a single, interpretable measure. This addresses the research gap identified in Section 1.2, where existing connectivity metrics were noted to not directly capture the relationship between node saturation and fault tolerance.

The validation of TC1 is subject to certain limitations. The evaluation was conducted exclusively on simulated network topologies rather than on operational blockchain or distributed networks, and the simulated networks were limited to a single scale of 100 nodes. While this scale was chosen to balance computational feasibility with the ability to observe meaningful structural patterns across a range of edge densities, the generalizability of the findings to substantially larger or smaller networks, or to networks with non-random connectivity patterns, remains to be established in future work.

4.2 Addressing Research Goal RG2: Load Balancing

RG2 sought to develop a method for node load balancing in distributed networks to enhance system robustness, based on greedy algorithms. This goal was addressed through the node load balancing method presented as TC2, evaluated alongside TC1 in Paper A using the same set of simulated 100-node networks. The results obtained in Paper A support hypothesis H2, which proposed that greedy graph coloring applied to networks with varying edge densities produces measurably more even load distribution across nodes as edge density increases up to a certain threshold, beyond which further increases yield diminishing improvements relative to the additional resource overhead incurred. Across the simulated networks, increasing edge density from 1000 to 2000 edges produced a clear improvement in color class evenness across all three greedy variants, indicating progressively better load distribution. Beyond approximately 2000 edges, this improvement became marginal: the additional connectivity provided by 2500 edges yielded only a small further gain in evenness, while imposing greater resource and communication overhead on the network. This pattern identifies an edge density of approximately 2000 edges as a practical operating point [62], balancing load distribution quality against network resource cost, and confirms the threshold-based behavior predicted by H2. Since TC2 repurposes the color classes produced by the same coloring procedure used for TC1, the evenness of color class distribution serves a dual function: it indicates topological stability, as discussed in Section 4.1, and simultaneously indicates

a balanced allocation of load across the network, since nodes within the same color class share no direct connection and can therefore be treated as non-conflicting candidates for parallel task assignment. This dual interpretation of a single coloring outcome reflects the methodological efficiency motivating the use of greedy algorithms throughout this thesis, as discussed in Section 2.2. Compared to conventional load balancing approaches in distributed systems, which predominantly rely on metaheuristic optimization methods such as genetic algorithms, simulated annealing, and ant colony optimization [59], the method proposed in this thesis trades a degree of solution optimality for a substantial reduction in computational cost. Metaheuristic methods are capable of exploring a larger solution space and may converge to better-balanced configurations, but they require multiple iterations over the network graph and are therefore considerably more demanding in terms of processing time and memory. The greedy coloring procedure underlying TC2 processes each node exactly once, making it suitable for deployment on resource-constrained nodes that cannot support iterative optimization, even though the resulting load distribution may not be globally optimal.

Whereas metaheuristic methods such as genetic algorithms require multiple iterations over the full network graph and converge over generations of candidate solutions, the greedy coloring procedure underlying TC2 processes each vertex exactly once in $O(n + m)$ time, where n is the number of nodes and m the number of edges, making it directly deployable on devices with constrained processing budgets without any iterative overhead.

This trade-off directly addresses the research gap identified in Section 1.2, where existing load balancing approaches were noted to impose computational overhead impractical for constrained nodes. A further observation concerns the relationship between the two ordering directions evaluated in this thesis. Since Greedy Descending corresponds to the Welsh-Powell algorithm and prioritizes high-degree nodes early in the coloring process, while Greedy Ascending reverses this order, the two variants were expected to produce different color class distributions. The comparison of these variants, together with DSATUR, provides insight into which ordering strategy yields more balanced load distribution under varying network densities, a question of direct relevance to the practical deployment of TC2 in heterogeneous distributed networks. The validation of TC2 is subject to the same limitations discussed for TC1 in Section 4.1, since both contributions were evaluated on the same simulated network topologies. In addition, the load balancing method proposed in this thesis assumes that nodes within a color class are computationally interchangeable for the purpose of task assignment, an assumption that does not account for the heterogeneous processing capacities discussed in Section 2.1. Extending the method to incorporate node-specific capacity weights remains an open direction for future work.

4.3 Addressing Research Goal RG3: Entropy Harvesting

RG3 sought to develop an offline entropy harvesting method from wearable device sensor data for application in cryptographic operations within distributed systems. This goal was addressed through the offline entropy harvesting method presented as TC3 and evaluated in Paper B.

The results obtained in Paper B support hypothesis H3, which proposed that sensor data collected from wearable devices via a motion-triggered offline sampling algorithm produces entropy of sufficient quality to serve as a viable randomness source for cryptographic key generation. The multisensor acquisition pipeline, evaluated under both the still and shake operating modes, produced binary sequences that satisfied the statistical quality criteria outlined in the NIST SP 800-90B entropy assessment methodology [52, 64] discussed in Section 2.3. The comparison between operating modes further indicated that motion-triggered acquisition introduces additional physical variability into the collected sensor readings, supporting the design rationale for including a deliberate motion-based sampling condition alongside passive, time-triggered sampling.

Compared to conventional approaches to cryptographic randomness generation, which rely on dedicated hardware random number generators or operating system entropy pools, the method proposed in this thesis demonstrates that commercially available wearable devices, already commonly used in IoT and blockchain-adjacent applications, can serve as self-contained entropy sources without requiring additional hardware. This directly addresses the research gap identified in Section 1.2, where resource-constrained nodes such as wearables were noted to generally lack dedicated entropy-generating hardware. It is important to note, as emphasised throughout this thesis, that the entropy source validated in Paper B is a single-source, single-modality acquisition method drawing exclusively from one wearable device. This does not constitute a multi-source entropy architecture as defined by NIST SP 800-90B, and the contribution of this thesis should accordingly be understood as establishing the viability of a single wearable device as one such source, rather than as a complete multi-source entropy framework.

The relevance of TC3 is reinforced by the supporting publications included in this thesis. The predictive analysis of quantum computing development presented in Paper C indicates that the cryptographic primitives currently relied upon by blockchain systems may become vulnerable within a projected timeframe, underscoring the importance of entropy sources that can support the larger key sizes typically required by post-quantum cryptographic schemes, as discussed in Section 2.4. Paper F extends this concern to business intelligence systems built on blockchain infrastructure, proposing a migration strategy toward post-quantum security that similarly depends on the availability of sufficient, locally generated randomness. Paper D, while primarily concerned with the broader problem of consensus mechanism selection in heterogeneous distributed networks, further illustrates the dependence

of distributed system design on node-level resource constraints, a dependence that TC3 addresses specifically with respect to cryptographic randomness.

The validation of TC3 is subject to notable limitations. The experimental evaluation was conducted using a single smartwatch model and a single participant, which constrains the generalizability of the results across different hardware platforms and user populations. Furthermore, the NIST SP 800-90B evaluation conducted in Paper B applied a subset of the full recommended test battery, sufficient to establish the statistical quality of the entropy source for the purposes of this thesis, but not equivalent to a complete formal certification process. These limitations, together with the single-source nature of the entropy acquisition method discussed above, define the boundaries within which the contribution of TC3 should be interpreted.

4.4 Cross-Cutting Themes and Synthesis

While the three research goals of this thesis address distinct technical problems, topology stability, load balancing, and entropy generation, the contributions developed to address them share a common methodological foundation and respond to a shared underlying motivation: the increasing heterogeneity of distributed networks, exemplified throughout this thesis by blockchain systems, and the resource constraints this heterogeneity imposes on a growing proportion of participating nodes.

The first cross-cutting theme concerns the methodological choice of greedy algorithms. TC1 and TC2 both rely directly on greedy graph coloring, while TC3 was motivated by the same underlying principle of computational efficiency, favoring a lightweight, locally executable acquisition and processing pipeline over more resource-intensive alternatives. Across all three contributions, the recurring design decision is to favor a method that is computationally inexpensive and deployable on resource-constrained nodes, even where this involves a trade-off against the solution quality achievable by more demanding heuristic, metaheuristic, or hardware-based alternatives. This shared design philosophy reflects the central premise established in Section 1.1: that heterogeneous distributed networks increasingly include nodes, such as wearables, embedded sensors, and other low-power devices, for which computationally demanding methods are simply not viable.

The second cross-cutting theme concerns the relationship between structural and security-oriented optimization. TC1 and TC2 address the structural robustness of a distributed network, namely its resilience to node saturation and its ability to distribute computational load evenly. TC3 addresses a distinct but related concern: the security of cryptographic operations performed by individual nodes within such a network. These two concerns, robustness and security, correspond directly to the two pillars named in the title of this thesis, “Optimisation Strategies for Enhancing Robustness and Security in Distributed Systems”.

TC1 and TC2 contribute to robustness by improving the structural and operational properties of the network as a whole, while TC3 contributes to security by enabling individual resource-constrained nodes to participate safely in cryptographic protocols that the network as a whole depends upon. The supporting publications, examined in Section 4.3, further situate this security contribution within the broader and longer-term security context of consensus mechanism design and post-quantum cryptographic readiness.

The third cross-cutting theme concerns the role of blockchain networks as the unifying motivating context for this thesis. Although the three contributions are formulated in terms of general heterogeneous distributed networks and are not restricted to blockchain applications, blockchain systems served throughout this research as the concrete domain in which the consequences of node heterogeneity were first identified and examined, as discussed in Section 1.1 and Section 2.1. This grounding in a specific, practically significant application domain distinguishes the contributions of this thesis from purely theoretical treatments of graph coloring or entropy generation, while the general formulation of TC1, TC2, and TC3 ensures that the proposed methods remain applicable beyond blockchain systems, to any heterogeneous distributed network exhibiting comparable structural and resource constraints.

Taken together, TC1, TC2, and TC3 represent three complementary optimisation strategies, unified by a common methodological commitment to computational efficiency and motivated by a common observation regarding the structural and security consequences of node heterogeneity in modern distributed networks.

4.5 Limitations

While the contributions presented in this thesis address the research goals defined in Chapter 1, several limitations should be acknowledged in order to correctly interpret the scope and applicability of the findings.

The first limitation concerns the simulation-based nature of the experiments underlying TC1 and TC2. The topology stability assessment model and the load balancing method were evaluated exclusively on synthetically generated network topologies of 100 nodes, with edge densities of 1000, 1500, and 2000 edges. While this range was selected to represent a realistic operating envelope for heterogeneous distributed networks, as discussed in Section 3.1, the simulated networks do not capture the dynamic behavior of an operational blockchain network, including node churn, asynchronous communication delays, and adversarial conditions. The generalizability of the findings to live distributed networks, or to networks of substantially different scale, has not been established and remains a direction for future research.

The second limitation concerns the load balancing method presented as TC2, which assumes that nodes within a given color class are computationally interchangeable for the

purpose of task assignment. This assumption does not account for the heterogeneous processing capacities discussed in Section 2.1, where nodes within a single network may range from resource-constrained wearables to high-performance computing infrastructure. Incorporating node-specific capacity weights into the coloring-based load balancing procedure would more accurately reflect the heterogeneous nature of the networks motivating this thesis but was not addressed within the scope of the present work.

The third limitation concerns the experimental validation of TC3, which was conducted using a single smartwatch model and a single participant. This constrains the generalizability of the results across different hardware platforms, sensor configurations, and user populations. In addition, the entropy source validated in Paper B draws from a single device and therefore constitutes a single-source, single-modality acquisition method, which does not satisfy the multi-source independence criteria defined by NIST SP 800-90B for entropy sources intended for high-assurance cryptographic applications. The NIST SP 800-90B evaluation conducted in Paper B applied a subset of the full recommended test battery, sufficient to establish the statistical quality of the entropy source for the purposes of this thesis, but not equivalent to a complete formal certification process.

The fourth limitation concerns the supporting publications addressing quantum computing threats and post-quantum migration, presented in Papers C and F. The predictive models developed in these publications are based on the projected capacity of quantum computing hardware under defined fault-tolerance scenarios, rather than on the demonstrated execution of cryptographically relevant quantum algorithms against operational systems. These models therefore provide a capacity-based projection of future risk rather than an empirically validated account of current vulnerability, and the projected timelines are subject to the inherent uncertainty of forecasting hardware development trends.

Taken together, these limitations indicate that the contributions of this thesis should be understood as establishing the feasibility and validity of the proposed methods under controlled experimental conditions, rather than as a complete, production-ready solution for deployment in operational heterogeneous distributed networks. Addressing these limitations through expanded experimental validation, multi-source entropy architectures, and capacity-weighted load balancing constitutes a natural continuation of the research presented in this thesis, as discussed in Chapter 5.

5

Conclusions and Future Work

This chapter summarizes the main findings of the thesis and reflects on the extent to which the research goals defined in Chapter 1 have been achieved. Section 5.1 presents the conclusions drawn from the three original scientific contributions, situating them within the broader context of heterogeneous distributed systems research. Section 5.2 outlines four directions for future research that build directly on the methods and findings presented in this thesis.

5.1 Conclusions

This thesis set out to develop optimization strategies for enhancing robustness and security in heterogeneous distributed systems, addressed through three original scientific contributions, each corresponding to one of the three research goals defined in Chapter 1. With respect to RG1, TC1 established a topology stability assessment model for heterogeneous distributed networks based on node color distribution. Evaluated on simulated 100-node networks across a range of edge densities, the model demonstrated that the number of colors required and the evenness of color class distribution serve as practical, computationally inexpensive indicators of structural balance, with denser networks consistently requiring fewer colors relative to their node count and exhibiting more even color class distributions, in line with hypothesis H1 [62]. With respect to RG2, TC2 extended the same greedy coloring procedure to node load balancing, repurposing color classes as non-conflicting groups for parallel task assignment. The evaluation further showed that increasing edge density from 1000 to 2000 edges produced a measurable improvement in load distribution evenness, while further increases yielded only

marginal gains relative to the additional resource overhead incurred, identifying an edge density of approximately 2000 as a practical operating point for the simulated 100-node networks studied, in line with hypothesis H2 [62]. Together, TC1 and TC2 demonstrate that a single, computationally inexpensive greedy coloring operation can inform both the structural stability and the load balancing of a heterogeneous distributed network, without requiring separate, more computationally demanding optimization procedures. With respect to RG3, TC3 established an offline entropy harvesting method from wearable device sensor data, implemented and validated on a commercial smartwatch under two acquisition modes. The shake (motion-triggered) mode achieved a Shannon entropy of 0.997 and a min-entropy of 0.918, outperforming the still (passive) mode, which achieved 0.991 and 0.851 respectively, confirming hypothesis H3 and demonstrating that commercially available wearable devices can serve as practical, self-contained entropy sources for cryptographic key generation without reliance on dedicated hardware or continuous network connectivity. Consistent with the principles of single-source entropy evaluation under NIST SP 800-90B, the contribution of TC3 is understood as establishing the viability of a single wearable device as one such source, rather than as a complete multi-source entropy architecture [64]. The three contributions are unified by a shared methodological commitment to computational efficiency and a shared motivation in the increasing heterogeneity of modern distributed networks, examined throughout this thesis primarily through the lens of blockchain systems. TC1 and TC2 address the robustness pillar named in the title of this thesis, while TC3 addresses its security pillar, together demonstrating that meaningful improvements in the structural and security properties of heterogeneous distributed networks can be achieved using methods that remain computationally tractable for resource-constrained nodes. The four supporting publications included in this thesis, addressing quantum computing threat timelines, consensus mechanism selection, smart city applications of heterogeneous networks, and post-quantum migration strategies, situate these three contributions within a broader and longer-term security context, reinforcing their relevance to the continued evolution of distributed systems toward greater node heterogeneity and toward post-quantum cryptographic readiness.

Beyond their individual contributions, TC1, TC2, and TC3 collectively address a fundamental tension in the design of heterogeneous distributed networks: the need to maintain system-level robustness and security while accommodating nodes with severely constrained computational resources. By demonstrating that graph-theoretic and sensor-based methods can deliver practically useful results within these constraints, this thesis contributes to a growing body of work that challenges the assumption that robust and secure distributed systems necessarily require homogeneous, high-capacity infrastructure. The findings suggest that lightweight, locally executable methods, grounded in well-understood algorithmic principles, can meaningfully extend the security and resilience of distributed networks to include the full diversity of their participating nodes.

5.2 Future Work

FW1: The topology stability and load balancing models developed as TC1 and TC2 were evaluated exclusively on simulated network topologies. Future work will implement an actual private blockchain network based on the conclusions of this thesis, in particular the identified edge-density operating point, and validate its performance under real-world conditions, including fault tolerance, latency, and throughput, on live, operational infrastructure. **FW2:** The entropy harvesting method developed as TC3 was validated using a single smartwatch model and a single participant. Future work will extend this evaluation to multiple users and multiple smartwatch models, apply the full NIST SP 800-90B test battery beyond the subset used in this thesis, and conduct adversarial testing to assess the robustness of the proposed method against deliberate attempts to predict or manipulate the harvested entropy. **FW3:** Building on the entropy harvesting capacity demonstrated in TC3 and its compatibility with post-quantum key generation discussed in Section 2.4, future work will explore the development of a proof-of-concept smartwatch prototype capable of generating post-quantum secure digital signatures directly on-device. This prototype will integrate post-quantum cryptographic primitives, in particular the Module-Lattice-Based Digital Signature Algorithm (ML-DSA, CRYSTALS-Dilithium), seeded using the offline entropy harvesting method developed in this thesis, thereby extending the offline, hardware-independent design principles established for TC3 toward practical post-quantum key generation on resource-constrained wearable devices. **FW4:** The greedy coloring-based methods developed as TC1 and TC2 were evaluated on static network topologies of fixed size. Future work will apply these optimization methods to dynamic networks with changing topology, node churn, and larger network configurations, examining whether the structural stability and load balancing properties established in this thesis are maintained as networks evolve over time and scale beyond the 100-node networks studied.

References

[1] Ant Colony Optimization - an overview | ScienceDirect Topics. URL https://www.sciencedirect.com/topics/engineering/ant-colony-optimization.

[2] Distributed Algorithms: | Guide books | ACM Digital Library. URL https://dl.acm.org/doi/book/10.5555/2821576.

[3] ISO/IEC 18031:2025(en), Information technology — Security techniques — Random bit generation. URL https://www.iso.org/obp/ui/en/#iso:std:iso-iec:18031:ed-3:v1:en.

[4] NP-hard Problem - an overview | ScienceDirect Topics. URL https://www.sciencedirect.com/topics/mathematics/np-hard-problem.

[5] K. A. Ajlan, T. Alsboui, O. Alshaikh, I. Inuwa-Dute, S. Khan, and S. Parkinson. The Emerging Challenges of Wearable Biometric Cryptosystems. *Cryptography*, 8(3):27, Sept. 2024. ISSN 2410-387X. doi: 10.3390/cryptography8030027. URL https://www.mdpi.com/2410-387X/8/3/27. Number: 3.

[6] M. Aslan and N. Baykan. A Performance Comparison of Graph Coloring Algorithms. *International Journal of Intelligent Systems and Applications in Engineering*, 4:1–1, Dec. 2016. doi: 10.18201/ijisae.273053.

[7] L. Atzori, A. Iera, and G. Morabito. The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805, Oct. 2010. ISSN 1389-1286. doi: 10.1016/j.comnet.2010.05.010. URL https://www.sciencedirect.com/science/article/pii/S1389128610001568.

[8] S. M. H. Bamakan, A. Motavali, and A. Babaei Bondarti. A survey of blockchain consensus algorithms performance evaluation criteria. *Expert Systems with Applications*, 154:113385, Sept. 2020. ISSN 09574174. doi: 10.1016/j.eswa.2020.113385. URL https://linkinghub.elsevier.com/retrieve/pii/S0957417420302098.

[9] L. Barenboim and M. Elkin. *Distributed Graph Coloring: Fundamentals and Recent Developments*. Synthesis Lectures on Distributed Computing Theory. Springer

International Publishing, Cham, 2013. ISBN 978-3-031-00881-8 978-3-031-02009-4. doi: 10.1007/978-3-031-02009-4. URL https://link.springer.com/10.1007/978-3-031-02009-4.

[10] E. Barker and J. Kelsey. Recommendation for Random Number Generation Using Deterministic Random Bit Generators. Technical Report NIST Special Publication (SP) 800-90A Rev. 1, National Institute of Standards and Technology, June 2015. URL https://csrc.nist.gov/pubs/sp/800/90/a/r1/final.

[11] Y. Baseri, A. Hafid, Y. Shahsavari, D. Makrakis, and H. Khodaiemehr. Blockchain Security Risk Assessment in Quantum Era, Migration Strategies and Proactive Defense. *IEEE Communications Surveys & Tutorials*, 28:2925–2964, 2026. ISSN 1553-877X, 2373-745X. doi: 10.1109/COMST.2025.3621113. URL http://arxiv.org/abs/2501.11798. arXiv:2501.11798 [cs].

[12] D. J. Bernstein. Introduction to post-quantum cryptography. In D. J. Bernstein, J. Buchmann, and E. Dahmen, editors, *Post-Quantum Cryptography*, pages 1–14. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-88701-0 978-3-540-88702-7. doi: 10.1007/978-3-540-88702-7_1. URL http://link.springer.com/10.1007/978-3-540-88702-7_1.

[13] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten. SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies. In *2015 IEEE Symposium on Security and Privacy*, pages 104–121, May 2015. doi: 10.1109/SP.2015.14. URL https://ieeexplore.ieee.org/document/7163021.

[14] P. Borowiecki. Computational aspects of greedy partitioning of graphs. *JOURNAL OF COMBINATORIAL OPTIMIZATION*, 35(2):641–665, Feb. 2018. ISSN 1382-6905, 1573-2886. doi: 10.1007/s10878-017-0185-2. URL https://www.webofscience.com/wos/woscc/full-record/WOS:000424670700022. Num Pages: 25 Place: Dordrecht Web of Science ID: WOS:000424670700022.

[15] M. Castro. Practical Byzantine Fault Tolerance. Apr. 2001.

[16] R. Cohen and S. Havlin. *Complex Networks: Structure, Robustness and Function*. Aug. 2010. ISBN 978-0-521-84156-6. doi: 10.1017/CBO9780511780356. Journal Abbreviation: Complex Networks: Structure, Robustness and Function Publication Title: Complex Networks: Structure, Robustness and Function.

[17] Y. Dodis, D. Pointcheval, S. Ruhault, D. Vergnaud, and D. Wichs. Security Analysis of Pseudo-Random Number Generators with Input: /dev/random is not Robust, 2013. URL https://eprint.iacr.org/2013/338. Publication info: Published elsewhere.

Minor revision. ACM Conference on Computer and Communication Security (CCS), November 2013.

[18] D. E. Eastlake 3rd, S. Crocker, and J. I. Schiller. Randomness Requirements for Security. Request for Comments RFC 4086, Internet Engineering Task Force, June 2005. URL https://datatracker.ietf.org/doc/rfc4086. Num Pages: 48.

[19] A. Fedorov, E. Kiktenko, and A. Lvovsky. Quantum computers put blockchain security at risk. *Nature*, 563:465–467, Nov. 2018. doi: 10.1038/d41586-018-07449-z.

[20] Z. Gutterman, B. Pinkas, and T. Reinman. Analysis of the Linux random number generator. In *2006 IEEE Symposium on Security and Privacy (S&P'06)*, pages 15 pp.–385, May 2006. doi: 10.1109/SP.2006.5. URL https://ieeexplore.ieee.org/document/1624027.

[21] S. Islam, M. J. Islam, M. Hossain, S. Noor, K.-S. Kwak, and S. M. R. Islam. A Survey on Consensus Algorithms in Blockchain-Based Applications: Architecture, Taxonomy, and Operational Issues. *IEEE Access*, 11:39066–39082, 2023. ISSN 2169-3536. doi: 10.1109/ACCESS.2023.3267047. Conference Name: IEEE Access.

[22] R. Janczewski, M. Kubale, K. Manuszewski, and K. Piwakowski. The smallest hard-to-color graph for algorithm DSATUR. *Discrete Mathematics*, 236(1-3):151–165, June 2001. ISSN 0012365X. doi: 10.1016/S0012-365X(00)00439-8. URL https://linkinghub.elsevier.com/retrieve/pii/S0012365X00004398.

[23] G. Jayabalasamy, C. Pujol, and K. Latha Bhaskaran. Application of Graph Theory for Blockchain Technologies. *Mathematics*, 12(8):1133, Jan. 2024. ISSN 2227-7390. doi: 10.3390/math12081133. URL https://www.mdpi.com/2227-7390/12/8/1133. Number: 8.

[24] A. Kristoffersson and M. Lindén. A Systematic Review of Wearable Sensors for Monitoring Physical Activity. *Sensors*, 22(2):573, Jan. 2022. ISSN 1424-8220. doi: 10.3390/s22020573. URL https://www.mdpi.com/1424-8220/22/2/573. Number: 2.

[25] N. Kshetri. Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy*, 41(10):1027–1038, Nov. 2017. ISSN 0308-5961. doi: 10.1016/j.telpol.2017.09.003. URL https://www.sciencedirect.com/science/article/pii/S0308596117302483.

[26] I. Lukić, M. Köhler, Z. Krpić, and M. Švarcmajer. Advancing Smart City Sustainability Through Artificial Intelligence, Digital Twin and Blockchain Solutions. *Technologies*, 13(7), July 2025. ISSN 2227-7080. doi: 10.3390/technologies13070300. URL https://www.mdpi.com/2227-7080/13/7/300.

[27] N. Lv, T. Chen, and Y. Ma. Analysis on Entropy Sources based on Smartphone Sensors. In *Proceedings of the 2020 10th International Conference on Communication and Network Security*, ICCNS '20, pages 21–31, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 978-1-4503-8903-7. doi: 10.1145/3442520.3442528. URL https://dl.acm.org/doi/10.1145/3442520.3442528.

[28] Y. Ma, M. Sun, W. Wang, T. Chen, N. Lv, and D. Han. Efficient and Accurate Min-entropy Estimation Based on Decision Tree for Random Number Generators. In *2024 IEEE 23rd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 707–716, Dec. 2024. doi: 10.1109/TrustCom63139.2024.00110. URL https://ieeexplore.ieee.org/document/10945165. ISSN: 2324-9013.

[29] A. Mansuri, V. Gupta, and R. S. Chandel. Coloring Programs in Graph Theory.

[30] N. Meghanathan, editor. *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks*. Information Science Reference, Hershey, Pennsylvania (701 E. Chocolate Avenue, Hershey, Pennsylvania, 17033, USA), 1st edition edition, July 2017. ISBN 978-1-5225-2814-2.

[31] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. URL https://bitcoin.org/bitcoin.pdf.

[32] C. T. Nguyen, D. T. Hoang, D. N. Nguyen, D. Niyato, H. T. Nguyen, and E. Dutkiewicz. Proof-of-Stake Consensus Mechanisms for Future Blockchain Networks: Fundamentals, Applications and Opportunities. *IEEE Access*, 7:85727–85745, 2019. ISSN 2169-3536. doi: 10.1109/ACCESS.2019.2925010. URL https://ieeexplore.ieee.org/document/8746079/.

[33] N. Niroumandrad, N. Lahrichi, and A. Lodi. Learning tabu search algorithms: A scheduling application. *Computers & Operations Research*, 170:106751, Oct. 2024. ISSN 0305-0548. doi: 10.1016/j.cor.2024.106751. URL https://www.sciencedirect.com/science/article/pii/S0305054824002235.

[34] K. Nurmalika, R. Prihandini, D. Jannah, I. Makhfurdloh, A. Agatha, and Y. Wulandari. Graph Coloring Application Using Welch Powell Algorithm On Radio Frequency In Australia. June 2024.

[35] S. Olariu. An optimal greedy heuristic to color interval graphs. *Information Processing Letters*, 37(1):21–25, Jan. 1991. ISSN 00200190. doi: 10.1016/0020-0190(91)90245-D. URL https://linkinghub.elsevier.com/retrieve/pii/002001909190245D.

[36] S. Olariu and J. Randall. Welsh-Powell opposition graphs. *Information Processing Letters*, 31(1):43–46, Apr. 1989. ISSN 0020-0190. doi: 10.1016/0020-0190(89)90107-5. URL https://www.sciencedirect.com/science/article/pii/0020019089901075.

[37] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkeviciūtė, C. Suo, and A. Fornito. Consistency and differences between centrality measures across distinct classes of networks. *PLOS ONE*, 14(7):e0220061, 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0220061. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0220061.

[38] A. Ometov, V. Shubina, L. Klus, J. Skibińska, S. Saafi, P. Pascacio, L. Flueratoru, D. Q. Gaibor, N. Chukhno, O. Chukhno, A. Ali, A. Channa, E. Svertoka, W. B. Qaim, R. Casanova-Marqués, S. Holcer, J. Torres-Sospedra, S. Casteleyn, G. Ruggeri, G. Araniti, R. Burget, J. Hosek, and E. S. Lohan. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. *Computer Networks*, 193:108074, July 2021. ISSN 1389-1286. doi: 10.1016/j.comnet.2021.108074. URL https://www.sciencedirect.com/science/article/pii/S1389128621001651.

[39] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson. The Rise of Consumer Health Wearables: Promises and Barriers. *PLOS Medicine*, 13(2):e1001953, 2016. ISSN 1549-1676. doi: 10.1371/journal.pmed.1001953. URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001953.

[40] J. Postigo, J. Soto-Begazo, V. R. Fiorela, G. M. Picha, R. Flores-Quispe, and Y. Velazco-Paredes. Comparative Analysis of the main Graph Coloring Algorithms. In *2021 IEEE Colombian Conference on Communications and Computing (COLCOM)*, pages 1–6, May 2021. doi: 10.1109/COLCOM52710.2021.9486299. URL https://ieeexplore.ieee.org/document/9486299/?arnumber=9486299.

[41] F. Răstoceanu, R. Rughiniş, Ş.-D. Ciocîrlan, and M. Enache. Sensor-Based Entropy Source Analysis and Validation for Use in IoT Environments. *Electronics*, 10(10):1173, Jan. 2021. ISSN 2079-9292. doi: 10.3390/electronics10101173. URL https://www.mdpi.com/2079-9292/10/10/1173.

[42] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, N. Heckert, J. Dray, S. Vo, and L. Bassham. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Technical Report NIST Special Publication (SP) 800-22 Rev. 1, National Institute of Standards and Technology, Apr. 2010. URL https://csrc.nist.gov/pubs/sp/800/22/r1/upd1/final.

[43] P. San Segundo. A new DSATUR-based algorithm for exact vertex coloring. *Computers & Operations Research*, 39(7):1724–1733, July 2012. ISSN 0305-0548. doi: 10.1016/j.cor.2011.10.008. URL https://www.sciencedirect.com/science/article/pii/S0305054811002966.

[44] E. Sazonov and M. Neuman. *Wearable Sensors: Fundamentals, Implementation and Applications*. Jan. 2014. Pages: 634.

[45] K. Schärer and M. Comuzzi. The quantum threat to blockchain: summary and timeline analysis. *Quantum Machine Intelligence*, 5, Apr. 2023. doi: 10.1007/s42484-023-00105-4.

[46] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948. ISSN 0005-8580. doi: 10.1002/j.1538-7305.1948.tb01338.x. URL https://ieeexplore.ieee.org/document/6773024.

[47] C. Shi, Y. Li, J. Zhang, Y. Sun, and P. S. Yu. A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering*, 29(1):17–37, Jan. 2017. ISSN 1558-2191. doi: 10.1109/TKDE.2016.2598561. URL https://ieeexplore.ieee.org/document/7536145.

[48] P. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, Nov. 1994. doi: 10.1109/SFCS.1994.365700. URL https://ieeexplore.ieee.org/document/365700.

[49] P. W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Review*, 41:303–332, Jan. 1999. ISSN 0036-1445. doi: 10.1137/S0036144598347011. URL https://ui.adsabs.harvard.edu/abs/1999SIAMR..41..303S. ADS Bibcode: 1999SIAMR..41..303S.

[50] C. C. Sobin. A Survey on Architecture, Protocols and Challenges in IoT. *Wireless Personal Communications*, 112(3):1383–1429, June 2020. ISSN 1572-834X. doi: 10.1007/s11277-020-07108-5. URL https://doi.org/10.1007/s11277-020-07108-5.

[51] B. Sunar, W. J. Martin, and D. R. Stinson. A Provably Secure True Random Number Generator with Built-In Tolerance to Active Attacks. *IEEE Transactions on Computers*, 56(1):109–119, Jan. 2007. ISSN 1557-9956. doi: 10.1109/TC.2007.250627. URL https://ieeexplore.ieee.org/document/4016501.

[52] M. Sönmez Turan, E. Barker, J. Kelsey, K. McKay, M. Baish, and M. Boyle. Recommendation for the Entropy Sources Used for Random Bit Generation. Technical

Report NIST Special Publication (SP) 800-90B, National Institute of Standards and Technology, Jan. 2018. URL https://csrc.nist.gov/pubs/sp/800/90/b/final.

[53] A. S. Tanenbaum and M. Van Steen. *Distributed Systems: Principles and Paradigms*. Prentice Hall PTR, Upper Saddle River, NJ, United States, 2001. ISBN 978-0-13-088893-8.

[54] N. I. o. S. a. Technology. Module-Lattice-Based Key-Encapsulation Mechanism Standard. Technical Report Federal Information Processing Standard (FIPS) 203, U.S. Department of Commerce, Aug. 2024. URL https://csrc.nist.gov/pubs/fips/203/final.

[55] N. I. o. S. Technology (NIST), , T. Dang, J. Lichtinger, Y.-K. Liu, C. Miller, D. Moody, R. Peralta, R. Perlner, and A. Robinson. Module-Lattice-Based Digital Signature Standard. *NIST*, Aug. 2024. URL https://www.nist.gov/publications/module-lattice-based-digital-signature-standard. Last Modified: 2024-08-13T13:08-04:00.

[56] Y. Wang. Review on greedy algorithm. *Theoretical and Natural Science*, 14:233–239, Nov. 2023. doi: 10.54254/2753-8818/14/20241041.

[57] W. S. Werner Killmann. A Proposal for Functionality Classes for Random Number Generators, 2011. URL https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/AIS_31_e.

[58] R. J. Wilson. *Introduction to Graph Theory*. Longman, 2010. ISBN 978-0-273-72889-4. Google-Books-ID: wwxTRAAACAAJ.

[59] H. Xue, K. T. Kim, and H. Y. Youn. Dynamic Load Balancing of Software-Defined Networking Based on Genetic-Ant Colony Optimization. *Sensors*, 19(2), Jan. 2019. ISSN 1424-8220. doi: 10.3390/s19020311. URL https://www.mdpi.com/1424-8220/19/2/311.

[60] H. Yu and Y. Kim. True Random Number Generator Using Bio-related Signals in Wearable Devices. In *2018 International SoC Design Conference (ISOCC)*, pages 231–232, Nov. 2018. doi: 10.1109/ISOCC.2018.8649804. URL https://ieeexplore.ieee.org/document/8649804/.

[61] S. Zhou, K. Li, L. Xiao, J. Cai, W. Liang, and A. Castiglione. A Systematic Review of Consensus Mechanisms in Blockchain. *Mathematics*, 11(10):2248, Jan. 2023. ISSN 2227-7390. doi: 10.3390/math11102248. URL https://www.mdpi.com/2227-7390/11/10/2248. Number: 10.

[62] M. Švarcmajer, D. Ivanović, T. Rudec, and I. Lukić. Application of Graph Theory and Variants of Greedy Graph Coloring Algorithms for Optimization of Distributed Peer-to-Peer Blockchain Networks. *Technologies*, 13(1), Jan. 2025. ISSN 2227-7080. doi: 10.3390/technologies13010033. URL https://www.mdpi.com/2227-7080/13/1/33.

[63] M. Švarcmajer, T. Krčmar, D. Šabanović, and I. Lukić. Designing Post-Quantum Secure Blockchain Infrastructure for Business Intelligence Systems: Migration Strategies and Cost Optimization. In *2025 International Symposium ELMAR*, pages 245–248, Sept. 2025. doi: 10.1109/ELMAR66948.2025.11193746. URL https://ieeexplore.ieee.org/document/11193746. ISSN: 2835-3781.

[64] M. Švarcmajer, M. Köhler, Z. Krpić, and I. Lukić. Entropy Extraction from Wearable Sensors for Secure Cryptographic Key Generation in Blockchain and IoT Systems. *Sensors*, 25(17):5298, Jan. 2025. ISSN 1424-8220. doi: 10.3390/s25175298. URL https://www.mdpi.com/1424-8220/25/17/5298.

[65] M. Švarcmajer, M. Köhler, J. Marinčić, and J. Perić. Utilizing Faculty Computers for Cryptocurrency Validation and Mining During Idle Time. In H. Glavaš, M. Hadzima-Nyarko, N. Ademović, and T. Hanák, editors, *33rd International Conference on Organization and Technology of Maintenance (OTO 2024)*, pages 121–134, Cham, 2025. Springer Nature Switzerland. ISBN 978-3-031-80597-4. doi: 10.1007/978-3-031-80597-4_10.

[66] M. Švarcmajer, M. Kohler, Z. Krpić, and I. Lukić. Consensus-Driven Framework for Data-Driven Optimization of Distributed Systems Through Blockchain Consensus Mechanism Selection. *Big Data and Cognitive Computing*, 10(5), May 2026. ISSN 2504-2289. doi: 10.3390/bdcc10050154. URL https://www.mdpi.com/2504-2289/10/5/154.

[67] M. Švarcmajer, D. Šabanović, M. Köhler, and I. Lukić. Predictive Modeling of Quantum Computing Development and its Impact on Blockchain Security. *IEEE Access*, 14: 68624–68636, 2026. ISSN 2169-3536. doi: 10.1109/ACCESS.2026.3689963. URL https://ieeexplore.ieee.org/document/11505837.

List of Publications

Main Contributing Publications

M. Švarcmajer, D. Ivanović, T. Rudec, and I. Lukić, "Application of Graph Theory and Variants of Greedy Graph Coloring Algorithms for Optimization of Distributed Peer-to-Peer Blockchain Networks," *Technologies*, vol. 13, no. 1, p. 33, Jan. 2025.

Author's contribution: First author. Conceived and implemented the graph coloring methodology, designed and executed the simulations, analysed the results, and wrote the manuscript.

M. Švarcmajer, M. Köhler, Z. Krpić, and I. Lukić, "Entropy Extraction from Wearable Sensors for Secure Cryptographic Key Generation in Blockchain and IoT Systems," *Sensors*, vol. 25, no. 17, p. 5298, Aug. 2025.

Author's contribution: First author. Developed the entropy harvesting method, implemented the data collection application, conducted measurements, performed the NIST SP 800-90B evaluation, and wrote the manuscript.

Supporting Publications

M. Švarcmajer, D. Šabanović, M. Köhler, and I. Lukić, "Predictive Modeling of Quantum Computing Development and Its Impact on Blockchain Security," *IEEE Access*, vol. 14, pp. 68624–68636, May 2026.

Author's contribution: First author. Developed the predictive model, collected and processed data, conducted the analysis, and wrote the manuscript.

M. Švarcmajer, M. Köhler, Z. Krpić, and I. Lukić, "Consensus-Driven Framework for Data-Driven Optimization of Distributed Systems Through Blockchain Consensus Mechanism Selection," *Big Data and Cognitive Computing*, vol. 10, no. 5, p. 154, May 2026.

Author's contribution: First author. Designed the CD-7 framework, developed the selection methodology, validated the framework through application scenarios, and wrote the manuscript.

I. Lukić, M. Köhler, Z. Krpić, and M. Švarcmajer, "Advancing Smart City Sustainability Through Artificial Intelligence, Digital Twin and Blockchain Solutions," *Technologies*, vol. 13,

no. 7, p. 300, 2025.

Author's contribution: Third author. Contributed to the blockchain architecture section and participated in manuscript review.

M. Švarcmajer, T. Krčmar, D. Šabanović, and I. Lukić, "Designing Post-Quantum Secure Blockchain Infrastructure for Business Intelligence Systems: Migration Strategies and Cost Optimization," in *Proc. 2025 International Symposium ELMAR*, pp. 245–248, IEEE, Sep. 2025.

Author's contribution: First author. Proposed the post-quantum migration framework, developed the cost optimization model, and wrote the manuscript.

A

Paper A: Application of Graph Theory and Variants of Greedy Graph Coloring Algorithms for Optimization of Distributed Peer-to-Peer Blockchain Networks

Technologies, vol. 13, no. 1, p. 33, January 2025 | DOI: 10.3390/technologies13010033

Miljenko Švarcmajer, Denis Ivanović, Tomislav Rudec and Ivica Lukić
Faculty of Electrical Engineering, Computer Science and Information Technology, Josip Juraj Strossmayer University of Osijek, Kneza Trpimira 2b, 31000 Osijek, Croatia

Abstract

This paper investigates the application of graph theory and variants of greedy graph coloring algorithms for the optimization of distributed peer-to-peer networks, with a special focus on private blockchain networks. The graph coloring problem, as an NP-hard problem, presents a challenge in determining the minimum number of colors needed to efficiently allocate resources within the network. The paper deals with the influence of different graph density, i.e., the number of links, on the efficiency of greedy algorithms such as DSATUR, Descending, and Ascending. Experimental results show that increasing the number of

links in the network contributes to a more uniform distribution of colors and increases the resistance of the network, whereby the DSATUR algorithm achieves the most uniform color saturation. The optimal configuration for a 100-node network has been identified at around 2000 to 2500 links, which achieves stability without excessive redundancy. These results are applied in the context of a private blockchain network that uses optimal connectivity to achieve high resilience and efficient resource allocation. The research findings suggest that adapting network configuration using greedy algorithms can contribute to the optimization of distributed systems, making them more stable and resilient to loads.

Keywords: greedy algorithms; graph coloring; DSATUR; distributed systems; peer-to-peer networks; connectivity optimization; private blockchain network; graph theory

1 Introduction

The question of the complexity of algorithms is one of the key ones in the theory of computing, and all problems that can be solved in polynomial time belong to class P [1]. On the other hand, the set of all problems for which we can check the correctness of the solution in polynomial time is denoted by NP. One of the most important open questions in the theory of computing is the following: is P = NP? Among the most famous NP problems is the problem of coloring graphs, which asks the following question: with how many colors can we color the vertices of the graph so that no edge has ends of the same color? The smallest number of colors satisfying this condition is called the chromatic number of the graph [2].

Graph coloring is widely used in computer science, especially in optimization and resource scheduling. For example, when allocating registers in compilers, graph coloring enables the efficient allocation of registers to variables in order to minimize the number of required registers and avoid collisions between variables [3]. In task scheduling, graph coloring helps to allocate tasks to a limited number of processors without conflicts, which is crucial for the optimization of parallel systems. Also, in network theory, graph coloring enables the allocation of frequencies to base stations in telecommunication networks to avoid interference between neighboring stations, thus optimizing spectrum utilization. It is also widely used in distributed computer networks [4] such as blockchain networks [5]. Furthermore, graph coloring is also used in exam scheduling, ensuring that students taking the same subjects do not overlap in terms [6].

One of the basic approaches to solving graph coloring problems is greedy algorithms, which work by assigning to each vertex the first available color that does not cause a conflict with neighboring vertices. Greedy algorithms are popular due to their simplicity and speed, which makes them suitable for application in situations where a fast response is required, such as task scheduling or register allocation [3,6]. However, they often use more colors than

is optimal, making them suboptimal when trying to minimize the number of colors needed to color a graph [7].

Existing blockchain solutions such as VeDB [8] and LedgerDB [9] enable high resistance to data changes and superior performance through advanced mechanisms such as Trusted Execution Environment (TEE) and verification structures such as VSA (Verifiable Shrubs Array); their implementation often requires specialized hardware and significant resource capacities. Since the goal is to develop a cost-effective private blockchain network of small to medium size, the paper focuses on the application of graph theory and variants of greedy coloring algorithms to optimize connectivity and stability in distributed systems. Those variants are used because they are widely recognized for their simplicity and computational efficiency. The goal of the research is to also identify the optimal connectivity configuration that allows for even resource distribution, reduced congestion, and increased network resilience to failures. The results have practical applications in the design of private blockchain networks, where the stability and efficiency of communication between nodes play a key role in maintaining a reliable and scalable network.

While this study provides valuable insights into the optimization of distributed peer-to-peer blockchain networks using greedy graph coloring algorithms, several limitations should be noted. Greedy algorithms, while computationally efficient, may not always achieve the theoretically minimal number of colors, potentially leading to suboptimal resource allocation in some cases. Additionally, the results presented here are based on simulations for a network of 100 nodes with a specific range of link densities, which may limit their generalizability to networks of different sizes or configurations. Finally, the conclusions are derived from theoretical models, and real-world implementation might introduce unforeseen challenges related to dynamic network behavior or external factors.

In the second chapter of this paper, the greedy algorithm will be described, and its basic algorithm will be given, and then, the most frequently mentioned types will be processed. In the third chapter, related work will be described, that is, a comparison of algorithms. In the fourth chapter, the results of testing subtypes of greedy algorithms on graphs of different density will be analyzed. The fifth chapter will describe how to apply graphs in network design, followed by the sixth chapter, which will describe how to determine the optimal number of links in computer networks using variants of greedy algorithms.

The main contributions of this paper can be summarized as follows:

1. The paper investigates how different variants of greedy graph coloring algorithms, including DSATUR, Descending, and Ascending, can optimize the stability and resilience of distributed peer-to-peer networks;

2. An optimal connectivity threshold for networks of 100 nodes (2000–2500 edges) has been identified, which balances resilience and efficiency without excessive redundancy;

3. The concept of evenness of color distribution as a metric of network stability is introduced, providing a new way to assess node saturation and connectivity;

4. The paper contributes to a better understanding of the efficiency of various greedy algorithms in distributed systems, extending their applications to blockchain technology.

2 Greedy Algorithms

The graph coloring problem is NP-hard, which means that computing the optimal solution is very complex, especially for larger graphs [10]. One of the simplest and most widely used approaches to solving the graph coloring problem is the greedy algorithm. This algorithm processes vertices one by one, assigning them the smallest possible color that is not already used by neighboring vertices. The greedy algorithm for solving the graph vertex coloring problem works as follows:

1. The vertices of the graph are arranged in an arbitrary sequence;

2. The first vertex in the series is colored with color 1;

3. Other vertices are colored with the color of the smallest number for which it is valid that this vertex is not connected to another vertex of the same color;

4. It returns to step 3 until all vertices are painted.

Although the greedy algorithm is simple, it often uses more colors than is optimal, which means that it does not always find the best solution. However, its speed and ease of use make it useful in many practical situations. In practice, the order in which the vertices are processed significantly affects the efficiency of the greedy algorithm. Some strategies include sorting the vertices by degree (the number of edges connecting them) in ascending or descending order or using a random order [3]. Sorting vertices in descending order is also called the Welsh–Powell algorithm [11]. In addition to this, the DSATUR algorithm is also popular, which colors vertices according to the degree of saturation, that is, it looks at how many neighboring vertices are already colored [12].

2.1. Descending (Welsh–Powell) Algorithm

Welsh–Powell is a variant of the greedy algorithm that uses descending order to color the vertices of a graph. The vertices are first sorted according to their degree, i.e., the number of neighbors they have, so that the vertices with the most neighbors are colored first. After the vertices are sorted, coloring takes place in the standard greedy way: each vertex is assigned the first available color that does not cause a conflict with its neighbors.

The advantage of this algorithm is that it often uses fewer colors than the basic greedy algorithm, because it colors the most connected vertices first, which reduces the number of colors needed for the rest of the vertices. Its main disadvantage is that it requires additional time to sort the vertices before coloring, but overall gives better results compared to random or unordered greedy algorithms [11].

2.2. Ascending Algorithm

In the ascending greedy variant, the vertices are sorted according to the ascending order of degrees, that is, the number of neighbors each vertex has. The vertices with the smallest number of neighbors are colored first. After sorting, coloring takes place according to the standard greedy algorithm, where each vertex is assigned the first available color that does not cause a conflict with neighboring vertices.

The advantage of this variant is its simplicity and quick application. Vertices with a lower degree are processed first, which can facilitate coloring in specific situations. The disadvantage is using more colors than other variants because it leaves more complex vertices for later, when most colors are already occupied, which can lead to suboptimal coloring [13].

2.3. DSATUR Algorithm

The DSATUR (Degree of Saturation) algorithm also uses a greedy method but relies on the degree of saturation of the peak. The degree of saturation measures how many neighbors of a given vertex are already colored with different colors. In each step of the algorithm, the vertex with the highest degree of saturation is selected for coloring, i.e., the vertex with the most colored neighbors. If several vertices have the same degree of saturation, then the vertex with the highest degree (number of neighbors) is selected.

The advantage of this algorithm is that it often gives better solutions than other variants of greedy algorithms because it intelligently chooses vertices with the highest degree of saturation, which reduces conflicts and enables more efficient coloring. The disadvantage of this algorithm is that it requires more computation due to monitoring the degree of saturation, which makes it somewhat more complex and slower than classical greedy algorithms [12].

2.4. Random Algorithm

In this variant of the greedy algorithm, instead of sorting the vertices according to some strategy (such as the degree of the vertex in ascending or descending order), the vertices are colored in a random order. After the order of vertices is randomly selected, the algorithm proceeds with coloring using the classic greedy method: each vertex is assigned the first available color that does not cause a conflict with its neighbors. The advantage of this

method is that it is very simple to implement and can quickly generate solutions. Due to the random selection of the order of vertices, the random greedy algorithm often gives suboptimal results and uses more colors than variants that use strategic ordering (e.g., Welsh–Powell) and this is the biggest flaw. However, due to its simplicity, it can be useful in situations where speed is more important than optimality [14].

3 Related Work

The graph coloring problem has wide applications in computer science, including resource optimization, task scheduling, and frequency allocation. Over the years, scientists have developed numerous methods and algorithms to improve the efficiency of the solution. The research of different approaches and variants of greedy algorithms is a key part of this work.

Some papers, such as [15] from 2013, focus on developing new heuristics for specific types of graphs, such as interval graphs. This paper introduces an optimal greedy heuristic that ensures that graphs are colored with the optimal number of colors using a properly chosen coloring order. The authors conclude that such optimization enables better performance in specific graph structures, thereby significantly reducing the complexity of the algorithm in real scenarios. Similarly, Ref. [3] explores the complexity of graph partitioning using the greedy approach. The authors conclude that greedy methods are particularly useful in large and complex graphs where speed optimization is key and suggest their application in graphs with specific partitioning requirements.

Other papers, such as [16] from 2024, explore variants of greedy algorithms such as b-greedy and z-greedy. The authors conclude that these variants provide better results in graphs with a high degree of connectivity, where classical greedy algorithms are not efficient. These variants help in the more accurate coloring of complex graph structures, where the balance between the number of colors and the density of vertices plays a key role. Also, the paper [17] from 2016 compares several variants of greedy algorithms, including DSATUR and Welsh–Powell. It was concluded that the DSATUR algorithm is most suitable for graphs with high density, while Welsh–Powell achieves the best results in situations where the speed of the algorithm is key.

Papers like [7] from 2021 provide a comprehensive comparison of various methods, including First Fit, DSATUR, and Welsh–Powell, and analyze their performance in different graphs. The authors conclude that DSATUR is the most efficient for complex graphs with high density, while the Welsh–Powell algorithm is the fastest for solving simpler graphs. This paper emphasizes the importance of choosing an appropriate algorithm in accordance with the characteristics of graphs, thereby contributing to the understanding of the relationship between graph structure and algorithm efficiency. On the other hand, the paper [18] offers an overview of different algorithms on graphs from the DIMACS benchmark, and concludes

that DSATUR consumes the least colors, while Welsh–Powell is the fastest. The authors recommend using these algorithms depending on priorities such as time optimization or the number of colors.

In addition to these studies, the paper [19] from 1992 offers a deeper insight into the limitations of greedy algorithms, especially in the context of finding complex structures such as cliques in large graphs. The authors conclude that greedy algorithms, although fast, fail to effectively detect such structures, which indicates the need for more complex methods in specific cases.

In conclusion, these papers provide a comprehensive overview of different variants of greedy algorithms, highlighting their applicability in different contexts, from specialized types of graphs, such as interval graphs, to more complex structures with a large number of vertices. The papers also emphasize the importance of comparison between algorithms in terms of efficiency and applicability on different types of graphs.

4 Testing Variants of Greedy Algorithm and Analysis of Results

To test the impact of graph density on the efficiency of graph coloring with greedy methods, 10 cases were used. In all 10 cases, a graph with 100 vertices was used, while the number of edges, i.e., their density, increased. The smallest number of edges is 500. In each subsequent iteration, the number of edges increased by 500, and this method of incrementing reached 4500. The largest number of edges between 100 vertices is 4950, and this is singled out as a special case.

The following variants of the greedy algorithm were used: Descending (Welsh–Powell), Ascending, Random, and DSATUR. For each edge density, each variant was tested a thousand times. The average of the colors was calculated, and the results are included in Table 1.

Table 1: Average number of colors used by greedy algorithm variants.

Edges	**Desc.**	**Asc.**	**Random**	**DSATUR**	**Best**	**Worst**	**Best/Worst**
500	6.41	8.07	7.31	6.42	DESC	ASC	20.56%
1000	9.80	11.84	10.83	9.97	DESC	ASC	17.19%
1500	13.02	15.35	14.19	13.32	DESC	ASC	15.18%
2000	16.41	19.05	17.72	16.75	DESC	ASC	13.81%
2500	19.93	22.94	21.46	20.43	DESC	ASC	13.10%
3000	24.00	27.33	25.75	24.58	DESC	ASC	12.19%
3500	28.95	32.68	30.86	29.59	DESC	ASC	11.41%
4000	35.22	39.47	37.31	35.84	DESC	ASC	10.78%
4500	44.91	50.41	47.77	45.76	DESC	ASC	10.91%
4950	100	100	100	100	–	–	0.00%

After processing the data from the table, some of the test conclusions are presented graphically on the graph shown in Figure 1.

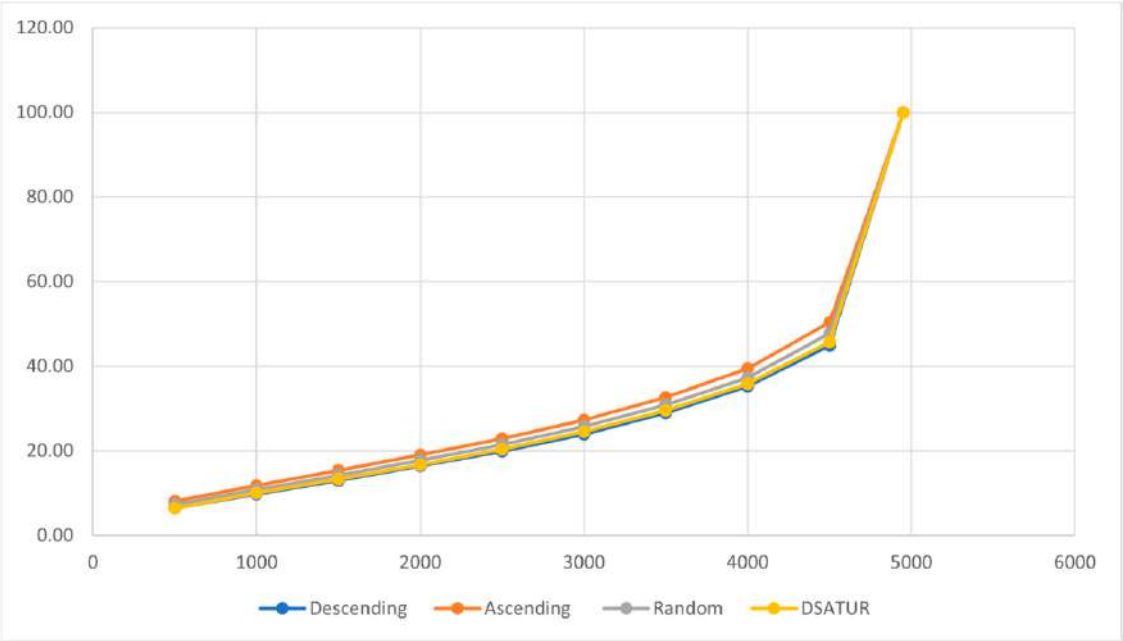

Figure 1: Impact of graph density on greedy algorithm variants.

Based on the presented graph, several key conclusions can be drawn about the influence of graph density on variants of greedy algorithms:

1. General trend and linear growth: All curves show an almost linear growth in the number of colors needed to color the graph as the number of edges increases, that is, as the graph becomes denser. This linear growth indicates that all algorithms have a proportional increase in the number of required colors with increasing graph density. As the number of adjacent vertices increases, more colors need to be used to avoid conflicts.

2. Descending consistently outperforms other variants by requiring fewer colors, as prioritizing highly connected vertices minimizes conflicts. This strategy is efficient because the most connected vertices receive colors before they are assigned to their neighbors, thus reducing the need for additional colors.

3. The Ascending Algorithm is the least efficient: The Ascending Algorithm almost always uses the largest number of colors. Sorting vertices in ascending order, where the vertices with the smallest number of neighbors are colored first, turns out to be an inefficient approach because it leaves more complex vertices for later, when most of the colors are already occupied.

4. Random and DSATUR algorithms: Random and DSATUR algorithms show average results, being placed between the Descending and Ascending variants. The DSATUR

algorithm is generally closer to Descending in efficiency because it uses the degree of saturation to select vertices, which allows it to better optimize the number of colors. The Random algorithm, although less consistent, also offers relatively good results.

5. Extreme density: It is interesting to note that all algorithms use the maximum number of colors (100% coloring) at a very high density of the graph, that is, when the graph has 4950 edges. This means that, in such a dense graph, each vertex must have its own unique color because all vertices are directly or indirectly connected.

Although the Ascending variant turned out to be the worst, its percentage reduction is still visible with regard to the best variant for a particular case. In the first variant, with the sparsest graph, the difference was 20.56%, while at the end of the test, with the densest graph (except for the special case with 4950 edges), that difference dropped to 10.91%. Figure 2 shows a drop in the percentage difference, but it can be seen that this drop is not linear.

It can be concluded that the Descending algorithm is the most efficient in most cases, while the Ascending algorithm is the least efficient. Random and DSATUR algorithms offer a balance between the speed and number of colors, depending on the density of the graph. The almost linear growth of the curves for all algorithms shows that the number of colors needed for graph coloring is proportional to the density of the graph, but also that the type of greedy algorithm can significantly affect the number of colors needed for efficient coloring.

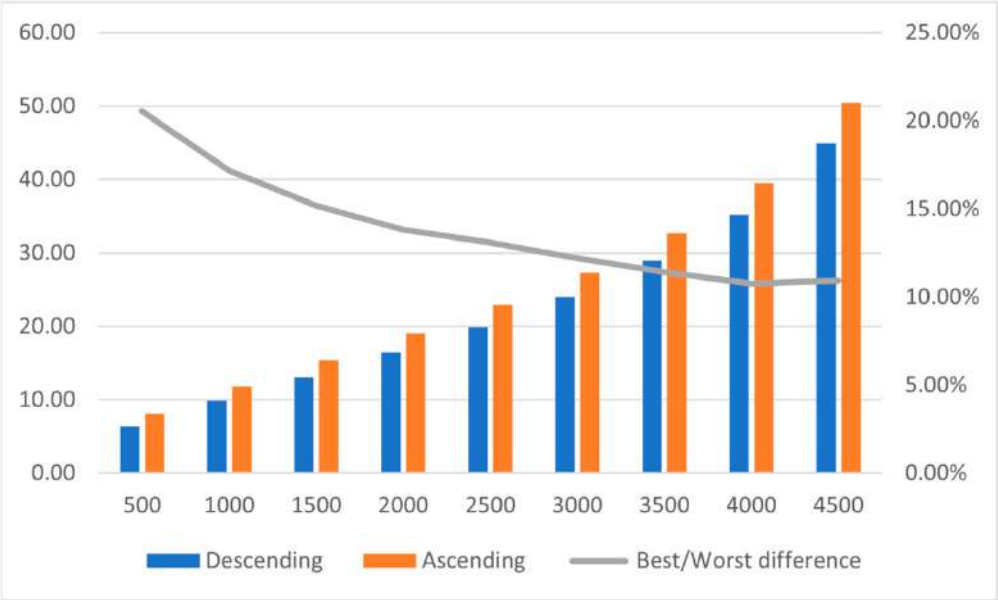

Figure 2: Percentage difference between best and worst case.

5 Application of Graphs in Blockchain Network Design

The private distributed network of 100 nodes represents a blockchain system in which each node has the possibility of direct communication with other nodes, and the purpose of

the network is to enable the safe and efficient exchange of information. In the blockchain architecture, each node participates in the maintenance of a distributed database, ensuring high fault tolerance and preserving data integrity. This kind of network is designed as a peer-to-peer (P2P) system without a central server, which achieves decentralization and increases security [20].

In this chapter, it will be explained how graph theory and vertex coloring with variants of the greedy algorithm can help in modeling such a network. Nodes are represented as vertices, and connections between nodes as edges of the graph. According to [21] the maximum number of edges between nodes is given by Equation (1):

$$\text{Number of edges} = \frac{n(n-1)}{2} \tag{1}$$

5.1. The Effect of the Number of Edges on the Performance of the Blockchain Network

In blockchain networks, there is no strictly defined minimum number of links (edges) that would automatically make the network decentralized and distributed. Decentralization in the blockchain architecture results from the independence of nodes and the distribution of data among nodes in the network. A network can be considered distributed and decentralized when nodes function independently of each other, data and tasks are evenly distributed, and nodes are sufficiently connected to allow information exchange and fault tolerance [22].

5.1.1. Problems with Too Few Links

Too few links in a network can significantly affect blockchain performance, creating communication bottlenecks, overloading certain nodes, and reducing fault tolerance. Namely, when nodes do not have enough links to other nodes, the risk of congestion increases because all information is transmitted through a limited number of channels, which reduces efficiency and increases network latency [23]. In a distributed network with too few links, the load becomes unevenly distributed, with some nodes overloaded by the need for multiple data transfers, while other nodes remain underutilized. This imbalance can compromise data security and integrity, as a lack of links makes it difficult to transfer data quickly and synchronize between nodes [22]. Figure 3 shows the worst-case scenario of the connected network. The minimum number of links was used in the network, which for this network of 100 nodes is 99. In the picture, the nodes with only one link are colored red. In such a network, there is no redundancy, and it is not resistant to failures.

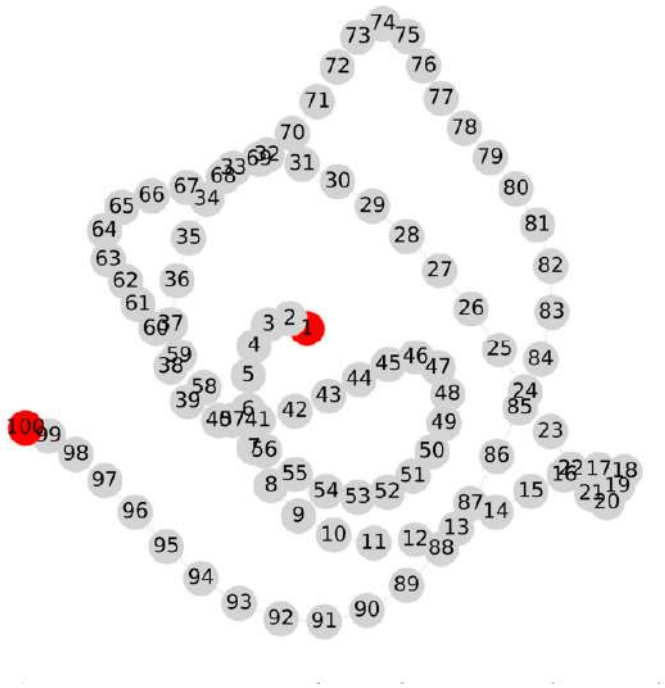

Figure 3: A network with 100 nodes and 99 links.

5.1.2. Advantages and Disadvantages of Too Many Links

The excessive number of links in the blockchain network has its advantages, but it also brings certain disadvantages. On the one hand, a larger number of links increases the network's resistance to failures because data can be redirected through alternative channels in the event of a failure of certain nodes. More links enable better load balancing between nodes, which avoids bottlenecks and reduces data transfer latency [24].

However, networks with too many links increase the complexity of data routing and burden the system with unnecessary communications. Increased connectivity can also mean greater exposure to security risks, as more links provide more potential points of attack. In private blockchain networks, which require security and resilience, an excessive number of links can create additional security vulnerabilities and increase network maintenance costs due to the exponential growth of the number of links with the growth of nodes [22]. Figure 4 shows a fully connected network with 100 nodes and 4950 links.

5.1.3. Optimal Number of Links

In order to achieve the optimal number of links, it is necessary to balance network resilience with performance economy. Graph theory suggests that a network with a moderate number of links can maintain fault tolerance and security without unnecessary redundancy. Partial mesh or small-world topologies in practice enable a relatively small number of edges that ensure a high resistance and efficiency of data transmission. In these topologies, the nodes are connected enough that data can circulate quickly through the network, but there are

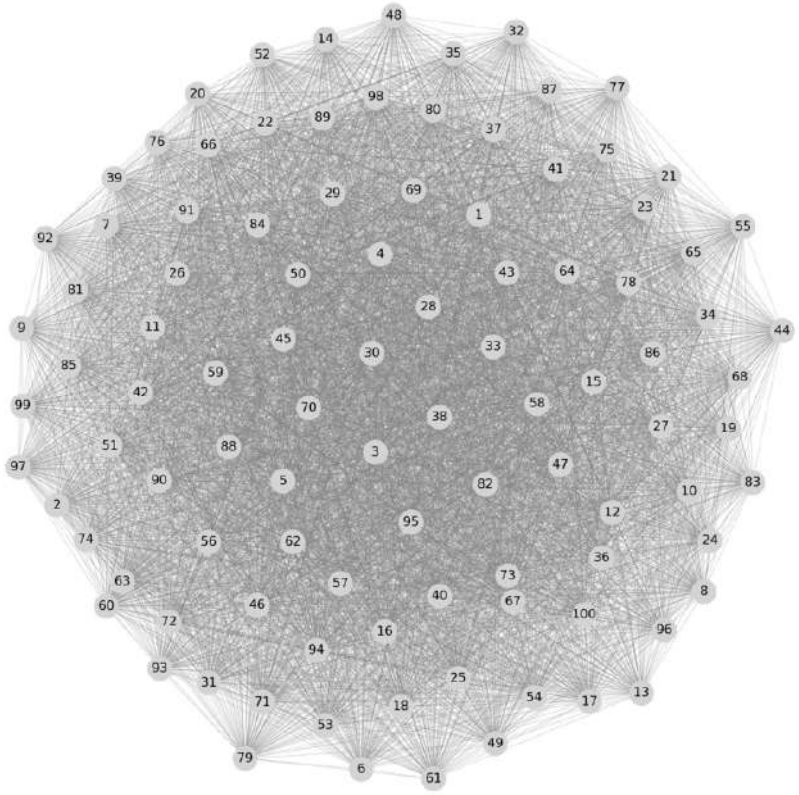

Figure 4: Fully connected network with 100 nodes and 4950 links.

not an excessive number of edges to create congestion or increase complexity. The optimal number of links can vary. According to [25,26], for a network of 100 nodes, 4950 links are not needed. According to [27], an approximation is 30–50% of that number, which is about 1500–2500 links. This is a rough recommendation that often provides a good balance between resilience and avoiding excessive redundancy in large networks, but the exact optimal number of edges depends on the specific requirements of the network, such as security, throughput, and resilience. The next chapter will reduce this rough approximation to more concrete numbers.

6 Determining the Optimal Number of Links Using Variants of the Greedy Algorithm

While existing systems such as VeDB and LedgerDB offer exceptional performance and tamper resistance through advanced mechanisms like Trusted Execution Environments (TEEs) and verifiable structures such as the Verifiable Shrubs Array (VSA), their implementation often requires specialized hardware and substantial resource investments. These requirements make them less suitable for small to medium-sized private blockchain networks, where cost-efficiency and scalability are critical. This study focuses on a simpler and more cost-effective approach by leveraging graph coloring algorithms, specifically variants of greedy algorithms,

to optimize node connectivity and resource distribution. The results demonstrate that it is possible to achieve network stability and efficient resource allocation without relying on advanced hardware infrastructures.

Greedy graph coloring algorithms, such as Descending and DSATUR, can effectively optimize the distribution of links among nodes by allocating links based on the current load and degree of connectivity of the nodes. This optimal configuration enhances network resilience, ensures efficient data transmission, and minimizes the likelihood of bottlenecks. These variants are particularly relevant because greedy algorithms are widely recognized for their simplicity and computational efficiency, making them well-suited for scenarios requiring rapid optimization, such as distributed systems. This study leverages their ability to optimize node connectivity and resource distribution within distributed peer-to-peer blockchain networks.

6.1. Insights into Network Structure Properties from Graph Coloring Algorithms

In networks that use different coloring strategies, some variants of greedy algorithms have specific effects on the color distribution and structure of the network. DSATUR minimizes conflicts in blockchain networks by prioritizing nodes with the most differently colored neighbors, ensuring balanced color distribution. If all the colors used in the DSATUR algorithm appear an equal number of times throughout the network, this means that each node has a similar number of neighbors with different colors, which suggests balanced connectivity and a consistent distribution of colors throughout the network. Such uniform color saturation helps stabilize the network and increases resilience to overloads, reducing the risk of bottlenecks.

In blockchain networks, the Descending algorithm efficiently colors densely connected nodes first, reducing conflicts and optimizing central node connectivity. If all the colors used in the Descending algorithm occur equally, this means that the links in the network are evenly distributed. Uniformity in the repetition of colors in this case indicates that there are no nodes with a dominantly large number of one color, which allows the network to achieve a balanced degree of connectivity and avoid congestion in densely connected nodes. This pattern of color distribution contributes to the stability of the network, making it more resistant to overload and congestion.

The Ascending algorithm allocates fewer colors to sparsely connected nodes, balancing color usage across the network. After coloring the nodes with fewer neighbors, the algorithm moves on to nodes with more neighbors, which then have a limited range of colors due to the already colored neighbors. The ascending approach results in a greater variety of colors in the edges of the network, while the central parts, with more links, remain balanced because they have enough options to avoid conflicts with the colors of their neighbors.

The application of the DSATUR algorithm ensures network stability because evenly distributed color saturation allows for balanced connectivity and reduces the risk of overload. The descending algorithm provides additional stability in networks with densely connected central parts because it allows for optimal color distribution among nodes with a high number of links. The ascending approach is useful for networks that require optimal color distribution in the edge parts, thus reducing the load on the central nodes. The combination of these algorithms can ensure network stability and resilience, adapting to different configurations and requirements of network systems, such as distributed and blockchain networks.

6.2. Coloring of a Network with 100 Nodes with Variants of the Greedy Algorithm for the Purpose of Optimization

For this paper, a computer-generated network of 100 nodes was used, and different amounts of edges were used. For this testing, a vertex represents a node in the blockchain network, while an edge acts as a link. Different link densities were tested.

6.2.1. 1500 Links

Figure 5 shows a graph with 1500 links colored according to the DSATUR variation. Coloring by saturation shows diversity in the use of colors, but this concentration of colors is not uniform throughout the network. Nodes with higher saturation use specific colors at the beginning. Since such vertices are colored first, these colors are first in the list. The colors are relatively uniform, but there are parts with a concentration of specific colors, indicating an uneven saturation of the network.

Figure 6, colored using the Descending variant, shows more colors in the very center of the network. Since this approach colors the nodes with the most neighbors first, it is evident that the nodes in the middle have more neighbors, that is, that the central nodes are more densely connected.

Figure 7 shows a larger number of different colors on the edge parts which can lead to a less uniform distribution in densely connected parts. In this example, the ascending variant of coloring was used.

These examples most clearly demonstrate the possibility of improving the network by pulling more links from peripheral nodes towards the center of the network.

6.2.2. 2000 Links

As the number of links increases, the DSATUR approach in Figure 8 shows a more uniform distribution of colors throughout the network. This approach manages to keep the colors more evenly distributed in the central and peripheral parts of the network, indicating stability and resilience in the network.

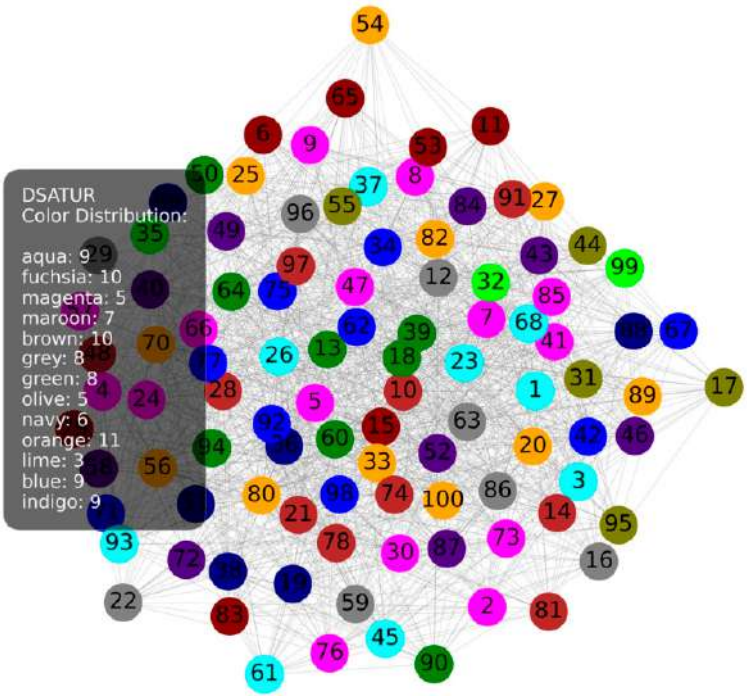

Figure 5: DSATUR 100 nodes 1500 links.

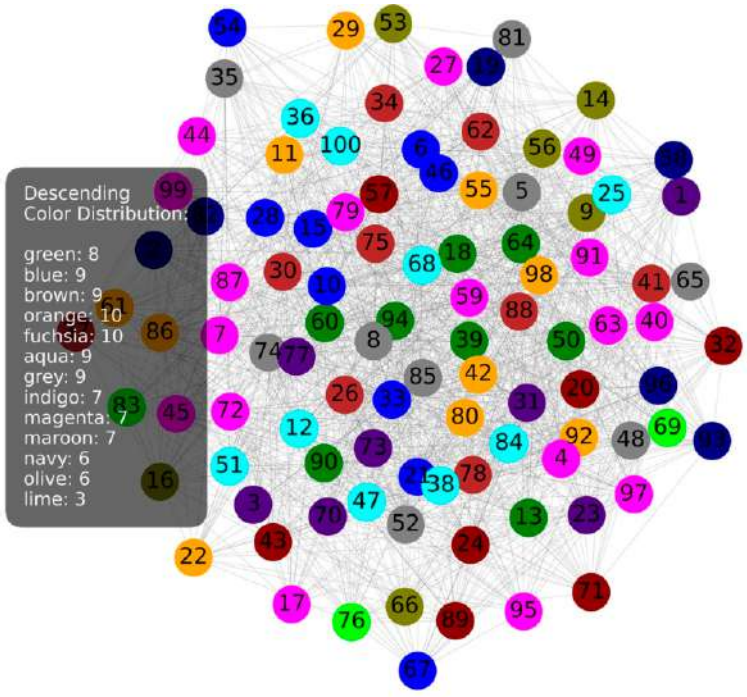

Figure 6: Descending 100 nodes 1500 links.

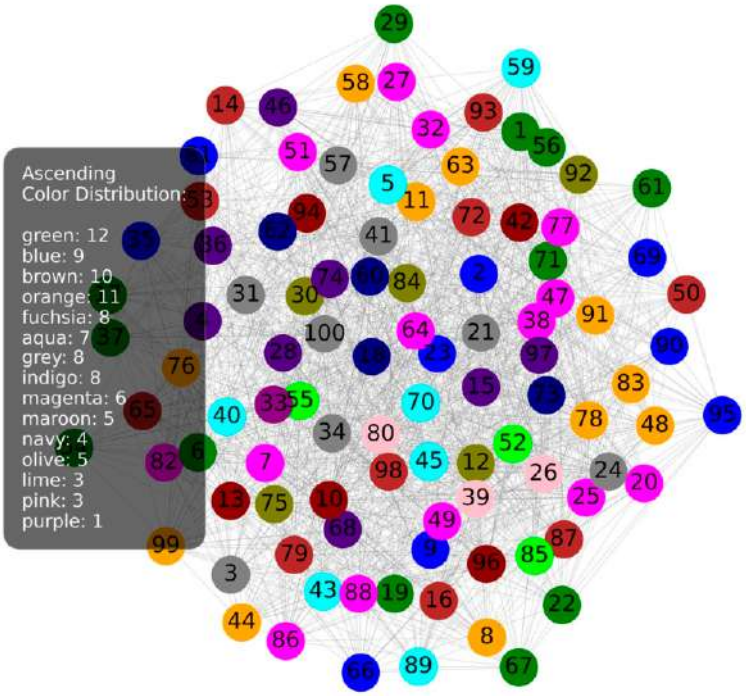

Figure 7: Ascending 100 nodes 1500 links.

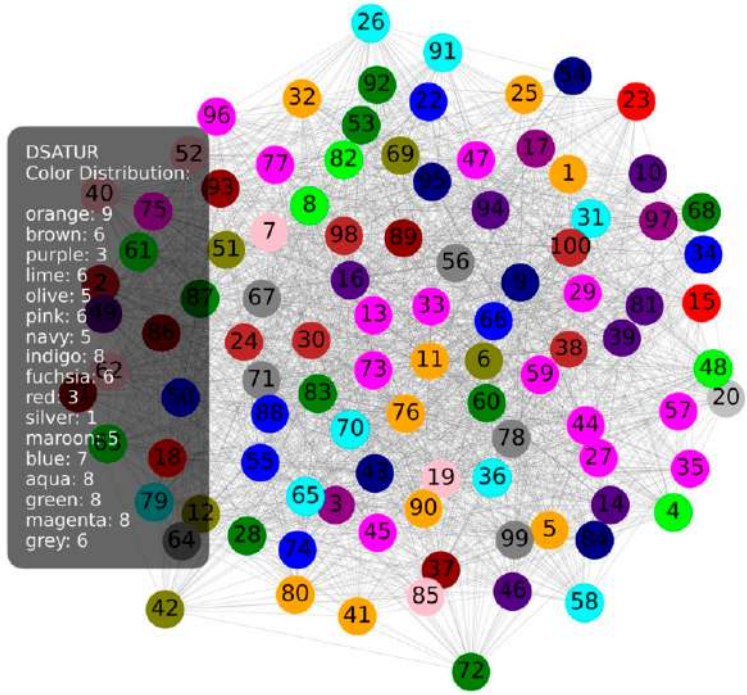

Figure 8: DSATUR 100 nodes 2000 links.

The descending variant in Figure 9 shows greater color uniformity compared to the previous case. This approach now uses fewer colors in the central part of the network, which would mean that the network is more uniformly loaded.

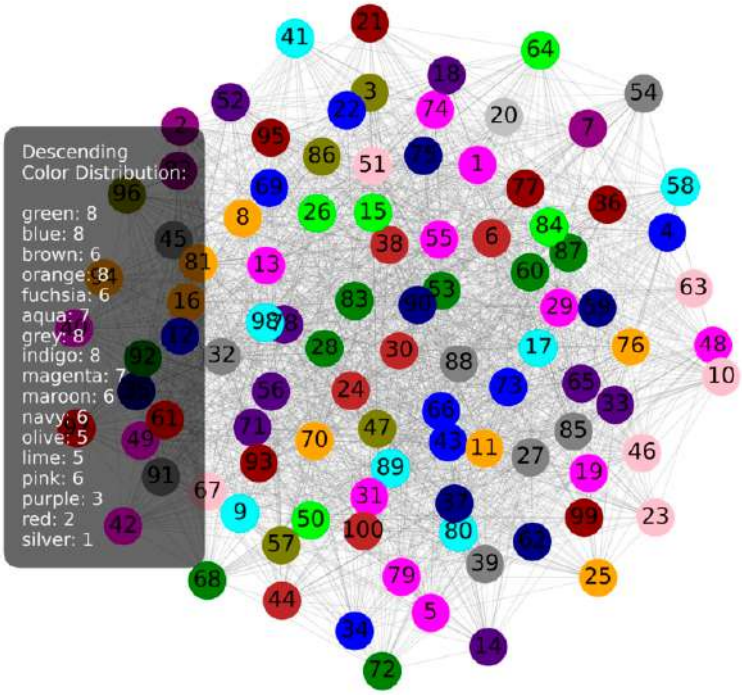

Figure 9: Descending 100 nodes 2000 links.

In Figure 10, the larger deviations in the use of colors at the edges of the network shown by the ascending variant are still visible. A greater variety of color usage on the peripheral parts indicates a weaker distribution towards those parts of the network.

When using 2000 links, slight improvements are visible in the form of less saturation of the central part of the network. The central part of the network is more evenly connected, which means that very few nodes stand out with a higher number of links than others. The peripheral part of the network still shows weaker connectivity.

6.2.3. 2500 Links

Increasing the number of links further improves the uniformity of coloring with the DSATUR algorithm. In Figure 11, the color is distributed even more evenly throughout the network, with balanced color usage in all parts. This case shows the highest stability.

In the configuration shown in Figure 12, the Descending variant, even fewer colors are used in the central nodes. The color distribution is relatively uniform, with less clustering of specific colors in individual zones. This case shows the lowest level of saturation.

The ascending variant in the case of Figure 13 also shows better uniformity in the color distribution compared to the smaller number of links. The colors are present in a larger

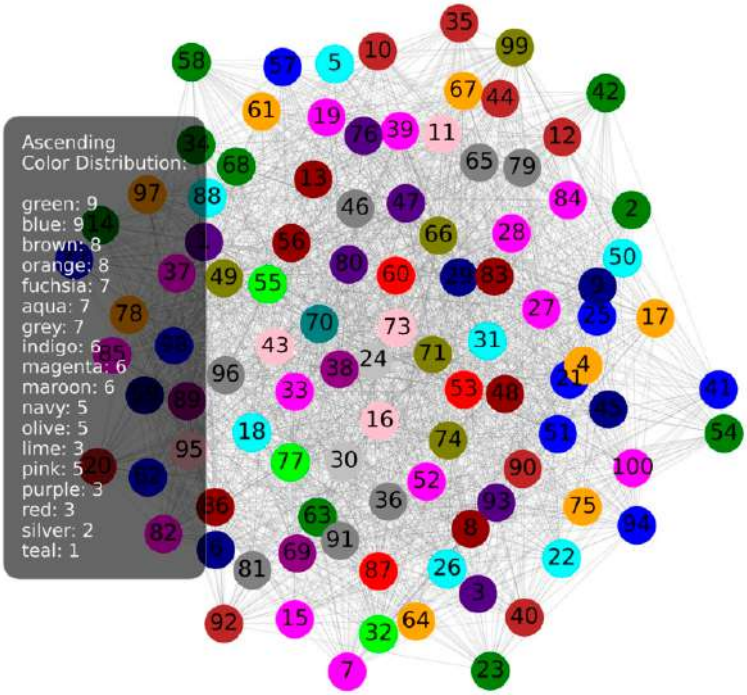

Figure 10: Ascending 100 nodes 2000 links.

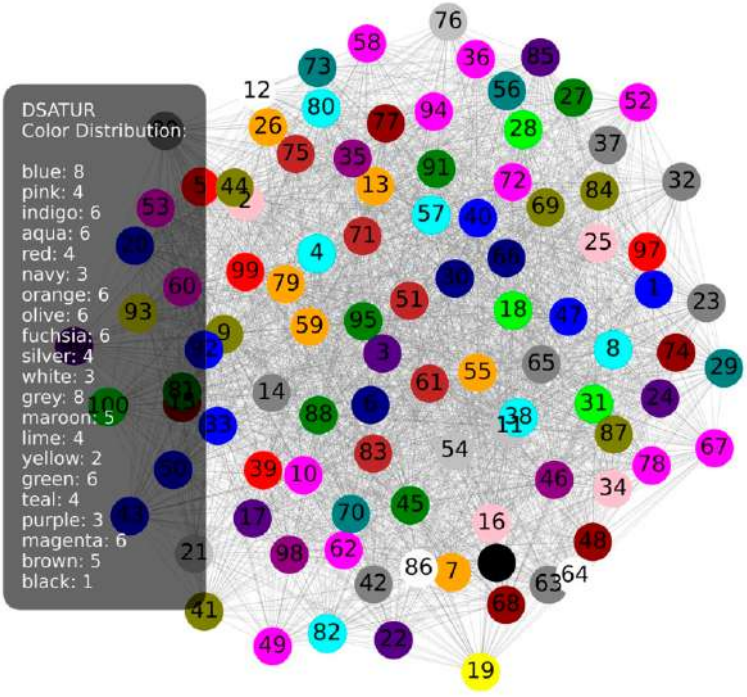

Figure 11: DSATUR 100 nodes 2500 links.

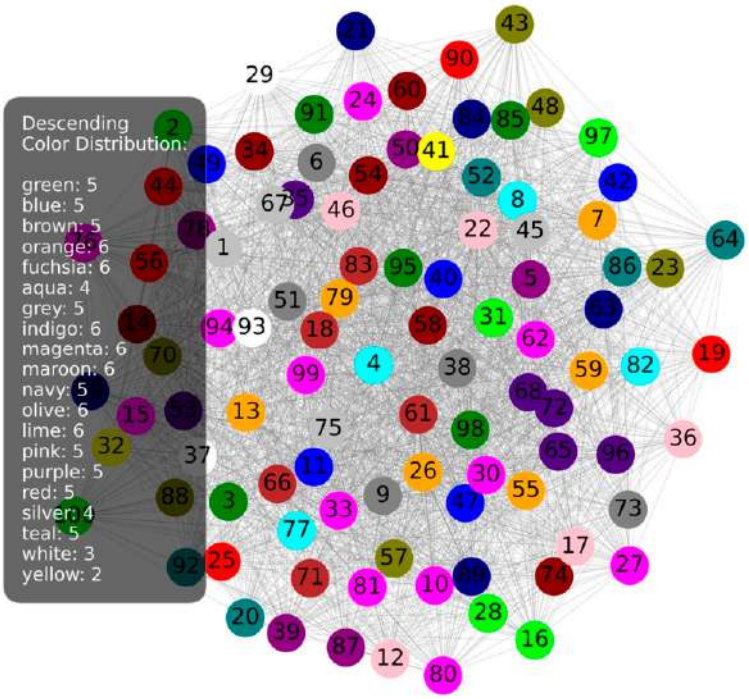

Figure 12: Descending 100 nodes 2500 links.

number of edge parts, but now with reduced contrast between the central and edge parts of the network.

6.3. Analysis of Results

Based on the analysis of each coloring approach for different numbers of links, several conclusions are drawn about the effects of using a larger number of links in the network and the optimal variants of the coloring algorithms. Using 2500 links results in the most balanced color distribution, especially when the DSATUR algorithm is applied. This approach achieves network stability and uniform color saturation among nodes, which indicates optimal connectivity without overloading individual nodes. In the DSATUR algorithm, all nodes have an equal number of neighbors with different colors, which reduces the risk of congestion and ensures stability.

A number of 2000 links can already provide enough connectivity for a 100-node network and be very close to optimal in many cases, especially when the goal is to balance resilience and efficiency. With 2000 links, the network should achieve high resilience and relatively good coloring uniformity, especially for algorithms like DSATUR and Descending.

From a coloring perspective, increasing to 2500 links does not provide a significant improvement over 2000 links. An additional 500 links may improve color uniformity somewhat, but this is usually a minimal improvement. Moreover, a larger number of links may increase redundancy and additional resource burden without a significant positive effect on coloring stability.

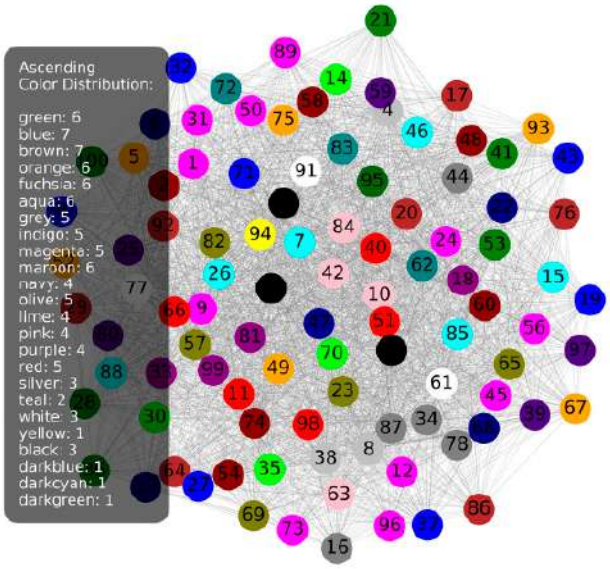

Figure 13: Ascending 100 nodes 2500 links.

Thus, the number of 2000 links can be considered an optimal compromise, providing high resistance and efficiency, while 2500 links can be an option for networks that require an additional safety net, but with minimal improvement in terms of coloring. This conclusion will be taken into account when creating the actual network.

Although the focus of this study is on private blockchain networks, the proposed methodology can be applied to other types of networks with similar topological characteristics. Examples include transport networks [28], where resource allocation and route optimization are critical, or power distribution networks [29], where balanced load distribution and fault tolerance are essential. This flexibility highlights the broader applicability of the presented approach.

7 Conclusions

This paper investigates the impact of graph density on the efficiency of greedy vertex coloring algorithms in solving the NP-hard graph coloring problem, with a special emphasis on applications in distributed peer-to-peer networks such as private blockchain systems. The results show that increasing the number of edges in the network can significantly improve the uniformity of color distribution, which directly affects the stability and resilience of the network to failures. Among the tested algorithms, DSATUR achieves the most uniform distribution of colors, especially in networks with a higher density of edges, while the Descending algorithm shows the best results in terms of the minimum number of colors used.

The analysis showed that the optimal number of links for a network of 100 nodes is between 2000 and 2500, where the network achieves a good balance between stability, resource efficiency, and resistance to congestion. This range of values represents a compromise

between the need for redundancy, which increases resilience, and the desire to avoid excessive complexity, which can negatively affect the performance of the network.

The application of graph theory and variants of the greedy algorithm has proven to be an effective approach for modeling and optimizing distributed peer-to-peer networks, including private blockchain networks. The proposed methodology allows for a simple and computationally efficient way to determine the optimal network configuration, without the need for complex and resource-intensive infrastructure.

8 Future Work

In this paper, the foundations for the creation of a real private blockchain network were laid, whereby the simulation results were used to define the optimal configurations. In the future, a private blockchain network based on these conclusions will be implemented. After the network is built, tests will be conducted under real conditions to evaluate performance, including fault tolerance, latency, and throughput.

Author Contributions: Conceptualization, T.R. and M.Š.; methodology, T.R. and M.Š.; software, D.I. and M.Š.; validation, T.R. and I.L.; formal analysis, M.Š.; investigation, M.Š.; resources, D.I. and M.Š.; writing—original draft preparation, M.Š.; writing—review and editing, I.L., T.R. and D.I.; supervision and funding acquisition, I.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded and is a part of research for the activities to be carried out for the project "Researching advanced algorithms and innovative business intelligence solutions in the cloud—NPOO.C3.2.R3-I1.04.0128".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in FigShare at `https://figshare.com/s/446f3f4d5933e37dae4b` (accessed on 6 November 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References (as published in the original article)

1. Kapron, B.M. *Logic, Automata, and Computational Complexity: The Works of Stephen A. Cook*; Morgan & Claypool: San Rafael, CA, USA, 2023; ISBN 9798400707803.

2. Erickson, J. *Algorithms*; University of Illinois, Urbana-Champaign: Champaign, IL, USA, 2019.

3. Borowiecki, P. Computational Aspects of Greedy Partitioning of Graphs. *J. Comb. Optim.* **2018**, 35, 641–665.

4. Barenboim, L.; Elkin, M. *Distributed Graph Coloring: Fundamentals and Recent Developments*; Synthesis Lectures on Distributed Computing Theory; Springer International Publishing: Cham, Switzerland, 2013; ISBN 978-3-031-00881-8.

5. Jayabalasamy, G.; Pujol, C.; Latha Bhaskaran, K. Application of Graph Theory for Blockchain Technologies. *Mathematics* **2024**, 12, 1133.

6. Nurmalika, K.; Prihandini, R.; Jannah, D.; Makhfurdloh, I.; Agatha, A.; Wulandari, Y. Graph Coloring Application Using Welch Powell Algorithm on Radio Frequency in Australia. Available online: `https://www.researchgate.net/publication/381511913_Graph_Coloring_Application_Using_Welch_Powell_Algorithm_On_Radio_Frequency_In_Australia` (accessed on 26 November 2024).

7. Postigo, J.; Soto-Begazo, J.; Fiorela, V.R.; Picha, G.M.; Flores-Quispe, R.; Velazco-Paredes, Y. Comparative Analysis of the Main Graph Coloring Algorithms. In Proceedings of the 2021 IEEE Colombian Conference on Communications and Computing (COLCOM), Cali, Colombia, 26–28 May 2021; pp. 1–6.

8. Yang, X.; Zhang, R.; Yue, C.; Liu, Y.; Ooi, B.C.; Gao, Q.; Zhang, Y.; Yang, H. VeDB: A Software and Hardware Enabled Trusted Relational Database. *Proc. ACM Manag. Data* **2023**, 1, 1–27.

9. Yang, X.; Zhang, Y.; Wang, S.; Yu, B.; Li, F.; Li, Y.; Yan, W. LedgerDB: A Centralized Ledger Database for Universal Audit and Verification. *Proc. VLDB Endow.* **2020**, 13, 3138–3151.

10. Ryan, C. Computer Algorithms. In *Encyclopedia of Physical Science and Technology*, 3rd ed.; Academic Press: San Diego, CA, USA, 2003; pp. 507–523.

11. Olariu, S.; Randall, J. Welsh-Powell Opposition Graphs. *Inf. Process. Lett.* **1989**, 31, 43–46.

12. San Segundo, P. A New DSATUR-Based Algorithm for Exact Vertex Coloring. *Comput. Oper. Res.* **2012**, 39, 1724–1733.

13. Wang, Y. Review on Greedy Algorithm. *Theor. Nat. Sci.* **2023**, 14, 233–239.

14. Dobzinski, S.; Mor, A. On the Greedy Algorithm for Combinatorial Auctions with a Random Order. *arXiv* **2015**, arXiv:1502.02178.

15. Olariu, S. An Optimal Greedy Heuristic to Color Interval Graphs. *Inf. Process. Lett.* **1991**, 37, 21–25.

16. Costa Ferreira da Silva, J.; Havet, F. On B-Greedy Colourings and z-Colourings. *Discret. Appl. Math.* **2024**, 359, 250–268.

17. Aslan, M.; Baykan, N. A Performance Comparison of Graph Coloring Algorithms. *Int. J. Intell. Syst. Appl. Eng.* **2016**, 4, 1–7.

18. Wadsungnoen, B.; Puphasuk, D.P. Permutation-Based Optimization Method for Solving Graph Coloring Problems. In Proceedings of PMO25 Conference, Khon Kaen University, Thailand, 2019; pp. 1–12.

19. Jerrum, M. Large Cliques Elude the Metropolis Process. *Random Struct. Algorithms* **1992**, 3, 347–359.

20. Blockchain for Committing Peer-to-Peer Transactions Using Distributed Ledger Technologies. Available online: `https://www.researchgate.net/publication/352455679_Blockchain_for_committing_peer-to-peer_transactions_using_distributed_ledger_technologies` (accessed on 25 October 2024).

21. Wilson, R.J. *Introduction to Graph Theory*; Longman: London, UK, 2010; ISBN 978-0-273-72889-4.

22. Wu, T.; Ren, H.; Li, P.; Leskovec, J. Graph Information Bottleneck. In *Advances in Neural Information Processing Systems*; Curran Associates Inc.: New York, NY, USA, 2020; Volume 33, pp. 20437–20448.

23. Meghanathan, N. (Ed.) *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks*, 1st ed.; Information Science Reference: Hershey, PA, USA, 2017; ISBN 978-1-5225-2814-2.

24. Alon, U.; Yahav, E. On the Bottleneck of Graph Neural Networks and Its Practical Implications. *arXiv* **2021**, arXiv:2006.05205.

25. Easley, D.; Kleinberg, J. *Networks, Crowds, and Markets.* Available online: `https://www.cs.cornell.edu/home/kleinber/networks-book/` (accessed on 30 October 2024).

26. Van Steen, M. *Graph Theory and Complex Networks: An Introduction*; ISBN 9789081540612.

27. Cohen, R.; Havlin, S. *Complex Networks: Structure, Robustness and Function*; Cambridge University Press: Cambridge, UK, 2010; ISBN 978-0-521-84156-6.

28. Rindone, C.; Russo, A. A Network Analysis for HSR Services in the South of Italy. In Proceedings of the Computational Science and Its Applications—ICCSA 2024 Workshops, Hanoi, Vietnam, 1–4 July 2024; pp. 217–232.

29. Beecroft, D.; Restrepo, J.G.; Angulo-Garcia, D. Greedy Optimization for Growing Stable Power Grids. *arXiv* **2021**, arXiv-2109.

B

Paper B: Entropy Extraction from Wearable Sensors for Secure Cryptographic Key Generation in Blockchain and IoT Systems

Sensors, vol. 25, no. 17, p. 5298, August 2025 | DOI: 10.3390/s25175298

Miljenko Švarcmajer, Mirko Köhler, Zdravko Krpić and Ivica Lukić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Osijek 31000, Croatia

Highlights

What are the main findings?

- Entropy generated from smartwatch sensors in shake mode reached Shannon entropy of 0.997 and min-entropy of 0.918, approaching levels of software-based RNGs.
- The proposed method enables fully offline, multisensor entropy harvesting on commercial Wear OS devices using only standard APIs.

What is the implication of the main finding?

- Smartwatches can serve as practical, user-controlled entropy sources for local cryptographic key generation, including blockchain transaction signing.

- The approach demonstrates a low-cost, open-source alternative to specialized hardware wallets, aligned with decentralization and self-sovereign identity principles.

Abstract

The increasing demand for decentralized and user-controlled cryptographic key management in blockchain ecosystems has created interest in alternative entropy sources that do not rely on dedicated hardware. This study investigates whether commercial smartwatches can generate sufficient entropy for secure local key generation by utilizing their onboard sensors. An open-source Wear OS application was developed to harvest sensor data in two acquisition modes: still mode, where the device remains stationary, and shake mode, where data collection is triggered by motion events exceeding a predefined acceleration threshold. A total of 4800 still-mode and 4800 shake-mode samples were collected, each producing 11,400 bits of sensor-generated data. Entropy was evaluated using statistical metrics commonly employed in entropy analysis, including Shannon entropy, min-entropy, Markov dependency analysis, and compression-based redundancy estimation. The shake mode achieved Shannon entropy of 0.997 and min-entropy of 0.918, outperforming the still mode (0.991 and 0.851, respectively) and approaching the entropy levels of software-based random number generators. These results demonstrate that smartwatches can act as practical entropy sources for cryptographic applications, provided that appropriate post-processing, such as cryptographic hashing, is applied. The method offers a low-cost, transparent, and user-friendly alternative to specialized hardware wallets, aligning with the principles of decentralization and self-sovereign identity.

Keywords: wearable sensors; smartwatch entropy generation; cryptographic key generation; Samsung Galaxy Watch; blockchain security; physiological and environmental signals; biometric entropy; random number generation

1 Introduction

Wearable devices, and particularly smartwatches, have become integral to health monitoring, activity recognition, and context-aware computing due to their continuous operation and integration of diverse sensors [1,2]. These sensors generate high-resolution physiological, motion, and environmental data, which, beyond their primary applications, have recently been explored as potential sources of physical entropy in IoT and mobile security contexts [3]. Ensuring high-quality entropy is critical not only for conventional cryptographic applications but also for emerging decentralized systems. Sensor-derived entropy offers an attractive alternative to traditional hardware or software random number generators, as it leverages naturally occurring physical variability that is difficult to replicate or predict externally [4,5].

The secure generation and management of cryptographic keys are foundational to the trust model of blockchain systems [6]. The integrity and authenticity of transactions rely on digital signatures derived from private keys, which must be generated using high-entropy sources to ensure cryptographic strength [7,8]. While traditional methods employ hardware security modules (HSMs), trusted platform modules (TPMs), or cryptographically secure pseudo-random number generators (CSPRNGs), these solutions are often bound to centralized infrastructure, proprietary implementations, or specialized external devices [9,10].

In contrast, users of blockchain-based systems typically favor decentralized, user-controlled, and open-source solutions [11,12]. Privacy, autonomy, and the minimization of trust in third parties are core principles in this domain. Consequently, there is growing interest in leveraging everyday personal devices as secure cryptographic platforms. Among these, the smartwatch presents a compelling opportunity. It is continuously worn, contains diverse motion and environmental sensors, and can operate independently of centralized services [13,14].

The key motivation behind this work lies in bridging the gap between user expectations in blockchain ecosystems and practical cryptographic security. By using entropy derived from physical interactions captured by a smartwatch (e.g., shaking motion), it is hypothesized that sufficient randomness can be harvested to generate private keys suitable for signing blockchain transactions. The approach promotes decentralization, avoids cloud-based key handling, and uses widely available hardware.

To illustrate the alignment between blockchain user values and the proposed approach, a comparison is provided in Table 1.

Table 1: Comparison of blockchain user preferences with traditional approaches and the proposed smartwatch-based method.

User Preferences	**Traditional Approach**	**Proposed (Smartwatch)**
Decentralized key control	HSM/Cloud dependent	Local on device
Privacy/no central entity	May rely on servers	No network dependency
Open-source compatibility	Often proprietary	Fully open implementation
No extra hardware	Requires special wallet	Uses existing smartwatch
Always-on/mobile	Desktop or specialized hardware required	Wrist-worn, always present

There is a clear disparity between the expectations of blockchain users and the constraints imposed by traditional key management solutions. Users of decentralized systems consistently demonstrate a preference for local control over cryptographic keys, as reliance on centralized

HSMs or cloud-based services introduces risks that conflict with the core principles of decentralization and self-sovereignty [6,11]. The smartwatch-based method addresses this by enabling local key generation and entropy harvesting on the user's own device.

Privacy and autonomy are strengthened by removing any dependency on external servers. This also eliminates the need for network connectivity during entropy acquisition and key generation. This aligns with the design goals of self-sovereign identity systems, which prioritize the minimization of external trust assumptions. In addition, the use of open-source software remains a critical criterion for adoption in the blockchain community, where transparency and verifiability are highly valued [15]. The proposed method, being platform-accessible and implementable on open devices such as those running Wear OS, supports this requirement.

The physical characteristics of smartwatches—being always-on, body-coupled, and already integrated into daily routines—offer an ergonomic and secure environment for cryptographic operations. Compared to smartphones, smartwatches typically run fewer third-party applications and operate in a more restricted software environment, reducing the potential attack surface and exposure to malicious software. This overcomes the usability and portability limitations of traditional hardware wallets, which often require additional devices and exhibit lower user acceptance [12,16].

Consequently, the smartwatch is positioned as a viable and user-aligned entropy source that addresses both technical requirements and ideological expectations in blockchain ecosystems.

Beyond meeting ideological expectations, the proposed approach derives its technical viability from the unique sensing capabilities of modern smartwatches. These devices include sensors such as accelerometers, gyroscopes, barometers, heart rate monitors, and ambient light sensors. They continuously capture multi-dimensional, time-varying signals. Such signals, particularly when elicited through deliberate user motion (e.g., hand shaking), introduce high levels of physical variability that are difficult to reproduce or predict externally. When properly sampled and conditioned, this sensor data has been shown to contain sufficient entropy to support secure cryptographic key generation [17–19].

This reliance on real-world, body-coupled sensor input—rather than computational pseudo-randomness—differentiates the proposed method from conventional approaches, situates it within the broader research agenda of secure entropy harvesting from physical environments. Due to their widespread adoption, ergonomic design, and advanced sensors, smartwatches are a promising but underexplored platform for cryptographic operations in decentralized systems.

The main contributions of this work are as follows:

- Systematic entropy analysis on a commercial smartwatch—to the best of our knowledge, this is the first study to systematically evaluate sensor-derived entropy on a commercial smartwatch in the context of cryptographic key generation for blockchain transaction

signing.

- Multisensor entropy harvesting with randomized sensor selection—a novel method is proposed that increases unpredictability by dynamically selecting a subset of sensors in each acquisition session.
- Quantitative evaluation using standardized metrics—entropy is analyzed using statistical measures recommended by NIST SP 800-90B [4], including Shannon entropy, min-entropy, Markov dependency analysis, and compression-based redundancy estimation.
- Comparison with established RNGs—results are compared against hardware- and software-based random number generators, demonstrating competitive entropy levels, particularly in motion-triggered acquisition.
- Practical implementation on a commercially available device—the proposed method is implemented as an open-source Wear OS application running fully offline on a Samsung Galaxy Watch 7 (Samsung Electronics Co., Ltd., Suwon, South Korea), demonstrating that no specialized hardware is required.

The study also has several limitations:

- Data were collected from a single smartwatch model and a single participant, limiting generalizability.
- Only a subset of NIST-recommended tests was applied, with full SP 800-90B validation left for future work.
- Per-sensor entropy contribution was not analyzed in detail, and motion sensors dominate the entropy pool due to their higher sampling rates.

The remainder of this paper is organized as follows. Section 2 reviews related work, providing a categorization of hardware-, software-, and sensor-based entropy generation methods. Section 3 presents the proposed method, describing the system architecture, sensor selection criteria, entropy acquisition protocol, and measurement procedure. Section 4 reports the results of entropy evaluation for still and shake acquisition modes, includes a comparative analysis with established random number generation methods, and discusses the implications of the findings. Finally, Section 5 concludes the study and outlines potential directions for future research.

2 Related Work

Reliable cryptographic key generation requires access to high-quality entropy sources that are resistant to prediction, manipulation, and external influence. The generation of such

entropy has been the focus of significant research across multiple domains, ranging from low-level hardware implementations to software-based pseudorandom generators and emerging approaches utilizing physiological or environmental data. Each of these methods exhibits distinct trade-offs in terms of security, transparency, hardware dependency, and suitability for mobile or decentralized contexts.

In this section, existing approaches to entropy generation are categorized into three main groups as shown in Figure 1. Section 2.1 reviews hardware-based methods, which typically rely on secure elements, TPMs, and hardware wallets equipped with true random number generators (TRNGs). Section 2.2 examines software-based methods, including cryptographically secure pseudorandom number generators (CSPRNGs), deterministic key derivation using mnemonic phrases, and entropy extraction from biometric data. Finally, Section 2.3 focuses on wearable and sensor-based approaches, which exploit human motion, environmental variability, and embedded sensing capabilities to derive entropy directly from physical interactions.

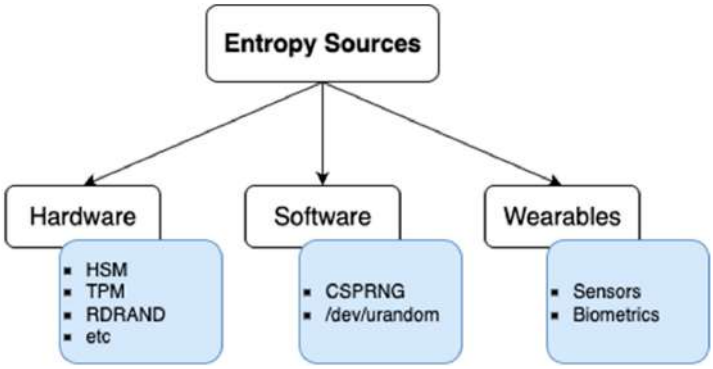

Figure 1: Classification of entropy sources.

2.1. Hardware-Based Entropy Generation Methods

Hardware wallets and secure elements represent the most established approach to cryptographic key generation in the blockchain domain, typically achieving Common Criteria EAL5+ standards [22,23]. During device initialization, entropy is gathered internally and processed via deterministic key derivation functions to generate BIP39-compliant mnemonic phrases [24,25]. The seed remains confined to the secure element, thereby minimizing attack surfaces from the host system. Similarly, TPMs available in modern computing platforms can generate entropy internally and securely store cryptographic keys for applications such as full disk encryption or TLS [26].

Another example of a hardware-based entropy source is Intel's RDRAND instruction, available on modern Intel processors. It implements a digital random number generator (DRNG) based on an on-chip TRNG seeding an AES-based deterministic random bit generator (DRBG), following the NIST SP 800-90A standard [9,10]. Although RDRAND provides cryptographically strong randomness in theory, concerns have been raised regarding the

transparency and auditability of its internal design, including potential backdoors [7]. As such, its use in security-critical applications is often debated, particularly within communities that prioritize verifiability and open-source principles.

While these approaches offer high entropy quality and resistance to tampering, they present several limitations when viewed from the perspective of decentralization and user autonomy. Most hardware wallets are proprietary, closed-source systems, limiting independent verification of the entropy generation process [12]. Furthermore, they require users to purchase and carry additional devices, which may not align with usability preferences in mobile-first or minimal-hardware contexts. HSMs and TPMs, while suitable for enterprise or server environments, are impractical for personal or wearable deployment due to their physical and infrastructural requirements.

In summary, hardware-based entropy sources are well-established and cryptographically robust, but their dependence on specialized components and limited transparency restricts their suitability for open, user-centric, and mobile cryptographic systems—particularly within decentralized ecosystems such as blockchain.

2.2. Software-Based Entropy Generation Methods

Software-based methods for entropy generation and random number derivation are widely employed in cryptographic systems due to their ease of deployment and platform independence. These methods generally rely on deterministic algorithms that are seeded with entropy gathered from various system-level events, user interactions, or hardware noise. Once initialized, such generators produce cryptographically secure pseudorandom numbers (CSPRNGs) that meet statistical criteria for unpredictability and uniformity [10].

One of the most widely adopted software-based approaches is the /dev/urandom interface in Unix-like operating systems, which combines system entropy pools with cryptographic mixing functions to generate pseudorandom output [27]. Although efficient and practical, /dev/urandom is not guaranteed to be forward- or backward-secure without sufficient initial entropy, particularly in early boot stages [28]. OpenSSL's RAND_bytes() and similar functions provide CSPRNG capabilities based on NIST-recommended deterministic random bit generators (DRBGs), but inherit similar limitations regarding their dependence on the quality and availability of initial entropy sources [29].

Stream cipher-based CSPRNGs, such as ChaCha20 as implemented in the libsodium library, have gained popularity due to their performance, simplicity, and resistance to side-channel attacks [30]. These generators are often used for ephemeral key generation, nonce construction, and secure messaging protocols. While their internal state transitions are cryptographically robust, their unpredictability is only as strong as the entropy present in their initial seeding.

In addition to traditional software approaches, biometric and behavioral data—such as

keystroke dynamics, heart rate variability, and touchscreen interactions—have been proposed as alternative entropy sources [31]. Although promising, these methods typically require user-specific calibration and may exhibit limited entropy under constrained conditions or in passive collection scenarios.

Overall, software-based methods provide accessible and efficient mechanisms for entropy generation. However, they are fundamentally limited by the quality and trustworthiness of their initial seeding sources, which may themselves be hardware-dependent or opaque. This reinforces the importance of evaluating entropy sources that are both transparent and physically grounded, as explored in wearable sensor-based methods.

2.3. Wearable and Sensor-Based Approaches

Recent advances in wearable technology have enabled the integration of a wide range of sensors into compact, body-coupled devices. These sensors, originally intended for health monitoring or activity tracking, have been increasingly explored as entropy sources for cryptographic purposes. The motivation lies in their capacity to capture inherently unpredictable and user-specific physical phenomena, such as hand movements, gait, temperature fluctuations, and cardiovascular signals [32].

Prior research has examined the feasibility of using accelerometer and gyroscope data to generate random sequences by exploiting the stochastic nature of human motion [33]. Other studies have explored entropy extraction from photoplethysmography (PPG) signals, electrodermal activity (EDA), and skin temperature, aiming to leverage biometric variability for key generation [34]. These approaches offer the advantage of continuous, real-time entropy harvesting, which aligns with the operational patterns of modern wearable devices.

Practical deployments remain limited. Many proposed systems rely on custom hardware, controlled environments, or data post-processing steps that hinder real-time applicability. Additionally, some platforms, such as proprietary smartwatches with closed firmware, restrict access to raw sensor data and reduce transparency and reproducibility. Previous studies were primarily focused on biometric authentication, secure pairing, or device identification rather than explicit cryptographic key generation, and to the best of our knowledge, no prior work has systematically evaluated entropy from commercial smartwatch sensors for this purpose.

In contrast, the present work focuses on entropy harvesting from commercially available smartwatches running open platforms (e.g., Wear OS) using native sensors and real-world user interactions. By analyzing motion data in both active (shake) and passive (still) conditions, a quantifiable entropy model is established and compared against traditional entropy sources. This sensor-based method is not only portable and decentralized but also verifiable and open for audit, traits that are particularly valued in blockchain ecosystems.

3 Proposed Method

In this section, the proposed approach for entropy harvesting from smartwatch sensors is presented. The method is designed to enable the local generation of cryptographic entropy on commercially available wearable devices, without the need for external hardware, cloud infrastructure, or platform-specific dependencies. Emphasis is placed on leveraging existing smartwatch onboard sensors to maximize compatibility, portability, and openness of the solution, as shown in Figure 2. The figure illustrates the complete entropy acquisition sequence: the user initiates the process through a shake gesture, multiple onboard sensors capture motion, environmental, and physiological data, the collected signals are vectorized into numerical representations, and these are subsequently processed to produce cryptographic key material.

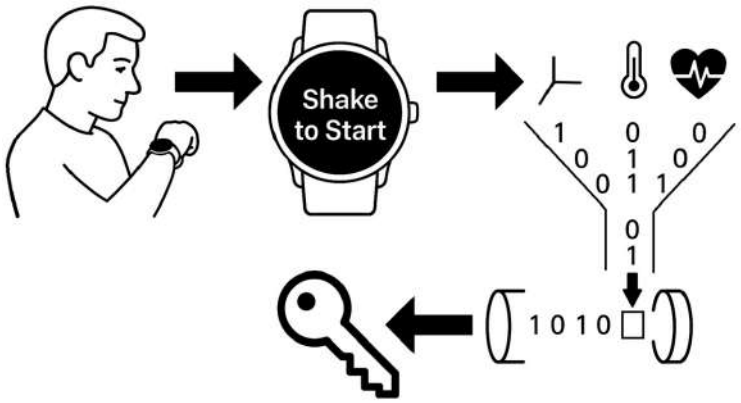

Figure 2: Visual representation of the proposed entropy generation process on a smartwatch.

3.1. System Overview

The system operates locally on a commercial smartwatch and collects entropy exclusively from onboard sensors. Two acquisition modes are supported:

- Still mode—data are recorded while the device remains stationary to capture environmental and physiological variability.
- Shake mode—acquisition is triggered by motion events exceeding a predefined acceleration threshold, collecting short bursts of high-frequency data from motion sensors.

In both modes, sensor data are transformed into numerical vectors and processed locally. Each acquisition produces a fixed-size entropy representation and a SHA-256 digest, enabling further analysis and potential cryptographic use. No network connectivity or external hardware is required.

3.2. Sensor Selection

The selection of sensors was based on their availability on commercial smartwatches, their capacity to produce non-deterministic data streams, and their relevance to both environmental and physiological variability. Sensors were grouped into three categories according to their origin and signal dynamics:

1. Motion sensors, including the accelerometer, gyroscope, rotation vector, and linear acceleration, capture rapid, user-dependent variations during wrist movement. These sensors exhibit high temporal resolution and significant entropy potential, especially when motion is spontaneous or randomized.

2. Body sensors, such as the heart rate monitor and step counter, provide user-specific physiological signals. While typically more stable than motion data, their variability across time and individuals can still contribute valuable entropy, particularly in multimodal fusion contexts.

3. Environmental sensors, such as the barometric pressure sensor, measure external physical conditions. These signals vary subtly over time and can introduce low-frequency noise that increases unpredictability in passive (still) acquisition scenarios.

In this study, a total of eight sensors were selected based on their prevalence across modern smartwatches (especially those running Wear OS), their accessibility via standard APIs, and a preliminary review of device hardware specifications from manufacturers such as Samsung and Google. The selected sensors were as follows: (1) Accelerometer; (2) Gyroscope; (3) Rotation Vector Sensor; (4) Linear Acceleration Sensor; (5) Gravity Sensor; (6) Barometer (Pressure Sensor); (7) Heart Rate Sensor; (8) Step Counter.

These sensors were chosen due to their ubiquity, dynamic behavior, and low power requirements. According to the official Android developer documentation and public specifications of devices such as the Samsung Galaxy Watch 5/6/7 and Google Pixel Watch 2 (Google LLC, Mountain View, CA, USA), these are consistently available and produce sufficient signal entropy under both still and active conditions [35,36]. An overview of their typical sampling frequencies and the expected number of measurements within a single 2.2 s acquisition window is provided in Table 2.

All onboard sensors are factory-calibrated according to Samsung manufacturer specifications [35], ensuring baseline measurement accuracy without the need for user-performed calibration.

The decision to utilize a randomized subset of active sensors during each session was made to increase entropy diversity and to mitigate predictability. This dynamic selection further prevents adversaries from replicating sensor access patterns or tailoring attacks to fixed data sources.

All selected sensors are accessible via the Android Wear OS SensorManager API without requiring elevated permissions, ensuring user-level deployment feasibility. Access to these

Table 2: Typical sampling frequencies and expected number of measurements per 2.2 s acquisition window for each selected Galaxy Watch 7 sensor.

Sensor	**Typical Frequency (Hz)**	**Expected Samples**
Accelerometer	100	220
Gyroscope	100	220
Rotation Vector	50	110
Linear Acceleration	50	110
Gravity	50	110
Barometer (Pressure)	5	11
Heart Rate	1	2
Step Counter	Event-driven ($\sim$0.5 Hz)	$\sim$1

sensors was managed entirely through the standard Android/Wear OS permission framework. At runtime, the system prompts the user to explicitly grant the necessary permissions (e.g., BODY_SENSORS for heart rate, ACTIVITY_RECOGNITION for motion data, and general sensor access for inertial and environmental sensors). These permissions are requested through standard Android dialog boxes, and data collection cannot proceed without user approval. Once granted, the application registers listeners through the Android SensorManager API, which delivers real-time data streams only from the sensors authorized by the user. This guarantees that all acquisitions remain subject to explicit consent, enforced at the OS level, and that no restricted or undocumented system functions are bypassed. Importantly, the combination of diverse sensor types enhances entropy quality through data fusion, as statistical correlations between sensor modalities are minimal, especially during short sampling windows.

The relative contribution of each sensor to the total amount of data collected in a single 2.2 s acquisition window is shown in Figure 3. Motion sensors such as the accelerometer and gyroscope dominate the data volume due to their high sampling rates, whereas environmental and physiological sensors (barometer, heart rate, and step counter) contribute significantly less.

Despite their lower sampling rates, the inclusion of these sensors is recommended because they provide heterogeneous, physically independent signal sources. This diversity increases overall entropy by introducing additional stochastic variability and reducing the risk of correlated patterns across sensor modalities.

3.3. Entropy Acquisition Protocol

The entropy acquisition process is governed by a lightweight protocol implemented entirely on the smartwatch, designed to operate autonomously and without network connectivity. Two distinct acquisition modes were defined to capture different entropy profiles:

1. Still mode: Sensor data are collected over a fixed time window starting immediately

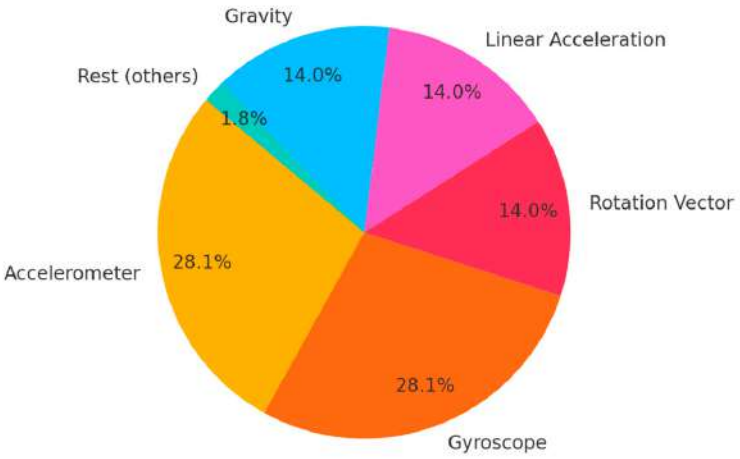

Figure 3: Expected percentage contribution of each sensor type to the total data volume collected during a 2.2 s acquisition window, based on typical sampling frequencies listed in Table 2.

upon application launch, while the user remains stationary. This mode targets slow-changing environmental and physiological inputs such as pressure, heart rate, and gravity.

2. Shake mode: Acquisition is event-driven and triggered by a motion-detection threshold. In this study, the threshold for acceleration magnitude was empirically set to 8.0 m/s^2 based on preliminary experiments to balance sensitivity and noise rejection. This value ensured reliable detection of intentional shake gestures while minimizing false triggers from normal wrist movements. Upon detecting a change in acceleration magnitude, a short burst of high-frequency data is collected from selected motion sensors.

Each acquisition session begins with the randomized selection of a subset of sensors from the list defined in Section 3.2. This randomization introduces additional unpredictability and hinders potential adversaries from anticipating sensor combinations.

Once triggered, the application performs the following steps, shown in Figure 4:

1. Sampling: Raw sensor data are collected using the Android SensorManager API, with typical sampling rates ranging from 50 Hz to 200 Hz, depending on sensor type. The acquisition window was set to 2.2 s to ensure that each selected sensor provides at least one complete measurement cycle, as determined from the manufacturer's specifications for update frequencies. An additional 0.3-s pause between sessions was introduced to ensure that sensor buffers were flushed before the next acquisition.

2. Vectorization: Each data frame is transformed into a structured 1D numerical vector and concatenated with a timestamp and sensor identifier. This produces a uniform and reproducible representation of heterogeneous sensor streams.

3. Preprocessing: Basic normalization is applied to account for sensor scaling, and values are discretized to 8-bit resolution to align with common entropy evaluation practices (e.g., NIST SP 800-90B), reducing high-frequency sensor noise and enabling direct applicability of standardized statistical tests.

4. Hashing: The final vectorized output is processed through a cryptographic hash function (SHA-256) to produce a fixed-length digest, which is treated as the entropy-bearing seed or candidate cryptographic key.

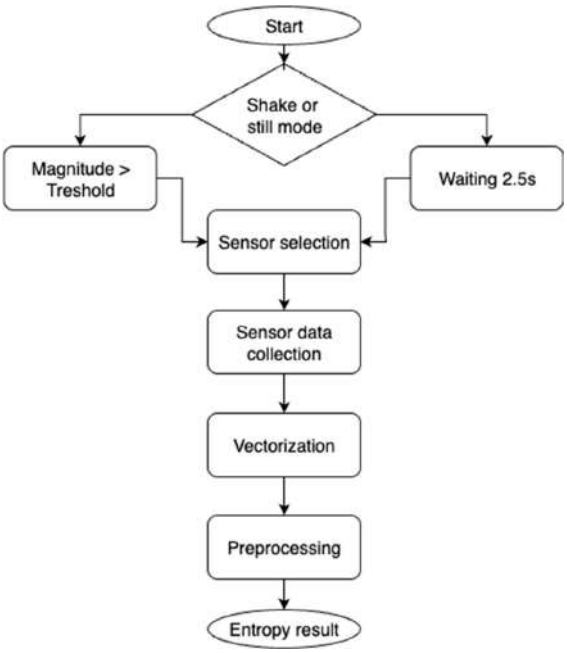

Figure 4: Flowchart of the proposed entropy acquisition pipeline for both shake and still modes.

All steps are performed entirely offline, and no data leaves the device at any point. Logging is used internally to record entropy metrics for evaluation purposes but remains sandboxed within the application environment. This approach ensures privacy, replicability, and full user control over the entropy generation process.

3.4. Implementation Details

The proposed entropy acquisition system was implemented as a standalone application for smartwatches running Wear OS. The application was developed using Kotlin (version 1.9.23) and Jetpack Compose (version 1.7.0), following modern Android development practices and targeting API level 30 and above. This ensured a responsive and efficient user interface suitable for low-power wearable environments.

The application's lifecycle is minimal and fully event-driven. Upon first launch, the app detects all available onboard sensors using the SensorManager API and performs a one-time

random shuffle to define the acquisition sequence. This randomized order remains fixed throughout each session to ensure consistency in entropy sampling.

For user interaction, a minimal interface is presented, displaying the message "Shake to Start" once initialization is complete. When the device detects a motion exceeding a predefined acceleration threshold, a short data acquisition window is triggered. During this window, data are collected asynchronously from the selected subset of sensors.

Each sampling session produces a single entry stored in a CSV file within the application sandbox. This entry includes: (1) a sequence index; (2) a transformed numerical representation of the collected sensor data (as a single entropy vector or derived scalar); (3) timestamp (optional).

The application operates fully locally on the smartwatch, ensuring user privacy and full control over the collected data.

For entropy assessment and cryptographic relevance, each collected sensor vector is hashed using the SHA-256 algorithm, generating a fixed-length digest suitable for seeding key material. These digests are used exclusively for evaluation and remain confined to local storage.

The application was evaluated on a Samsung Galaxy Watch 7, which provided full compatibility with the Wear OS platform and offered unrestricted access to all required sensor interfaces. This smartwatch was chosen as the test device due to its exceptional suitability for entropy-related experimentation: it supports open app deployment, features a wide array of onboard sensors, and is well-documented within the Android development ecosystem.

From a security perspective, the potential attack surface is inherently reduced since all entropy harvesting, processing, and key generation steps are executed entirely on the smartwatch. No raw sensor data or intermediate entropy values are transmitted to external devices or networks, which significantly limits opportunities for interception or manipulation during the acquisition process.

The full source code is publicly available on GitHub, enabling independent verification of the data acquisition process and reproducibility of the experiments without exposing sensitive raw sensor data.

Figure 5 provides an overview of the proposed system architecture, showing the interactions between on-watch physical sensors, acquisition mode logic, and processing modules. This complements the procedural details described in Sections 3.3 and 3.4.

3.5. Measurement Procedure

The measurement procedure was conducted in two phases: an initial phase aimed at application refinement and a final phase dedicated to systematic entropy evaluation.

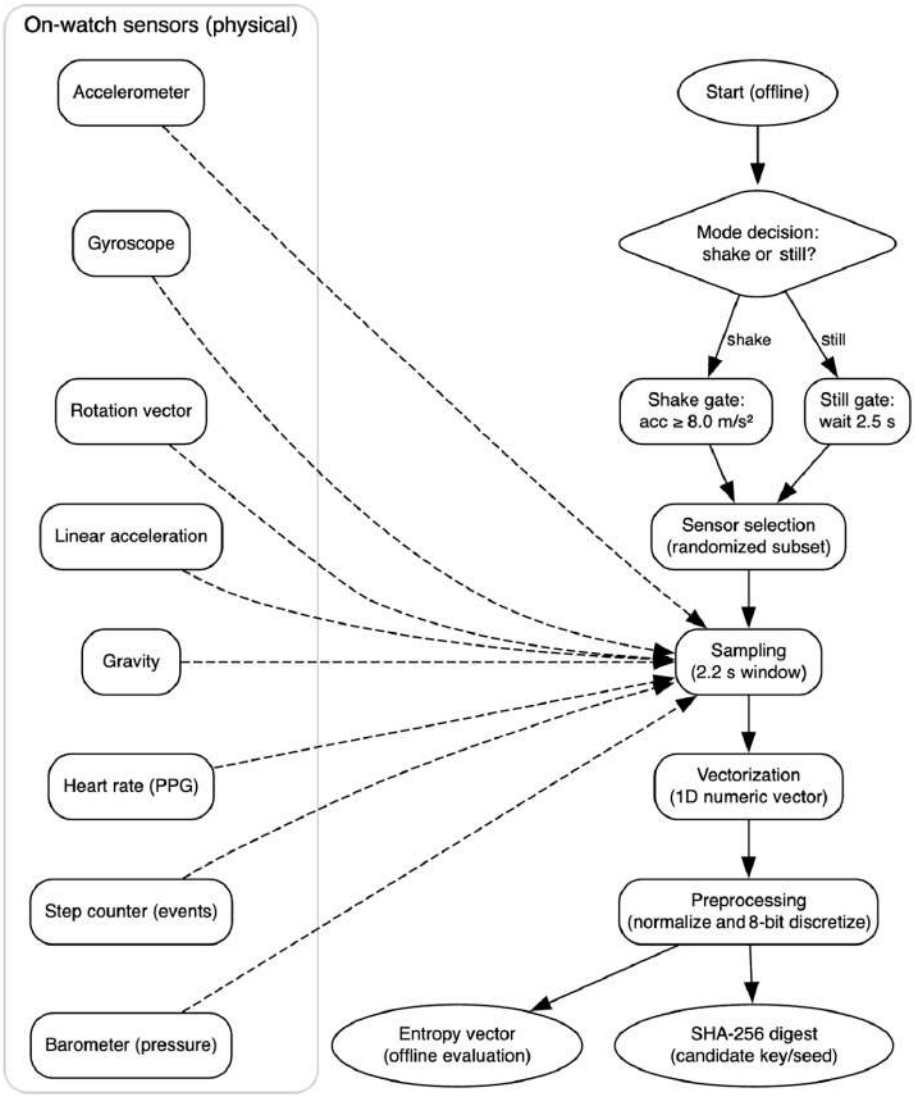

Figure 5: System architecture for entropy acquisition and processing on a commercial smartwatch.

3.5.1. Preliminary Testing

Several iterations of the application were tested to refine sensor handling, data transformation, and motion detection logic. Approximately 1000 preliminary samples were collected during this phase. These were used exclusively for debugging and early entropy estimation and were excluded from the final statistical analysis.

3.5.2. Final Measurements—Still Mode

The first final measurement phase targeted the still mode, with the smartwatch placed motionless on a desk to minimize user-induced variability. Entropy in this mode was therefore primarily derived from sensor noise and subtle environmental micro-variations.

- A total of 5000 measurements were initially recorded.
- Each measurement lasted 2.2 s and produced 11,400 bits of sensor-derived data.
- Successive measurements were automatically triggered with a 0.3-s delay.
- Before the completion of the 5000 still-mode measurements, the smartwatch was inadvertently displaced on the desk. To ensure consistency, the last 200 measurements were excluded from the analysis, leaving 4800 valid samples, which was more than sufficient for statistical testing.

3.5.3. Final Measurements—Shake Mode

In the second phase, the smartwatch was worn on the wrist of a single participant performing daily activities. Data acquisition was event-driven and triggered only when wrist motion exceeded a predefined acceleration threshold (as described in Section 3.3). Each trigger produced one entropy sample under the same sampling and formatting conditions as in still mode. A total of 4800 samples were collected in this mode.

All measurements were stored in CSV format, with one line per sample containing the sequence index, the numerical entropy vector, and its locally computed SHA-256 digest. No preprocessing was applied beyond normalization and discretization (Section 3.3), ensuring that entropy analysis reflected raw sensor variability. Datasets were exported manually for offline statistical evaluation. The experimental workflow is summarized in Figure 6.

4 Analysis of Results

This section presents the entropy analysis of the proposed smartwatch-based method for still and shake acquisition modes. Several statistical metrics commonly used in entropy evaluation, including measures recommended by NIST SP 800-90B, were applied to assess the unpredictability of the collected data.

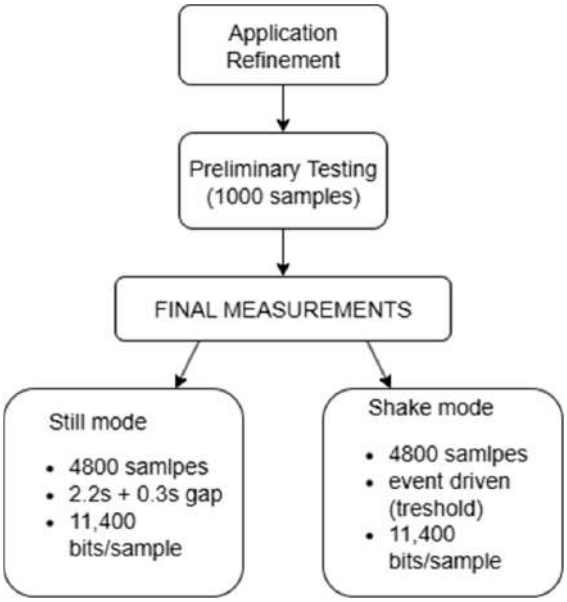

Figure 6: Experimental workflow for still and shake entropy acquisition.

The obtained results are compared with established hardware- and software-based entropy sources to evaluate the cryptographic suitability of the method.

4.1. Entropy Evaluation Methodology

The statistical evaluation of the smartwatch sensor data was conducted in accordance with widely accepted practices for assessing entropy sources, with reference to the guidelines outlined in NIST SP 800-90B [4]. Since the goal of the study was to determine whether sensor-derived data could provide sufficient unpredictability for cryptographic key generation, multiple complementary metrics were employed to capture different aspects of randomness and statistical independence.

Four primary evaluation methods were applied:

1. Shannon Entropy. Shannon entropy was computed according to Equation (1) to quantify the average information content per symbol in the collected sequences. This metric is widely used as an indicator of randomness quality, as it reflects how uniformly the observed symbols are distributed [37]. Although NIST SP 800-90B does not recommend it as a primary security metric due to its sensitivity to distributional assumptions, it provides a valuable initial assessment of entropy sources.

$$H_{\text{Shannon}} = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{1}$$

Here, p_i represents the empirical probability of symbol i, and n is the total number of distinct symbols. Higher entropy values indicate a more uniform distribution and therefore greater unpredictability.

2. Min-Entropy. Min-entropy was calculated according to Equation (2) to estimate the worst-case unpredictability of the sensor-derived sequences. Unlike Shannon entropy, which measures the average information content, min-entropy focuses on the most probable symbol and therefore provides a conservative lower bound on the true entropy of a source. This makes it particularly relevant for cryptographic applications, as recommended by NIST SP 800-90B [38].

$$H_{\min} = -\log_2(\max_i P_i) \tag{2}$$

Here, $\max_i P_i$ is the highest empirical probability observed among all symbols. Lower maximum probabilities correspond to higher min-entropy values, indicating reduced predictability and greater suitability for secure key generation.

3. Markov Dependency Analysis. To account for potential temporal correlations between successive samples, a first-order Markov model was used to estimate the transition probabilities between states [39]. This method is recommended for non-IID (non-independent and identically distributed) sources, which are typical for physical sensor measurements.

4. Compression-Based Redundancy Estimation (GZIP Test). Compression-based methods are often used as heuristic indicators of structural regularities in data [40]. In this study, a standard GZIP algorithm based on LZ77 compression was applied; sequences with higher entropy are expected to show lower compression ratios. While not formally equivalent to the compression test described in NIST SP 800-90B, the underlying principle is comparable [41].

The selected metrics were chosen because they capture both global distributional properties (Shannon, GZIP) and conservative worst-case guarantees (min-entropy, Markov). More advanced NIST-defined tests, such as the most common value (MCV) test or restart/adaptive proportion tests, were not applied in this initial study and are planned for future work.

All metrics were computed on raw sensor-derived vectors exported directly from the smartwatch, without post-processing beyond the basic normalization described in Section 3.3.

4.2. Results—Smartwatch Entropy (Still vs. Shake)

The entropy of the collected data was evaluated separately for still and shake modes. Four statistical metrics were applied: Shannon entropy, min-entropy, Markov dependency analysis, and GZIP compression ratio. These metrics provide complementary perspectives on randomness, capturing both average information content and worst-case predictability. The results are summarized in Table 3.

The shake mode consistently outperformed the still mode across all metrics, reflecting the higher variability and reduced temporal correlation introduced by spontaneous wrist motion. Figure 7 shows the distribution of Shannon entropy values (computed according to

Table 3: Comparative entropy evaluation metrics for still and shake acquisition modes.

Metric	**Still Mode**	**Shake Mode**
Shannon entropy (H)	0.991	0.997
Min-entropy (H_{min})	0.851	0.918
Markov dependency (avg. state prob.)	Moderate correlation	Low correlation
GZIP compression ratio	0.72	0.95

Equation (1)) for both modes, with the shake mode exhibiting a slightly higher median and a narrower interquartile range. This indicates that motion-triggered acquisition not only improves the average information content but also provides more stable randomness quality across individual samples.

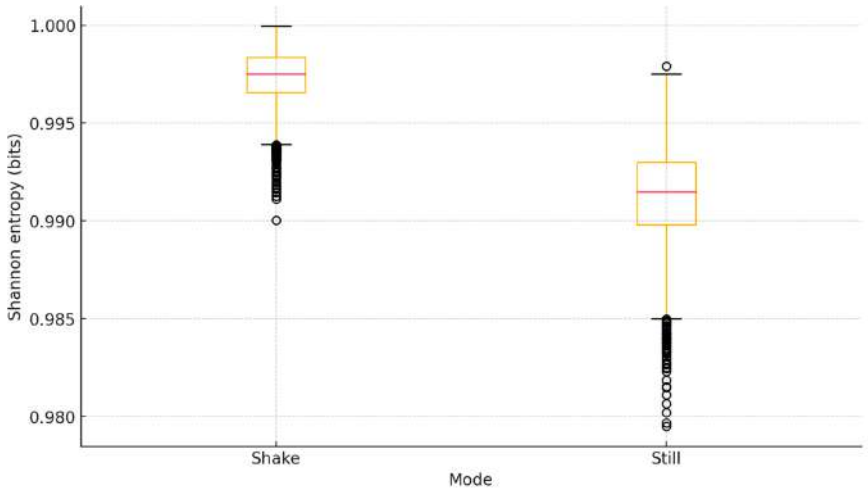

Figure 7: Boxplot comparison of Shannon entropy values for shake and still acquisition modes.

The min-entropy distribution (Figure 8) reveals a more pronounced difference between the two modes. Still-mode samples show a wider spread and a lower median value, indicating stronger predictability and the presence of repeated patterns. In contrast, shake-mode samples cluster around higher values, demonstrating improved worst-case unpredictability, which is critical for cryptographic applications where the weakest entropy segment determines overall security.

Notably, the min-entropy of 0.918 (computed according to Equation (2)) approaches levels typically observed in hardware-based random sources, indicating potential suitability for cryptographic applications. The lower GZIP compression ratio in still mode further confirms the presence of more regular patterns when the device is stationary.

These findings suggest that motion-driven acquisition provides significantly stronger entropy, while still-mode data may serve as a supplementary or fallback source under passive conditions.

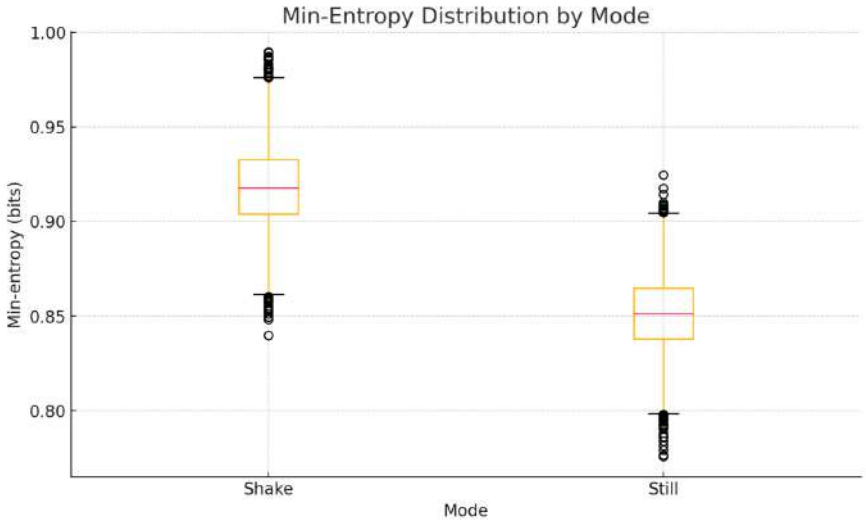

Figure 8: Boxplot comparison of min-entropy values for shake and still acquisition modes.

4.3. Comparative Analysis with Other Methods

To assess the cryptographic relevance of the smartwatch-based entropy source, the obtained results were compared with representative hardware- and software-based random number generators. The comparison includes TRNGs, CSPRNGs, and commonly used operating system RNGs. Metrics for these established methods were taken from publicly available evaluations and previous studies and compared in Table 4.

Table 4: Comparative evaluation of the proposed smartwatch-based entropy generation approach against established entropy sources.

Method/Source	**Shannon Entropy**	**Min-Entropy**	**Notes**
TRNG (e.g., Intel RDRAND)	≈0.999	≈0.998	Hardware, closed-source [9]
CSPRNG (ChaCha20)	≈0.999	≈0.997	Software, deterministic but cryptographically secure [30]
OS RNG (/dev/urandom)	≈0.995	≈0.950	Mixed hardware/software seeding [28]
Smartwatch—Still mode	0.991	0.851	Passive, environmental, and body sensors (this work)
Smartwatch—Shake mode	0.997	0.918	Motion-triggered, multisensor fusion (this work)

The shake mode results are competitive with operating system RNGs, showing only a minor deficit in min-entropy compared to /dev/urandom. While still below TRNG or CSPRNG levels, the smartwatch-based method demonstrates strong unpredictability given

its reliance solely on locally harvested physical signals.

The still mode, although weaker, remains within acceptable entropy levels for non-critical randomness applications. Both modes benefit from decentralization, transparency, and the absence of external dependencies, which are key requirements for privacy-sensitive blockchain ecosystems.

These results demonstrate that a commodity smartwatch, without any dedicated cryptographic hardware, can deliver entropy levels approaching those of widely trusted OS RNGs and even nearing the lower range of TRNG performance. This finding is significant because it proves that secure, high-quality randomness can be sourced from devices that many blockchain users already own, thereby reducing cost barriers and removing the need for specialized external modules. The implications extend to decentralized systems and privacy-sensitive environments, where local and verifiable entropy generation directly on the user's device mitigates supply-chain trust issues and centralization risks inherent in proprietary RNGs.

5 Discussion

The results indicate that smartwatch sensors, particularly in motion-triggered (shake) mode, can serve as a viable source of cryptographic entropy. The observed Shannon entropy (0.997) and min-entropy (0.918) are comparable to software-based random number generators such as /dev/urandom and approach the lower bound of hardware-based TRNGs. This level of unpredictability is considered sufficient for many cryptographic operations, including key generation for blockchain transaction signing, provided that additional post-processing (e.g., cryptographic hashing or key-stretching) is applied [1,2].

The still mode produced weaker results, with a noticeably lower min-entropy (0.851) and higher data regularity, as confirmed by the compression analysis. Nevertheless, even this mode can supplement entropy pools in scenarios where active user interaction is not possible. Its inclusion reflects the practical need for passive entropy harvesting in wearables.

The Markov analysis confirmed that shake-mode data exhibit lower temporal correlations than still-mode data, validating the assumption that user-driven motion introduces additional randomness. This is particularly relevant for non-IID entropy sources, which are common in physical systems [3].

Although human body motion or physiological indicators may seem externally observable, the entropy acquisition process is based on short 2.2-s sampling windows combined with randomized sensor subset selection. Within such a narrow time frame, it is practically infeasible to reproduce identical micro-variations in wrist acceleration, rotational dynamics, barometric pressure changes, and heart rate fluctuations. The unpredictability is further amplified by the fact that the active set of sensors differs across acquisition sessions and that their outputs are concatenated into entropy vectors according to the order in which

measurements are received by the system. This results in subtle but irreproducible variations in vector structure. Consequently, even if gross user movements could be monitored from a distance, it would remain extremely difficult to replicate the exact multidimensional and temporally ordered conditions necessary to obtain the same entropy output.

Although the smartwatch-based method does not yet match the theoretical security of dedicated TRNGs or mature CSPRNGs, it offers unique advantages aligned with blockchain user requirements:

- Decentralization and local control—entropy is generated and used entirely on the user's own device, without reliance on centralized hardware or servers.
- Transparency and auditability—open-source implementation allows public verification, unlike closed systems such as Intel RDRAND.
- Practicality and accessibility—the method leverages existing consumer hardware, removing the need for specialized wallets or external RNG modules.

Future work will include a full NIST SP 800-90B evaluation, extended multi-user testing, and optimization of sensor fusion techniques to further improve min-entropy. These steps are expected to establish the method as a credible alternative for lightweight, user-centric cryptographic applications.

6 Conclusions

This study demonstrated that commercially available smartwatches can serve as practical entropy sources for cryptographic applications, including blockchain transaction signing. By harvesting data from multiple onboard sensors, two acquisition modes—still and shake—were evaluated. The shake mode achieved Shannon entropy of 0.997 and min-entropy of 0.918, values comparable to established software-based random number generators, while the still mode, although weaker, provided a consistent supplementary entropy source.

The method offers several advantages relevant to blockchain ecosystems: it operates fully offline, requires no specialized hardware, and supports open-source implementation, aligning with the principles of decentralization and self-sovereign identity. Although the entropy levels do not yet reach those of dedicated TRNGs, the results indicate that, with appropriate post-processing (e.g., cryptographic hashing), smartwatch-generated entropy is sufficient for secure local key generation.

Several limitations should be acknowledged. The current evaluation was conducted using a single smartwatch model and participant, restricting the generalizability of the results. Only a subset of NIST SP 800-90B tests was applied, and no direct comparison with medical-grade measurement devices was performed, although the smartwatch sensors are factory-calibrated according to manufacturer specifications. A detailed security analysis was not

conducted; however, the potential attack surface is significantly reduced by the fact that entropy harvesting and key generation occur entirely on the smartwatch without transmitting raw data externally.

Future work will expand testing to multi-user datasets, analyze entropy contributions of individual sensors, apply the complete NIST SP 800-90B framework, and include adversarial testing to evaluate robustness against targeted attacks. Additionally, a comparative analysis of power consumption and computational efficiency will be conducted to assess the practicality of long-term use. These steps aim to establish smartwatch-based entropy generation as a credible, user-friendly, and energy-efficient alternative to traditional key management solutions in security-critical domains.

Author Contributions: Conceptualization, M.Š. and M.K.; methodology, M.Š. and Z.K.; software, M.Š.; validation, I.L., M.K. and Z.K.; formal analysis, M.Š. and I.L.; writing—original draft preparation, M.Š.; writing—review and editing, I.L. and Z.K.; visualization, M.K.; supervision, M.K.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to its non-invasive nature and the fact that the only participant was the lead author.

Informed Consent Statement: Written informed consent has been obtained from the author, who was the only participant in the study.

Data Availability Statement: The data and source code supporting the findings of this study are publicly available at `https://github.com/msvarcmajer/Entropy_samsung` (accessed on 25 August 2025). The repository contains all sensor datasets in CSV format and the Wear OS application source code used for data acquisition and entropy evaluation.

Conflicts of Interest: The authors declare no conflicts of interest.

References (as published in the original article)

1. Kristoffersson, A.; Lindén, M. A Systematic Review of Wearable Sensors for Monitoring Physical Activity. *Sensors* **2022**, 22, 573.

2. Piwek, L.; Ellis, D.A.; Andrews, S.; Joinson, A. The Rise of Consumer Health Wearables: Promises and Barriers. *PLoS Med.* **2016**, 13, e1001953.

3. Yang, W.; Wang, S.; Sahri, N.M.; Karie, N.M.; Ahmed, M.; Valli, C. Biometrics for Internet-of-Things Security: A Review. *Sensors* **2021**, 21, 6163.

4. Sönmez Turan, M.; Barker, E.; Kelsey, J.; McKay, K.; Baish, M.; Boyle, M. *Recommendation*

for the Entropy Sources Used for Random Bit Generation; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2018.

5. Răstoceanu, F.; Rughiniş, Răzvan; Ciocîrlan, Ştefan-Dan; Enache, M. Sensor-Based Entropy Source Analysis and Validation for Use in IoT Environments. *Electronics* **2021**, 10, 1173.

6. Antonopoulos, A.M. *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*; O'Reilly Media: Sebastopol, CA, USA, 2014; ISBN 978-1-4919-0264-6.

7. Narayanan, A.; Bonneau, J.; Felten, E.; Miller, A.; Goldfeder, S. *Bitcoin and Cryptocurrency Technologies*; Princeton University Press: Princeton, NJ, USA, 2016; ISBN 978-0-691-17169-2.

8. Kshetri, N. Blockchain's Roles in Strengthening Cybersecurity and Protecting Privacy. *Telecommun. Policy* **2017**, 41, 1027–1038.

9. Intel Corporation. Intel® Digital Random Number Generator (DRNG) Software Implementation Guide. Available online: `https://www.intel.com/content/www/us/en/developer/articles/guide/intel-digital-random-number-generator-drng-software-html` (accessed on 16 July 2025).

10. Barker, E.; Kelsey, J. *Recommendation for Random Number Generation Using Deterministic Random Bit Generators*; NIST Special Publication 800-90A Rev. 1; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2015.

11. Eskandari, S.; Moosavi, S.; Clark, J. SoK: Transparent Dishonesty: Front-Running Attacks on Blockchain. In *Financial Cryptography and Data Security, FC 2020*; Springer: Cham, Switzerland, 2020; pp. 170–189.

12. Krombholz, K.; Mayer, W.; Schmiedecker, M.; Weippl, E. "I Have No Idea What I'm Doing": On the Usability of Deploying HTTPS. In Proceedings of the 26th USENIX Security Symposium, Vancouver, BC, Canada, 16–18 August 2017; pp. 1339–1356.

13. Das, A.; Borisov, N.; Caesar, M. Do You Hear What I Hear? Fingerprinting Smart Devices Through Embedded Acoustic Components. In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, AZ, USA, 3–7 November 2014; pp. 441–452.

14. Silva-Trujillo, A.G.; González González, M.J.; Rocha Pérez, L.P.; García Villalba, L.J. Cybersecurity Analysis of Wearable Devices: Smartwatches Passive Attack. *Sensors* **2023**, 23, 5438.

15. Gudgeon, L.; Perez, D.; Harz, D.; Livshits, B.; Gervais, A. The Decentralized Financial Crisis. In Proceedings of the 2020 Crypto Valley Conference on Blockchain Technology, Rotkreuz, Switzerland, 11–12 June 2020; pp. 1–15.

16. Ilieva, G.; Yankova, T.; Dzhabarova, Y.; Ruseva, M.; Angelov, D.; Klisarova-Belcheva, S. Customer Attitude toward Digital Wallet Services. *Systems* **2023**, 11, 185.

17. Cornelius, C.; Kotz, D. Recognizing Whether Sensors Are on the Same Body. In *Pervasive Computing, Pervasive 2011*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 332–349.

18. Rukhin, A.; Soto, J.; Nechvatal, J.; Smid, M.; Barker, E.; Leigh, S.; Levenson, M.; Vangel, M.; Banks, D.; Heckert, N.; et al. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*; NIST Special Publication 800-22, Rev. 1a; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2010.

19. Yu, H.; Kim, Y. True Random Number Generator Using Bio-Related Signals in Wearable Devices. In Proceedings of the 2018 International SoC Design Conference, Daegu, Republic of Korea, 12–15 November 2018; pp. 231–232.

20. Killmann, W.; Schindler, W. *A Proposal for Functionality Classes for Random Number Generators*; Bundesamt für Sicherheit in der Informationstechnik (BSI): Bonn, Germany, 2011.

21. ISO/IEC 18031:2025; *Information Technology—Security Techniques—Random Bit Generation.* International Organization for Standardization/International Electrotechnical Commission: Geneva, Switzerland, 2025.

22. CCMB. *Common Methodology for Information Technology Security Evaluation*; Version 3.1, Revision 5; CCMB-2017-04-004; Common Criteria Management Board: Brussels, Belgium, 2017.

23. Ledger. Ledger's Security Model: How Are Ledger Devices Secured? Available online: `https://www.ledger.com/academy/basic-basics/ledgers-ecosystem/why-is-ledger-nano` (accessed on 18 July 2025).

24. Sarfaraz, A.; Chakrabortty, R.K.; Aslam, S.; Rafsanjani, A.S. Optimized Key Recovery for Blockchain Wallets in Sustainable Supply Chains. In Proceedings of the 2024 IEEE International Conference on Industrial Engineering and Engineering Management, Bangkok, Thailand, 15–18 December 2024; pp. 343–347.

25. Bitcoin Community. BIP-0039: Mnemonic Code for Generating Deterministic Keys. Available online: `https://github.com/bitcoin/bips/blob/master/bip-0039.mediawiki` (accessed on 14 July 2025).

26. Trusted Computing Group. *TPM 2.0 Library Specification*; Trusted Computing Group: Beaverton, OR, USA, 2019.

27. Eastlake, D.E., 3rd; Crocker, S.; Schiller, J.I. *Randomness Requirements for Security*; RFC 4086; Internet Engineering Task Force (IETF): Wilmington, NC, USA, 2005.

28. Gutterman, Z.; Pinkas, B.; Reinman, T. Analysis of the Linux Random Number Generator. In Proceedings of the 2006 IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 21–24 May 2006; pp. 371–385.

29. OpenSSL Software Foundation. RAND_bytes—OpenSSL Documentation. Available online: `https://docs.openssl.org/1.1.1/man3/RAND_bytes/` (accessed on 16 July 2025).

30. Bernstein, D.J. ChaCha, a Variant of Salsa20. Workshop Record of SASC 2008: The State of the Art of Stream Ciphers. 2008, pp. 1–14.

31. Song, K.; Zhu, Z.; Yang, H.; Ni, T.; Xu, W. MobileKey: A Fast and Robust Key Generation System for Mobile Devices. In Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2022; pp. 427–431.

32. Lv, N.; Chen, T.; Ma, Y. An Analysis on Entropy Sources Based on Smartphone Sensors. In Proceedings of the 2020 10th International Conference on Communication and Network Security, Tokyo, Japan, 27–29 November 2020; pp. 21–31.

33. Moosavi, S.R. PPG-KeyGen: Using Photoplethysmogram for Key Generation in Wearable Devices. *Procedia Comput. Sci.* **2021**, 184, 291–298.

34. Ajlan, K.A.; Alsboui, T.; Alshaikh, O.; Inuwa-Dute, I.; Khan, S.; Parkinson, S. The Emerging Challenges of Wearable Biometric Cryptosystems. *Cryptography* **2024**, 8, 27.

35. Samsung Electronics. Health Sensor Data Specifications. Available online: `https://developer.samsung.com/health/sensor/guide/data-specifications.html` (accessed on 16 July 2025).

36. Google. Sensor|API Reference. Available online: `https://developer.android.com/reference/android/hardware/Sensor` (accessed on 16 July 2025).

37. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, 27, 379–423.

38. Thinh, L.P.; Sheridan, L.; Scarani, V. Bell Tests with Min-Entropy Sources. *Phys. Rev. A* **2013**, 87, 062121.

39. Mettle, F.O.; Quaye, E.N.B.; Laryea, R.A. A Methodology for Stochastic Analysis of Share Prices as Markov Chains with Finite States. *SpringerPlus* **2014**, 3, 657.

40. Lee, T.-H.; He, J.-D. Entropy-Based Profiling of Network Traffic for Detection of Security Attack. In Proceedings of the TENCON 2009 IEEE Region 10 Conference, Singapore, 23–26 January 2009; pp. 1–5.

41. Klein, S.T.; Shapira, D. On the Randomness of Compressed Data. *Information* **2020**, 11, 196.

C

Paper C: Predictive Modeling of Quantum Computing Development and Its Impact on Blockchain Security

IEEE Access, vol. 14, pp. 68624–68636, May 2026 | DOI: 10.1109/ACCESS.2026.3689963

Miljenko Švarcmajer, Dina Šabanović, Mirko Köhler, and Ivica Lukić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, J. J. Strossmayer University of Osijek, 31000 Osijek, Croatia

Abstract

Quantum computers pose a long-term threat to classical cryptographic systems, including those used in blockchain platforms. While prior work has estimated the number of logical qubits required to compromise cryptographic schemes such as RSA and ECDSA, relatively few studies have modeled when such quantum capacity might realistically be achieved. In this paper, we construct a time-series dataset of publicly documented gate-based quantum processors developed between 2016 and 2025. We estimate the number of logical qubits under three fault-tolerance scenarios and apply forecasting models based on exponential trend analysis. Our results provide year-by-year projections for when quantum computers may reach logical qubit thresholds associated with widely used cryptographic schemes. These

projections are mapped to specific blockchain security mechanisms, offering a practical timeline for transition planning toward post-quantum systems. These projections should be interpreted as indicators of logical qubit availability rather than precise estimates of full cryptanalytic execution timelines.

Index Terms: Blockchain security, cryptographic thresholds, ECDSA, fault tolerance, logical qubits, post-quantum cryptography, quantum computing, quantum forecasting, quantum hardware growth, time-series modeling.

1 Introduction

Quantum computing is emerging as one of the most consequential developments in modern information science. Unlike classical systems that rely on deterministic binary logic, quantum machines operate using the principles of quantum mechanics, enabling new forms of computation based on superposition, entanglement, and interference. This paradigm allows quantum algorithms to solve certain classes of problems more efficiently than their classical counterparts.

The theoretical foundation of quantum computing was first outlined by Richard Feynman in 1982, who proposed that classical systems are inherently inefficient at simulating quantum processes [1]. David Deutsch later formalized the model of a universal quantum computer [2], establishing the theoretical framework for a machine capable of performing any computation that follows the laws of physics.

A turning point occurred in 1994 when Peter Shor introduced an efficient algorithm for factoring large integers [3], directly threatening the security of the RSA cryptosystem. Grover's algorithm followed, reducing the complexity of unstructured search and impacting symmetric encryption schemes [4]. Together, these results demonstrated that quantum computers could mathematically outperform classical systems on specific cryptographic tasks, a finding that has since been rigorously verified and widely accepted in the field.

Although physical implementations of quantum processors remain limited in scale and stability, progress has accelerated. Devices with hundreds or even thousands of physical qubits are now publicly documented. The transition from noisy intermediate-scale quantum (NISQ) devices to fault-tolerant quantum computers, however, requires significant overhead in error correction. A single logical qubit may demand hundreds or thousands of physical qubits [5], creating a substantial resource gap between theoretical capability and practical threat.

This gap is particularly relevant for blockchain systems, which rely on the long-term hardness of classical cryptographic primitives such as ECDSA, EdDSA, and SHA-256. Because these networks are publicly auditable and append-only, any compromise in the cryptographic layer could retroactively impact their integrity. As quantum computing capability grows, so does the urgency to understand when such attacks may become feasible.

Most existing work focuses on estimating the number of qubits needed to break specific cryptosystems or optimizes quantum circuits for such attacks. However, relatively few studies attempt to forecast when quantum hardware will actually reach these thresholds based on empirical development trends. This paper aims to fill that gap by constructing a time-series dataset of real quantum devices and using it to project the likely emergence of cryptographically dangerous quantum capacity.

The contributions of this work are summarized as follows:

- We construct a curated and reproducible time-series dataset of gate-based quantum processors, focusing on frontier systems that define the leading edge of hardware capabilities.
- We develop a capacity-based modeling approach for estimating logical qubit availability using conservative, baseline, and optimistic fault-tolerance assumptions drawn from published literature [6], [7].
- We integrate empirical hardware trends with cryptographic resource estimates, linking projected quantum capacity to the specific requirements of ECDSA-256 and blockchain security models.
- We provide a multi-scenario forecasting analysis that incorporates uncertainty through bootstrap resampling, sensitivity analysis, and leave-one-out validation, offering a structured assessment of projection robustness.
- We translate hardware-level projections into system-level implications for blockchain platforms, including actionable timelines for post-quantum migration planning.

The remainder of this paper is organized as follows. Section II reviews the cryptographic impact of quantum computing and the structure of blockchain cryptography. Section III describes the dataset construction and filtering process. Section IV outlines the forecasting methodology. Section V presents the forecasting results, while Section VI discusses the implications and transition strategies. Section VII concludes the paper with final observations and future directions.

2 Background and Related Work

This section provides the necessary technical background on quantum cryptographic threats and blockchain cryptography. It also reviews recent literature that quantifies the quantum resources required to compromise classical systems and forecasts the timeline for practical attacks.

A. Quantum Computing and Cryptographic Vulnerabilities

Not all quantum computers are capable of undermining modern cryptographic protections. Many experimental systems, such as those developed by D-Wave Systems and used by institutions like Lockheed Martin, are built for specialized optimization tasks and rely on quantum annealing [8]. These systems are not universal quantum computers and cannot execute general-purpose quantum algorithms.

To pose a credible threat to cryptographic systems, a quantum computer must meet strict architectural and operational criteria. A viable quantum computer must satisfy the DiVincenzo criteria, including scalable qubits, reliable initialization, long coherence times, universal gate operations, and efficient measurement [9].

Only gate-based quantum processors that fulfill (or aim to fulfill) these criteria are relevant for cryptographic applications. Furthermore, they must support fault-tolerant computation through quantum error correction. This requires constructing stable logical qubits from multiple physical qubits, often at an overhead of hundreds to thousands of physical qubits per logical one [7].

Once fault tolerance is achieved, these devices can theoretically implement two quantum algorithms that directly compromise classical cryptography: Shor's algorithm enables polynomial-time factorization and discrete logarithm computation, effectively breaking RSA, Diffie–Hellman, and elliptic-curve cryptography [3]. Grover's algorithm provides a quadratic speedup for brute-force search, halving the effective key length in symmetric systems such as AES or SHA-2 [4].

While Grover's attack can be mitigated by doubling key or digest sizes (e.g., using AES-256 instead of AES-128), the threat posed by Shor's algorithm requires complete cryptographic migration. Table 1 summarizes the quantum vulnerability of standard cryptographic primitives.

Table 1: Quantum threat model for classical cryptography.

Scheme	**Algorithm Type**	**Quantum Threat**	**Impact**
RSA-2048	Asymmetric	Shor	Broken
Elliptic Curve (e.g., secp256k1)	Asymmetric	Shor	Broken
Diffie–Hellman (DH)	Asymmetric	Shor	Broken
AES-128	Symmetric	Grover	64-bit effective
AES-256	Symmetric	Grover	Acceptable
SHA-256	Hash Function	Grover	Reduced Security

In the context of blockchain systems, this quadratic speedup reduces the effective security

of hash functions (e.g., SHA-256) by approximately half, effectively lowering a 256-bit security level to 128 bits. While this does not render hash-based mechanisms immediately insecure, it increases the feasibility of brute-force and preimage attacks, particularly in long-term threat scenarios. As a result, hash-based components of blockchain systems may require parameter adjustments, such as increased digest sizes, to maintain adequate security margins.

This paper focuses only on gate-based quantum processors with the theoretical capacity to threaten public-key cryptosystems. The following sections evaluate the current and projected capabilities of such devices and estimate the timeframe in which they may reach cryptographically significant thresholds.

B. Blockchain Cryptography

Modern blockchain systems rely heavily on public-key and hash-based cryptographic mechanisms to ensure integrity, authentication, and consensus. These primitives form the foundation of address generation, transaction validation, and block creation. Since public blockchains are fully transparent and immutable, any cryptographic compromise affects all past and future data.

The most widely adopted blockchain platforms—Bitcoin, Ethereum, and Solana—use elliptic-curve cryptography (ECC) for digital signatures and cryptographic hash functions for mining and data integrity.

Bitcoin uses the secp256k1 elliptic curve for digital signatures through the Elliptic Curve Digital Signature Algorithm (ECDSA). The integrity of transactions is ensured via SHA-256, which also forms the basis of Bitcoin's proof-of-work mechanism [10].

Ethereum (pre-2.0 and post-Merge) also uses ECDSA over secp256k1, though some applications within Ethereum have experimented with alternative schemes such as EdDSA. Keccak-256 (a variant of SHA-3) is used for hashing, particularly in address generation and state commitment [11].

Solana, a newer high-throughput blockchain, uses ed25519 signatures based on Edwards curves and SHA-256 for hashing. Ed25519 is regarded as more efficient and resistant to certain implementation vulnerabilities compared to ECDSA, but it is still fundamentally based on elliptic-curve hardness assumptions.

Table 2 summarizes the main cryptographic components used by these platforms.

Table 2: Cryptographic mechanisms in selected blockchain platforms.

Platform	**Signature Scheme**	**Curve/Algorithm**	**Hash Function**
Bitcoin	ECDSA	secp256k1	SHA-256
Ethereum	ECDSA	secp256k1	Keccak-256
Solana	EdDSA	ed25519	SHA-256

All of these platforms rely on assumptions that are vulnerable to quantum algorithms.

In particular, ECDSA and EdDSA can be broken using Shor's algorithm, since they rely on the intractability of the elliptic-curve discrete logarithm problem. Hash-based mechanisms are more resilient, but Grover's algorithm still weakens their effective security by a square root factor.

While hash functions can be strengthened by increasing digest length, replacing elliptic-curve cryptography in decentralized systems is significantly more challenging due to protocol rigidity, backward compatibility, and the lack of widely adopted post-quantum alternatives. Ethereum developers and research groups such as the Ethereum Foundation [12] and Cardano's IOHK [13] have begun investigating quantum-resilient cryptography, but no major blockchain platform currently offers native post-quantum security.

This makes blockchain networks an attractive long-term target for adversaries with delayed decryption strategies—harvesting encrypted public data today, and decrypting it once quantum hardware becomes capable.

C. Related Work

Recent years have seen an intensification of research on the cryptographic impact of large-scale quantum computing. Several lines of work focus on estimating the quantum resources required to compromise existing cryptographic systems, projecting the growth rate of quantum hardware, and assessing the specific implications for blockchain security.

From an algorithmic perspective, Gidney and Ekerå [14] introduced an optimized implementation of Shor's algorithm for factoring RSA-2048 using approximately 20 million physical qubits and one million T-gates. This work was among the first to quantify in detail the space-time cost of a cryptographic quantum attack using realistic fault-tolerant assumptions. In a significant follow-up, Gidney [15] further improved the resource estimates, showing that the same cryptographic goal could be achieved with just one million physical qubits (or 1400 logical qubits) in under a week. This nearly 20× reduction in hardware requirements over just three years highlights how algorithmic advances alone can substantially accelerate the quantum threat.

Complementary studies by Webber et al. [16] and Gheorghiu and Mosca [17] provide detailed estimates of the number of physical and logical qubits required to break asymmetric and symmetric cryptographic schemes under various timing constraints. Webber et al., for example, calculated that breaking Bitcoin's ECDSA key within one hour would require over 300 million physical qubits, assuming state-of-the-art error correction. These findings emphasize that not just qubit counts but also fidelity, gate time, and architectural constraints define real-world feasibility.

Several works have also modeled the growth trajectory of quantum hardware. Sevilla and Riedel [18] created a dataset of 50+ quantum processors and fit historical data to predict when cryptographically relevant quantum computers might emerge. Their forecast

suggested less than a 5% chance of an RSA-breaking quantum device by 2039 under standard trends. Ezratty [19] extended this analysis with a critique of Moore-style projections, showing that qubit count alone does not guarantee useful scaling due to fidelity limitations and architectural bottlenecks.

In the context of blockchain, Schärer and Comuzzi [20] and Khodaiemehr et al. [21] analyzed how quantum threats apply to signature schemes such as ECDSA and Ed25519, which are foundational to Bitcoin, Ethereum, and Solana. They noted that once public keys are revealed on-chain, they become retroactively vulnerable to quantum attacks. This threat is especially critical when quantum attacks are possible within short timeframes: any system using RSA-based authentication, for instance, could be hijacked in real time if factoring RSA-2048 takes under 24 hours. Such a capability would allow adversaries to compute private keys from exposed public keys during the validation window, effectively stealing control of funds or system credentials.

Finally, Švarcmajer et al. [22] and Baseri et al. [23] proposed hybrid migration strategies and risk assessments for blockchain networks under quantum pressure. They emphasize the need to upgrade consensus, signature, and identity layers well before quantum hardware reaches relevant thresholds.

This paper complements prior work by shifting focus from cryptanalytic resource estimation to a predictive modeling approach. We use documented hardware trajectories to estimate when quantum systems may reach the critical number of logical qubits needed to compromise current cryptographic primitives, enabling a year-by-year forecast of risk horizons across blockchain platforms.

3 Data Collection and Preprocessing

This section describes the construction of the quantum computing dataset, the criteria used to filter and standardize system characteristics, and the methodology employed to translate reported physical-qubit counts into estimated logical-qubit capacity.

A. Quantum Computing Dataset

The dataset was compiled from publicly available sources describing representative quantum processors across multiple generations, including peer-reviewed publications, preprints, and official technical disclosures. Only systems with verifiable technical specifications were included [24]–[38].

For each quantum computing system or processor, the dataset records four principal attributes: the year of initial announcement or public release, the official system or processor designation, the number of physical qubits, and the foundational quantum technology employed. The dataset encompasses diverse technological paradigms, including superconducting qubits

with an emphasis on transmon-based architectures, trapped-ion platforms incorporating ytterbium ions, and related variants. To ensure data integrity and reproducibility, all numerical values and system characteristics were independently corroborated against primary technical literature or authoritative hardware documentation.

While different quantum computing technologies exhibit distinct scaling characteristics and engineering constraints, the empirical progression of frontier systems shows a consistent increase in achievable qubit counts over time. This study therefore focuses on modeling the observed frontier trend, rather than deriving architecture-specific scaling laws.

B. Filtering and Selection Criteria

The dataset is restricted to gate-based, programmable quantum processors capable of implementing universal quantum circuits. This restriction reflects the architectural requirements for executing Shor's algorithm and related fault-tolerant cryptanalytic workloads, and is aligned with the DiVincenzo criteria for viable quantum computation [9]. These criteria include scalable qubits, long coherence times, universal gate sets, qubit initialization, and efficient measurement.

Photonic systems, quantum annealers, and educational or simulator platforms were excluded from consideration. In numerous photonic implementations, reported qubit counts frequently refer to optical modes, time bins, or sampling complexity metrics, which are not directly analogous to physical qubits in error-corrected, circuit-based architectures. Furthermore, quantum annealers do not possess the universal gate model necessary for the implementation of general-purpose quantum algorithms.

While quantum annealing systems may offer advantages in specific optimization problems and could potentially impact certain aspects of blockchain systems (e.g., heuristic optimization or mining-related processes), their computational model is not universal and does not support the implementation of circuit-based quantum algorithms such as Shor's algorithm.

Similarly, photonic and sampling-based quantum systems lack the fault-tolerant, gate-based architecture required for scalable cryptanalysis of public-key cryptographic schemes. Although these platforms may demonstrate advantages in specialized domains, their impact on digital signature security in blockchain systems is currently limited.

In addition to these established platforms, emerging architectures such as neutral atom arrays have recently demonstrated promising scalability and controllability characteristics. While these systems are still under active development, they may contribute to future advances in large-scale, fault-tolerant quantum computing [39].

This filtering guarantees that all systems included in the dataset are architecturally pertinent to fault-tolerant cryptanalysis and that qubit counts correspond to physically instantiated, gate-addressable qubits.

C. Logical Qubit Estimation

The progression from noisy intermediate-scale quantum (NISQ) devices to fully fault-tolerant quantum computers requires encoding quantum information within logical qubits protected by quantum error-correction (QEC) codes. Logical qubits are constructed using quantum error correction (QEC), which distributes quantum information across multiple physical qubits to suppress errors. The required overhead depends on physical error rates, QEC protocols, and target logical error probabilities [7], [40]. Consequently, physical qubit counts alone do not directly reflect cryptanalytic capability, necessitating explicit modeling of QEC overhead.

To evaluate when quantum computers may present a credible threat to current cryptographic systems, this analysis applies physical-to-logical qubit ratios from existing literature to representative hardware growth trajectories. The study is formulated as a capacity-threshold analysis, estimating when sufficient logical qubits may become available rather than the wall-clock time required to execute a complete cryptanalytic attack, which also depends on gate speed, T-state distillation throughput, and architectural parallelism constraints.

For surface codes, which represent the most mature QEC architecture for superconducting quantum platforms, the number of physical qubits required per logical qubit scales approximately as

$$N_{\text{phys}} = 2(d+1)^2, \tag{1}$$

where d denotes the code distance [14]. Foundational studies suggest that achieving logical error rates suitable for large-scale, fault-tolerant quantum computation typically demands between 10^3 to 10^4 physical qubits for each logical qubit. The exact number depends on the physical error rates of the system and various architectural factors [7]. Importantly, these estimates capture not only the overhead of encoding, but also additional system-level requirements—including qubit routing, allocation of auxiliary qubits, syndrome extraction processes, and the synthesis of fault-tolerant gates. At the scales relevant for cryptanalysis, these system-level costs can become especially significant.

Recent experimental advances by Google Quantum AI demonstrated the realization of a distance-7 surface-code logical qubit implemented using 101 physical qubits, achieving a logical error rate of 0.143% per cycle and surpassing the lifetime of the best constituent physical qubit by a factor of 2.4 [41]. This achievement confirms operation below the surface-code threshold and represents a significant milestone toward scalable quantum computation. Nevertheless, cryptographically relevant computations require substantially larger code distances than those demonstrated in proof-of-concept experiments. Accordingly, small-distance demonstrations should be interpreted as evidence of threshold operation rather than as direct indicators of imminent large-scale algorithmic capability.

Gidney and Ekerå [14] provide a comprehensive resource estimate for factoring RSA-2048

using Shor's algorithm, requiring approximately 20 million physical qubits and about 14,000 logical qubits, corresponding to an effective ratio of 1400:1. This estimate incorporates routing, distillation, and architectural constraints, making it a realistic benchmark for large-scale cryptanalysis.

In the context of elliptic curve cryptography, Roetteler et al. [42] estimate that solving the elliptic curve discrete logarithm problem over a 256-bit prime field, corresponding to ECDSA-256 employed in Bitcoin and Ethereum, requires approximately 2,330 logical qubits. While the mathematical structures underpinning RSA and elliptic curve cryptography are distinct, the resource analysis for breaking RSA-2048 currently offers the most detailed and widely accepted architectural model for evaluating large-scale cryptanalytic workloads. As such, we use the RSA-2048 benchmark as a practical reference point for estimating the physical-to-logical qubit overhead required. This approach enables us to estimate the hardware scale at which quantum cryptanalysis may become feasible under surface-code-based fault tolerance. It is important to note, however, that future advances in circuit design or error correction may shift these requirements.

Recent advances in quantum algorithm optimization suggest that the required number of logical qubits for elliptic-curve cryptanalysis may be further refined under certain assumptions. For example, improvements in circuit design and resource estimation have been reported in more recent studies such as Yan et al. [43]. However, such refinements primarily affect the exact threshold value rather than the overall structure of the capacity-based modeling approach adopted in this work.

Accordingly, the threshold of 2,330 logical qubits is retained as a representative and widely cited reference point for ECDSA-256 cryptanalysis, providing a consistent baseline for comparative analysis, while acknowledging that future improvements may lead to moderate adjustments of this value.

Recent advances in quantum low-density parity-check (qLDPC) codes offer the potential for substantial reductions in overhead. Bravyi et al. [44] demonstrated high-threshold, low-overhead quantum memory constructions capable of encoding 12 logical qubits in 288 physical qubits. These constructions achieve significantly higher encoding rates than conventional surface codes under favorable noise assumptions. Nevertheless, they require non-local qubit connectivity, which is not currently available on most hardware platforms, and large-scale experimental validation remains ongoing. Notably, IBM has publicly outlined a roadmap to experimentally realize such fault-tolerant architectures in the latter half of this decade [45]. For this reason, our primary scenarios remain anchored in experimentally validated surface-code assumptions, while qLDPC developments motivate the optimistic end of the physical-to-logical ratio range explored in Section IV-A.

4 Methodology

This section outlines the quantitative approach used to estimate when quantum hardware may reach cryptographically significant thresholds. The methodology comprises four main components: estimating logical qubit capacity from physical qubit data, constructing and filtering the dataset of quantum processors, applying multiple forecasting models to fit the data, and evaluating uncertainty through resampling and sensitivity analysis. All models are built around the assumption that only fault-tolerant, gate-based quantum processors are relevant for cryptanalytic applications.

A. Modeling Logical Qubit Conversion

To estimate logical qubit capacity, we apply physical-to-logical ratios based on findings from peer-reviewed, large-scale resource analyses. Because error-correction overhead is multiplicative, logical capacity is modeled as

$$Q_{\text{logical}} = \frac{Q_{\text{phys}}}{\rho}, \tag{2}$$

where ρ represents the effective physical-to-logical conversion ratio.

Three scenarios are considered to span plausible architectural trajectories:

- **Optimistic**: $\rho = 1000{:}1$, reflecting continued improvement in physical error rates (approaching 10^{-4}), optimized surface-code implementations, and partial realization of projected overhead reductions from emerging QEC schemes.
- **Baseline**: $\rho = 1500{:}1$, consistent with end-to-end RSA-2048 resource estimates incorporating routing and magic-state distillation overhead [14].
- **Conservative**: $\rho = 3000{:}1$, reflecting slower fidelity improvements, higher required code distances, and less efficient distillation pipelines.

Taken together, these scenarios span an uncertainty range of approximately a factor of three in overhead. It is important to emphasize that this conversion addresses only the availability of logical qubits; it does not account for algorithm runtime, T-gate throughput, or constraints on architectural parallelism. As such, these projections should be interpreted as logical qubit capacity thresholds, not as comprehensive time-to-break estimates.

In particular, the baseline scenario incorporates overhead associated with magic-state distillation, which represents a dominant resource cost in fault-tolerant implementations of Shor's algorithm. This ensures that the conversion ratios reflect system-level requirements beyond idealized logical qubit encoding.

It should be noted that the physical-to-logical qubit conversion ratio is not a fixed constant in practice, but depends on factors such as physical error rates, code distance,

decoding efficiency, and architectural constraints. In this work, this variability is approximated through a scenario-based approach, where optimistic, baseline, and conservative ratios are used to capture a plausible range of fault-tolerance overheads reported in the literature.

To clarify the evaluation framework, the proposed forecasting model is based on several key assumptions. First, only gate-based, fault-tolerant quantum computers are considered, as these are required for executing Shor's algorithm. Second, the model assumes that physical qubit growth observed in frontier systems can be extrapolated under sustained technological progress, without abrupt discontinuities caused by unforeseen physical or engineering limitations. Third, it is assumed that physical error rates improve sufficiently over time to support scalable surface-code error correction, allowing the physical-to-logical conversion ratios to remain within the defined scenario bounds (1:1000 to 1:3000).

The evaluation is therefore conducted under a capacity-based framework, where projections reflect the availability of logical qubits rather than full cryptanalytic execution feasibility. All reported results should be interpreted within these assumptions and corresponding modeling constraints.

B. Dataset Construction

This analysis draws on publicly reported qubit counts from 18 quantum computing systems announced between 2016 and 2024. For comparative purposes, we constructed two datasets:

- The full dataset ($n = 18$) includes all reported systems, representing the breadth of quantum hardware development across multiple manufacturers and technologies.
- The frontier dataset ($n = 8$) comprises the maximum qubit count observed in each year, subject to a monotonicity constraint to reflect the cumulative maximum. This approach captures the progression of leading-edge capability, which is pertinent to questions regarding the timeline for potential cryptographic threats.

Because the central goal of this research is to identify the earliest point at which quantum computers might pose a cryptographic threat, we use the frontier dataset as the primary basis for forecasting. The full dataset serves as a secondary robustness check, offering further context about the wider field of quantum hardware development. While the frontier dataset includes only eight data points, this limited sample is intentional and methodologically sound: each entry represents the most advanced gate-based quantum processor publicly disclosed in a given year, as documented in peer-reviewed or technical sources. Incorporating additional systems with lower qubit counts or notable experimental constraints would dilute the forecast's specificity and relevance. Thus, the small sample size reflects the current realities of public reporting rather than a methodological shortcoming.

C. Physical-to-Logical Qubit Conversion

Physical qubit counts are converted to logical qubit estimates using the three physical-to-logical ratios defined in Section IV-A.

Table 3: Full dataset of gate-based quantum processors (2016–2024).

Year	System / Processor	Physical Qubits	Technology
2016	IBM Quantum Experience	5	Superconducting
2018	Intel Tangle Lake	49	Superconducting
2019	Google Sycamore	54	Superconducting
2020	Quantinuum System Model H1	10	Trapped-ion
2020	Rigetti Aspen-7	32	Superconducting
2021	IBM Eagle	127	Superconducting
2021	USTC Zuchongzhi 2.0	66	Superconducting
2021	Rigetti Aspen-11	40	Superconducting
2022	Quantinuum System Model H1 (upgrade)	20	Trapped-ion
2022	Rigetti Aspen-M	80	Superconducting
2022	IBM Osprey	433	Superconducting
2022	IonQ Aria	25	Trapped-ion
2023	IBM Condor	1,121	Superconducting
2023	Quantinuum System Model H2	32	Trapped-ion
2024	Quantinuum System Model H2 (upgrade)	56	Trapped-ion
2024	IonQ Forte	36	Trapped-ion
2024	Origin Quantum Wukong	72	Superconducting
2024	Google Willow	105	Superconducting

For each projected physical qubit count Q_{phys}, logical qubit capacity is computed as:

$$Q_{\text{logical}} = \frac{Q_{\text{phys}}}{\rho}, \tag{3}$$

where $\rho \in \{1000, 1500, 3000\}$ corresponds to the optimistic, baseline, and conservative scenarios, respectively.

The cryptographic threshold is set at 2,330 logical qubits, consistent with published resource estimates for solving the elliptic-curve discrete logarithm problem underlying ECDSA-256 [42].

Table 4: Frontier dataset (leading-edge systems per year).

Year	Frontier System	Physical Qubits
2016	IBM Quantum Experience	5
2018	Intel Tangle Lake	49
2019	Google Sycamore	54
2020	Google Sycamore (frontier retained)	54
2021	IBM Eagle	127
2022	IBM Osprey	433
2023	IBM Condor	1,121
2024	IBM Condor (frontier retained)	1,121

D. Curve Fitting Model

Given historical evidence of sustained exponential growth in frontier qubit counts, logical capacity is modeled as

$$Q(t) = ae^{bt}, \tag{4}$$

where t denotes years since 2016, a is a scale parameter, and b is the continuous growth rate.

Estimation is performed via ordinary least squares on the log-transformed model:

$$\ln Q(t) = \alpha + bt. \tag{5}$$

This specification assumes multiplicative growth dynamics and homoscedastic residuals in log-space. The implied doubling time is

$$T_d = \frac{\ln(2)}{b}. \tag{6}$$

Threshold-crossing time is computed analytically as

$$t^* = \frac{\ln(Q^*/a)}{b}. \tag{7}$$

Since the conversion ratio ρ appears as a multiplicative factor, it results solely in a shift of the intercept α without influencing the estimated growth rate b. Therefore, uncertainty associated with sustained growth predominates over uncertainty arising from overhead assumptions.

E. Uncertainty and Robustness Analysis

1) Bootstrap Confidence Intervals. To quantify forecast uncertainty, nonparametric bootstrap resampling with 2,000 iterations is employed [46]. In each iteration, the frontier dataset is resampled with replacement, the log-linear model is re-estimated, and the year in which the threshold is predicted to be crossed is determined analytically. The 95% confidence intervals are defined by the empirical 2.5th and 97.5th percentiles of the bootstrap distribution. Because the analysis relies exclusively on resampling without introducing additional parameter perturbations, the resulting intervals reflect sampling variability arising from the limited historical data. Given the limited size of the frontier dataset, these confidence intervals should be interpreted cautiously, as bootstrap resampling over a small number of observations may not fully capture the underlying uncertainty. Therefore, the reported intervals are indicative of sampling variability rather than strict probabilistic bounds.

2) Growth Rate Sensitivity. To assess the sensitivity of the projections to assumptions about sustained growth, the estimated growth rate parameter, b, is systematically varied by applying multiplicative adjustments ranging from $b \times 0.7$ to $b \times 1.3$. These adjustments correspond to doubling times from approximately 0.8 to 1.5 years. During this procedure, the scale parameter a is held fixed while b is modified. This approach isolates the effect of changes in long-term growth rates on the projected years in which thresholds may be crossed. This analysis is primarily illustrative. Empirical bootstrap estimates indicate a strong negative correlation between a and b (approximately -0.88).

3) Leave-One-Out Stability. Leave-one-out (LOO) cross-validation is used to determine whether any single observation has an outsized impact on the projections. In this approach, each data point in the frontier dataset is removed one at a time, the model is re-estimated, and the projected threshold-crossing year is recalculated. The variability across these LOO estimates offers insight into the sensitivity of the results to individual data points.

On a logarithmic scale, frontier logical qubits exhibit approximately linear growth, consistent with sustained exponential scaling. The estimated growth rate is $b = 0.672 \pm 0.066$ (standard error), corresponding to a doubling time of 1.03 years. The model achieves good fit ($R^2 = 0.85$, $p < 0.001$).

Projected crossing years for 2,330 logical qubits are:

- Optimistic (1:1000): 2035 [95% CI: 2033–2039]
- Baseline (1:1500): 2036 [95% CI: 2033–2040]
- Conservative (1:3000): 2037 [95% CI: 2034–2041]

The narrow two-year spread across scenarios reflects the fact that the growth rate b is identical for all conversion ratios; only the intercept differs. Consequently, uncertainty in the physical-to-logical ratio contributes less to forecast variance than uncertainty in sustained growth.

4) Historical Context: Moore's Law Comparison. To contextualize sustained exponential scaling, frontier qubit growth is compared with historical transistor scaling under Moore's Law (1971–2017) [47]. Both series are normalized to their initial values, and doubling times are estimated via log-linear regression. This comparison is heuristic and intended to contextualize growth rates rather than imply structural equivalence between semiconductor scaling and quantum hardware development. The quantum computing series spans only eight years, whereas the Moore's Law reference spans 46 years; extrapolation should be interpreted cautiously.

Importantly, quantum hardware scaling is governed by fundamentally different physical constraints than semiconductor lithography, including coherence limitations, error rates, and architectural overhead associated with fault tolerance. As such, the analogy to Moore's Law should be interpreted as an illustrative comparison of growth rates rather than a predictive model of long-term scaling behavior.

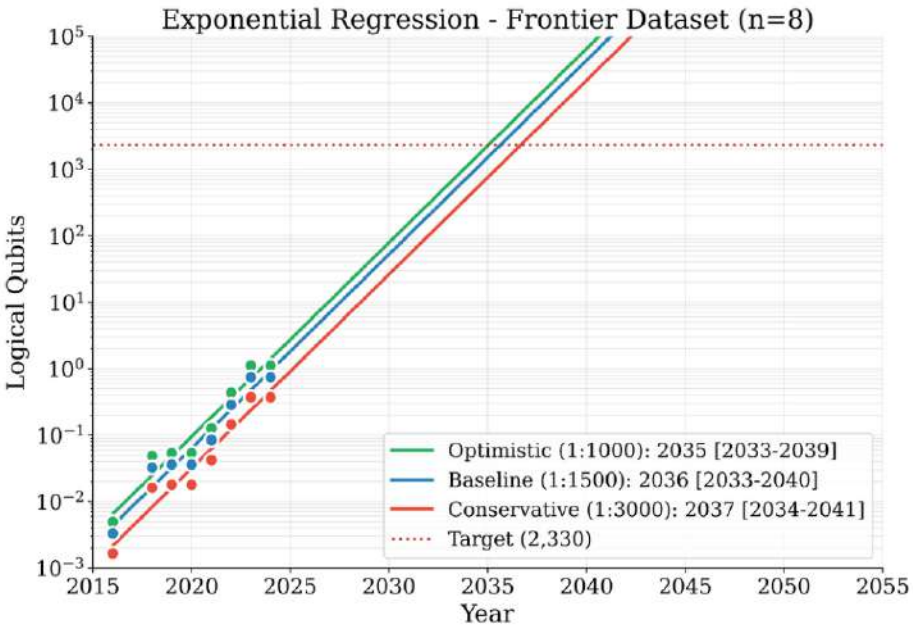

Figure 1: Exponential regression on the frontier dataset ($n = 8$). Green: Optimistic (1:1000), Blue: Baseline (1:1500), Red: Conservative (1:3000). The dashed line indicates the 2,330 logical qubit threshold. Legend shows crossing years with 95% bootstrap confidence intervals.

5 Results

This section presents the outcomes of the forecasting models applied to the curated dataset of quantum processors. The analysis focuses on extrapolating logical qubit growth under exponential assumptions and estimating when critical cryptographic thresholds may be reached. Results are reported for optimistic, baseline, and conservative fault-tolerance scenarios, using the frontier dataset as the primary input. Additional robustness checks

are also conducted to assess sensitivity to growth rates and individual data points.

A. Exponential Growth Model

Figure 1 presents the exponential regression fitted to the frontier dataset under three physical-to-logical conversion scenarios.

B. Bootstrap Uncertainty Quantification

Figures 2 and 3 display the uncertainty estimates derived from the bootstrap analysis. Across all scenarios, the resulting bootstrap distributions are centered in the mid-2030s, with 95% confidence intervals spanning roughly 6 to 7 years. The distributions exhibit mild right skew, indicating greater uncertainty toward later crossing years. This asymmetry arises because slower growth rates that delay the threshold crossing are generally more plausible than scenarios involving dramatic accelerations in progress.

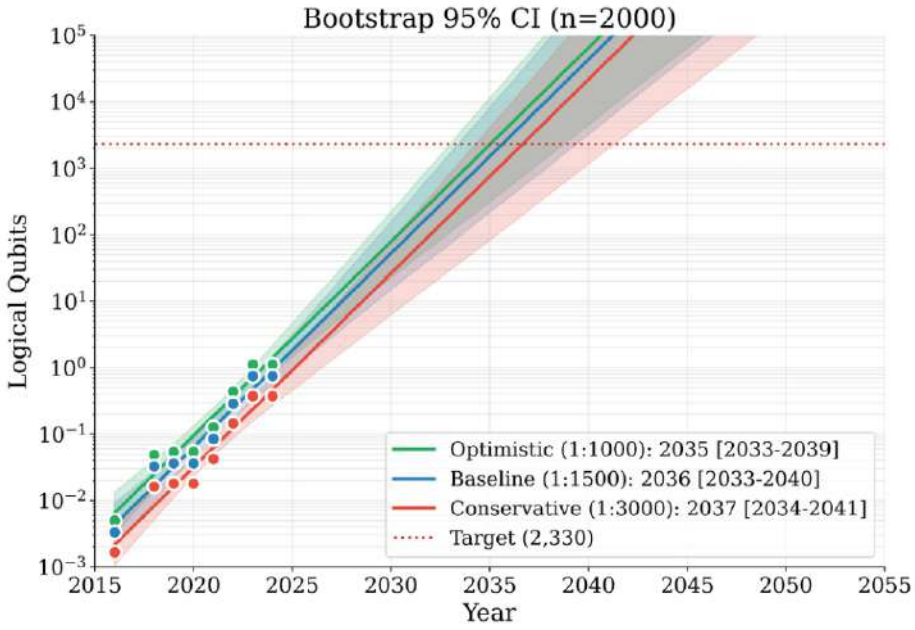

Figure 2: Bootstrap 95% confidence bands around the exponential trajectory. Shaded regions represent uncertainty from pure resampling across 2,000 iterations.

C. Growth Rate Sensitivity

Figure 4 evaluates how changes in the sustained growth rate affect crossing-year projections. Under the baseline growth rate ($b \times 1.0$), the crossing year is 2036 with a doubling time of 1.03 years. A 30% reduction in growth rate ($b \times 0.7$) delays the crossing to 2044, while a 30% increase ($b \times 1.3$) advances it to 2031. This 13-year range demonstrates the structural sensitivity of long-horizon forecasts to sustained growth assumptions. Because the scale

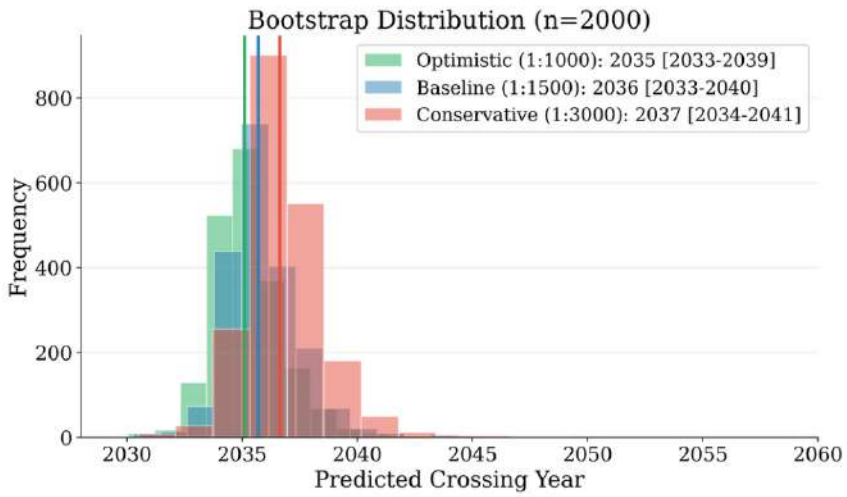

Figure 3: Empirical distribution of projected crossing years from bootstrap resampling. All three scenarios show tight clustering around 2035–2037, with right-skewed tails extending toward 2045.

parameter a is held fixed while b varies, and a and b are negatively correlated in practice, these bounds should be interpreted as illustrative rather than exact probabilistic limits.

D. Leave-One-Out Robustness

Figure 5 displays leave-one-out stability analysis. Removing any single frontier observation shifts the projected crossing year by at most ± 1 year:

- Optimistic: 2035 [LOO range: 2034–2036]
- Baseline: 2036 [LOO range: 2035–2037]
- Conservative: 2037 [LOO range: 2036–2038]

The standard deviation across LOO estimates is approximately 0.6 years. This indicates that the forecast is robust to individual data points and is not driven by any single observation, such as the 2023 IBM Condor announcement (1,121 qubits).

E. Moore's Law Comparison

Figure 6 compares quantum computing growth with historical transistor scaling. The frontier qubit trajectory exhibits a doubling time of 1.03 years, approximately twice as fast as Moore's Law transistor scaling (2.14 years). While this comparison does not guarantee multi-decade continuation—and the quantum computing series is substantially shorter than the Moore's Law reference—it demonstrates that sustained exponential hardware growth over decades is

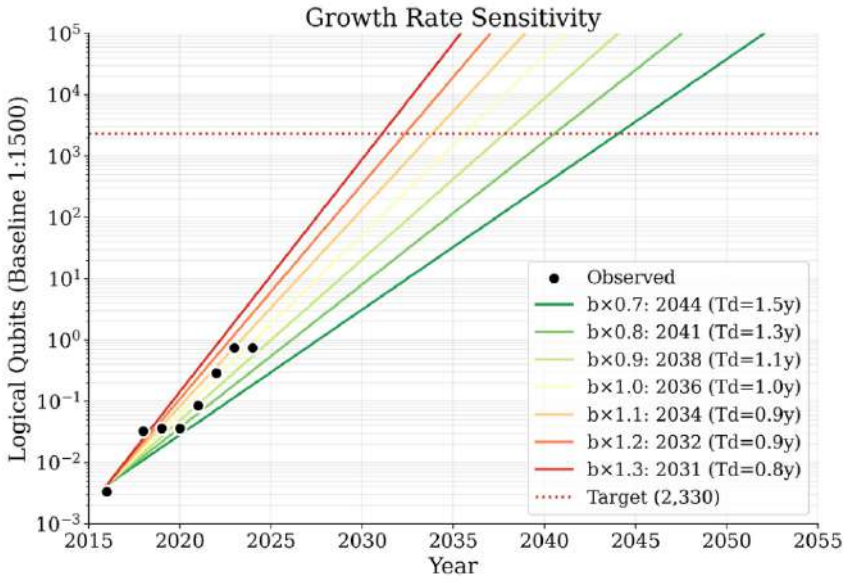

Figure 4: Growth rate sensitivity analysis. Multipliers from 0.7 to 1.3 applied to the estimated growth rate ($b = 0.672$) produce crossing years ranging from 2031 to 2044. T_d = doubling time.

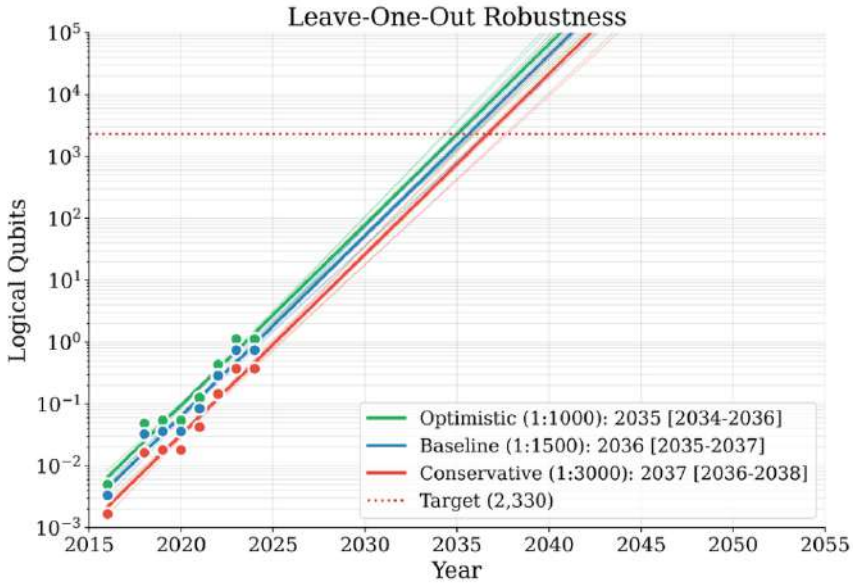

Figure 5: Leave-one-out robustness. Bold lines show the full-model fit; faded lines show fits with each observation removed. Legend displays the full-model crossing year and the LOO range.

historically precedented in computing technology. This comparison is intended as illustrative context rather than empirical evidence or a predictive model, and serves only to provide a historical reference for sustained exponential growth patterns in computing systems.

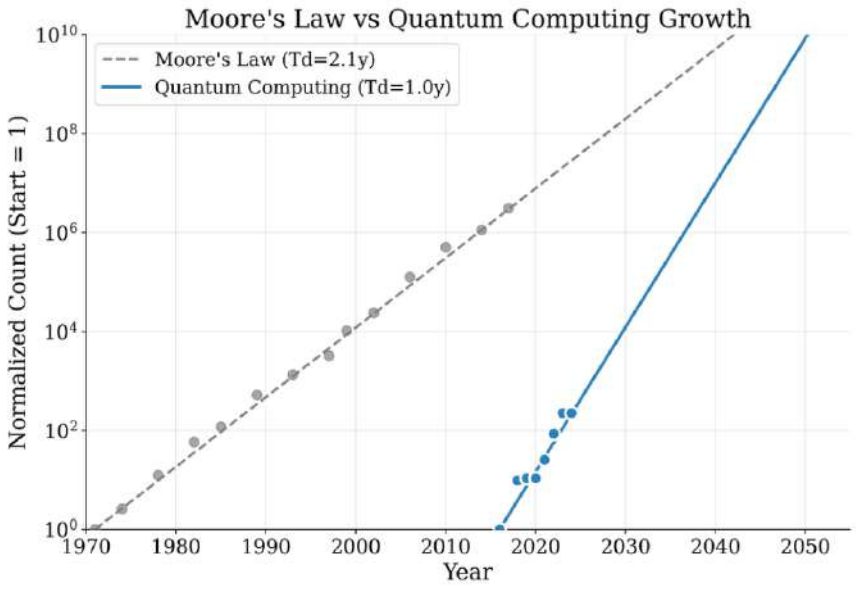

Figure 6: Normalized growth comparison: Moore's Law transistor counts (1971–2017, $n = 14$) versus frontier physical qubits (2016–2024, $n = 8$). Quantum computing exhibits a doubling time of approximately 1.0 year, compared to 2.1 years for transistors.

F. Frontier vs Full Dataset Comparison

Figure 7 compares forecasts derived from frontier and full datasets. The two datasets yield substantially different results:

Table 5: Frontier vs full dataset comparison (baseline scenario).

Metric	**Frontier** ($n = 8$)	**Full** ($n = 18$)
Growth rate (b)	0.672	0.260
Doubling time	1.03 years	2.67 years
R^2	0.85	< 0
Crossing year	2036	2064
95% CI	[2033–2040]	[2043–2078]

The full dataset contains quantum computing systems that fall significantly behind the technological frontier. For instance, it includes 32-qubit ion trap devices from years when superconducting systems with 1,121 qubits were already available. Including these lagging systems weakens the observed growth trend, reduces the model's fit to nearly zero (with negative R^2 values indicating performance worse than a simple horizontal trend), and pushes projected threshold-crossing years back by about 28 years.

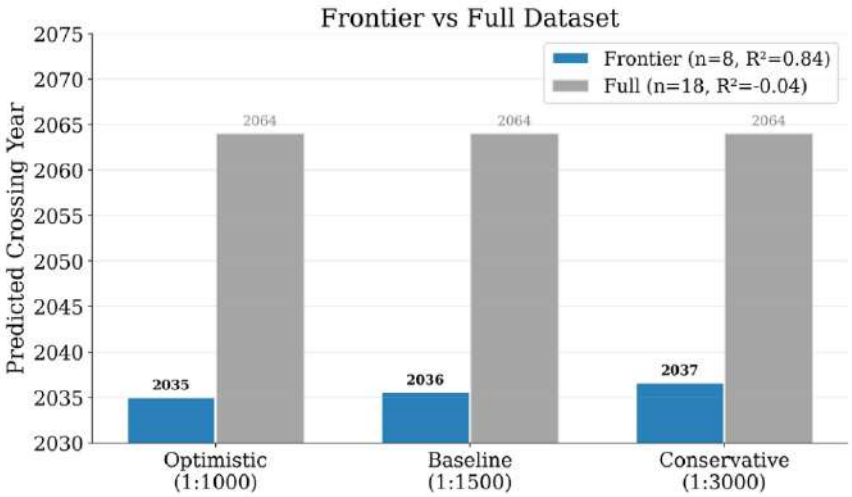

Figure 7: Comparison of crossing-year projections from frontier ($n = 8$) and full ($n = 18$) datasets. The frontier dataset achieves substantially higher model fit ($R^2 = 0.85$ vs $R^2 < 0$) and earlier crossing years.

Since cryptographic risk depends on the first system to reach the required threshold, the frontier dataset provides the most relevant basis for forecasting, whereas the full dataset reflects broader but less critical development trends.

G. Blockchain Cryptographic Implications

Under the baseline frontier scenario, the threshold of 2,330 logical qubits sufficient for ECDSA-256 cryptanalysis is projected to be reached around 2036 [95% CI: 2033–2040]. Blockchain systems relying on ECDSA may therefore become vulnerable once such capacity is achieved, enabling adversaries to derive private keys from exposed public keys and forge transaction signatures.

These projections suggest a limited migration window of approximately 10–12 years for transitioning toward post-quantum cryptographic standards. The "harvest now, decrypt later" threat model further emphasizes urgency, as data exposed today may become vulnerable once sufficient quantum capacity is available.

6 Discussion

This section interprets the forecasting results in the context of blockchain cryptographic security. We examine the robustness of the predictions, highlight methodological constraints, and discuss their practical implications for migration planning and protocol design.

A. Summary and Forecast Interpretation

The forecasting models developed in this study suggest that quantum computers capable of executing fault-tolerant cryptanalytic workloads may emerge as early as the mid-2030s. Under the baseline exponential growth model, the projected crossing point for 2,330 logical qubits—sufficient to break ECDSA-256—is the year 2036, with a 95% confidence interval spanning from 2033 to 2040. Conservative and optimistic scenarios shift this window by approximately ±3 years.

These projections remain robust under multiple forms of perturbation:

- Leave-one-out (LOO) analysis shows year-to-year forecasts are stable to individual data point removal (±1 year impact).
- ±30% variation in annual growth rate yields crossing dates between 2031 and 2044.
- Logistic model variants were evaluated but did not indicate observable saturation.
- Growth patterns observed in gate-based quantum hardware remain broadly consistent with early Moore-style trends in classical computing.

Logistic model variants were evaluated to assess potential saturation effects in hardware scaling. However, due to the limited temporal span of the dataset and the absence of observable plateauing in frontier qubit counts, logistic fits did not yield stable or meaningful parameter estimates. As a result, exponential models were retained as the primary forecasting approach, as they more accurately reflect the observed empirical trend within the available data.

The models do not extrapolate based on theoretical designs or speculative architectures, but rely exclusively on realized processors with verified specifications.

B. Model Scope and Design Justification

The following methodological constraints were adopted to maximize empirical relevance and analytical rigor:

- **Device Filtering:** Only gate-based quantum processors were included, as these are the only architectures suitable for running Shor's and Grover's algorithms.
- **Frontier-Based Selection:** One frontier processor per year was selected to represent the upper bound of available quantum hardware at the time.
- **Logical Qubit Estimation:** Physical qubits were converted to logical equivalents under three fault-tolerance scenarios: optimistic, baseline, and conservative, using ratios sourced from peer-reviewed literature.

- **Data Provenance:** All qubit counts and device specifications were extracted from verifiable publications, industry announcements, or official technical documentation.
- **Physical Realism:** No extrapolations were based on unbuilt or purely theoretical machines.
- **Exclusion of Non-Universal Devices:** Quantum annealers, boson samplers, and photonic demonstrators were excluded due to lack of general-purpose computational capability.
- **Cryptographic Threshold Anchoring:** The forecast target (2,330 logical qubits) was based on resource estimates for breaking ECDSA-256, drawn from recent high-impact cryptanalytic studies.
- **Forecast Horizon Limiting:** The model restricts projections to a 30-year horizon to mitigate long-term extrapolation risk.

C. Limitations

The empirical basis of the model consists of eight to ten frontier quantum devices spanning less than a decade. As such, the underlying trend is fitted to a limited dataset, and assumptions regarding growth continuity, error-correction overhead, and physical scalability may not hold indefinitely. Differences among hardware technologies (e.g., superconducting vs. ion trap) are also abstracted into unified scaling assumptions, introducing modeling simplifications.

Additionally, the evaluation is conducted under a capacity-based framework, which estimates the availability of logical qubits rather than full cryptanalytic execution feasibility. Consequently, factors such as algorithm runtime, T-gate throughput, and architectural parallelism may significantly affect the practical realization of quantum attacks and are not explicitly modeled in this study.

Furthermore, uncertainty estimates derived from bootstrap resampling should be interpreted as indicative ranges, as the limited number of observations constrains the statistical representativeness of the resampled distributions.

D. Security Implications

Regardless of precise timing, the analysis suggests that quantum vulnerability windows for classical cryptosystems are measured in years rather than decades. Blockchain platforms and cryptographic systems dependent on ECDSA, EdDSA, or other elliptic-curve schemes should anticipate the crossing of critical thresholds within the 2030s.

We recommend that organizations:

- Begin planning migration to NIST-approved post-quantum algorithms.
- Inventory cryptographic assets and public keys that may be vulnerable to "harvest now, decrypt later" attacks.
- Design systems with crypto-agility in mind to enable phased upgrades and hybrid deployments.

These recommendations are aligned with ongoing standardization efforts led by the National Institute of Standards and Technology (NIST) [48], which has recently selected a set of post-quantum cryptographic algorithms for standardization, including lattice-based schemes such as CRYSTALS-Kyber and CRYSTALS-Dilithium.

In practice, upgrading blockchain systems to post-quantum cryptography presents significant challenges due to protocol rigidity and the need for backward compatibility. For example, Bitcoin relies on soft-fork mechanisms for protocol upgrades [50], which require broad consensus and gradual adoption, while Ethereum explores more flexible approaches such as account abstraction [51] to enable cryptographic agility.

Several blockchain ecosystems have already initiated research toward quantum-resistant upgrades. For instance, Cardano (IOHK) has investigated quantum-resistant signature schemes and long-term migration strategies, highlighting both the technical and governance challenges associated with integrating post-quantum cryptography into decentralized systems. These developments provide a concrete and practical foundation for transitioning existing cryptographic systems toward quantum-resistant alternatives.

It is important to note that reaching the logical qubit threshold does not directly imply immediate cryptanalytic feasibility. The execution of Shor's algorithm requires not only sufficient logical qubits, but also substantial computational time, which depends on factors such as circuit depth, T-gate distillation throughput, and available parallelism. Existing resource estimates suggest that large-scale cryptographic attacks may require substantial execution time (e.g., hours to days under certain assumptions), depending on architectural design and error-correction efficiency. Therefore, practical attack feasibility depends not only on qubit capacity but also on system-level performance characteristics, which are not explicitly modeled in this study.

E. Recommendations for Blockchain Developers

Protocol developers should consider forward-compatible upgrades that allow post-quantum signature schemes to be integrated without full consensus rewrites. Migration roadmaps should be tied to qubit-based threat forecasting, with provisions for opt-in adoption and legacy account sunset policies. Cross-chain coordination and open standardization efforts will be critical to ensure coherent and timely global response across ecosystems.

In addition to algorithmic migration toward post-quantum cryptography, recent research has explored quantum-secure communication mechanisms as complementary approaches for enhancing blockchain resilience. For example, quantum key distribution (QKD) and hybrid quantum-classical communication models have been proposed to enhance secure data exchange and transaction validation in distributed systems (e.g., [52], [53]). These approaches provide an additional layer of defense, although their large-scale deployment remains subject to practical and infrastructural constraints.

7 Conclusion

Quantum computing presents a growing and measurable threat to the cryptographic assumptions underlying blockchain systems. Unlike previous studies that concentrate on algorithmic cost estimates, this work offers a forward-looking analysis based on actual hardware development.

We assembled a curated dataset of gate-based quantum processors built between 2016 and 2025. From this, we estimated logical qubit capacity under three error-correction scenarios and applied multiple modeling approaches to predict when quantum systems could exceed the 2,330 logical qubit threshold associated with ECDSA-256 cryptanalysis. The results suggest that such capabilities could realistically emerge between 2033 and 2040.

These projections are grounded in verifiable device specifications, resilient to variations in data selection, and consistent with recent trends in quantum hardware scaling. Although uncertainties remain—particularly in error correction and architectural progress—the analysis shows that the timeline for quantum disruption is narrowing. These projections should therefore be interpreted as indicators of when sufficient logical qubit capacity may become available, rather than precise estimates of when complete cryptanalytic attacks can be executed.

This study provides a practical framework for anticipating cryptographic risk based on technology growth, not speculation. For blockchain platforms and developers, the findings reinforce the need to plan and execute post-quantum migrations well in advance of hardware breakthroughs. Continued monitoring, refinement of predictive models, and cross-domain coordination will be critical to ensure that the infrastructure of decentralized systems remains secure in the presence of emerging quantum capabilities.

References (as published in the original article)

1. R. P. Feynman, “Simulating physics with computers,” *Int. J. Theor. Phys.*, vol. 21, nos. 6–7, pp. 467–488, Jun. 1982.

2. D. Deutsch, “Quantum theory, the church–turing principle and the universal quantum computer,” *Proc. Roy. Soc. London. A. Math. Phys. Sci.*, vol. 400, no. 1818, pp.

97–117, Jul. 1985.

3. P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Rev.*, vol. 41, no. 2, pp. 303–332, Jan. 1999.

4. L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," *Phys. Rev. Lett.*, vol. 79, no. 2, pp. 325–328, Jul. 1997.

5. C. Monroe and J. Kim, "Scaling the ion trap quantum processor," *Science*, vol. 339, no. 6124, pp. 1164–1169, Mar. 2013.

6. D. J. Bernstein and T. Lange, "Post-quantum cryptography," *Nature*, vol. 549, no. 7671, pp. 188–194, Sep. 2017.

7. A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation," *Phys. Rev. A*, vol. 86, no. 3, Sep. 2012, Art. no. 032324.

8. Z. Merali, "First sale for quantum computing," *Nature*, vol. 474, no. 7349, p. 18, Jun. 2011.

9. D. P. DiVincenzo, "Topics in quantum computers," 1996, arXiv:cond-mat/9612126.

10. "Bitcoin: The Cryptoanarchists' Answer To Cash," *IEEE Spectrum*. Available online: `https://spectrum.ieee.org/bitcoin-the-cryptoanarchists-answer-to-cash`.

11. Ethereum Whitepaper. Available online: `https://ethereum.org`.

12. M. Swayne, "Ethereum Foundation Elevates Post-Quantum Security to Top Strategic Priority," Jan. 2026.

13. "Research Program To Work on Hardening Cardano Against Quantum Computers," Feb. 2018.

14. C. Gidney and M. Ekerå, "How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits," *Quantum*, vol. 5, p. 433, Apr. 2021.

15. C. Gidney, "How to factor 2048 bit RSA integers with less than a million noisy qubits," 2025, arXiv:2505.15917.

16. M. Webber, V. Elfving, S. Weidt, and W. K. Hensinger, "The impact of hardware specifications on reaching quantum advantage in the fault tolerant regime," *AVS Quantum Sci.*, vol. 4, no. 1, Jan. 2022, Art. no. 013801.

17. V. Gheorghiu and M. Mosca, "Quantum resource estimation for large scale quantum algorithms," *Future Gener. Comput. Syst.*, vol. 162, Jan. 2024, Art. no. 107480.

18. J. Sevilla and C. Riedel, "Forecasting timelines of quantum computing," 2020, arXiv:2009.05045.

19. O. Ezratty, "Is there a Moore's law for quantum computing?" 2023, arXiv:2303.15547.

20. K. Schärer and M. Comuzzi, "The quantum threat to blockchain: Summary and timeline analysis," *Quantum Mach. Intell.*, vol. 5, 2023, Art. no. 19.

21. H. Khodaiemehr, K. Bagheri, and C. Feng, "Navigating the quantum computing threat landscape for blockchains: A comprehensive survey," *Comput. Sci. Rev.*, vol. 59, Feb. 2026, Art. no. 100846.

22. M. Švarcmajer, T. Krčmar, D. Šabanović, and I. Lukić, "Designing post-quantum secure blockchain infrastructure for business intelligence systems: Migration strategies and cost optimization," in *Proc. Int. Symp. ELMAR*, Sep. 2025, pp. 245–248.

23. Y. Baseri, A. Hafid, Y. Shahsavari, D. Makrakis, and H. Khodaiemehr, "Blockchain security risk assessment in quantum era, migration strategies, and proactive defense," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 2925–2964, 2026.

24. A. C. Santos, "O computador quântico da IBM e o IBM quantum experience," *Revista Brasileira de Ensino de Física*, vol. 39, no. 1, Sep. 2017, Art. no. e1301.

25. IBM Quantum, "Five Years Ago Today, We Put the First Quantum Computer on the Cloud. Here's How We Did It," May 2021.

26. J. Hsu, "CES 2018: Intel's 49-Qubit Chip Shoots for Quantum Supremacy," Jan. 2018.

27. F. Arute et al., "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

28. "Get To Know Honeywell's Latest Quantum Computer (System Model H1)," Oct. 2020.

29. Rigetti Computing, "Rigetti Computing Expands Global Presence With U.K. Quantum Computer Launch," Jun. 2022.

30. IBM Newsroom, "IBM Unveils Breakthrough 127-Qubit Quantum Processor," Nov. 2021.

31. Quantinuum, "Quantinuum Completes Hardware Upgrade; Achieves 20 Fully Connected Qubits," Jun. 2022.

32. Rigetti Computing, "Rigetti Computing Announces Commercial Availability of 80-Qubit Aspen-M System," Feb. 2022.

33. IBM Newsroom, "IBM Unveils 400 Qubit-Plus Quantum Processor and Next-Generation IBM Quantum System Two," Nov. 2022.

34. IonQ, "IonQ Aria: Versatile Quantum Computer," 2022.

35. IBM Quantum, "IBM Quantum System Two: The Era of Quantum Utility is Here," Dec. 2023.

36. Quantinuum, "For the First Time Ever, Quantinuum's New H2 Quantum Computer Has Created Non-Abelian Topological Quantum Matter," May 2023.

37. "Quantinuum Launches Industry-First, Trapped-ion 56-Qubit Quantum Computer," Jun. 2024.

38. "How We Achieved Our 2024 Performance Target of #AQ 35," Jan. 2024.

39. A. Browaeys and T. Lahaye, "Many-body physics with individually controlled Rydberg atoms," *Nature Phys.*, vol. 16, no. 2, pp. 132–142, Feb. 2020.

40. B. M. Terhal, "Quantum error correction for quantum memories," *Rev. Modern Phys.*, vol. 87, no. 2, pp. 307–346, Apr. 2015.

41. R. Acharya et al., "Quantum error correction below the surface code threshold," *Nature*, vol. 638, no. 8052, pp. 920–926, 2024.

42. M. Roetteler, M. Naehrig, K. M. Svore, and K. Lauter, "Quantum resource estimates for computing elliptic curve discrete logarithms," in *Proc. Adv. Cryptol.–ASIACRYPT*, vol. 2017, 2017, pp. 241–270.

43. B. Yan et al., "Factoring integers with sublinear resources on a superconducting quantum processor," 2022, arXiv:2212.12372.

44. S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall, and T. J. Yoder, "High-threshold and low-overhead fault-tolerant quantum memory," *Nature*, vol. 627, no. 8005, pp. 778–782, Mar. 2024.

45. IBM Quantum, "IBM Lays Out Clear Path to Fault-Tolerant Quantum Computing," Jun. 2025.

46. B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1993.

47. R. Schaller, "Moore's law: Past, present and future," *IEEE Spectr.*, vol. 34, no. 6, pp. 52–59, Jun. 1997.

48. NIST, "Module-Lattice-Based Key-Encapsulation Mechanism Standard," Aug. 2024.

49. NIST, "Module-Lattice-Based Digital Signature Standard," Aug. 2024.

50. I. Stewart, D. Ilie, A. Zamyatin, S. Werner, M. F. Torshizi, and W. J. Knottenbelt, "Committing to quantum resistance: A slow defence for Bitcoin against a fast quantum computing attack," *Roy. Soc. Open Sci.*, vol. 5, no. 6, Jun. 2018, Art. no. 180410.

51. Ethereum Improvement Proposals, "ERC-4337: Account Abstraction Using Alt Mempool."

52. E. O. Kiktenko et al., "Quantum-secured blockchain," 2017, arXiv:1705.09258.

53. Z.-Z. Sun et al., "Quantum blockchain relying on quantum secure direct communication network," *IEEE Internet Things J.*, vol. 12, no. 10, pp. 14375–14385, May 2025.

D

Paper D: Consensus-Driven Framework for Data-Driven Optimization of Distributed Systems Through Blockchain Consensus Mechanism Selection

Big Data and Cognitive Computing, vol. 10, no. 5, p. 154, May 2026 | DOI: 10.3390/bdcc10050154

Miljenko Švarcmajer, Mirko Köhler, Zdravko Krpić and Ivica Lukić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, 31000 Osijek, Croatia

Abstract

Modern data-driven distributed systems increasingly rely on blockchain technologies to ensure trust, transparency, and decentralized coordination. However, the rapid proliferation of consensus mechanisms has created a complex design space, making the selection of an appropriate protocol a non-trivial architectural and decision-making challenge. Different consensus mechanisms rely on distinct security resources, validator admission models, and agreement architectures, leading to diverse trade-offs between scalability, decentralization,

performance, and governance. Existing studies primarily focus on classification or performance comparison of consensus mechanisms, while the problem of systematic, requirement-driven selection remains insufficiently addressed. In particular, there is a lack of structured approaches that integrate multiple system requirements into a unified decision framework suitable for real-world environments. To address this gap, this paper proposes a consensus-driven, layered framework for blockchain consensus mechanism selection, formulated as a multi-criteria decision problem. The framework organizes the consensus design space across key architectural dimensions and analyzes 32 consensus mechanisms, enabling systematic comparison and supporting data-driven decision-making. The approach is further demonstrated through five representative use-case scenarios, showing its applicability in optimizing distributed system design.

Keywords: blockchain consensus mechanisms; consensus selection framework; multi-criteria decision making; blockchain architecture; distributed systems; decision support systems

1 Introduction

Blockchain technology was initially introduced as the under- lying infrastructure supporting cryptocurrencies, most notably Bitcoin. Early blockchain systems, often referred to as the first generation of blockchain technologies, focused primarily on secure and decentralized digital currency transactions. These systems relied on the Proof of Work (PoW) consensus mechanism to ensure agreement among distributed participants and to protect the network from malicious behavior.

The evolution of blockchain technology led to the emergence of second-generation platforms that introduced programmable smart contracts, enabling the execution of application logic directly within the blockchain environment. Platforms such as Ethereum extended the applicability of blockchain systems beyond digital currencies by enabling decentralized applications and automated contract execution [1].

In recent years, blockchain technology has evolved further and is increasingly used as a foundational infrastructure for a wide range of distributed systems, including supply chain management, digital identity platforms, smart city infrastructures, data integrity systems, and enterprise information systems. The adoption of blockchain in such diverse application domains is largely motivated by its core characteristics, including decentralization, tamper-resistant data storage, and transparent verification of transactions [2].

However, the growing diversity of blockchain applications has also revealed limitations of early consensus mechanisms. Proof of Work, while highly secure and robust in open environments, is often considered inefficient in terms of energy consumption and system performance when applied to many modern blockchain use cases. Consequently, a large

number of alternative consensus mechanisms have been proposed in the literature. These mechanisms attempt to address various limitations of earlier approaches by modifying the security resource, validator selection process, or agreement architecture [3].

As a result, the consensus design space has become increasingly complex. Different consensus mechanisms exhibit distinct advantages and trade-offs with respect to security, scalability, decentralization, energy efficiency, and system governance. Selecting an appropriate consensus mechanism therefore represents a critical architectural decision when designing a blockchain-based system. However, the lack of systematic guidelines for mapping system requirements to consensus properties makes this decision challenging for system designers.

Despite the large number of proposed consensus mechanisms, their selection is often performed without a structured methodology that links system requirements to consensus properties. In many cases, consensus protocols are adopted based on popularity or historical precedent rather than on a systematic evaluation of system characteristics and operational constraints [4].

This paper addresses this challenge by proposing a structured framework for selecting blockchain consensus mechanisms. In the context of this work, the term "data-driven" refers to the use of structured system requirements and measurable design parameters such as performance constraints, security assumptions, and governance models as inputs to the consensus selection process. The proposed framework does not rely on empirical datasets or machine learning techniques. Instead, it uses formalized system characteristics to support a multi-criteria, requirement-driven decision process. The framework organizes the consensus design space along several orthogonal dimensions that capture key architectural properties of blockchain systems. By linking consensus mechanisms to validator governance models, security re- sources, agreement architectures, and performance characteristics, the proposed framework provides a methodology for aligning consensus design with the requirements of specific application domains.

The main contributions of this paper are as follows:

- A structured analysis of blockchain consensus mechanisms that identifies key architectural design dimensions influencing security, scalability, decentralization, and system governance.
- A 7-layer consensus selection framework (CD-7) that transforms consensus design from an ad hoc choice into a systematic architectural decision process.
- A conceptual consensus design space that visualizes the relationships between identity-based and resource-based security models across different validator admission structures.
- A validation of the proposed framework through five representative application scenarios, including institutional governance, infrastructure monitoring, and post-quantum secure systems

The remainder of this paper is organized as follows. Section 2 reviews existing literature on blockchain consensus mechanisms and their classification approaches. Section 3 provides a structured overview of the selected consensus mechanisms included in this study. Section 4 introduces the proposed consensus selection framework and describes its conceptual foundations. Section 5 demonstrates the applicability of the framework through several representative application scenarios. Section 6 discusses the implications and limitations of the proposed approach. Finally, Section 7 concludes the paper and outlines potential directions for future research.

2 Related Work

Blockchain technology has become a key component of modern distributed systems, with consensus mechanisms playing a fundamental role in ensuring the security, scalability, and reliability of decentralized networks. Over the past decade, a significant body of research has analyzed different consensus algorithms and their properties. However, most existing studies focus either on specific protocol designs or on performance comparisons, rather than providing a systematic framework for selecting consensus mechanisms based on application requirements.

One of the most widely cited surveys of blockchain technology is presented by Zheng et al., who provide a comprehensive overview of blockchain architecture, consensus mechanisms, application domains, and open research challenges. Their work analyzes the fundamental components of blockchain systems, including block structures, cryptographic primitives, and distributed consensus protocols, while also highlighting key limitations such as scalability constraints and energy consumption in Proof-of-Work systems [5]. Although this survey provides a broad overview of blockchain technology, it does not address the problem of selecting consensus mechanisms according to system design requirements.

Several studies have focused specifically on the classification and comparison of consensus algorithms. The IEEE survey on blockchain consensus mechanisms provides an extensive taxonomy of protocols such as Proof-of-Work (PoW), Proof-of-Stake (PoS), and Byzantine Fault Tolerant (BFT) mechanisms. These works analyze the fundamental principles behind consensus protocols and discuss their advantages and limitations. However, such surveys typically emphasize protocol categorization rather than the systematic selection of consensus mechanisms for specific system architectures [6].

Another important research direction concerns the evaluation of consensus mechanisms according to performance and security criteria. Bamakan et al. analyze consensus protocols using multiple evaluation metrics, including throughput, energy efficiency, decentralization level, and resistance to security attacks [6]. Similarly, Gervais et al. investigate the security and performance trade-offs of Proof-of-Work systems and demonstrate how protocol parameters

influence the stability and robustness of blockchain networks [7]. While these works provide valuable insights into the performance characteristics of consensus algorithms, they primarily focus on analytical comparisons rather than decision-support methodologies for selecting appropriate mechanisms.

From an architectural perspective, Vukolić investigates the fundamental differences between classical Byzantine Fault Tolerant protocols and Nakamoto-style consensus mechanisms. This line of research highlights how different consensus paradigms influence the scalability, communication complexity, and fault tolerance of distributed ledger systems [8]. However, these architectural analyses generally focus on specific protocol families rather than providing a unified framework that connects consensus properties with system-level requirements.

In addition to protocol analysis, several studies investigate blockchain applications across different domains. For example, Kaur et al. provide a survey of blockchain applications in financial systems, Internet of Things environments, and decentralized autonomous organizations [9]. These studies illustrate the diversity of blockchain use cases, but they typically focus on application domains rather than on the methodological selection of consensus protocols.

Recent research has also explored alternative consensus mechanisms that utilize useful computational work instead of purely cryptographic puzzles. Toulemonde et al. propose a Useful Work protocol in which participants perform verifiable computations for externally submitted tasks, such as scientific or mathematical problems [10]. Similar approaches investigate the integration of machine learning workloads or other useful computations into consensus processes. These emerging mechanisms aim to address the energy inefficiency of traditional Proof-of-Work systems, but they remain focused on individual protocol designs rather than providing a general decision framework.

Despite the large body of research on blockchain consensus mechanisms, two main limitations remain in the existing literature. First, most studies analyze consensus protocols individually without providing a structured methodology for selecting an appropriate mechanism based on system requirements. Second, many surveys focus primarily on traditional consensus models and do not systematically incorporate emerging paradigms such as useful-work consensus, machine-learning-based consensus, or post-quantum security considerations. A comparative overview of representative studies and their research focus is summarized in Table 1.

To address these limitations, this paper proposes a structured framework for the systematic selection of blockchain consensus mechanisms. The proposed approach organizes the consensus design space according to several fundamental dimensions, including validator admission models, security resource models, agreement architectures, and performance trade-offs. By mapping consensus mechanisms into this multi-dimensional design space, the framework provides a practical methodology for aligning consensus protocols with the architectural and operational requirements of real-world blockchain systems.

The criteria used for comparison are derived from recurring design dimensions identified across the literature, particularly those related to validator participation, security assumptions,

Table 1. Comparison of related work on blockchain consensus mechanisms.

Paper	Consensus Analysis	Perf. Metrics	Applications	Selection Framework
Toulemonde et al. [11]	∼	✗	✗	✗
Zheng et al. [5]	∼	✗	✓	✗
Kaur et al. [10]	∼	✗	✓	✗
IEEE survey [6]	✓	∼	✗	✗
Vukolić [9]	✓	∼	✗	✗
Bamakan et al. [7]	✓	✓	✗	✗
Gervais et al. [8]	✓	✓	✗	✗
This paper	✓	✓	✓	✓

✓fully addressed; ∼ partially addressed; ✗not addressed.

agreement mechanisms, and system performance. These dimensions form the basis for the structured framework proposed in this work.

3 Selection Scope and Overview of Mechanisms

Blockchain consensus mechanisms have developed into a broad and heterogeneous field that includes substantially different validation logics, governance structures, and assumptions about the resources required to achieve agreement. Over time, a large number of new proposals have been introduced, ranging from foundational models to highly specialized or application-oriented designs. Simply listing these mechanisms without a clear selection rationale would offer limited analytical insight. For this reason, the present section clarifies the criteria used to determine which mechanisms are included in this study and then provides a structured overview of the selected approaches.

3.1. Selection Scope and Design Rationale

The purpose of this study is not to provide an exhaustive catalogue of all consensus mechanisms proposed in the literature, but rather to construct a representative and structurally diverse design landscape that enables the development of a generalized selection framework. Given the large number of published mechanisms, many of which represent incremental optimizations or application-specific adjustments, a selective and principled approach was necessary.

The mechanisms included in this study were chosen according to several guiding criteria.

First, foundational paradigms were selected in order to capture the core architectural shifts that have shaped blockchain consensus research. These include Proof of Work (PoW) as the first widely deployed blockchain consensus mechanism, Proof of Stake (PoS) as a fundamental alternative to computation based validation, Practical Byzantine Fault Tolerance (PBFT) as the classical Byzantine agreement model adapted to distributed ledgers, and

Directed Acyclic Graph (DAG) as a structural alternative to linear blockchain architectures.

Second, the selection was guided by resource-based diversity. Consensus mechanisms fundamentally differ in the primary resource used to establish agreement [11]. The selected mechanisms collectively represent validation models based on computational power, virtual stake, identity and authority, storage capacity and time commitment, reputation and trust relationships, activity and contribution metrics, delegated and federated governance, hybrid constructions, machine learning effort, and quantum cryptographic principles. This diversity ensures that the subsequent classification is not limited to a single consensus philosophy.

Third, representative mechanisms introducing structurally meaningful variations in validation resources, governance structures, or agreement architectures were included. Mechanisms were selected if they alter the governance structure, the underlying validation resource, or the fault tolerance model in a conceptually significant way. Minor parameter adjustments and rebranded implementations without structural novelty were excluded.

Finally, both practically deployed mechanisms and research-level proposals were incorporated. Including mechanisms that exist primarily in scientific literature allows the framework to account for emerging directions in consensus design, particularly in areas such as machine learning integration and quantum-resilient architectures.

The selection of the 32 consensus mechanisms was guided by the objective of capturing a representative and structurally diverse set of designs, covering distinct validation resources, governance models, and agreement architectures, while excluding purely incremental variations.

Through this structured selection strategy, the analyzed mechanisms span the principal consensus design dimensions required to evaluate adaptability across heterogeneous operational environments. This representational completeness forms the analytical foundation for the seven-layer consensus design framework developed in Section 4. Consensus mechanisms form the operational backbone of blockchain systems, enabling distributed nodes to agree on transaction validity without centralized coordination. This section presents the analyzed mechanisms in a structured descriptive manner, providing historical context, operational principles, implementation status, and distinguishing characteristics.

3.2. Consensus Mechanism Overview

While this section provides a descriptive overview of individual consensus mechanisms, their systematic categorization and comparative analysis are performed in Section 4 through the proposed seven-layer consensus design framework. This subsection provides a structured overview of the consensus mechanisms selected according to the criteria defined in Section 3.1. The goal is not to present exhaustive technical specifications, but to establish a consistent descriptive baseline that supports the comparative categorization developed in Section 4.

Each mechanism is described using a uniform structure that includes the year of introduction, core operational principle, deployment status, and its distinguishing characteristics. Where

applicable, the implementation status of each mechanism is explicitly indicated, distinguishing between production-deployed systems and research-stage proposals. Mechanisms that currently exist primarily at the research level are explicitly identified as such.

The mechanisms are presented in a logically organized sequence, beginning with early consensus paradigms that shaped blockchain development, followed by mechanisms introducing alternative security resources, governance structures, or agreement models, and concluding with emerging and research-oriented approaches. This progression reflects both historical development and conceptual diversification within the consensus design space.

The following subsections introduce each mechanism individually.

3.2.1. Proof of Work

Proof of Work (PoW), introduced by Nakamoto in 2008 [12], represents the first widely deployed blockchain consensus mechanism. It requires network participants (miners) to solve computationally intensive cryptographic puzzles in order to propose new blocks. Security is achieved through economic cost: an adversary must control a majority of the network's computational power to manipulate the ledger.

PoW is implemented in Bitcoin and several early-generation blockchains. Its defining characteristic is the transformation of computational power into a security resource. While it established strong probabilistic security and decentralization, its substantial energy consumption and limited scalability motivated the development of alternative models.

3.2.2. Proof of Stake

Proof of Stake (PoS), proposed in 2012 by King and Nadal [13], replaces computational competition with economic commitment. Validators are selected proportionally to the amount of cryptocurrency they lock as stake, shifting security guarantees from energy expenditure to financial incentives.

PoS and its variants (e.g., Ouroboros, Casper, Tendermint) are now widely implemented across contemporary blockchain platforms [14]. Its distinctiveness lies in significantly reducing energy requirements while maintaining economic deterrence. However, token concentration introduces governance and centralization concerns that remain an active research topic.

3.2.3. Delegated Proof of Stake

Delegated Proof of Stake (DPoS), introduced by Larimer in 2014 [16], modifies the PoS paradigm by incorporating a representative voting layer. Token holders elect a limited number of delegates responsible for block production, thereby increasing throughput and reducing confirmation time.

DPoS has been implemented in platforms such as EOS and BitShares [16]. Its primary innovation is the separation between economic stake and operational block validation. While

this model enhances performance, it introduces structured governance and potential concentration of influence among elected delegates.

3.2.4. Proof of Authority

Proof of Authority (PoA), formalized in 2017 [17], is designed for permissioned blockchain environments. Instead of relying on stake or computation, PoA assigns block validation rights to pre-approved validators whose real-world identities are known.

This model is commonly used in enterprise and consortium blockchains. Its distinguishing feature is replacing economic or computational security with identity-based trust. While highly efficient and energy conservative, it trades off decentralization for operational control.

3.2.5. Proof of History

Proof of History (PoH), introduced by Yakovenko in 2018 [18], establishes a cryptographic time-ordering mechanism using sequential hashing. Rather than functioning as a standalone consensus protocol, PoH provides a verifiable time source that enables parallel transaction processing.

PoH is implemented within the Solana blockchain architecture. Its innovation lies in decoupling transaction ordering from consensus voting, enabling high throughput. However, its effectiveness depends on precise hardware synchronization and integration with additional consensus layers.

3.2.6. Practical Byzantine Fault Tolerance

Practical Byzantine Fault Tolerance (PBFT), developed by Castro and Liskov in 1999 [19], is a deterministic consensus algorithm designed for asynchronous distributed systems. PBFT tolerates up to one-third of malicious nodes by employing multi-phase message exchange to ensure agreement on transaction order.

PBFT forms the foundation of numerous permissioned blockchain systems and enterprise platforms. Its defining characteristic is deterministic finality without probabilistic confirmation. Nevertheless, communication complexity grows quadratically with network size, limiting scalability in large public environments.

3.2.7. Directed Acyclic Graph

Directed Acyclic Graph (DAG) architectures emerged in blockchain applications around 2014–2015 through projects such as IOTA and Nano [20]. Instead of grouping transactions into blocks, DAG-based systems allow each transaction to validate previous ones, enabling parallel processing.

DAG models emphasize scalability and reduced latency by eliminating global block intervals. Their distinguishing feature is the absence of traditional block mining. However, coordination, security analysis, and attack resistance mechanisms remain areas of ongoing development.

3.2.8. Honey Badger BFT

HoneyBadgerBFT, proposed in 2016 by Miller et al. [21], represents the first practical asynchronous Byzantine Fault Tolerant (BFT) protocol that operates without timing assumptions. Unlike partially synchronous BFT systems, HoneyBadgerBFT guarantees safety and liveness even under unpredictable network delays by combining reliable broadcast, erasure coding, and binary Byzantine agreement.

The protocol is designed for fully asynchronous environments and has been experimentally validated in distributed settings with large numbers of nodes. Its defining characteristic is complete independence from network synchrony assumptions, which enhances robustness in adversarial or unreliable network conditions. However, this robustness comes at the cost of increased cryptographic overhead and potentially higher latency compared to partially synchronous BFT systems.

3.2.9. FIBFT

FIBFT, introduced in 2023 by Gao et al. [22], extends Byzantine consensus models to address scalability challenges in edge computing environments. The protocol integrates clustering techniques, specifically K-medoids, to partition nodes into subnets based on performance characteristics. Each subnet then executes an enhanced Byzantine agreement protocol.

This architecture reduces communication complexity and improves scalability by limiting consensus rounds within smaller groups. Its distinctive feature lies in combining machine learning-based clustering with consensus logic. While promising for large-scale and heterogeneous environments, the approach introduces additional computational overhead related to clustering and dynamic subnet management.

3.2.10. Ripple Protocol Consensus Algorithm

The Ripple Protocol Consensus Algorithm (RPCA), deployed in 2012 by Ripple Labs [23], operates using a Unique Node List (UNL) model. Each validator maintains a list of trusted nodes and reaches consensus when a supermajority (typically 80%) agrees on transaction validity.

RPCA is optimized for financial settlement systems and is actively used within the Ripple network for cross-border transactions. Its defining feature is the reliance on partially overlapping trusted validator sets rather than open participation. This design enables

rapid confirmation and low energy usage, but introduces structured trust assumptions that differentiate it from fully permissionless networks [24].

3.2.11. Stellar Consensus Protocol

The Stellar Consensus Protocol (SCP), introduced in 2015 by Mazières [25], is based on the Federated Byzantine Agreement (FBA) model. In SCP, each node selects its own quorum slice—trusted peers whose agreement it requires for consensus. Consensus emerges from the intersection of these quorum slices across the network.

SCP is implemented in the Stellar network and supports decentralized financial infrastructure. Its distinguishing property is decentralized quorum configuration, allowing open membership without predefined validator sets. While highly efficient and energy conservative, the security of SCP depends on careful quorum slice selection to prevent network fragmentation.

3.2.12. Proof of Elapsed Time

Proof of Elapsed Time (PoET), introduced by Intel in 2016 [26], leverages Trusted Execution Environments (TEE), such as Intel SGX, to generate verifiable random wait times for leader selection. Validators compete by generating secure random timers, and the node with the shortest valid wait time proposes the next block.

PoET has been implemented within the Hyperledger Sawtooth platform and is primarily intended for permissioned enterprise blockchains. Its defining innovation is shifting consensus fairness enforcement from economic or computational resources to harware-based secure randomness. This dependence on trusted hardware introduces potential risks related to enclave vulnerabilities and manufacturer trust assumptions.

3.2.13. Proof of Burn

Proof of Burn (PoB), proposed in 2012 by Stewart [27], replaces computational mining with the irreversible destruction of cryptocurrency tokens. Participants send coins to verifiably unspendable addresses, thereby demonstrating long-term economic commitment to the network.

PoB has been implemented in projects such as Slimcoin and serves as an alternative resource-based consensus approach. Its defining feature is the transformation of financial sacrifice into mining probability. Although it reduces energy expenditure relative to PoW, it introduces irreversible capital loss and requires careful economic modeling to prevent inefficiencies.

3.2.14. Proof of Capacity

Proof of Capacity (PoC), popularized by Burstcoin in 2014 [28], utilizes storage space as the primary security resource. Participants precompute and store cryptographic data (plotting),

which is later used to determine eligibility for block generation.

PoC enables energy-efficient consensus by replacing active computation with disk storage utilization. Its distinguishing characteristic is reliance on storage capacity rather than computational power or financial stake. However, ensuring honest space commitment and preventing manipulation through storage optimization techniques remains an ongoing challenge.

3.2.15. Proof of Space

Proof of Space (PoSpace), formally described in 2015 by Dziembowski et al. [29], uses disk storage capacity as the primary resource for achieving consensus. Participants allocate a specified amount of storage to commit cryptographic data that can later be efficiently verified by the network.

Unlike Proof of Work, PoSpace minimizes continuous computational effort by relying on precomputed storage commitments. The model has influenced systems such as Chia and related storage-based blockchains. Its distinguishing characteristic lies in transforming persistent storage allocation into a security primitive, thereby reducing operational energy consumption. However, maintaining reliable proof verification and preventing storage manipulation remain technical challenges.

3.2.16. Proof of Space-Time

Proof of Space-Time (PoST), introduced in 2016 by Moran and Orlov[30], extends storage-based consensus by requiring participants to prove that they have retained specific data over a defined time interval. Rather than merely demonstrating disk allocation, PoST ensures continuous storage commitment through periodic cryptographic verification.

The mechanism emphasizes long-term data persistence as a trust resource and has influenced decentralized storage networks. Its defining feature is coupling spatial commitment with temporal guarantees, thereby strengthening reliability assurances. Nevertheless, ensuring secure long-term storage verification without introducing new attack vectors remains an area of active research.

3.2.17. Proof of Activity

Proof of Activity (PoAct), proposed in 2014 by Bentov et al. [31], combines elements of Proof of Work and Proof of Stake. The protocol begins with a PoW-based empty block generation phase, followed by a PoS-based validator selection process that finalizes the block through digital signatures from selected stakeholders.

This hybrid design aims to integrate computational security with economic stake-based validation. Its distinctive contribution is the layered security structure that distributes responsibility across two resource models. While enhancing robustness, the dual-phase architecture increases implementation complexity and partially preserves PoW energy costs.

3.2.18. Proof of Engagement

Proof of Engagement (PoE), introduced in 2021 by Xu et al. [32], evaluates nodes based on their active participation and contribution within the network. Engagement metrics, including transaction activity and computational contribution, determine the probability of being selected as a block proposer.

The mechanism remains primarily conceptual and described in academic literature. Its defining feature is shifting consensus eligibility toward behavioral contribution rather than static resource ownership. By attempting to mitigate wealth concentration effects, PoE introduces dynamic participation metrics, although safeguarding against artificial engagement inflation presents a significant design challenge.

3.2.19. Proof of Importance

Proof of Importance (PoI), deployed within the NEM blockchain in 2015 [33], extends stake-based validation by incorporating transaction graph analysis into validatorselection. Accounts are assigned an importance score derived from token holdings and network activity metrics.

PoI is implemented in production and emphasizes economic participation alongside ownership. Its distinguishing feature is the integration of graph-theoretic influence modeling into consensus probability. While reducing passive wealth dominance, the computational complexity of importance scoring and potential manipulation of activity metrics require careful governance mechanisms.

3.2.20. Proof of Stake Velocity

Proof of Stake Velocity (PoSV), introduced by Ren in 2014 for the Reddcoin network [34], combines stake ownership with transactional activity as determinants of validator selection. In addition to holding tokens, participants must demonstrate active participation to maintain influence.

PoSV has been deployed in cryptocurrency systems emphasizing transactional throughput. Its defining characteristic is integrating monetary velocity into the consensus model, thereby incentivizing circulation rather than passive holding. Although improving network dynamism, maintaining balanced incentive structures without enabling artificial transaction generation is a persistent challenge.

3.2.21. Proof of Vote

Proof of Vote (PoV), proposed in 2017 by Li et al. [34], is designed for consortium blockchain environments. Block validation is achieved through structured voting among authorized participants, separating decision authority from block creation roles.

The model exists primarily in academic proposals. Its distinctive aspect is formalizing consensus as an explicit voting protocol rather than probabilistic resource competition. While energy-efficient and suitable for closed governance structures, PoV inherently relies on predefined trust relationships among consortium members.

3.2.22. Proof of Trust

Proof of Trust (PoT), introduced in 2018 by Bahri and Girdzijauskas [35], incorporates trust metrics into validator selection. Nodes accumulate trust scores based on historical behavior, and higher-trust nodes face reduced validation difficulty or increased selection probability.

The mechanism is primarily described in research literature and has not achieved widespread deployment. Its distinguishing contribution lies in formalizing trust graphs as a consensus resource. While potentially reducing computational effort, designing tamper-resistant trust evaluation frameworks presents substantial complexity.

3.2.23. Proof of Reputation

Proof of Reputation (PoR), proposed in 2020 by Dehez-Clementi et al. [36], determines validator selection based on accumulated reputation scores derived from historical behavior. Nodes gain influence through consistent participation and correct validation activity rather than computational power or economic stake.

PoR remains primarily an academic proposal. Its defining feature is formalizing social or behavioral capital as a consensus resource. While potentially reducing energy consumption and improving accountability, designing objective, manipulation-resistant reputation metrics presents a central implementation challenge.

3.2.24. Proof of Contribution

Proof of Contribution (PoCon), hereafter abbreviated as PoCon to distinguish it from Proof of Capacity (PoC), was introduced in 2018 by (Xue et al., 2018) [37]. It modifies the Proof of Work paradigm by integrating contribution-based weighting into mining difficulty. Nodes that have previously contributed successfully to the network may receive adjusted validation conditions.

The mechanism exists primarily in the research literature and seeks to balance fairness and energy efficiency. Its distinctive characteristic lies in linking historical productive behavior to consensus probability. However, accurately measuring contribution without introducing exploitable incentive distortions remains a complex design challenge.

3.2.25. Proof of Luck

Proof of Luck (PoL), proposed in 2016 by Milutinović [38], utilizes Trusted Execution Environments (TEE) to generate secure random values used for leader selection. Validators produce cryptographically verifiable "luck" values, and the highest value determines block proposer eligibility.

PoL is described primarily in academic work and has not achieved broad deployment. Its defining innovation is replacing computational competition with hardware-enforced randomness. The model significantly reduces energy requirements but introduces dependency on secure enclave integrity and hardware trust assumptions.

3.2.26. Yuma Consensus

The Yuma Consensus Protocol, introduced in 2021 within the Bittensor network [39], evaluates contributions such as machine learning model outputs through validator scoring mechanisms. Participants submit models, which are assessed using transparent performance metrics to determine reward allocation.

Yuma is actively deployed within the Bittensor AI ecosystem. Its distinguishing characteristic is embedding incentive aligned evaluation of AI contributions directly into consensus logic. While promoting fairness and measurable productivity, the system depends on carefully calibrated scoring functions and resistant commit-reveal mechanisms to prevent model copying or manipulation.

3.2.27. Proof of Learning

Proof of Learning (PoLearn), introduced in 2019 by Bravo-Marquez et al. [40], integrates machine learning training tasks into consensus. Validators compete by training models on published datasets, and performance on unseen evaluation data determines block validation rights.

The mechanism remains largely theoretical. Its defining feature is transforming computational effort into socially useful work rather than arbitrary cryptographic puzzles. Although conceptually attractive for sustainability, coordinating dataset integrity, evaluation fairness, and ongoing task supply presents nontrivial operational complexity.

3.2.28. Proof of Karma

Proof of Karma (PoK), proposed in 2022 by Biswas et al. [41], assigns leadership probability based on accumulated behavioral scores referred to as "karma." Nodes that consistently behave according to protocol rules increase their likelihood of selection.

The model is described in academic literature and is intended for consortium or controlled environments. Its distinctive element lies in integrating behavioral reinforcement directly

into leader election. Ensuring resistance against artificial karma inflation and maintaining decentralized fairness remain significant design considerations.

3.2.29. Proof of Endorse Contracts

Proof of Endorse Contracts (PoEC), introduced in 2022 [42], facilitates cross-chain consensus transfer by allowing a supervisory blockchain to endorse state decisions executed within a subordinate chain. Validators reference externally confirmed contract states to validate blocks.

PoEC is primarily conceptual and targets interoperability between blockchain systems. Its distinguishing contribution is decoupling consensus validation from native resource models by leveraging external verification layers. While potentially scalable, it introduces dependency on cross-chain communication integrity and supervisory network reliability.

3.2.30. Quantum Teleportation-Based Consensus

The quantum teleportation-based consensus mechanism, proposed in 2022 by Wen et al. [43], explores the use of quantum entanglement and state projection to validate blockchain transactions. By transferring quantum states rather than classical messages, the protocol aims to achieve high-security verification independent of classical computational constraints.

This approach remains theoretical due to the current limitations of quantum hardware. Its defining characteristic is leveraging quantum mechanical properties such as entanglement to achieve theoretically unconditional security. Practical deployment is constrained by qubit coherence management and large-scale quantum infrastructure feasibility.

3.2.31. Quantum Zero-Knowledge-Based Consensus

Quantum zero-knowledge consensus, introduced in 2022 [44], integrates quantum zero-knowledge proofs to validate transactions without revealing sensitive information. The protocol applies quantum cryptographic principles to replace classical probabilistic verification models.

The mechanism is currently conceptual. Its distinguishing property lies in combining privacy-preserving validation with post-quantum security assumptions. Although energy-efficient in classical terms, it depends entirely on the maturity and scalability of quantum computing systems.

3.2.32. Post-Quantum Delegated Proof of Luck

Post-Quantum Delegated Proof of Luck (PQ-DPoL), proposed in 2023 by Kim et al. [46], integrates post-quantum cryptographic primitives such as CRYSTALS-Dilithium into a delegated consensus structure. Validators are elected through a representative model similar to DPoS, while cryptographic security is reinforced against quantum adversaries.

PQ-DPoL is currently described in research literature. Its defining contribution is combining delegated governance with post-quantum resilience. While improving long-term cryptographic robustness, it inherits governance concentration risks typical of delegated systems.

3.3. Diversification of the Consensus Design Space

Since the introduction of the first blockchain consensus mechanisms, the design space of distributed agreement protocols has expanded significantly. Early paradigms such as Proof of Work and Byzantine fault-tolerant protocols established the fundamental principles of decentralized agreement. As blockchain systems began to be applied in different operational environments, numerous alternative consensus approaches were proposed in order to address limitations related to scalability, energy efficiency, governance structures, and validation resources.

Over time, these developments have resulted in a diverse ecosystem of consensus mechanisms. Many proposals introduce new validation resources, such as economic stake, storage capacity, behavioral metrics, or hardware-assisted randomness, while others modify the governance structure or communication architecture of existing protocols. As a consequence, the consensus landscape has evolved into a branching design space in which new mechanisms often build upon, extend, or conceptually derive from earlier paradigms.

Figure 1 provides a conceptual overview of this diversification. The diagram illustrates how foundational consensus paradigms gave rise to multiple design branches, including stake-based validation, storage-oriented mechanisms, behavioral and reputation-based models, machine-learning–integrated approaches, and quantum-oriented proposals. While not representing a strict genealogical relationship, the figure highlights the conceptual connections among the mechanisms analyzed in this study and illustrates how the consensus design space has progressively expanded over time.

This diversification of consensus mechanisms motivates the need for a structured analytical framework capable of systematically evaluating different designs across multiple architectural dimensions. The following section introduces the proposed seven-layer consensus design framework used to organize and analyze the mechanisms presented in this study.

4 The CD-7 Framework for Systematic Consensus Mechanism Selection

While Section 3 provided a structured overview of existing consensus mechanisms, descriptive analysis alone does not sufficiently support the systematic selection of a consensus model for a new blockchain system. In practice, blockchain architects must select a consensus mechanism based on a set of design constraints, including governance structure, validator

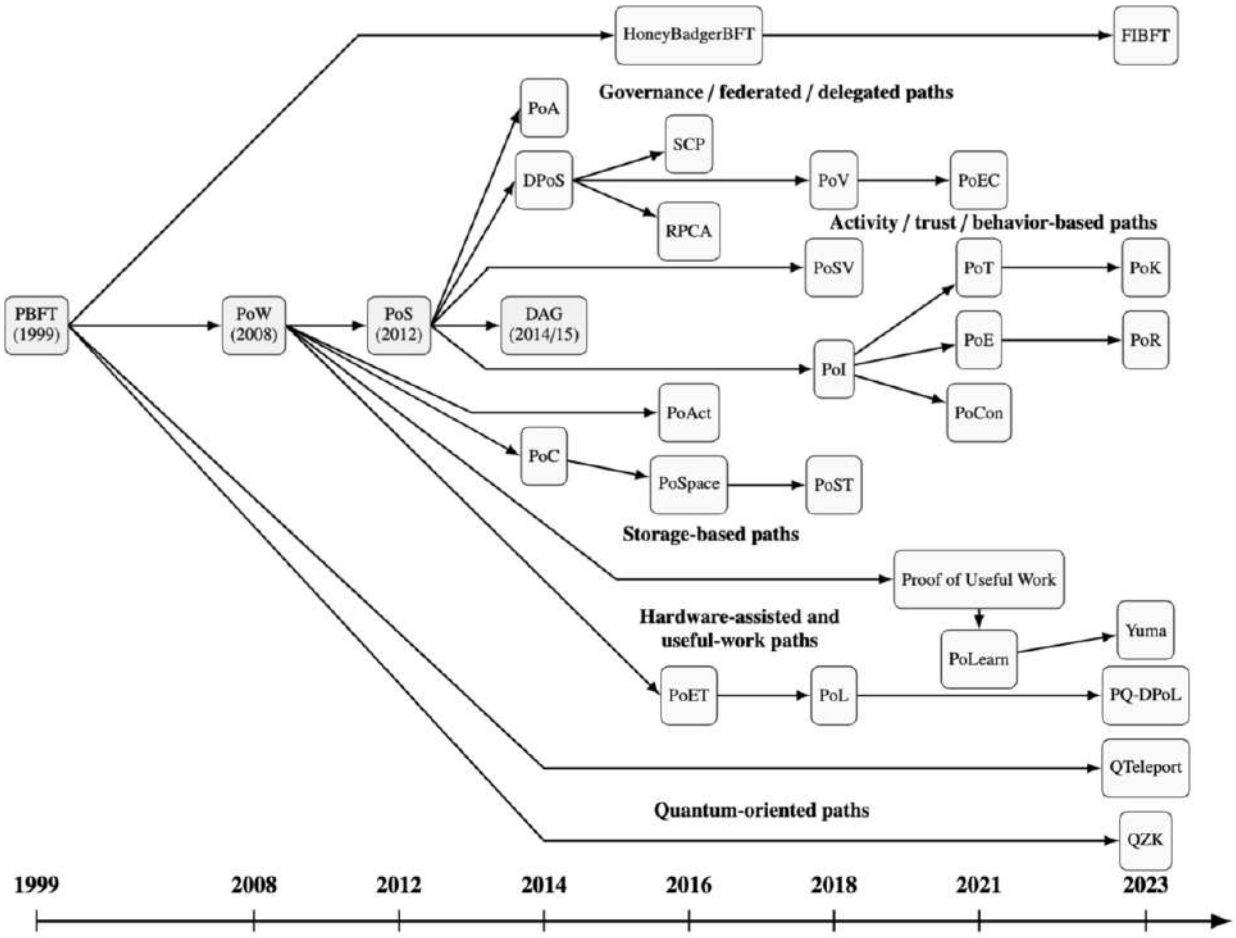

Figure 1. Diversification of the blockchain consensus design space. The diagram illustrates conceptual relationships among representative consensus mechanisms and highlights the emergence of multiple design branches based on different validation resources and governance models.

participation rules, available security resources, and desired performance characteristics.

Given the rapid proliferation of consensus proposals, selecting an appropriate mechanism has become a non-trivial design problem. Different consensus mechanisms rely on fundamentally different assumptions regarding security, trust, and validator coordination. Consequently, choosing an inappropriate consensus mechanism may lead to inefficiencies, security vulnerabilities, or governance conflicts within the network.

To address this challenge, this section introduces a structured framework for consensus mechanism selection. Rather than focusing on specific blockchain implementations, the framework analyzes intrinsic properties of consensus algorithms and organizes them along several orthogonal dimensions. These dimensions capture the essential design choices that define how consensus is achieved in distributed ledger systems.

The proposed framework consists of seven analytical layers that capture the key architectural dimensions of consensus design:

- L1: Accessibility layer (validator admission model)
- L2: Primary security resource
- L3: Communication topology
- L4: Agreement logic
- L5: Finality and settlement model
- L6: Governance and incentive structure
- L7: Application-specific performance metrics

The relationships between these layers and their role in the consensus selection process are illustrated in Figure 2. Together, these seven layers form a structured decision-oriented architecture that guides the systematic selection of consensus mechanisms by progressively narrowing the design space according to network requirements. Rather than evaluating individual consensus algorithms in isolation, system designers can identify suitable candidates by matching network requirements with the fundamental properties of the underlying protocol.

4.1. Accessibility Layer (Validator Admission Model)

The first and most fundamental design decision concerns validator participation. Consensus mechanisms differ significantly in how validators are admitted to the network and how the validator set is formed.

Three primary validator admission models can be distinguished:

Permissionless validation. In permissionless systems, any participant may attempt to become a validator without prior approval. Validator eligibility is determined solely by

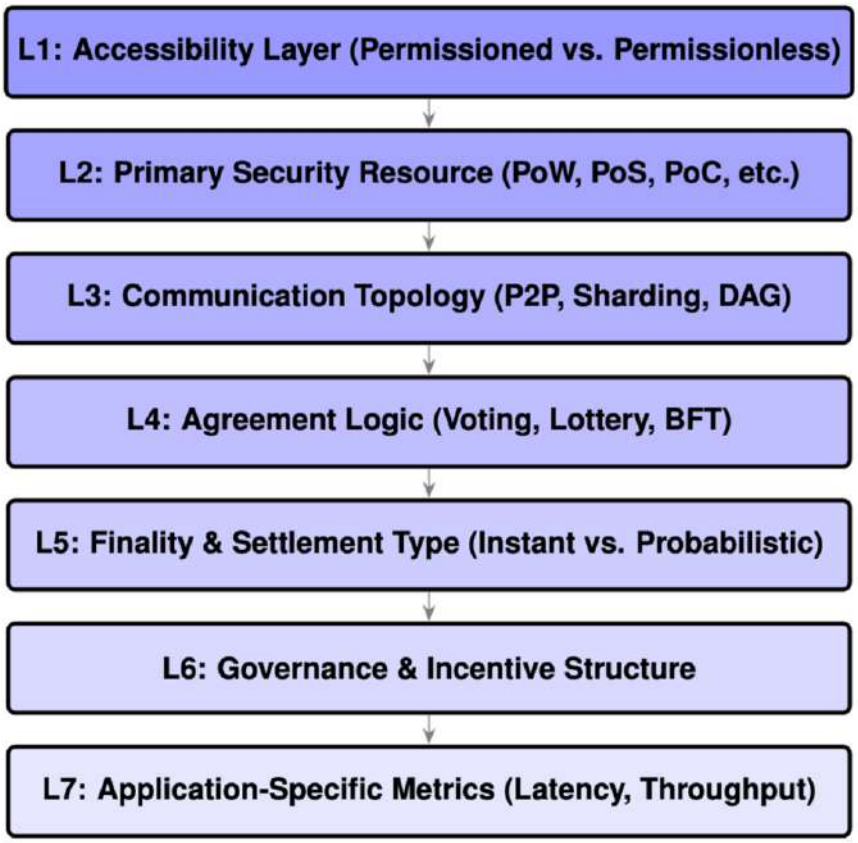

Figure 2. The proposed CD-7 decision-oriented framework for blockchain consensus mechanism selection. The framework organizes consensus design into seven architectural layers, enabling a systematic top-down evaluation of candidate mechanisms.

protocol-defined requirements such as computational work, token ownership, or storage capacity. This model maximizes openness and decentralization but requires mechanisms that prevent Sybil attacks.

Protocol-restricted validation. In this model, validator participation is determined by protocol rules rather than direct administrative approval. Validators may be selected through mechanisms such as delegation, federated trust, or weighted voting. While participation in the network may remain open, the validator set itself is restricted by protocol-defined mechanisms.

Permissioned validation. Permissioned systems rely on a predefined validator set. Validator nodes are explicitly authorized by network governance rules or administrative entities. Such systems are common in enterprise and consortium blockchain deployments where participant identities are known.

These three models define the institutional structure of validator participation and strongly influence the security assumptions of the consensus mechanism.

4.2. Primary Security Resource

The second dimension concerns the fundamental resource used to secure the network and prevent Sybil attacks. In decentralized systems, validators must demonstrate control over a scarce or verifiable resource in order to influence the consensus process.

Most consensus mechanisms can be categorized according to the primary resource they employ:

- Computational work (Proof of Work)
- Economic stake (Proof of Stake)
- Storage capacity (Proof of Space / Capacity)
- Identity or authority (Proof of Authority)
- Reputation or trust relationships
- Useful computational work such as machine learning tasks
- Quantum cryptographic primitives

From a conceptual perspective, these mechanisms can be grouped into two fundamental security paradigms.

The first paradigm is resource-based security, where validator influence is proportional to the quantity of a scarce resource controlled by the participant. Examples include Proof of Work and Proof of Stake.

The second paradigm is identity-based security, where consensus relies on known validators, trusted entities, or institutional authority. Byzantine Fault Tolerant protocols and authority-based systems typically fall within this category.

Emerging consensus models based on machine learning tasks or quantum cryptographic primitives extend the resource-based paradigm by introducing alternative forms of verifiable work or cryptographic validation.

4.3. Communication Topology and Agreement Logic

Once the validator set has been established, consensus mechanisms must provide a protocol for validators to reach agreement on the ordering and validity of transactions.

Several distinct agreement architectures have emerged in the literature:

Nakamoto consensus. This probabilistic consensus model relies on block production competitions among validators. Blocks are appended to the chain through resource-based leader selection mechanisms such as Proof of Work or Proof of Stake.

Byzantine Fault Tolerant (BFT) consensus. BFT protocols rely on deterministic voting among validators to reach agreement on the next block. These protocols provide strong safety guarantees but typically require a known validator set.

Federated consensus. Federated systems allow validators to define trust relationships with subsets of other validators. Agreement emerges through overlapping trust sets rather than a globally defined validator group.

DAG-based consensus. Directed acyclic graph (DAG) architectures replace the linear blockchain structure with graph-based transaction confirmation mechanisms. Validators confirm transactions by referencing multiple previous transactions rather than extending a single chain.

Each architecture represents a different trade-off between decentralization, communication complexity, and finality guarantees.

4.4. Finality, Governance, and Performance Trade-offs

Consensus mechanisms also differ in their operational performance characteristics. These characteristics reflect fundamental trade-offs between security, decentralization, scalability, and resource efficiency.

The most frequently analyzed properties include:

- Security robustness
- Scalability
- Energy efficiency

- Transaction throughput
- Degree of decentralization

No consensus mechanism simultaneously optimizes all of these properties. Instead, consensus protocols position themselves at different points within the design space defined by these trade-offs. Resource-based mechanisms such as Proof of Work typically prioritize decentralization and security, while BFT-based protocols often provide higher throughput and lower latency at the cost of reduced validator openness.

These trade-offs play a critical role when selecting an appropriate consensus mechanism for a specific application domain.

4.5. Mapping Consensus Mechanisms in the Design Space

The interaction between validator admission models and security paradigms defines a conceptual design space for blockchain consensus mechanisms. Figure 3 illustrates this design space by mapping representative consensus mechanisms according to these two dimensions.

Resource-based mechanisms such as Proof of Work and Proof of Stake typically operate in permissionless environments, whereas identity-based mechanisms such as PBFT and Proof of Authority are commonly deployed in permissioned networks. Protocol-restricted models occupy an intermediate position, combining elements of both paradigms.

4.6. Positioning of Consensus Mechanisms in the Design Space

The position of each consensus mechanism in the design space is determined by its dominant validator admission model and the primary security paradigm described in the literature.

The vertical axis represents the validator admission model, ranging from permissioned systems with predefined validator sets to permissionless systems where validator participation is open to any network participant.

The horizontal axis represents the dominant security paradigm. Resource-based mechanisms derive security from the control of scarce computational or economic resources, whereas identity-based mechanisms rely on known validators, institutional trust, or authority-based governance.

Some consensus mechanisms exhibit hybrid characteristics. In such cases, their placement reflects the dominant operational model reported in protocol specifications or academic analyses. Mechanisms positioned near the horizontal axis correspond to protocol-restricted validator admission models, where network participation may remain open but block production rights are limited by protocol-defined selection rules.

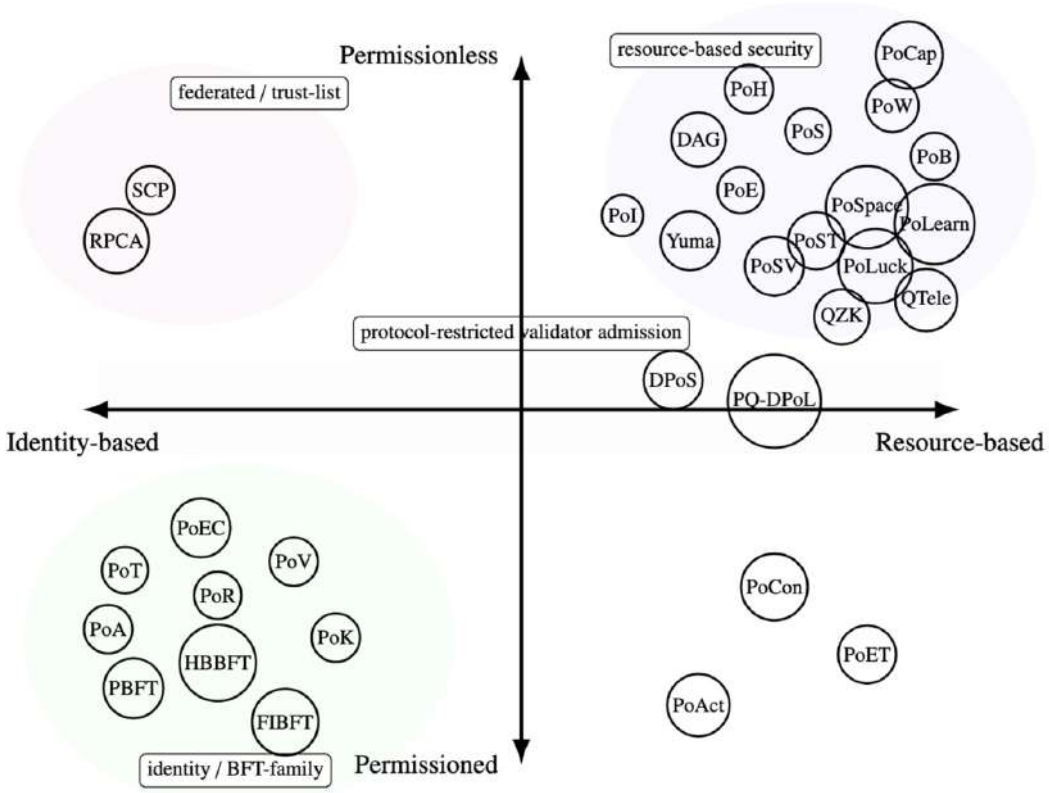

Figure 3. Consensus mechanism design space based on validator admission (permissioned vs. permissionless) and dominant security paradigm (identity-based vs. resource-based). Abbreviations: QTele = quantum teleportation-based consensus; QZK = quantum zero-knowledge-based consensus; HBBFT = HoneyBadgerBFT; PoA = Proof of Authority; PoAct = Proof of Activity; PoCon = Proof of Contribution; PoCap = Proof of Capacity

4.7. Framework-Based Prioritization and Comparative Evaluation

The proposed CD-7 framework enables the prioritization of consensus mechanisms through a sequential filtering process across its seven layers. Instead of ranking protocols based on a single criterion, the framework progressively constrains the set of feasible candidates by aligning system requirements with architectural properties.

At the initial stage, the validator admission model (L1) eliminates incompatible classes of consensus mechanisms by defining whether the system requires permissionless, protocol-restricted, or permissioned participation. This decision alone significantly reduces the candidate space. The second layer (L2) further refines the selection by identifying the dominant security resource, such as computational work, economic stake, storage, or identity.

Subsequent layers (L3–L5) determine the structural and operational characteristics of the system, including communication topology, agreement logic, and finality requirements. Finally, governance (L6) and performance constraints (L7) guide the prioritization of remaining candidates based on application-specific requirements.

Through this layered decision process, consensus mechanisms are not evaluated in isolation but are prioritized based on their compatibility with system-level constraints.

In contrast to existing approaches, which primarily provide descriptive taxonomies or isolated performance comparisons, the proposed framework introduces a decision-oriented structure that integrates multiple evaluation dimensions into a unified selection process.

Traditional surveys typically analyze consensus mechanisms using either qualitative classifications (e.g., PoW vs PoS vs BFT) or quantitative performance metrics such as throughput, latency, and scalability. While these analyses provide valuable insights, they do not directly support decision-making, as they treat performance characteristics independently of governance models and system constraints.

The CD-7 framework addresses this limitation by combining both qualitative and quantitative aspects within a single structured model. Qualitative properties such as decentralization, fault tolerance, and trust assumptions are captured in layers L1, L3, and L4, while quantitative considerations such as throughput, latency, and efficiency are incorporated in layer L7.

This integration enables a more comprehensive evaluation compared to traditional approaches. Instead of optimizing for a single metric, the framework supports multi-criteria decision-making by aligning consensus properties with application requirements.

Furthermore, unlike existing methods, the proposed framework explicitly models trade-offs between key system properties, such as decentralization, scalability, and security, thereby providing a practical tool for navigating the consensus design space.

4.8. Decision Framework for Consensus Selection

Based on the previously defined dimensions, the selection of a consensus mechanism can be formulated as a structured decision process.

The process begins by determining the validator admission model required by the network. This decision defines the institutional structure of validator participation. The next step involves selecting the security resource used to prevent Sybil attacks. Subsequently, an appropriate agreement architecture must be chosen to coordinate the validator set.

Finally, candidate consensus mechanisms are evaluated according to the performance requirements of the target application.

This sequential decision process transforms consensus selection from an ad-hoc choice into a structured architectural design problem. Instead of evaluating consensus algorithms individually, system architects can narrow the set of candidate mechanisms by progressively matching network requirements with the structural properties of consensus protocols.

The resulting framework provides a practical methodology for selecting blockchain consensus mechanisms across a wide range of application domains. This interpretation is consistent with the decision-oriented nature of the framework, where system parameters are used as structured input for selecting suitable consensus mechanisms.

5 Application Scenarios and Framework Validation

The selected application scenarios are intended to represent distinct regions of the consensus design space, rather than specific real-world implementations. Each scenario was chosen to reflect a different combination of validator admission models, security resources, and performance requirements, thereby enabling a systematic evaluation of the framework across heterogeneous system conditions. Specifically, the scenarios cover permissioned identity-based systems (waste management, institutional voting), protocol-restricted hybrid environments (smart city platforms), emerging security paradigms (post-quantum enterprise systems), and resource-based computation-driven models (cloud business intelligence platforms). This selection ensures that all major dimensions of the proposed framework are exercised, providing a representative validation of its applicability across diverse classes of distributed systems.

Rather than prescribing a single universally optimal consensus mechanism, the framework enables the identification of candidate consensus models that best align with the structural requirements of a given system. Each scenario is therefore analyzed by mapping system characteristics to the seven framework layers: validator admission model, security resource model, agreement architecture, and performance requirements. The decision process enabled by the CD-7 framework can be illustrated through a simplified example. Consider a blockchain system designed for a decentralized financial application that requires open validator participation, strong security guarantees, and moderate transaction throughput.

Following the framework, the first layer (L1) indicates that the network requires permissionless validator participation. The second layer (L2) determines the dominant security resource, such as computational work or economic stake. The subsequent layers (L3–L4) guide the

selection of an appropriate agreement architecture, for example a Nakamoto-style protocol or a BFT-based mechanism depending on the required finality guarantees.

The remaining layers refine the decision by considering settlement properties (L5), governance and incentive structures (L6), and application-specific performance requirements (L7). Through this layered filtering process, the set of candidate consensus mechanisms can be progressively narrowed to those that best match the requirements of the target system.

5.1. Blockchain-Based Waste Management Systems

Waste management infrastructures in smart cities increasingly rely on distributed data collection, regulatory compliance monitoring, and transparent auditing of waste flows. Blockchain technology has been proposed as a mechanism for ensuring tamper-resistant recording of waste collection, transportation, and disposal events [47]

In such systems, participating entities typically include municipal authorities, waste collection companies, recycling facilities, and regulatory agencies. The set of participants is therefore known in advance and controlled through institutional governance.

Within the proposed framework, this scenario corresponds to a permissioned validator admission model, as validator nodes are operated by authorized organizations participating in the waste management ecosystem. The dominant security resource is institutional identity, since validators represent regulated entities rather than anonymous participants.

Given the limited number of validators and the requirement for high data integrity, Byzantine Fault Tolerant (BFT) agreement architectures are particularly suitable. Protocols such as Practical Byzantine Fault Tolerance (PBFT), Flexible Byzantine Fault Tolerance (FIBFT), or authority-based models such as Proof of Authority (PoA) provide deterministic finality and predictable transaction confirmation times.

Performance requirements in this scenario prioritize reliability, auditability, and regulatory transparency rather than maximum throughput or full decentralization.

5.2. Institutional Blockchain Voting Systems

Blockchain-based voting systems have been explored as mechanisms for ensuring transparency, integrity, and verifiability in organizational decision-making processes. One example is the use of blockchain technology for recording and validating voting results in institutional governance bodies such as faculty councils [47].

In this context, all participants are known members of the institution, and validator nodes are typically operated by trusted organizational units or administrative infrastructure.

Consequently, the validator admission model is naturally permissioned, with validator identity serving as the primary security resource. The consensus mechanism must guarantee strong consistency, deterministic finality, and resistance to Byzantine faults.

Agreement architectures based on Byzantine Fault Tolerance are therefore well suited to this environment. Protocols such as PBFT or HoneyBadgerBFT enable reliable consensus among a limited set of validators while maintaining strong safety guarantees.

From a performance perspective, voting systems prioritize correctness and auditability over scalability. Transaction volumes are typically low, while the consequences of incorrect consensus are significant. As a result, consensus models emphasizing strong consistency are preferable to probabilistic mechanisms.

5.3. Smart City Infrastructure Platforms

Smart city platforms integrate data streams from numerous urban services, including transportation systems, environmental sensors, energy management infrastructure, and public services [48]. Blockchain technology has been proposed as a mechanism for coordinating data sharing and ensuring trust among multiple stakeholders.

Unlike purely institutional systems, smart city platforms often allow participation by various service providers, infrastructure operators, and external partners. While access to the network may be open, validator selection is typically governed by protocol rules or delegated governance structures.

This scenario therefore corresponds to a protocol-restricted validator admission model. Validator selection may be influenced by economic stake, service participation, or delegated authority within the platform ecosystem.

The dominant security resource in this context is often economic stake or participation weight. Consensus models such as Proof of Stake (PoS), Delegated Proof of Stake (DPoS), or activity-weighted mechanisms such as Proof of Importance (PoI) provide mechanisms for aligning validator influence with platform participation.

Agreement architectures in such systems often combine stake-weighted validator selection with probabilistic block production models similar to Nakamoto-style consensus.

Performance requirements emphasize scalability and transaction throughput, as smart city infrastructures may generate large volumes of data from heterogeneous sources.

5.4. Post-Quantum Secure Enterprise Systems

Enterprise blockchain deployments increasingly consider long-term cryptographic security as a critical design requirement. The potential emergence of large-scale quantum computers introduces risks to widely used public-key cryptographic schemes such as elliptic curve signatures [49].

Organizations seeking long-term integrity of blockchain records may therefore consider consensus mechanisms compatible with post-quantum cryptographic primitives.

In this scenario, validator participation is typically restricted to known organizational entities, resulting in a protocol restricted or permissioned validator model. The security

resource remains resource-based, but the consensus protocol must support post-quantum cryptographic operations.

An example of such an approach is Post-Quantum Delegated Proof of Luck (PQ-DPoL), which combines delegated validator selection with post-quantum signature schemes. By integrating quantum-resistant cryptographic primitives into validator authentication and block validation processes, such systems aim to preserve long-term security guarantees.

Performance requirements in enterprise environments prioritize predictability, security robustness, and operational stability over maximum decentralization.

5.5. Blockchain-Based Cloud Business Intelligence Platforms

Cloud-based business intelligence platforms increasingly integrate advanced analytics and machine learning capabilities for large- scale data processing. When such platforms incorporate blockchain technology to ensure data provenance, auditability, or decentralized coordination among service providers, consensus design may take advantage of the computational resources already present within the system.

In particular, platforms performing extensive machine learning training may consider consensus mechanisms that transform useful computational work into validator selection criteria.

Within the proposed framework, this scenario typically corresponds to a permissionless or protocol-restricted validator admission model, depending on whether external contributors may participate in training tasks. The dominant security resource becomes useful computational work, specifically machine learning model training and evaluation.

Proof of Learning (PoLearn) represents an example of such an approach, where validator selection is linked to the quality of trained models evaluated on unseen data. This mechanism allows computational effort to contribute both to consensus formation and to the improvement of analytical models used by the platform.

However, the practical applicability of PoLearn depends on the availability of reliable evaluation pipelines capable of verifying model quality while preventing manipulation or collusion among participants.

5.6. Comparative Scenario Analysis

The results presented in Figure 4 provide a comparative visualization of how different application scenarios map onto key consensus design requirements. Each radar chart illustrates the relative importance of fundamental system properties, including decentralization, scalability, throughput, finality, security, and resource efficiency, as derived from the corresponding use-case constraints.

A clear differentiation between scenarios can be observed. Institutional systems, such as voting platforms, emphasize security and deterministic finality while placing less importance

on scalability and throughput. In contrast, smart city platforms and cloud-based data systems prioritize scalability and throughput due to the high volume of data processing, while maintaining moderate security and decentralization requirements. Similarly, enterprise and infrastructure-oriented systems exhibit more balanced profiles, with increased emphasis on reliability, auditability, and long-term security.

The radar charts clearly demonstrate that no single consensus mechanism can optimally satisfy all criteria simultaneously, reinforcing the presence of inherent trade-offs within the consensus design space. Instead, each application domain corresponds to a distinct performance profile that reflects its operational priorities.

The proposed framework enables dynamic adjustment by allowing system designers to systematically prioritize different layers and performance dimensions according to application-specific requirements. Rather than relying on static comparisons of consensus protocols, the framework supports adaptive decision-making in which the relative importance of criteria such as security, decentralization, or efficiency can be adjusted based on system context. This capability represents a key advantage over traditional approaches, which typically evaluate consensus mechanisms using fixed or isolated metrics.

Table 2 summarizes the relationship between application scenarios and the corresponding framework dimensions, while Figure 4 provides a multi-criteria visual comparison of their performance requirements. The selection of candidate consensus mechanisms is derived directly from the constraints identified across the seven framework layers, ensuring traceability between system requirements and protocol selection.

Together, these results demonstrate that consensus mechanisms cannot be evaluated independently of system context. Instead, appropriate consensus models emerge from the interaction between validator governance, available security resources, agreement protocols, and performance objectives. The proposed framework therefore provides a structured and systematic methodology for aligning consensus design with the architectural and operational requirements of blockchain systems.

6 Discussion

The analysis presented in Sections 4 and 5 demonstrates that the selection of blockchain consensus mechanisms cannot be treated as a purely technical choice independent of system context. Instead, the appropriate consensus model emerges from the interaction between governance structure, available security resources, agreement architecture, and performance requirements. In addition, the framework does not prescribe a fully automated decision process, and the final selection of a consensus mechanism may still depend on expert judgment and context-specific considerations.

The proposed framework organizes these design factors into a structured decision process

Figure 4. Detailed comparative analysis of consensus design requirements across five specialized application domains.

Table 2. Consensus selection across representative blockchain scenarios.

Scenario	**Validator Model**	**Security Resource**	**Architecture**	**Candidate Consensus**
Waste management	Permissioned	Identity	BFT	PBFT, PoA
Faculty voting	Permissioned	Identity	BFT	PBFT
Smart city platform	Protocol-restricted	Stake/activity	Nakamoto-style	PoS, DPoS, PoI
Post-quantum enterprise	Protocol-restricted	Resource-based	Hybrid	PQ-DPoL
Cloud BI platform	Permissionless / restricted	ML computation	Resource-based	PoLearn

consisting of seven analytical layers (L1–L7).

One important observation from the analyzed scenarios is the vertical alignment between layers. For instance, in institutional governance (Section 5.2), the restriction at the Accessibility Layer (L1) directly dictates the deterministic nature of the Finality Layer (L5). This layered approach enables system designers to evaluate consensus mechanisms in relation to the structural properties of the target system rather than selecting protocols based solely on popularity or historical precedent.

A complementary observation is that different application domains naturally align with distinct regions of the consensus design space. Institutional governance systems, such as faculty voting platforms, tend to operate within permissioned environments where validator identity represents the primary security resource. In such cases, Byzantine Fault Tolerant consensus architectures provide deterministic finality and strong consistency guarantees.

Infrastructure-oriented systems, including waste management platforms and smart city data coordination systems, often involve multiple organizations with partially shared governance responsibilities. These environments frequently adopt consortium or protocol-restricted validator models, where authority or economic participation determines validator influence.

In contrast, large-scale open systems and distributed computational platforms tend to rely on resource-based security models in which validator influence is derived from measurable contributions such as computational work, economic stake, or storage capacity. In these contexts, consensus mechanisms based on probabilistic leader selection or delegated participation become more suitable.

The framework also highlights the increasing diversification of security resources used by modern consensus mechanisms. While early blockchain systems relied primarily on computational work or economic stake, recent proposals incorporate additional resources

such as storage capacity, reputation systems, trusted hardware, or useful computational tasks such as machine learning workloads. These developments indicate that consensus design is gradually expanding beyond traditional resource models toward more application-specific security primitives.

Another important observation concerns the role of emerging technological constraints, particularly the potential impact of quantum computing on cryptographic infrastructures. Recent studies have demonstrated that sufficiently advanced quantum computers could compromise widely used public-key cryptographic schemes, including those underlying blockchain systems, thereby posing significant security risks [51]. Enterprise systems requiring long-term data integrity may therefore increasingly consider post-quantum cryptographic compatibility when selecting consensus protocols.

Recent research has also explored quantum Byzantine agreement protocols, which leverage quantum mechanical properties to achieve higher fault tolerance and information-theoretic security. Such approaches have been shown to tolerate up to 50% faulty nodes while maintaining decentralization, representing a promising direction for future consensus designs[52].

Despite these advantages, several limitations of the proposed framework should be acknowledged. First, the categorization of consensus mechanisms relies on dominant design characteristics, while real-world implementations may combine multiple architectural features. Hybrid consensus architectures continue to emerge, blurring the boundaries between previously distinct design paradigms. Furthermore, the mapping within the CD-7 framework acknowledges the "Consensus Trilemma"; as shown in our multi-scenario analysis (Fig. 5), increasing decentralization (L1/L3) often necessitates trade-offs in throughput or finality (L5/L7).

Second, performance characteristics such as scalability, energy efficiency, and throughput were evaluated qualitatively rather than through quantitative benchmarking. Although this approach enables comparison across heterogeneous consensus models, it does not replace empirical performance evaluation within specific deployment environments. As a result, the framework should be interpreted as a decision-support tool rather than a fully objective optimization model.

Moreover, the abstraction level of the framework may omit certain low-level implementation details, which can influence performance and security in specific real-world deployments.

Finally, some recently proposed consensus mechanisms remain largely theoretical and have not yet been validated through large-scale production deployments. Their classification within the framework therefore reflects design assumptions rather than operational experience.

Nevertheless, the proposed framework provides a structured methodology for understanding the rapidly expanding design space of blockchain consensus mechanisms. Importantly, the contribution of this work lies not in proposing a new consensus mechanism, but in formalizing the decision process through which existing mechanisms can be systematically selected and aligned with system-level requirements. By linking consensus selection to system architecture and governance requirements, the framework contributes toward a more systematic approach

to blockchain system design.

7 Conclusion

This paper addressed the increasing complexity of the blockchain consensus landscape by proposing a structured framework for the systematic selection of consensus mechanisms. While numerous consensus protocols have been proposed in the literature, their selection is often performed in an ad hoc manner, without a clear methodological connection between system requirements and consensus design. The proposed CD-7 framework aims to reduce this ambiguity by organizing consensus selection around a seven-layer architectural decomposition, ranging from validator accessibility (L1) and primary security resources (L2) to agreement logic (L4), governance (L6), and application-specific metrics (L7).

By structuring the consensus design space along these dimensions, the framework provides a decision-oriented methodology that allows system architects to align consensus mechanisms with the institutional, technical, and operational constraints of the target blockchain system. Rather than evaluating consensus protocols in isolation, the framework emphasizes the relationship between system governance, available security resources, and the architectural properties of the consensus protocol.

The applicability of the framework was demonstrated through several representative application scenarios, including infrastructure monitoring systems, institutional governance platforms, smart city data coordination systems, enterprise environments requiring post-quantum security considerations, and cloud-based business intelligence platforms incorporating machine learning workloads. These scenarios illustrate that different application domains naturally align with different regions of the consensus design space.

The analysis further highlights the ongoing diversification of security resources employed by modern consensus mechanisms. While early blockchain systems primarily relied on computational work or economic stake, recent proposals increasingly incorporate additional resources such as storage capacity, reputation systems, trusted hardware, and application-specific useful computation. This trend suggests that future consensus designs may become increasingly specialized and closely integrated with the functional objectives of the underlying blockchain system.

Although the proposed framework provides a structured methodology for consensus selection, the rapidly evolving nature of blockchain technologies implies that new consensus paradigms will continue to emerge. Future research may therefore focus on extending the framework to incorporate hybrid consensus architectures, emerging cryptographic primitives, and empirical performance evaluations in large-scale deployments.

Overall, the framework presented in this study contributes toward a more systematic understanding of the blockchain consensus design space and provides practical guidance for

selecting consensus mechanisms that align with the governance models, security assumptions, and performance requirements of real-world blockchain systems.

Author Contributions: Conceptualization, M.S.; methodology, M.S.; validation, I.L., M.K. and Z.K.; formal analysis, M.S.; investigation, M.S.; resources, I.L, M.K.; data curation, M.S.; writing—original draft preparation, M.S.; writing—review and editing, M.K., Z.K; visualization, M.S., M.K.; funding acquisition, I.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded and is a part of research for the activities to be carried out for the project "Researching advanced algorithms and innovative business intelligence solutions in the cloud—NPOO.C3.2.R3-I1.04.0128".

Conflicts of Interest: The authors declare no conflicts of interest

References (as published in the original article)

1. Arooj, A.; Farooq, M.S.; Umer, T. Unfolding the Blockchain Era: Timeline, Evolution, Types and Real-World Applications. *J. Netw. Comput. Appl.* **2022**, 207, 103511.

2. Rawat, D.B.; Chaudhary, V.; Doku, R. Blockchain Technology: Emerging Applications and Use Cases for Secure and Trustworthy Smart Systems. *J. Cybersecur. Priv.* **2020**, 1, 4–18.

3. Rožman, N.; Corn, M.; Škulj, G.; Berlec, T.; Diaci, J.; Podržaj, P. Exploring the Effects of Blockchain Scalability Limitations on Performance and User Behavior in Blockchain-Based Shared Manufacturing Systems. *Appl. Sci.* **2023**, 13, 4251.

4. Jin, M.; Zhang, X.; Yue, X.; Zuo, Q.; Nie, J. Consensus Mechanism Selection in Web 3.0 Blockchain Platforms: Proof of Work vs. Proof of Stake. *Transp. Res. Part E Logist. Transp. Rev.* **2025**, 200, 104173.

5. Zheng, Z.; Xie, S.; Dai, H.-N.; Chen, X.; Wang, H. Blockchain Challenges and Opportunities: A Survey. *Int. J. Web Grid Serv.* **2018**, 14, 352.

6. Ramkumar, N.; Sudhasadasivam, G.; Saranya, K.G. A Survey on Different Consensus Mechanisms for the Blockchain Technology. In *Proc. 2020 Int. Conf. Communication and Signal Processing*; IEEE: Piscataway, NJ, USA, 2020; pp. 458–464.

7. Bamakan, S.M.H.; Motavali, A.; Babaei Bondarti, A. A Survey of Blockchain Consensus Algorithms Performance Evaluation Criteria. *Expert Syst. Appl.* **2020**, 154, 113385.

8. Gervais, A.; Karame, G.O.; Wüst, K.; Glykantzis, V.; Ritzdorf, H.; Capkun, S. On the Security and Performance of Proof of Work Blockchains. In *Proc. 2016 ACM SIGSAC*

Conf. Computer and Communications Security; ACM: New York, NY, USA, 2016; pp. 3–16.

9. Vukolić, M. The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication. In *Open Problems in Network Security*; Springer: Cham, Switzerland, 2016; pp. 112–125.

10. Kaur, S.; Chaturvedi, S.; Sharma, A.; Kar, J. A Research Survey on Applications of Consensus Protocols in Blockchain. *Secur. Commun. Netw.* **2021**, 2021, 6693731.

11. Toulemonde, A.; Besson, L.; Goubin, L.; Patarin, J. Useful Work: A New Protocol to Ensure Usefulness of PoW-Based Consensus for Blockchain. In *Proc. 2022 ACM Conf. Information Technology for Social Good*; ACM: New York, NY, USA, 2022; pp. 308–314.

12. Bano, S.; Sonnino, A.; Al-Bassam, M.; Azouvi, S.; McCorry, P.; Meiklejohn, S.; Danezis, G. SoK: Consensus in the Age of Blockchains. In *Proc. 1st ACM Conf. Advances in Financial Technologies*; ACM: New York, NY, USA, 2019; pp. 183–198.

13. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008.

14. King, S.; Nadal, S. *PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake*; Self-Published White Paper, 2012.

15. Amoussou-Guenou, Y.; del Pozzo, A.; Potop-Butucaru, M.; Sara, T.-P. *Correctness and Fairness of Tendermint-Core Blockchains*; UPMC Sorbonne Universités, 2018.

16. Yang, F.; Zhou, W.; Wu, Q.; Long, R.; Xiong, N.N.; Zhou, M. Delegated Proof of Stake With Downgrade. *IEEE Access* **2019**, 7, 118541–118555.

17. Kyriazis, N.A. A Survey on Volatility Fluctuations in the Decentralized Cryptocurrency Financial Assets. *J. Risk Financ. Manag.* **2021**, 14, 293.

18. Proof-of-Authority (PoA). Available online: `https://ethereum.org/hr/developers/docs/consensus-mechanisms/poa/` (accessed on 13 January 2025).

19. Yakovenko, A. *Solana: A New Architecture for a High Performance Blockchain v0.8*; Solana Labs, 2018.

20. Castro, M. Practical Byzantine Fault Tolerance. In *Proc. 3rd Symp. Operating Systems Design and Implementation*; USENIX: Berkeley, CA, USA, 2001.

21. Devarajan, A.; Karabulut, E. Directed Acyclic Graph Based Blockchain Systems. *arXiv* **2020**, arXiv:2312.09816.

22. Miller, A.; Xia, Y.; Croman, K.; Shi, E.; Song, D. The Honey Badger of BFT Protocols. In *Proc. 2016 ACM SIGSAC Conf. Computer and Communications Security*; ACM: New York, NY, USA, 2016; pp. 31–42.

23. Gao, N.; Huo, R.; Wang, S.; Huang, T. FIBFT: An Improved Byzantine Consensus Mechanism for Edge Computing. In *Proc. 2023 IEEE Wireless Communications and Networking Conf.*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.

24. Ripple XRP Whitepapers. Available online: `https://whitepaper.io/document/1/ripple-whitepaper` (accessed on 28 June 2023).

25. Amores-Sesar, I.; Cachin, C.; Mićić, J. Security Analysis of Ripple Consensus. In *Proc. 24th Int. Conf. Principles of Distributed Systems*; Schloss Dagstuhl, 2021; Vol. 184, pp. 10:1–10:16.

26. Mazières, D. *The Stellar Consensus Protocol: A Federated Model for Internet-Level Consensus*; Stellar Development Foundation, 2016.

27. Chen, L.; Xu, L.; Shah, N.; Gao, Z.; Lu, Y.; Shi, W. On Security Analysis of Proof-of-Elapsed-Time (PoET). In *Stabilization, Safety, and Security of Distributed Systems*; Springer: Cham, Switzerland, 2017; pp. 282–297.

28. Karantias, K.; Kiayias, A.; Zindros, D. Proof-of-Burn. In *Financial Cryptography and Data Security*; Springer: Cham, Switzerland, 2020.

29. Bada, A.; Damianou, A.; Angelopoulos, C.M.; Katos, V. Towards a Green Blockchain: A Review of Consensus Mechanisms and Their Energy Consumption. In *Proc. 2021 17th Int. Conf. Distributed Computing in Sensor Systems*; IEEE: Piscataway, NJ, USA, 2021.

30. Dziembowski, S.; Faust, S.; Kolmogorov, V.; Pietrzak, K. Proofs of Space. In *Advances in Cryptology—CRYPTO 2015*; Springer: Berlin/Heidelberg, Germany, 2015; Vol. 9216, pp. 585–605.

31. Moran, T.; Orlov, I. *Simple Proofs of Space-Time and Rational Proofs of Storage*; Springer: Cham, Switzerland, 2019; Vol. 11692, pp. 7–14.

32. Bentov, I.; Lee, C.; Mizrahi, A.; Rosenfeld, M. Proof of Activity: Extending Bitcoin's Proof of Work via Proof of Stake. *ACM SIGMETRICS Perform. Eval. Rev.* **2014**, 42, 34–37.

33. Xu, Y.; Yang, X.; Zhang, J.; Zhu, J.; Sun, M.; Chen, B. Proof of Engagement: A Flexible Blockchain Consensus Mechanism. *Wirel. Commun. Mob. Comput.* **2021**, 2021, 6185910.

34. NEM.io Foundation. *NEM Technical Reference*, 2018.

35. Ren, L. *Proof of Stake Velocity: Building the Social Currency of the Digital Age*; Reddcoin, 2014.

36. Li, K.; Hui, L.; Hou, H.; Li, K.; Chen, Y. Proof of Vote: A High-Performance Consensus Protocol Based on Vote Mechanism & Consortium Blockchain. In *Proc. 2017 IEEE 19th Int. Conf. High Performance Computing and Communications*; IEEE: Bangkok, Thailand, 2017; pp. 466–473.

37. Bahri, L.; Girdzijauskas, S. When Trust Saves Energy: A Reference Framework for Proof of Trust (PoT) Blockchains. In *Companion Proc. Web Conference 2018*; IW3C2: Geneva, Switzerland, 2018; pp. 1165–1169.

38. Dehez-Clementi, M.; Rabah, M.; Ghamri-Doudane, Y. A Privacy-Preserving Proof-of-Reputation. In *Proc. 2023 IFIP Networking Conf.*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–9.

39. Xue, T.; Yuan, Y.; Ahmed, Z.; Moniz, K.; Cao, G.; Wang, C. Proof of Contribution: A Modification of Proof of Work to Increase Mining Efficiency. In *Proc. 2018 IEEE 42nd Annual Computer Software and Applications Conf.*; IEEE: Tokyo, Japan, 2018; pp. 636–644.

40. Milutinovic, M.; He, W.; Wu, H.; Kanwal, M. Proof of Luck: An Efficient Blockchain Consensus Protocol. In *Proc. 1st Workshop on System Software for Trusted Execution*; ACM: Trento, Italy, 2016; pp. 1–6.

41. Yuma Consensus | Bittensor. Available online: `https://docs.bittensor.com/yuma-consensus` (accessed on 16 January 2025).

42. Bravo-Marquez, F.; Reeves, S.; Ugarte, M. Proof-of-Learning: A Blockchain Consensus Mechanism Based on Machine Learning Competitions. In *Proc. 2019 IEEE Int. Conf. Decentralized Applications and Infrastructures*; IEEE: Newark, CA, USA, 2019; pp. 119–124.

43. Biswas, A.; Yadav, R.; Baranwal, G.; Tripathi, A.K. Proof of Karma (PoK): A Novel Consensus Mechanism for Consortium Blockchain. *IEEE Trans. Serv. Comput.* **2023**, 16, 2908–2922.

44. Cheng, J.; Zhang, Y.; Yuan, Y.; Li, H.; Tang, X.; Sheng, V.; Hu, G. PoEC: A Cross-Blockchain Consensus Mechanism for Governing Blockchain by Blockchain. *Comput. Mater. Contin.* **2022**, 73, 1385–1402.

45. Wen, X.; Chen, Y.; Zhang, W.; Jiang, Z.L.; Fang, J. Blockchain Consensus Mechanism Based on Quantum Teleportation. *Mathematics* **2022**, 10, 2385.

46. Wen, X.-J.; Chen, Y.-Z.; Fan, X.-C.; Zhang, W.; Yi, Z.-Z.; Fang, J.-B. Blockchain Consensus Mechanism Based on Quantum Zero-Knowledge Proof. *Opt. Laser Technol.* **2022**, 147, 107693.

47. Kim, W.; Kang, Y.; Kim, H.; Jang, K.; Seo, H. PQ-DPoL: An Efficient Post-Quantum Blockchain Consensus Algorithm. In *Information Security Applications*; Springer Nature: Singapore, 2024; pp. 310–323.

48. Jiang, P.; Zhang, L.; You, S.; Fan, Y.V.; Tan, R.R.; Klemes, J.J.; You, F. Blockchain Technology Applications in Waste Management. *J. Clean. Prod.* **2023**, 421, 138466.

49. Köhler, M.; Švarcmajer, M.; Lukić, I.; Stipanić, T. Organization of a Digital Voting System Based on Blockchain Technology for the Faculty Council. In *Proc. 31st Int. Conf. Organization and Technology of Maintenance*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 25–36.

50. Lukić, I.; Köhler, M.; Krpić, Z.; Švarcmajer, M. Advancing Smart City Sustainability Through Artificial Intelligence, Digital Twin and Blockchain Solutions. *Technologies* **2025**, 13, 300.

51. Švarcmajer, M.; Krčmar, T.; Šabanović, D.; Lukić, I. Designing Post-Quantum Secure Blockchain Infrastructure for Business Intelligence Systems. In *Proc. 2025 Int. Symp. ELMAR*; IEEE: Piscataway, NJ, USA, 2025; pp. 245–248.

52. Xie, Y.-M.; Lu, Y.-S.; Weng, C.-X.; Cao, X.-Y.; Jia, Z.-Y.; Bao, Y.; Wang, Y.; Fu, Y.; Yin, H.-L.; Chen, Z.-B. Breaking the Rate-Loss Bound of Quantum Key Distribution with Asynchronous Two-Photon Interference. *PRX Quantum* **2022**, 3, 020315.

53. Weng, C.-X.; Gao, R.-Q.; Bao, Y.; Li, B.-H.; Liu, W.-B.; Xie, Y.-M.; Lu, Y.-S.; Yin, H.-L.; Chen, Z.-B. Beating the Fault-Tolerance Bound and Security Loopholes for Byzantine Agreement with a Quantum Solution. *Research* **2023**, 6, 0272.

E

Paper E: Advancing Smart City Sustainability Through Artificial Intelligence, Digital Twin and Blockchain Solutions

Technologies, vol. 13, no. 7, p. 300, 2025

Ivica Lukić, Mirko Köhler, Zdravko Krpić and Miljenko Švarcmajer
Faculty of Electrical Engineering, Computer Science and Information Technology, Josip Juraj Strossmayer University of Osijek, Kneza Trpimira 2b, 31000 Osijek, Croatia

Abstract

This paper presents an integrated Smart City platform that combines digital twin technology, advanced machine learning, and private blockchain network to enhance data-driven decision-making and operational efficiency in both public enterprises and small and medium-sized enterprises (SMEs). The proposed cloud-based business intelligence model automates Extract, Transform, Load (ETL) processes, enables real-time analytics, and secures data integrity and transparency through blockchain-enabled audit trails. By implementation of proposed solution Smart City and public service providers can significantly improve operational efficiency, including a 15% reduction in costs and a 12% decrease in fuel consumption for waste management, as well as increased citizen engagement and transparency in Smart City governance.

The digital twin component facilitated scenario simulations and proactive resource management, while the participatory governance module empowered citizens through transparent, immutable records of proposals and voting. The study also discusses technical, organizational, and regulatory challenges, such as data integration, scalability, and privacy compliance. The results indicate that the proposed approach offers a scalable and sustainable model for Smart City transformation, fostering citizen trust, regulatory compliance, and measurable environmental and social benefits.

Keywords: digital twin; blockchain; business intelligence; smart city; ETL; machine learning; participatory governance; urban sustainability

1 Introduction

In today's digital economy, the need for effective Business Intelligence (BI) systems is growing exponentially [1], especially given the huge amount of data generated every day. However, while business intelligence offers significant benefits in terms of data-driven decision-making, the implementation of these systems in small and medium-sized enterprises (SMEs) faces numerous challenges such as limited scalability, high implementation costs, outdated data, security threats, lack of management support, insufficient financial resources, ineffective project management, and resistance to change, all of which can hinder their successful adoption and use of advanced business intelligence systems. These threats make it difficult for SMEs to access business intelligence, while the cost is often being the biggest barrier.

Smart Cities can act as data hubs, offering ETL processes, analytics, and advisory services to their public enterprises and departments, thus supporting holistic digital transformation and operational excellence across the urban ecosystem [2]. In addition to challenges and opportunities for SMEs, this paper presents that the integration of digital twin technology, artificial intelligence to secure scalable data services, not only to private SMEs, but also to public enterprises and city administrations.

According to the research [3], SMEs face a number of obstacles when adopting business intelligence systems. One of the key challenges is exceeding the implementation budget since they do not have sufficient financial resources to launch projects. Also, ineffective BI project management and lack of management support are obstacles to the successful implementation of BI systems in SMEs. A lack of understanding of BI system capabilities among key stakeholders leads to uncertainty and resistance to change, globalization and a turbulent environment, where finding competitive advantages is a difficult task for both small and large enterprises [4]. Similar problems are identified in [5], where the authors state that the lack of training and support for users is another significant challenge. SMEs often do not have the necessary resources to train their employees in the use of BI tools, which can lead to a low adoption rate of the system. In addition, user resistance to change is often present, as

many employees believe that BI systems require significant changes in their work processes and decision-making methods. However, despite the challenges, the implementation of BI systems in SMEs brings numerous advantages. SMEs that successfully implement BI systems report greater decision-making efficiency. BI systems enable faster data analysis and better understanding of market trends, which enables SMEs to make better strategic decisions based on data.

Security and data protection are also key factors for SMEs when digitizing business using data science and artificial intelligence, and it is necessary to use advanced security measures to protect privacy [6]. In order to increase security, transparency, and data protection, this paper suggests connecting the BI system with a private blockchain network. This approach allows SMEs to protect themselves from threats, thereby increasing trust between partners and clients. Although the implementation of a BI system in SMEs can be challenging, considering financial, human and technical resources, the benefits it provides in the form of better data analysis, increased efficiency and data security make BI a key tool for digital transformation. By optimizing business processes, reducing costs, and improving decision-making, SMEs can achieve a competitive advantage in the global market.

This paper presents a cloud-based business intelligence model that integrates advanced machine learning algorithms into the ETL process, proposes a digital twin of the Smart City, and utilizes a private blockchain network. The functionalities of the proposed Smart City digital twin include automating data collection and preparation, while advanced analytics and predictive models based on machine learning are applied to optimize urban processes. At the same time, private blockchain technology is employed to ensure confidentiality, integrity, and transparency of data to address key security challenges in business analytics. The digital twin facilitates real-time simulation and proactive management of urban resources, supporting scenario analysis and data-driven decision-making for public enterprises and city administrations. Through blockchain-enabled audit trails, the integrity and traceability of all ETL process stages are maintained, supporting regulatory compliance. The functionalities of the proposed Smart City model are demonstrated in a pilot implementation, where measurable improvements in operational efficiency, cost reduction, environmental impact, and citizen engagement are achieved, confirming the potential of the integrated digital twin and blockchain approach for sustainable Smart City transformation.

While recent research has demonstrated the potential of digital twins, AI, and blockchain to enhance various aspects of smart city management—including real-time simulation, participatory governance, and secure data sharing—most existing solutions focus on isolated applications or lack integration across the full urban data lifecycle [7,8]. The approach presented in this paper distinguishes itself by delivering a unified, modular platform that tightly couples a real-time digital twin with advanced machine learning–driven ETL processes and a permissioned blockchain infrastructure for end-to-end data integrity, transparency, and participatory governance. Unlike prior studies that emphasize either digital twin analytics, blockchain-secured audit

trails, or AI-powered optimization in isolation, our model uniquely enables seamless interoperability between public services, SMEs, and citizens through open APIs and privacy-preserving identity management [8,9]. This comprehensive integration not only supports automated, secure, and transparent urban data workflows, but also empowers stakeholders—including city administrations, enterprises, and citizens—to collaboratively co-create, monitor, and govern smart city services in real time. The pilot deployment demonstrates measurable benefits in operational efficiency, citizen engagement, and environmental impact, positioning the proposed solution as a scalable and replicable model for next-generation sustainable smart cities.

Despite significant advances in Smart City technologies, several critical challenges remain unaddressed. First, data interoperability is a persistent issue, as urban data originates from diverse sources and different formats which is complicating integration and analysis. Second, security and privacy concerns arise due to the sensitive nature of urban and citizen data, necessitating robust mechanisms for data protection and compliance. Third, effective stakeholder engagement, including citizen participation and transparent governance remains limited by the lack of trustworthy and accessible digital platforms. This paper addresses these challenges by proposing an integrated Smart City platform that combines digital twin technology, advanced machine learning, and private blockchain infrastructure to ensure seamless data interoperability, enhanced security, and participatory governance.

The paper is structured as follows: Section 2 reviews related work on the integration of digital twin technology, AI, and blockchain in smart city. Section 3 introduces the proposed digital twin and blockchain architecture. Section 4 describes the implementation of the private blockchain network using the Hyperledger Besu. Section 5 presents the evaluation of the pilot implementation and discusses technical and organizational challenges as well as directions for future work. Section 6 concludes the paper.

2 Related Work

Recent research highlights the transformative potential of integrating real-time digital twins, blockchain-enabled participatory governance, and AI-powered resource management at the city scale [7,9]. By leveraging these technologies, cities are enabled in real time to simulate, monitor, and optimize core systems such as energy, water, waste management and mobility. Digital twins serve as dynamic, data-driven replicas of city environments, allowing for scenario modeling, predictive analytics, and proactive resource allocation [10,11]. Blockchain technology is being deployed to underpin participatory governance, providing immutable and transparent records of citizen proposals, voting, and administrative actions [12,13]. This ensures that all stakeholders can audit decision-making processes and fosters greater trust in public administration. AI-driven analytics further enhance operational efficiency by

enabling predictive maintenance, demand forecasting, and the automation of complex city processes, leading to measurable improvements in cost savings, environmental impact, and service quality. Open dashboards and transparent data platforms are used to engage citizens, offering real-time insights into city performance and enabling meaningful participation in city governance [10,14,15]. The combination of these technologies not only improves operational excellence but also supports regulatory compliance, sustainability goals, and citizen empowerment, positioning cities as resilient, data-driven ecosystems ready to address future challenges.

The ETL process is a key step in the data preparation for analytical systems, whereby data from various sources is collected, processed and loaded into data warehouses. Traditional ETL systems often face challenges such as handling large amounts of data, identifying anomalies, ensuring data quality, and supporting real-time updates. In this context, machine learning brings significant innovations, enabling the automation and optimization of these processes in order to improve the performance and accuracy of data processing [16,17].

Machine learning and decision-making needs a large volume of heterogeneous data in various complex formats. To solve this problem a hybrid swarm intelligence algorithm with tabu search for handling large amounts of data in a cloud-based ETL process in a data warehouse is proposed [18]. Furthermore, in [19] authors provide an overview of the existing ETL process and data quality approaches to enable efficiently data processing by sharing data in a distributed cloud architecture. The data is collected from various sources such as the private sector, companies, government sectors, and their analysis helps users to make better business and organizational decisions. The data collection can be optimized by determining the time intervals for data collection from external sources [20]. Another approach is proposed in [21], where authors propose various ETL tools for data processing when making business decisions. Business intelligence is widely applied across many domains, including healthcare for patient data analysis and decision support data [22], as well as for real-time data processing and secure analytics using relational algebra and other advanced techniques [23].

Traditional ETL processes often face several key challenges, such as lack of scalability where large amounts of data can slow down processing, which is especially problematic in big data environments. Traditional ETL systems are often not scalable enough to handle large volumes of data [24,25]. Also, handling and integration of heterogeneous data sources, which can be unstructured and semi-structured data, traditional systems cannot always process efficiently [26]. Manual anomaly detection and correction of errors in data is a time-consuming process, which can negatively affect data quality and analysis accuracy [27]. Real-time updates pose a challenge for many traditional ETL systems, which are optimized for periodic data loads [28]. The integration of machine learning into ETL processes enables significant improvements in the following stages:

- Data extraction: machine learning algorithms can automatically identify key information

from large and diverse data sources, including text documents, Internet of Things (IoT) devices, and social networks. Automating this step reduces the time required for data collection and increases process reliability [16,29].

- Data transformation: algorithms such as clustering and deviation analysis can recognize irregularities in data and automatically correct them, thereby improving data quality and standardization of data from heterogeneous sources [30]. Prediction of missing values by techniques such as regression, allow filling in missing data based on existing patterns, reducing the need for manual data entry or discarding.
- Data loading: by optimizing performance, machine learning algorithms can predict optimal time intervals for loading large amounts of data, ensuring minimal system load and faster processing [31]. Continuous updating and synchronization of data is enabled, which is crucial for applications that require updated information in real time [16].

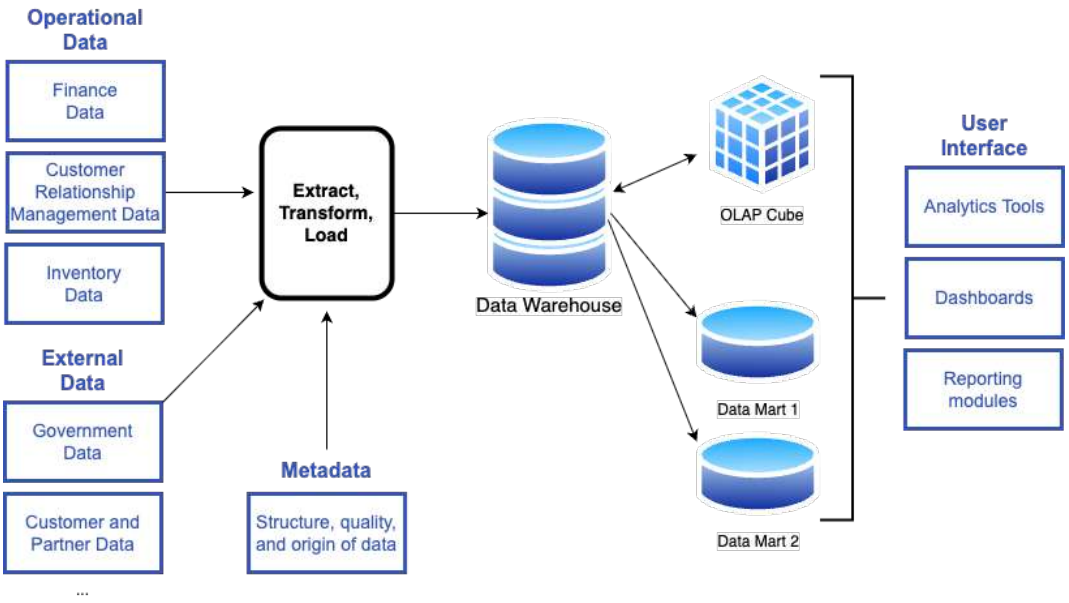

Figure 1. The ETL process and business intelligence architecture.

By introducing advanced machine learning algorithms, ETL processes become faster, more reliable, and more adaptable to dynamic business needs. Machine learning algorithms for optimizing ETL processes represent a key step towards creating business intelligence systems that enable informed decisions based on high-quality and up-to-date data. This technology, with integration with blockchain for additional security, has the potential to revolutionize the way organizations process and use data, enabling them to successfully face the challenges of the digital age.

Blockchain technology, best known for its application in cryptocurrencies, is becoming increasingly relevant in other industry sectors, including business intelligence. Key features

of blockchain, such as decentralization, non-repudiation, and transparency, make it ideal for solving security challenges in data analysis and processing systems. These advantages become even more pronounced in private networks because they allow better control over data within a closed ecosystem. The application of blockchain technology ensures the permanent recording of transactions or changes in data sets. Once data is entered into the blockchain, it cannot be changed without leaving a trace. This feature ensures that the history of data can be easily verified, thus eliminating the possibility of manipulation [32]. For business intelligence, this means that data-driven analysis and decisions are always based on accurate and credible information [33]. Blockchain uses advanced encryption mechanisms to ensure the confidentiality and integrity of data. Only authorized users can access certain parts of the data within the network [34]. Private blockchain networks further restrict access to only authorized participants, reducing the risk of unauthorized intrusion. In [35] authors discussed and benchmarked different blockchain platforms for data processing, and blockchain access control is explained to improve security of private blockchain networks [36]. Furthermore, all transactions within the blockchain are visible to relevant participants, which allows for complete transparency and easier tracking of activities [37]. For example, in business intelligence systems this can be used for audit trails to ensure that data is used and processed in accordance with regulations [12,38,39]. Unlike the centralized systems, blockchain stores data on a distributed network, eliminating a single point of vulnerability [32]. Even if one node within the network becomes compromised, the rest of the system remains intact and functional [34].

For all of the above, the advantages of using blockchain in business intelligence are obvious because it can ensure data non-repudiation during all stages of the ETL process. During extraction, the blockchain can record metadata about the data source to ensure its authenticity, and during transformation, every change in the data is stored, creating an audit trail. When loaded, the data can be stored in a warehouse that is connected to the blockchain for continuous validation. Data protection regulations are also supported, as private blockchain networks can help organizations comply with regulations such as GDPR and HIPAA by ensuring transparency in how data is managed and providing evidence of compliance through a transaction log. Fraud and data misuse in business intelligence systems are prevented because data stored in the blockchain network can automatically detect attempts at manipulation or unauthorized access. Transparent logs within the blockchain can serve as evidence in the event of any disputes. Due to the application of smart contracts [40], blockchain can be applied for authentication and trust management in business intelligence [41], solving security problems and challenges [42,43], collecting data from IoT systems [44,45].

3 Integrated Digital Twin and Blockchain Architecture for Smart City Transformation

The integration of private blockchain technology with business intelligence brings a new level of security and transparency in data transformation and analysis. Data non-repudiation, encryption, transparency and resistance to attacks enable Smart City organizations decision making with greater confidence. In combination with advanced machine learning algorithms and AI, blockchain becomes a key tool in the development of systems that provide reliable and secure business analytics, ready for the challenges of the digital age.

3.1. Business Intelligence Model and Digital Twin

The business intelligence model proposed in this paper has several key advantages compared to existing solutions:

- Increasing competitiveness: using advanced machine learning algorithms in the ETL process more accurate data analysis is enabled. Users are provided with an insight into market trends, optimized business processes and a better strategy, which achieves a competitive advantage on a global level [46].
- Digital business transformation: integrating business intelligence in the cloud allows SMEs to modernize and migrate business processes to digital platforms by connecting it with blockchain technology to ensure data transparency and security, which is key to digital transformation in industries such as finance [47,48], healthcare [49], or retail [50].
- Increased security and trust: blockchain networks offer a high level of security and transparency, which allows for easier verification of data and transactions. This increases trust among partners, customers and regulators, which is key to the global competitiveness of SMEs.
- Scalability and flexibility: a cloud solution enables scalability, which means it can grow with the business while reducing the need for large initial infrastructure investments. It also uses the flexibility that clouds offer in terms of speed of implementation and adaptability to market changes [51,52].
- Optimizing business decisions: the use of business intelligence enables faster making of informed decisions based on data. With the help of advanced analytics and machine-based recommendations, Smart City companies can better understand market opportunities, reduce risks, and optimize resources, thus increasing their efficiency and reducing costs [53].

These advantages can significantly improve business processes and make SMEs more competitive on the global market, which is crucial for long-term sustainability and success in the digital era [54]. The proposed model extends the traditional BI approach by embedding a real-time digital twin of the Smart City by integrating IoT data streams (e.g., energy, water, waste, air quality, mobility) and advanced visualization tools. This enables both Smart City officials and citizens to monitor urban metrics, participate in scenario planning, and create solutions for sustainable Smart City development. A digital twin enables the creation of a dynamic, real-time virtual replica of a city environment, process, or entire city and enables simulation of various scenarios to predict future outcomes. This virtual model is continuously synchronized with its real-world counterpart using data from sensors, IoT devices, and other sources, allowing for up-to-date monitoring, analysis, and simulation. By leveraging digital twins, organizations and city administration is enabled to: monitor and analyze the performance of assets and systems in real time, detecting inefficiencies or faults as they arise; run simulations and scenario analyses to predict the outcomes of changes, optimize operations, and support data-driven decision-making; test new designs, policies, or interventions virtually before implementing them in the physical environment, reducing risk and costs; enable predictive maintenance and proactive resource management, improving operational efficiency and reducing downtime; foster transparency and participatory governance by providing stakeholders—including citizens—with access to real-time data and interactive dashboards; and support sustainability goals by optimizing resource usage, reducing emissions, and enhancing overall system performance.

In the context of Smart Cities, digital twins serve as a strategy accelerator integrating static, historical, and real-time data to provide actionable insights into city administration, improve service delivery, and empower both decision-makers and citizens.

3.2. Private Blockchain Network

The implementation of a private blockchain network requires detailed planning and implementation of specific steps to ensure security, reliability and transparency. Private permissioned blockchain networks differ from public blockchain networks in that only authorized members have access to the network, which makes them an ideal solution for business purposes that require a high level of confidentiality and data control [35,36]. In the context of Smart Cities, blockchain is not only used for securing business intelligence data, but also for enabling participatory governance. In this paper a permissioned blockchain infrastructure where all Smart City initiatives are immutably recorded, ensuring transparency, auditability, and trust is proposed. Privacy-preserving digital identity management should be integrated to protect personal data and support secure authentication for both citizens and Smart City administration.

Key steps in implementation a private blockchain network are choosing the right blockchain

platform suitable for smart contracts [55]. The most commonly used platforms for this purpose include Hyperledger Fabric which offers a modular architecture, flexibility and support for smart contracts, making it suitable for complex business processes [56]. Corda focuses on data privacy and is used mainly in the financial sector [57]. Quorum is a version of the Ethereum blockchain network adapted for business use with an emphasis on privacy and transaction speed. The choice of platform depends on specific requirements such as performance, scalability and support for smart contracts [58]. Hyperledger Besu is an EVM-compatible blockchain client that enables the development of private and consortium blockchain networks with a focus on security [59].

In the next step is necessary to implement a Public Key Infrastructure (PKI) where certificates ensure that only authorized users can access the network. Authentication and authorization rules where users and network nodes are given specific roles and access rights, thus minimizing the risks of abuse [34]. For example, network administrators have full access, while other users have limited privileges, depending on their role. Configuration of network nodes with different roles and functions, where validation nodes verify and approve transactions and client nodes that enable users to access the network. Each node must be configured to communicate with other nodes via secure protocols. To achieve optimal performance, the network should be designed to ensure load balance between nodes [60]. Securing the network infrastructure is a key aspect of the security of a private blockchain network. This includes data encryption using asymmetric encryption to protect sensitive data [27]. Monitoring network activity using software tools to identify suspicious activities and unauthorized access attempts. Regular backups are required to ensure data integrity and enable rapid recovery after an incident.

After the network is secure smart contracts should be implemented to enable automation of business processes within the blockchain. For example, the authenticity of collected data can be verified where the data is checked and verified without the need for an intermediary. The blockchain stores metadata about all collected data in the ETL process, user queries, etc. Payment automation can be implemented after all conditions are met and the payment is executed automatically. Coding smart contracts requires thorough testing to prevent errors that could compromise the security or functionality of the network. Testing and optimization are key to uncovering weaknesses in the system. Performance testing checks the speed of transactions and the scalability of the network under different loads. It is necessary to provide security protocols that simulate potential attacks to ensure network security [61,62]. Compatibility with other systems participating in the model should also be checked. After testing and spotting potential flaws, the network is adjusted to achieve optimal performance. Maintenance and upgrades ensure security patches and regular software updates to eliminate security vulnerabilities. Continuous network monitoring to ensure reliability. Adding new features according to business needs. Maintenance requires an expert team to ensure that the network remains secure, scalable and compliant with regulatory requirements [63].

3.3. Specific Model Requirements for Business Intelligence Applications

Blockchain networks are particularly useful in the context of business intelligence, where it is crucial to ensure data integrity throughout the ETL process. During the ETL process, the collected data is stored in a data warehouse [64]. A key innovation of the proposed digital twin model is the application of blockchain based audit trails beyond the private sector, supporting compliance and transparency in public resource management (e.g., waste, mobility, energy). For example, smart waste bins equipped with IoT sensors, machine learning for route optimization, and blockchain for lifecycle tracking enable cities to reduce costs, improve recycling rates, and report transparently to citizens and regulators.

The metadata are also stored in the blockchain during all phases. During the data extraction phase, the following metadata is stored: record ID (a unique identifier for each data item collected); information about the source of the collected data (e.g., database, IoT device, API); timestamps for historical traceability; content of the collected data; data hash values calculated over key data sets to ensure their integrity; information about the users or applications that initiated the extraction process; and records of the quality and validation of the data source.

During the data transformation phase, the following metadata is stored: transformation operations (types of changes, such as aggregation, normalization, or data cleansing); data versions (the history of changes for each transformation); validation rules (the criteria applied to ensure accuracy and consistency); and error identification (any anomalies or errors detected and how they were handled).

During the data loading phase, the following metadata is stored: destinations (locations where the data is stored); load timestamps; data integrity controls; and audit trail records of all process steps.

This implementation allows for the creation of a transparent and secure ETL process, thus ensuring confidence in analytical results and compliance with regulatory requirements [12]. It enables secure, transparent, and efficient processing, while ensuring data confidentiality and integrity.

3.4. User Interaction With the Proposed Model Using Natural Language Processing

Once the data is stored in the data warehouse, users can ask the LLM Server queries related to the data in the warehouse, allow it to perform analysis from their own data, request advice, simulate possible scenarios, and more. In addition to SME users, the Smart City platform supports city administration, public enterprise managers, and citizens, enabling them to interact with the system via Natural Language Processing (NLP) queries, access real-time analytics, and participate in governance processes. All interactions are securely logged on

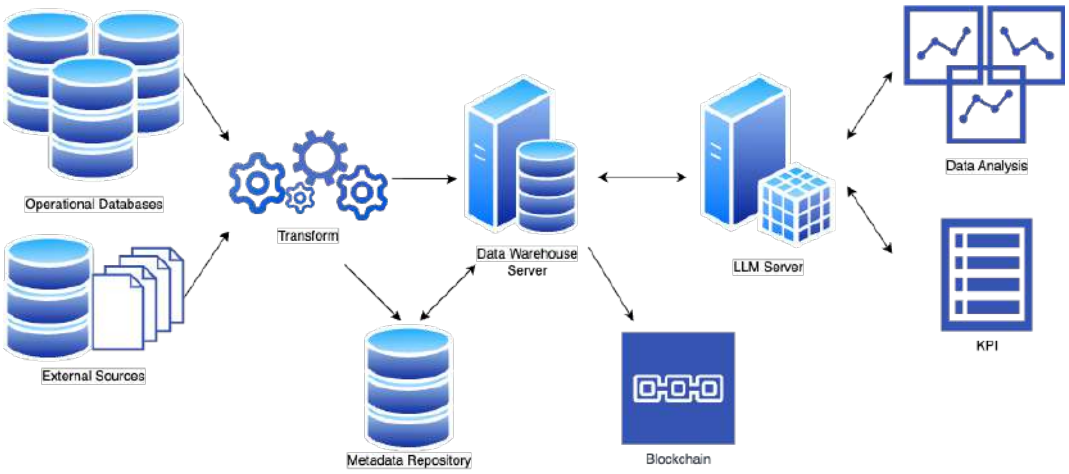

Figure 2. Proposed ETL process with LLM and blockchain integration.

the blockchain, ensuring auditability for both operational and participatory functions. This process includes additional security and functional measures: user data is not stored on the platform's central servers to preserve privacy, instead remaining on the user's computer with the LLM Server accessing it only temporarily during processing; recording all user queries and results in the blockchain enables authentication and transparency; the hash of results allows for later verification of accuracy; and an audit trail allows tracking of user activities.

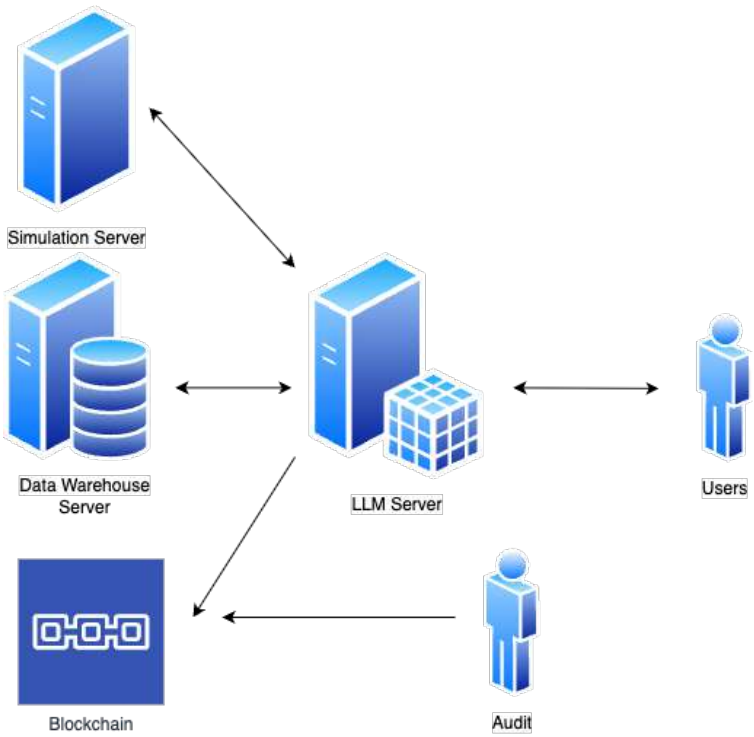

Figure 3. User communication with the proposed model.

To clarify the AI layer of the system, the platform integrates several machine learning algorithms tailored to specific smart city tasks. For predicting energy consumption, traffic management, waste generation and optimizing collection schedules, Long Short-Term Memory

(LSTM) neural networks can be used due to their proven ability to model temporal dependencies and non-linear trends in urban data streams. LSTM models are especially effective for capturing daily and weekly fluctuations in waste patterns, as confirmed in recent smart city studies [65]. Random Forest classifiers can be applied to identify anomalies in sensor data and to classify waste types, chosen for their robustness to noisy data and high performance in multi-class classification tasks [66]. Route optimization and dynamic resource allocation are addressed using a hybrid approach combining classical optimization algorithms with reinforcement learning, enabling adaptive, real-time decision-making in response to changing urban conditions.

The AI layer is enhanced by the integration of a dedicated LLM server, enabling advanced NLP capabilities for user interaction, analytics, and decision support. The LLM server processes queries, performs analyses, provides recommendations, and simulates scenarios based on real-time and historical data, while protecting user privacy by not storing data centrally. Each query and its result are hashed and recorded as part of the blockchain's immutable audit trail. The system additionally supports Large Language Model Agent integration (LAMA), leveraging generative AI for automated report generation, anomaly detection, predictive analytics, and scenario simulation. The LAMA agent is deployed on the LLM server, provisioned with robust hardware and containerized using Docker for scalability and reliability, with integration achieved via open APIs for interoperability between the LLM server, digital twin, blockchain, and other city services.

3.5. Comparative Analysis and Contextualization

Digital twins serve as dynamic virtual replicas of urban environments, enabling real-time monitoring, scenario simulation, and predictive analytics for city systems. Recent literature highlights their growing use in supporting decision and policy making, particularly for scenario modeling, sustainability planning, and adaptive governance. However, most implementations remain domain-specific and are not fully integrated with participatory governance or secure data management frameworks. Recent research demonstrates that integrating digital twin technologies with advanced data analytics at the neighborhood level significantly enhances urban resource management and resilience [10], providing a systematic review of digital twin applications in public sector decision support and emphasizing the need for context-specific models and integration with governance processes. The development of a smart city platform based on digital twin technology enables real-time monitoring and simulation of urban processes, supports decision-making, and enhances transparency and citizen engagement in city management [67]. The comparative analysis across different approaches to smart city digitalization is shown in Table 1.

The proposed cloud-based platform and business intelligence (BI) model integrates machine learning (ML) into the ETL process to automate data workflows, enhance data quality, and

Table 1: Key Features Across Different Approaches.

Feature	**Traditional Models**	**Recent Integrated Models**	**Proposed Platform**
Digital Twin Scope	Domain-specific, siloed	Multi-domain, limited integration	Unified, real-time, cross-domain
Blockchain Use	Audit trails, transactions	Data provenance, some participatory tools	End-to-end integrity, participatory governance
AI Integration	Standalone analytics	Predictive analytics, not fully coupled	Embedded in ETL, scenario simulation
Data Interoperability	Low-moderate	Improving, but still challenging	High, via modular/open APIs
Citizen Engagement	Limited, portal-based	Some participatory pilots	Blockchain-enabled, auditable

Table 2: Traditional ETL vs. Proposed ML-Enhanced ETL.

Feature	**Traditional ETL**	**Proposed ML-Enhanced ETL**
Transformation	Rule-based, manual scripting, static transformations	ML-driven automation predicts optimal transformations, detects anomalies, and imputes missing values dynamically
Scalability	Limited by on-premise infrastructure; batch-oriented	Cloud-native leverages modular architectures for elastic scaling
Data Quality	Manual error detection, reactive fixes	Proactive quality control due to ML models having anomaly data detection
Processing Speed	Hours/days for large datasets	Real-time capabilities enabled near-instantaneous processing of streaming data via AI/ML
Cost Efficiency	High upfront infrastructure costs	Cloud services reduce operational costs

enable real-time analytics. By leveraging ML-driven automation, the platform dynamically predicts optimal transformations, detects anomalies, and imputes missing values, resulting in more accurate and reliable data processing compared to rule-based, manual scripting used in conventional systems. This also enables real-time analytics, allowing for near-instantaneous processing of streaming data, a capability typically limited to batch-oriented, slower traditional ETL workflows.

Scalability is another key strength; while traditional ETL is constrained by on-premise infrastructure, the platform's cloud-native, modular architecture allows for elastic scaling to accommodate growing data volumes and complex urban analytics. This reduces operational costs and supports continuous, automated updates and integration with diverse data sources. Overall, the platform significantly enhances operational efficiency, data integrity, and decision-making speed for both public enterprises and SMEs, fostering a more responsive and sustainable Smart City ecosystem.

Despite significant progress, the literature identifies persistent challenges which our platform tries to solve: data standardization and interoperability remains a technical barrier as harmonizing data formats and integrating legacy systems is difficult [12]; privacy and compliance, particularly with blockchain-based systems, remains a critical concern [41]; and scalability and generalizability remain a concern as most integrated platforms are still at the pilot stage, with further research needed to validate scalability in larger or more complex urban environments [9].

The integrated platform described in the article advances the state-of-the-art by uniting digital twin, blockchain, and AI technologies into a cohesive, scalable, and participatory smart city solution. This approach addresses key limitations of traditional and current frameworks, as documented in recent scientific literature, and offers a replicable model for data-driven, transparent, and sustainable urban transformation.

4 Developing Digital Infrastructure for Smart Cities

4.1. Implementation of a Private Blockchain Network Using the Hyperledger Besu Platform

Implementation of a private blockchain network requires a thorough and multidisciplinary approach. Selecting the right platform, defining access rules, implementing security measures, and ongoing maintenance are key factors for successfully integrating blockchain technology into Smart City platform. This approach not only improves data security and transparency, but also allows organizations and companies to achieve additional value through automation and optimization of business processes.

We propose the deployment of an integrated digital twin, blockchain, and Artificial Intelligence (AI) platform in a smart City, involving city administration, waste management

companies, technology providers, and research institutions. This modular architecture enables the integration of additional city services (e.g., energy management, green mobility incentives) and supports open APIs for interoperability with other smart city solutions.

Three prominent platforms can be used to build a private blockchain network: Hyperledger Fabric, Hyperledger Besu, and Corda. Each of these platforms has unique characteristics and application areas, depending on business requirements, as shown in Table 3.

Table 3: The Comparison of Various Blockchain Platforms.

Feature	**Hyperledger Fabric**	**Hyperledger Besu**	**Corda**
Blockchain type	Permissioned	Permissioned or Public	Permissioned (Peer-to-peer)
Architecture	Modular, Layered	Ethereum compatible	Notary network using DHT for peer-to-peer transactions
Consensus	Kafka, Raft, Solo, BFT variants	IBFT, QBFT, PoA, PoW	Notary service
Smart contracts	Chaincode (Go, JavaScript, TypeScript)	Solidity (Ethereum EVM)	CorDapps (Kotlin, Java)
Transaction finalization	Very fast (no mining)	Depends on consensus mechanism (slower with PoW)	Fast
Privacy support	Private channels and Collection Policies	Private transactions via Tessera	Strictly enforced transaction privacy
Application area	Supply chain, logistics, healthcare	Enterprise Ethereum, fintech, banking	Banking, insurance, financial services
Interoperability	Supports interoperability with other blockchains	Compatible with Ethereum networks	Restricted interoperability

Hyperledger Fabric is a permissioned blockchain platform designed for business use, especially in industries that require high security, privacy and scalability [56]. Its biggest advantage is its modular architecture, enabling flexible network configuration, selection of consensus protocols and customization of data access rules. This platform is most often used in supply chains, logistics, healthcare and digital identity.

Hyperledger Besu is an EVM compatible blockchain client, making it an excellent choice for enterprises that want to use private networks but retain the ability to connect to the Ethereum blockchain [59]. Besu uses Solidity smart contracts, allowing development teams to port or reuse existing Ethereum applications, and supports IBFT, QBFT, PoA and PoW

consensus protocols [43,68]. It is most often used in the financial sector, fintech applications and private Ethereum networks.

Corda is a permissioned blockchain platform designed primarily for the financial sector and intercompany transactions [57]. Unlike Hyperledger Fabric and Besu, Corda uses a peer-to-peer model where transactions are shared only between involved parties, using a Notary service system rather than classic blockchain consensus.

In this paper we decided to use Hyperledger Besu as the private blockchain platform, based on several key advantages: EVM and Solidity compatibility allowing easy migration from existing Ethereum-based applications; flexible consensus mechanisms (IBFT, QBFT, PoA, PoW); the ability to operate in both public and private permissioned environments; private transactions via Tessera; and benefits from the mature Ethereum developer ecosystem.

While Hyperledger Fabric offers advanced modular architecture, comprehensive privacy controls, and high scalability, the choice of Hyperledger Besu was made due to its seamless integration with Ethereum standards, greater flexibility in consensus protocols, and the broad adoption of Ethereum technology in enterprise settings.

For the purpose of testing the proposed blockchain model, a private network was implemented using Hyperledger Besu deployed across 9 nodes, designed to balance scalability, fault tolerance, and realistic testing conditions. Four nodes operated as validators, participating in the consensus protocol, responsible for proposing, validating, and finalizing blocks, with each validator managed by a distinct organization. The remaining five nodes functioned as regular (non-validator) nodes, maintaining a copy of the blockchain, submitting transactions, interacting with smart contracts, and providing API endpoints for client applications. Besu's permission features were configured to restrict network participation to authorized nodes and accounts, with private transactions enabled via integration with a privacy manager such as Tessera.

The network infrastructure was provisioned with 9 workstations, each meeting or exceeding minimum hardware specifications: minimum 4 virtual CPUs, at least 16 GB RAM, and SSD drives for low latency and high I/O throughput. Necessary software components included Docker and Docker Compose for container orchestration, and the Node.js runtime environment. The network was configured by generating cryptographic key pairs for each node and defining node identities and permissions using Besu's permissioning configuration files. Communication between nodes was secured using TLS certificates. Smart contracts were developed in Solidity and tested using industry-standard tools such as Remix, Truffle, or Hardhat before deployment, validating the integrity and reliability of peer-to-peer communication, correctness and finality of transactions, system performance under simulated transaction loads, and fault tolerance to node failures or malicious activity.

To demonstrate the flexibility and cross-platform compatibility of the private blockchain network, the deployment utilized a heterogeneous environment: a subset of nodes was deployed on Windows OS inside Docker containers, another subset operated on native Linux

installations, and the remaining nodes ran within Docker containers on Linux hosts. This mixed deployment strategy showcased the ability of Hyperledger Besu to operate seamlessly across different operating systems and container orchestration environments.

4.2. Building the Foundation of a Smart City Digital Twin

To create a digital twin of a smart city, it is necessary to collect a wide range of information and data, creating the most accurate and functional digital replica of the physical urban environment. A digital twin of a city is not just a 2D/3D model, but a dynamic system that enables monitoring, simulation and decision-making in real time. Spatial and physical data about the city is needed, as well as historical and analytical data, and real-time data from sensors and other IoT devices, including traffic and environmental data.

In this paper, we present the beginnings of building a digital twin of a smart city. A database was created to collect a large amount of data about the city of Osijek. All geographic locations of house numbers in the city of Osijek were first stored in the database, displayed in a geocoordinate system. A web application was created to collect geo-locations using different online services for collecting geo-locations by searching for parameters such as the name of the city, street name and street number. The result is an .xml file with all geo-location data. To validate the collected data, geo-locations are plotted resulting in the contour of the city of Osijek.

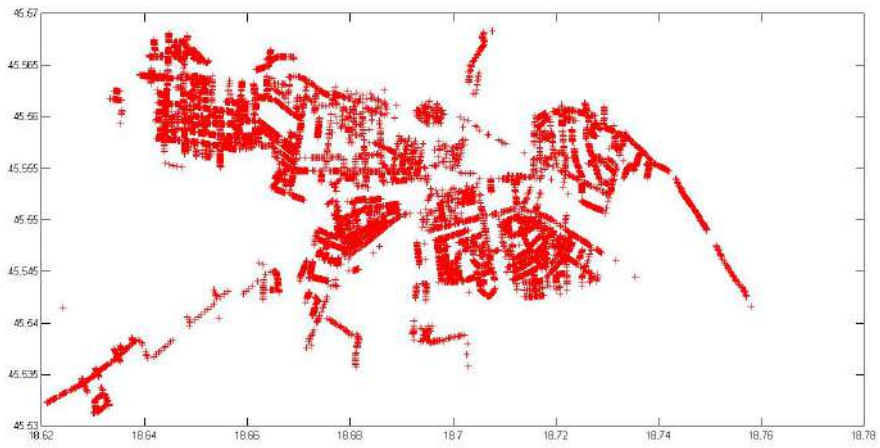

Figure 4. Geo-coordinates of the city of Osijek.

These geolocations need to be entered into the proposed blockchain. For each location, which represents the house number of a real-world building, it is possible to create an identity, and later add additional possibilities to that identity, such as invoices, construction permits, reporting of damage or malfunctions, etc., through smart contracts implemented in the blockchain network.

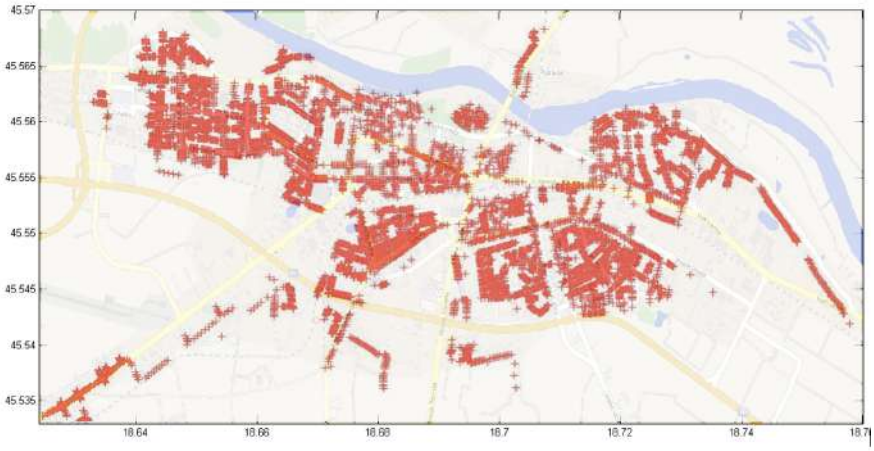

Figure 5. Geo-coordinates imposed on the map of the city of Osijek.

5 Evaluation and Future Work

The Smart City platform should be implemented in collaboration with a city administration, waste management companies, traffic companies, technology providers, research institutions and all other interested parties. The deployment focused on three main domains: digital twin integration, blockchain-based participatory governance, and AI-driven Smart City management.

Quantitative results from the implementation include AI-powered optimization of waste collection routes that could reduce operational costs by approximately 15% and decrease fuel consumption by 12%, as measured over a three-month period in the selected urban district [69,70]. Smart waste bins equipped with IoT sensors enabled more accurate forecasting of waste generation, reducing unnecessary collections by 18%. Environmental impact is gained through improved scheduling and route optimization contributing to an estimated 10% reduction in CO_2 emissions from waste collection vehicles. The Smart City platform also should facilitate up to a 9% increase in recycling rates, attributed to better monitoring and citizen engagement features. However, it should be noted that in the referenced case study [70], the reduction in unnecessary collections was reported as a 15% decrease in collection frequency, while the increase in recyclables collected reached 15%.

The blockchain-enabled participatory governance module is designed to increase citizen proposals, participation, trust, and willingness to engage, which should be confirmed by post-pilot surveys. Qualitative feedback highlighted increased transparency and trust in municipal services due to immutable blockchain records, greater citizen empowerment, and enhanced cross-departmental collaboration within the city administration. The digital twin component enabled real-time simulation of various urban scenarios, such as changes in waste generation patterns during public events or the impact of new recycling policies.

In future work the Smart City platform will be evaluated through a three-month pilot implementation in collaboration with city administration and waste management companies. Data will be collected from IoT-enabled waste bins, operational records of municipal services,

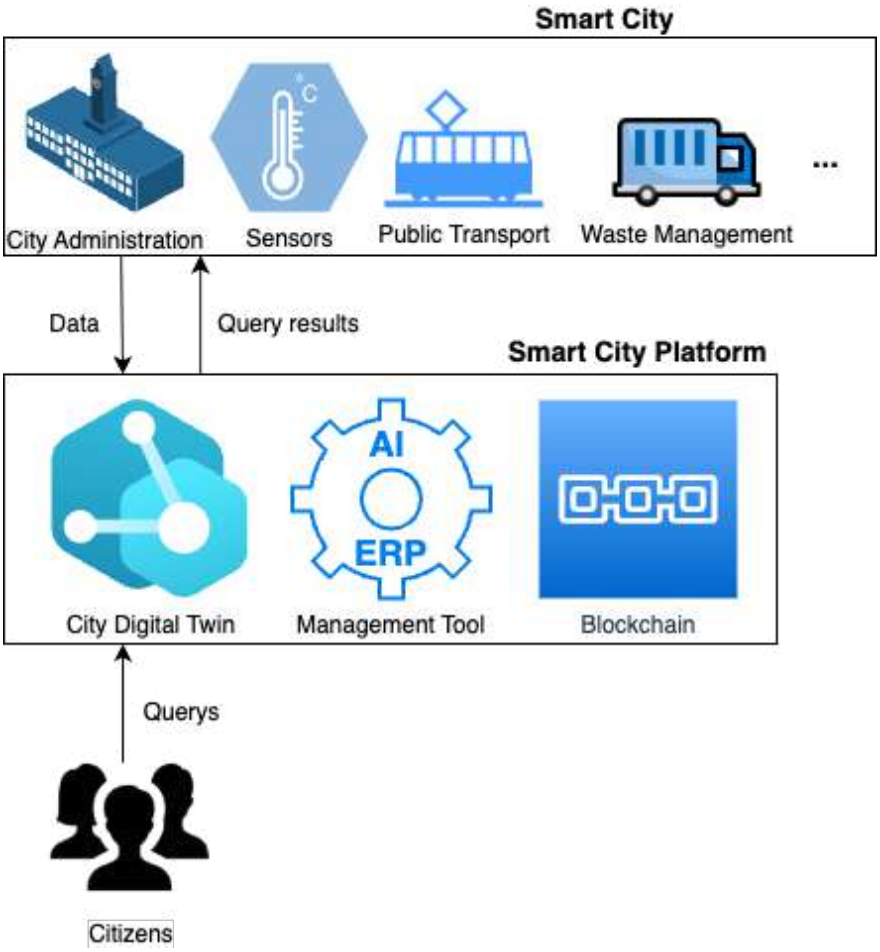

Figure 6. The proposed system architecture of Smart City platform.

blockchain transaction logs, and user participation modules. Key Performance Indicators (KPI) will be continuously monitored and visualized on public dashboards. The metrics will be standardized and KPIs defined in accordance with international frameworks such as the NIST Holistic KPI (H-KPI) Framework [71] and the digital twin evaluation framework [72]. Metrics included operational cost savings (%), reduction in fuel consumption (%), CO_2 emissions (tons/month), recycling rates (%), and citizen participation (number of proposals/votes).

Technical challenges to the proposed Smart City platform include integrating heterogeneous data sources (e.g., IoT devices, legacy systems, citizen apps), which requires significant effort in data normalization and semantic harmonization. Inconsistent data standards and incomplete datasets occasionally limited the accuracy of analytics and simulations. Ensuring the platform's scalability to support additional city services demanded a modular architecture and adherence to open standards. To maintain robust security across all layers, particularly for blockchain-based digital identities and sensitive urban data, continuous monitoring, regular audits, and compliance with GDPR and other regulations are required [73,74].

Organizational and regulatory challenges require strong coordination among city departments, public enterprises, and private partners. There is also a talent and skills gap, with the need for specialized skills in AI, IoT, blockchain, and data analytics within the city workforce, necessitating targeted training and capacity-building.

This paper shows that a Smart City integrated digital twin and blockchain platform can transform both public and private sector operations within the urban environment. The modularity and open standards of the proposed architecture ensure scalability and alignment with EU strategies such as the Green Deal and Digital Europe Programme.

5.1. Comprehensive User Survey

While the technical performance of the platform has been thoroughly evaluated, the platform is still in the development and pilot phase, and in future work a comprehensive survey of end users will be conducted. A user survey is planned to assess satisfaction, usability, and perceived impact of the platform among different stakeholder groups, covering topics such as ease of use, perceived transparency, trust in data security, and willingness to participate in digital governance modules. The target groups will include citizens, city staff, and representatives from municipal enterprises. Data will be collected anonymously and analyzed using standard statistical methods to identify trends and areas for improvement.

5.2. Platform Adaptability and Modularity

The platform's architecture is intentionally designed to support a wide range of urban environments, including smaller municipalities and cities with lower levels of digital maturity. This adaptability is achieved through modular implementation, interoperability and integration,

efficiency and cloud support. Cities can selectively implement platform components that align with their immediate needs and digital capabilities. The use of open APIs and standardized data formats facilitates seamless integration with existing legacy systems and third-party solutions. The platform supports cloud-based deployment and modular architecture, allowing cities with limited IT resources to leverage scalable solutions. Implementation plans include comprehensive training and technical support for local staff, as well as phased roll-out strategies. Smaller cities can leverage national and EU funding opportunities aimed at supporting digital transformation and smart city initiatives.

In future work we plan to pilot the proposed model in a single city and later apply given results in larger or more complex urban environments. We also plan platform scaling to additional Smart City services (e.g., energy management, water distribution) and enhancing interoperability with national and EU-level Smart City data ecosystems.

6 Conclusions

This work demonstrates that integrating digital twin technology, advanced machine learning, and private blockchain network provides a robust and scalable foundation for Smart City transformation. The proposed platform enables real-time Smart City data management, transparent participatory governance, and AI-powered optimization of public services, resulting in measurable improvements in operational efficiency, environmental impact, and citizen engagement. Pilot implementation showed significant reductions in operational costs and emissions, as well as increased recycling rates and community participation. The modular, interoperable architecture supports both public enterprises and companies, ensuring adaptability to diverse urban environments and alignment with EU digital and sustainability strategies. While challenges remain in data integration, stakeholder coordination, and regulatory compliance, the proposed platform offers a replicable pathway for cities seeking to enhance transparency, trust, and sustainability through data-driven innovation. Future work will focus on scaling the platform to additional city services and expanding participatory features to further empower Smart City communities.

Despite the demonstrated benefits of the integrated Smart City platform, it is important to acknowledge several limitations of this study. First, the pilot implementation will be conducted in a single urban district and over a limited time period, which may affect the generalizability of the results to larger or more diverse urban environments. Second, challenges related to data integration, interoperability, and legacy system compatibility remain significant barriers to full-scale deployment. Third, the platform's scalability and adaptability to different regulatory frameworks and organizational cultures require further validation in real-world settings. Finally, the current evaluation relies on a relatively small set of key performance indicators and may not fully capture all aspects of stakeholder satisfaction

or long-term sustainability. As part of future work, during the project implementation, it is planned to conduct further studies and pilot implementations in larger and more complex urban environments to validate and extend the findings of this research.

Author Contributions: Conceptualization, I.L.; methodology, I.L., M.K. and M.Š.; software, M.Š. and Z.K.; validation, I.L., M.K. and Z.K.; formal analysis, I.L.; investigation, M.Š.; resources, I.L.; data curation, Z.K. and M.Š.; writing—original draft preparation, I.L.; writing—review and editing, I.L.; visualization, M.K.; supervision, I.L.; project administration, I.L.; funding acquisition, I.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded and is a part of research for the activities to be carried out for the project "Researching advanced algorithms and innovative business intelligence solutions in the cloud—NPOO.C3.2.R3-I1.04.0128".

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References (as published in the original article)

1. Adewusi, A.O.; Okoli, U.I.; Adaga, E.; Olorunsogo, T.; Asuzu, O.F.; Daraojimba, D.O. Business Intelligence in the Era of Big Data: A Review of Analytical Tools and Competitive Advantage. *Comput. Sci. IT Res. J.* **2024**, 5, 415–431.

2. Hansen, M.M.; Koonsanit, K.; Kulmala, V. How Can Data Contribute to Smart City Innovation: A Study from Thailand's Smart City Initiatives. *Front. Sustain. Cities* **2025**, 6, 1473123.

3. Gudfinnsson, K.; Strand, M. Challenges with BI Adoption in SMEs. In *Proc. 2017 8th Int. Conf. Information, Intelligence, Systems & Applications*; IEEE: Larnaca, 2017; pp. 1–6.

4. Nenzhelele, T.E.; Pellissier, R. Competitive Intelligence Implementation Challenges of Small and Medium-Sized Enterprises. *Mediterr. J. Soc. Sci.* **2014**.

5. Alsibhawi, I.A.A.; Yahaya, J.B.; Mohamed, H.B. Business Intelligence Adoption for Small and Medium Enterprises: Conceptual Framework. *Appl. Sci.* **2023**, 13, 4121.

6. Tawil, A.-R.H.; Mohamed, M.; Schmoor, X.; Vlachos, K.; Haidar, D. Trends and Challenges towards Effective Data-Driven Decision Making in UK SMEs. *Big Data Cogn. Comput.* **2024**, 8, 79.

7. Rizzatto, M.; L'Erario, A.; Maganha De Almeida, E. Prototyping Smart City Solutions with Metaverse and Digital Twins. In *Proc. 27th Int. Conf. Enterprise Information Systems*; SCITEPRESS: Porto, 2025; pp. 193–200.

8. Huang, J.; Bibri, S.E.; Keel, P. Generative Spatial Artificial Intelligence for Sustainable Smart Cities. *Environ. Sci. Ecotechnology* **2025**, 24, 100526.

9. Bibri, S.E.; Huang, J.; Jagatheesaperumal, S.K.; Krogstie, J. The Synergistic Interplay of Artificial Intelligence and Digital Twin in Environmentally Planning Sustainable Smart Cities. *Environ. Sci. Ecotechnology* **2024**, 20, 100433.

10. Gkontzis, A.F.; Kotsiantis, S.; Feretzakis, G.; Verykios, V.S. Enhancing Urban Resilience: Smart City Data Analyses, Forecasts, and Digital Twin Techniques. *Future Internet* **2024**, 16, 47.

11. Aghazadeh Ardebili, A.; Zappatore, M.; Ramadan, A.I.H.A.; Longo, A.; Ficarella, A. Digital Twins of Smart Energy Systems. *Energy Inform.* **2024**, 7, 94.

12. Regueiro, C.; Seco, I.; Gutiérrez-Agüero, I.; Urquizu, B.; Mansell, J. A Blockchain-Based Audit Trail Mechanism: Design and Implementation. *Algorithms* **2021**, 14, 341.

13. Tavares, J.M.R.S.; Karri, C.; Machado, J.J.M.; Jain, D.K.; Dannana, S.; Gottapu, S.K.; Gandomi, A.H. Recent Technology Advancements in Smart City Management. *Comput. Mater. Contin.* **2024**, 81, 3617–3663.

14. Biasin, M.; Delle Foglie, A. Blockchain and Smart Cities for Inclusive and Sustainable Communities. *Sustainability* **2024**, 16, 6669.

15. Huzzat, A.; Anpalagan, A.; Khwaja, A.S.; Woungang, I.; Alnoman, A.A.; Pillai, A.S. A Comprehensive Review of Digital Twin Technologies in Smart Cities. *Digit. Eng.* **2025**, 4, 100040.

16. Mondal, K.C.; Biswas, N.; Saha, S. Role of Machine Learning in ETL Automation. In *Proc. 21st Int. Conf. Distributed Computing and Networking*; ACM: Kolkata, 2020; pp. 1–6.

17. Mantri, A. Advanced ML Techniques for Optimizing ETL Workflows with Apache Spark and Snowflake. *J. Artif. Intell. Cloud Comput.* **2023**, 2, 1–6.

18. Dinesh, L.; Devi, K.G. An Efficient Hybrid Optimization of ETL Process in Data Warehouse of Cloud Architecture. *J. Cloud Comput.* **2024**, 13, 12.

19. Souibgui, M.; Atigui, F.; Zammali, S.; Cherfi, S.; Yahia, S.B. Data Quality in ETL Process: A Preliminary Study. *Procedia Comput. Sci.* **2019**, 159, 676–687.

20. Zdravevski, E.; Apanowicz, C.; Stencel, K.; Slezak, D. Scalable Cloud-Based ETL for Self-Serving Analytics. In *Advances in Data Mining*; Springer: Cham, 2019; pp. 387–394.

21. Singh, M.M. Extraction Transformation and Loading (ETL) of Data Using ETL Tools. *Int. J. Res. Appl. Sci. Eng. Technol.* **2022**, 10, 4415–4420.

22. Mayo, C.; Kessler, M.L.; Feng, M.; Weyburn, G.; El Naqa, I.; Eisbruch, A.; Matuszak, M.M.; McShan, D.; Moran, J.M.; Anderson, C.J.; et al. Taming Big Data: Implementation of a Clinical Use-Case Driven Architecture. *Int. J. Radiat. Oncol.* **2016**, 96, E417–E418.

23. Santos, V.; Belo, O. Using Relational Algebra on the Specification of Real World ETL Processes. In *Proc. 2015 IEEE Int. Conf. Computer and Information Technology*; IEEE: Liverpool, 2015; pp. 861–866.

24. Zhang, X.; Wen, S.; Yan, L.; Feng, J.; Xia, Y. A Hybrid-Convolution Spatial–Temporal Recurrent Network For Traffic Flow Prediction. *Comput. J.* **2024**, 67, 236–252.

25. Li, B.; Zhou, X.; Ning, Z.; Guan, X.; Yiu, K.-F.C. Dynamic Event-Triggered Security Control for Networked Control Systems with Cyber-Attacks. *Inf. Sci.* **2022**, 612, 384–398.

26. Kim, S.-S.; Lee, W.-R.; Go, J.-H. A Study on Utilization of Spatial Information in Heterogeneous System Based on Apache NiFi. In *Proc. 2019 Int. Conf. Information and Communication Technology Convergence*; IEEE: Jeju Island, 2019; pp. 1117–1119.

27. Meng, L.; McWilliams, B.; Jarosinski, W.; Park, H.-Y.; Jung, Y.-G.; Lee, J.; Zhang, J. Machine Learning in Additive Manufacturing: A Review. *JOM* **2020**, 72, 2363–2377.

28. Biswas, N.; Biswas, S.; Mondal, K.C.; Maiti, S. Challenges and Solutions of Real-Time Data Integration Techniques by ETL Application. In *Advances in Business Information Systems and Analytics*; IGI Global, 2024; pp. 348–371.

29. Castellanos, M.; Simitsis, A.; Wilkinson, K.; Dayal, U. Automating the Loading of Business Process Data Warehouses. In *Proc. 12th Int. Conf. Extending Database Technology*; ACM: Saint Petersburg, 2009; pp. 612–623.

30. Meng, L.; McWilliams, B.; Jarosinski, W.; Park, H.-Y.; Jung, Y.-G.; Lee, J.; Zhang, J. Machine Learning in Additive Manufacturing: A Review. *JOM* **2020**, 72, 2363–2377.

31. Tawil, A.-R.; Mohamed, M.; Schmoor, X.; Vlachos, K.; Haidar, D. Trends and Challenges Towards an Effective Data-Driven Decision Making in UK SMEs. 2023.

32. Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*; bitcoin.org, 2008.

33. Salzano, F.; Marchesi, L.; Pareschi, R.; Tonelli, R. Integrating Blockchain Technology within an Information Ecosystem. *Blockchain Res. Appl.* **2024**, 5, 100225.

34. Crosby, M.; Pattanayak, P.; Verma, S.; Kalyanaraman, V. Blockchain Technology: Beyond Bitcoin. **2016**, 2, 6–19.

35. Dinh, T.T.A.; Wang, J.; Chen, G.; Liu, R.; Ooi, B.C.; Tan, K.-L. BLOCKBENCH: A Framework for Analyzing Private Blockchains. In *Proc. 2017 ACM Int. Conf. Management of Data*; ACM: Chicago, 2017; pp. 1085–1100.

36. Dinh, T.T.A.; Liu, R.; Zhang, M.; Chen, G.; Ooi, B.C.; Wang, J. Untangling Blockchain: A Data Processing View of Blockchain Systems. *IEEE Trans. Knowl. Data Eng.* **2018**, 30, 1366–1385.

37. Francisco, K.; Swanson, D. The Supply Chain Has No Clothes: Technology Adoption of Blockchain for Supply Chain Transparency. *Logistics* **2018**, 2, 2.

38. Marijan, D.; Lal, C. Blockchain Verification and Validation: Techniques, Challenges, and Research Directions. *Comput. Sci. Rev.* **2022**, 45, 100492.

39. Ullah, F.; He, J.; Zhu, N.; Wajahat, A.; Nazir, A.; Qureshi, S.; Pathan, M.S.; Dev, S. Blockchain-Enabled EHR Access Auditing. *Heliyon* **2024**, 10, e34407.

40. Drummer, D.; Neumann, D. Is Code Law? Current Legal and Technical Adoption Issues and Remedies for Blockchain-Enabled Smart Contracts. *J. Inf. Technol.* **2020**, 35, 337–360.

41. Asif, M.; Aziz, Z.; Bin Ahmad, M.; Khalid, A.; Waris, H.A.; Gilani, A. Blockchain-Based Authentication and Trust Management Mechanism for Smart Cities. *Sensors* **2022**, 22, 2604.

42. Mollajafari, S.; Bechkoum, K. Blockchain Technology and Related Security Risks: Towards a Seven-Layer Perspective and Taxonomy. *Sustainability* **2023**, 15, 13401.

43. Puthal, D.; Mohanty, S.P. Proof of Authentication: IoT-Friendly Blockchains. *IEEE Potentials* **2019**, 38, 26–29.

44. Alkhateeb, A.; Catal, C.; Kar, G.; Mishra, A. Hybrid Blockchain Platforms for the Internet of Things (IoT): A Systematic Literature Review. *Sensors* **2022**, 22, 1304.

45. Le, T.-V.; Hsu, C.-L.; Chen, W.-X. A Hybrid Blockchain-Based Log Management Scheme With Nonrepudiation for Smart Grids. *IEEE Trans. Ind. Inform.* **2022**, 18, 5771–5782.

46. Gao, H.; Kou, G.; Liang, H.; Zhang, H.; Chao, X.; Li, C.-C.; Dong, Y. Machine Learning in Business and Finance: A Literature Review and Research Opportunities. *Financ. Innov.* **2024**, 10, 86.

47. Javaid, M.; Haleem, A.; Singh, R.P.; Suman, R.; Khan, S. A Review of Blockchain Technology Applications for Financial Services. *BenchCouncil Trans. Benchmarks Stand. Eval.* **2022**, 2, 100073.

48. Sharma, J. Blockchain Technology Adoption in Financial Services: Opportunities and Challenges. In *Advances in Finance, Accounting, and Economics*; IGI Global, 2023; pp. 99–117.

49. Khezr, S.; Moniruzzaman, M.; Yassine, A.; Benlamri, R. Blockchain Technology in Healthcare: A Comprehensive Review and Directions for Future Research. *Appl. Sci.* **2019**, 9, 1736.

50. Xu, P.; Lee, J.; Barth, J.R.; Richey, R.G. Blockchain as Supply Chain Technology: Considering Transparency and Security. *Int. J. Phys. Distrib. Logist. Manag.* **2021**, 51, 305–324.

51. Sengar, R. The Changing Business Prospective with Cloud Computing: A Review. *SSRN Electron. J.* **2023**.

52. Dziembek, D.; Ziora, L. Cloud-Based Business Intelligence Solutions in the Management of Polish Companies. In *Advances in Information Systems Development*; Springer: Cham, 2023; Vol. 63, pp. 35–52.

53. Rane, N.L.; Paramesha, M.; Choudhary, S.P.; Rane, J. Artificial Intelligence, Machine Learning, and Deep Learning for Advanced Business Strategies: A Review. **2024**.

54. González-Varona, J.; López-Paredes, A.; Poza, D.; Acebes, F. Building and Development of an Organizational Competence for Digital Transformation in SMEs. *J. Ind. Eng. Manag.* **2021**, 14, 15.

55. Tern, S. Survey of Smart Contract Technology and Application Based on Blockchain. *Open J. Appl. Sci.* **2021**, 11, 1135–1148.

56. Androulaki, E.; Barger, A.; Bortnikov, V.; Cachin, C.; Christidis, K.; De Caro, A.; Enyeart, D.; Ferris, C.; Laventman, G.; Manevich, Y.; et al. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains. In *Proc. 13th EuroSys Conf.*; ACM: Porto, 2018; pp. 1–15.

57. Han, H.; Shiwakoti, R.K.; Chen, W. Unlocking Enterprise Blockchain Adoption: A R3 Corda Case Study. *J. Gen. Manag.* **2024**.

58. Baliga, A.; Subhod, I.; Kamat, P.; Chatterjee, S. Performance Evaluation of the Quorum Blockchain Platform. 2018.

59. Shahrukh, Md.R.H.; Rahman, Md.T.; Mansoor, N. Exploration of Hyperledger Besu in Designing Private Blockchain-Based Financial Distribution Systems. 2023.

60. Švarcmajer, M.; Ivanović, D.; Rudec, T.; Lukić, I. Application of Graph Theory and Variants of Greedy Graph Coloring Algorithms for Optimization of Distributed Peer-to-Peer Blockchain Networks. *Technologies* **2025**, 13, 33.

61. Serena, L.; D'Angelo, G.; Ferretti, S. Security Analysis of Distributed Ledgers and Blockchains through Agent-Based Simulation. 2021.

62. Fedotov, I.; Khritankov, A. Statistical Model Checking of Common Attack Scenarios on Blockchain. **2021**.

63. Zohar, A. Bitcoin: Under the Hood. *Commun. ACM* **2015**, 58, 104–113.

64. Vassiliadis, P. A Survey of Extract–Transform–Load Technology. *Int. J. Data Warehous. Min.* **2009**, 5, 1–27.

65. Herath, H.M.K.K.M.B.; Mittal, M. Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review. *Int. J. Inf. Manag. Data Insights* **2022**, 2, 100076.

66. Mrabet, M.; Sliti, M. Integrating Machine Learning for the Sustainable Development of Smart Cities. *Front. Sustain. Cities* **2024**, 6, 1449404.

67. Dani, A.A.H.; Supangkat, S.H.; Lubis, F.F.; Nugraha, I.G.B.B.; Kinanda, R.; Rizkia, I. Development of a Smart City Platform Based on Digital Twin Technology for Monitoring and Supporting Decision-Making. *Sustainability* **2023**, 15, 14002.

68. Oyinloye, D.P.; Teh, J.S.; Jamil, N.; Alawida, M. Blockchain Consensus: An Overview of Alternative Protocols. *Symmetry* **2021**, 13, 1363.

69. Addas, A.; Khan, M.N.; Naseer, F. Waste Management 2.0 Leveraging Internet of Things for an Efficient and Eco-Friendly Smart City Solution. *PLOS ONE* **2024**, 19, e0307608.

70. Kuzhin, M.F.; Joshi, A.; Mittal, V.; Khatkar, M.; Guven, U. Optimizing Waste Management through IoT and Analytics. *BIO Web Conf.* **2024**, 86, 01090.

71. Serrano, M.; Griffor, E.; Wollman, D.; Dunaway, M.; Burns, M.; Rhee, S.; Greer, C. *Smart Cities and Communities: A Key Performance Indicators Framework*; NIST: Gaithersburg, MD, 2022.

72. Tang, Z.; Zhuang, D.; Zhang, J. Evaluation Framework for Domain-Specific Digital Twin Platforms. *Sci. Rep.* **2025**, 15, 10544.

73. Alamri, B.; Crowley, K.; Richardson, I. Cybersecurity Risk Management Framework for Blockchain Identity Management Systems in Health IoT. *Sensors* **2022**, 23, 218.

74. Zafar, A. Reconciling Blockchain Technology and Data Protection Laws. *J. Cybersecurity* **2025**, 11, tyaf002.

F

Paper F: Designing Post-Quantum Secure Blockchain Infrastructure for Business Intelligence Systems: Migration Strategies and Cost Optimization

Proc. 2025 International Symposium ELMAR, pp. 245–248, IEEE, September 2025 | DOI: 10.1109/ELMAR66948.2025.11193746

Miljenko Švarcmajer, Tea Krčmar, Dina Šabanović, Ivica Lukić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Osijek, Croatia

Abstract

The growing reliance on cloud-based business intelligence systems has increased the demand for secure communication and verifiable data integrity. However, the emergence of quantum computing poses a significant threat to current cryptographic protocols, as widely adopted encryption schemes may become vulnerable to quantum-enabled attacks. In this paper, the vulnerabilities introduced by quantum advancements in contemporary business intelligence systems are examined. Recent approaches in the literature are evaluated, and a hybrid

security model is proposed that integrates quantum-resistant cryptographic components — such as post-quantum digital signatures and Post-Quantum Delegated Proof of Ledger (PQ-DPoL) consensus. A security model based on permissioned blockchain architecture is presented as a practical and scalable solution for long-term resilience, offering enhanced protection in anticipation of future quantum capabilities.

Keywords: Blockchain; Business Systems; Consensus Protocols; System Security; Quantum Computing

I. Introduction

Business systems increasingly rely on digital technology for transaction management, data protection and operational security. In a modern business environment, data security and integrity are essential for financial institutions, healthcare systems, supply chains and other sectors that rely on digital networks for the transmission and storage of sensitive information. In this context, the blockchain has emerged as a key technology to ensure the secure recording and management of transactions in a distributed network.

A key component of blockchain's security framework is elliptic-curve cryptography (ECC) or Rivest-Shamir-Adleman (RSA), both of which are asymmetric cryptographic methods used for securing transactions through encryption and digital signatures. ECC is known for its efficiency and relatively small key size, while RSA relies on the difficulty of factoring large integers.

However, rapid advancement in quantum computing has been identified as a significant threat to classical cryptographic algorithms that secure modern business systems. ECC encryption can be efficiently compromised by Shor's algorithm due to its lower key size and Shor's utilisation of quantum parallelism, while RSA suffers the same fate due to Shor's drastic reduction in computational effort required for factorisation. To mitigate these risks, the development and implementation of post-quantum security mechanisms have been deemed necessary.

II. Quantum Threats and Future Development of Quantum Computers

Quantum computing is a rapidly advancing field that uses the principles of quantum mechanics to solve problems that are potentially beyond the capabilities of classical computers. Unlike classical computing, where data is encoded in bits that can exist as either 0 or 1, quantum computing relies on qubits, which can exist in a superposition of both states simultaneously.

Quantum supremacy was claimed in 2019 by Google, as its 53-qubit Sycamore processor was reported to have completed a specific computational task in approximately 200 seconds

that would take 10,000 years on a classical supercomputer. Quantum computing is currently in the Noisy Intermediate-Scale Quantum (NISQ) era, where processors are still constrained by noise and high error rates.

The cryptographic security of blockchain networks is seriously threatened by the development of quantum computing. One major concern is that public keys, which are derived from private keys, are exposed after a transaction is broadcast but before it is confirmed in a block. During the waiting period (typically around 10 minutes for Bitcoin), a quantum computer could potentially derive the private key from the exposed public key.

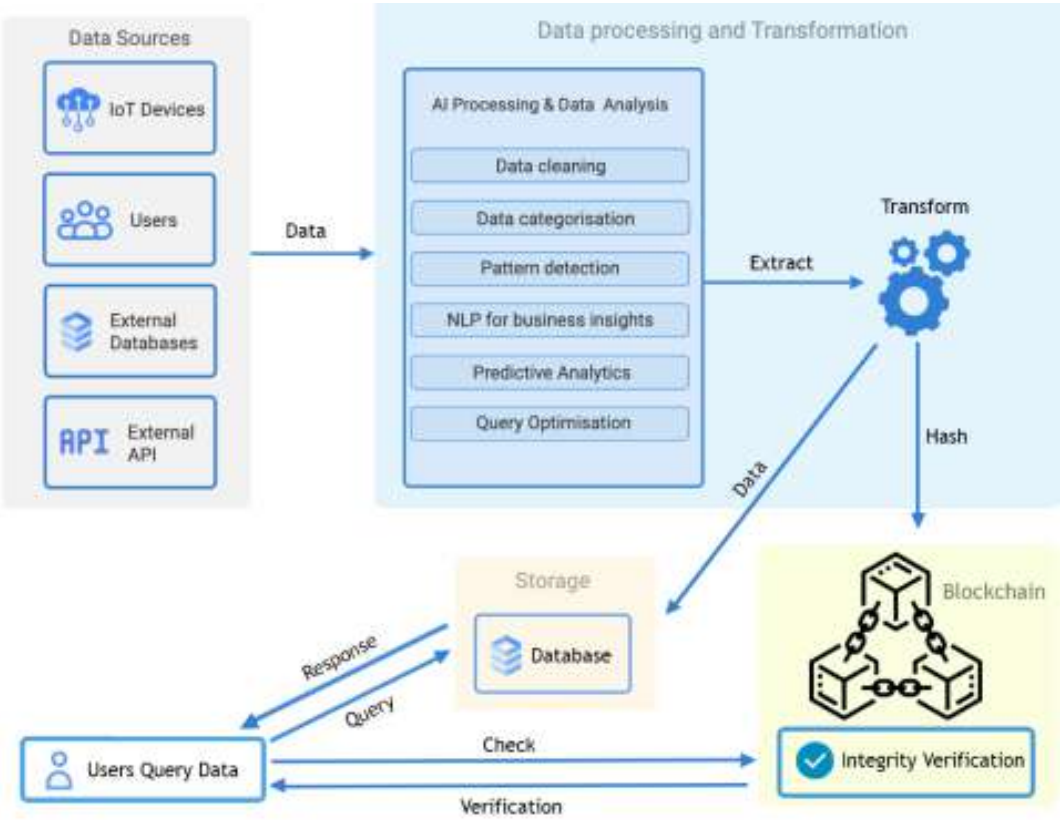

Figure 1. Detailed flow of information in a business system with blockchain.

III. Enhancing System Security with a Post-Quantum Blockchain System

In a business system secured by blockchain technology, data is sourced from Internet of Things (IoT) devices, users, external databases and external APIs. Security enhancements include the implementation of advanced cryptographic methods to ensure protection against future quantum computing attacks:

- Post-quantum digital signatures, such as Dilithium and SPHINCS+, provide security that is resistant to quantum computers, ensuring data authenticity and integrity.

- Post-quantum encryption, such as Kyber and NTRU, enable secure communication and data storage even in the presence of quantum computers.
- A quantum-secure consensus mechanism, PQ-DPoL, ensures network reliability and decentralization.
- The use of SHA-3 as a hash function further increases security by making it impervious to the Grover attack.

Permissioned blockchain offers additional benefits in this context, as it allows for easier adaptation without significant costs and provides greater control over the algorithms and security protocols used.

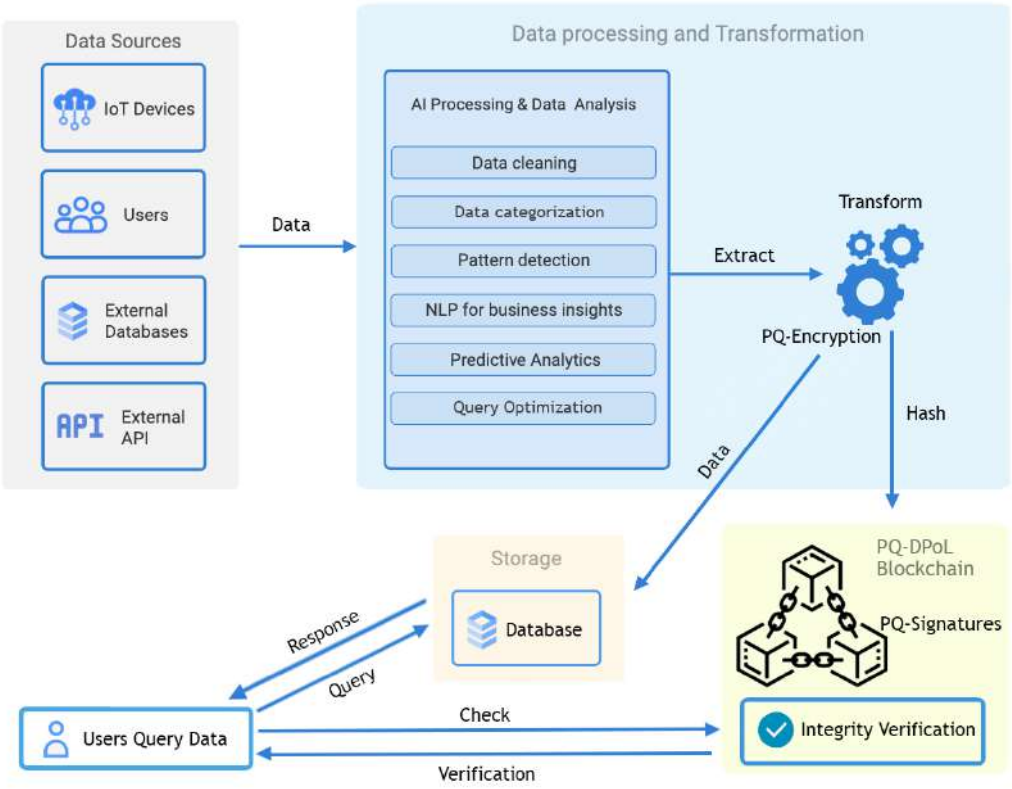

Figure 2. Architecture of a post-quantum secure business intelligence system using blockchain and classical storage.

IV. Cost-Efficient Implementation of Post-Quantum Security

The migration to post-quantum cryptographic (PQC) security is considered a complex and high-cost process that requires careful financial planning, strategic deployment, and long-term adaptation to future security needs. Based on prior cryptographic transitions, full-scale implementation is expected to span at least five years.

One of the primary technical challenges is hardware compatibility. Post-quantum encryption algorithms demand greater computational resources, often exceeding the capabilities of existing cryptographic processors and HSMs. For organizations with limited resources, cost-reduction strategies play a critical role. A hybrid cryptographic model — where classical and post-quantum encryption operate concurrently — has been proposed as a transitional framework. This model enables incremental implementation, maintaining backward compatibility while limiting initial investment.

To maintain security over the next four decades, a structured and proactive approach to PQC migration is recommended: comprehensive inventory of cryptographic assets, adoption of NIST-standardized PQC algorithms, emphasis on cryptographic agility, and continuous monitoring of advancements in quantum computing and cryptanalysis.

V. Conclusion and Future Work

Post-quantum security is increasingly recognized as an essential component of future-resilient system architectures. The use of permissioned blockchain infrastructure, supported by PQ-DPoL consensus mechanisms and post-quantum digital signatures, has been identified as an effective approach to ensuring long-term security and system resilience. Organizations that adapt their infrastructures to incorporate post-quantum safeguards at this stage are expected to gain a strategic advantage in future digital ecosystems.

In subsequent work, comprehensive performance evaluations of PQ-DPoL will be conducted within permissioned blockchain networks in order to assess operational efficiency, latency, and scalability under realistic conditions. Additionally, hybrid models combining classical and post-quantum cryptographic methods will be examined to determine the most effective balance between security, performance, and deployment feasibility in diverse enterprise environments.

Acknowledgment

This work was conducted as part of the project Research of Advanced Algorithms and Solutions for Innovative Business Intelligence in the Cloud — NPOO.C3.2.R3-I1.04.0128.

References

1. H. Taherdoost and M. Madanchian, "Blockchain-based new business models: A systematic review," Electronics, vol. 12, p. 1479, 03 2023.

2. V. Koolwal, S. Kumar, and K. Mohbey, The Role of Blockchain Technology to Make Business Easier and Effective The Role of Blockchain Technology to Make Business

Easier and Effective. 11 2019.

3. A. Arabiun and A. Moghadasi, "The application of blockchain in businesses," 01 2024.

4. X.-J. Wen, Y.-Z. Chen, X.-C. Fan, W. Zhang, Z.-Z. Yi, and J.-B. Fang, "Blockchain consensus mechanism based on quantum zero-knowledge proof," Optics Laser Technology, vol. 147, p. 107693, 03 2022.

5. Y. Long, Y. Gong, W. Huang, J. Cai, N. Xu, and K.-C. Li, Cryptography of Blockchain, pp. 340–349. 03 2023.

6. M. Kumar, "Post-quantum cryptography algorithm's standardization and performance analysis," Array, vol. 15, p. 100242, 08 2022.

7. H. Kim, W. Kim, Y. L. Kang, H. Kim, and H. Seo, "Post-quantum delegated proof of luck for blockchain consensus algorithm," Applied Sciences, 2024.

8. S. Singh and E. Sakk, "Implementation and analysis of shor's algorithm to break rsa cryptosystem security," 01 2024.

9. M. Vidaković and M. Kruno, "Performance and applicability of post-quantum digital signature algorithms in resource-constrained environments," Algorithms, vol. 16, p. 518, 11 2023.

10. P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in Proceedings 35th Annual Symposium on Foundations of Computer Science, pp. 124–134, 1994

11. A. Adedoyin and O. Oladele, "Quantum computing and its implications for data security laws," 01 2025.

12. L. K. Grover, "A fast quantum mechanical algorithm for database search," in Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, STOC '96, (New York, NY, USA), p. 212–219, Association for Computing Machinery, 1996.

13. W. Khan, Q. Arshad, M. Raza, and M. Imran, "Quantum cryptography a real threat to classical blockchain: Requirements and challenges," 10 2022.

14. H. Gharavi, J. Granjal, and E. Monteiro, "Post-quantum blockchain security for the internet of things: Survey and research directions," IEEE Communications Surveys Tutorials, vol. 26, no. 3, pp. 1748–1774, 2024.

15. H. A. Bhat, F. A. Khanday, B. K. Kaushik, F. Bashir, and K. A. Shah, "Quantum computing: Fundamentals, implementations and applications," IEEE Open Journal of Nanotechnology, vol. 3, pp. 61–77, 2022.

16. J. Preskill, “Quantum computing and the entanglement frontier,” 2012.

17. F. Arute, K. Arya, R. Babbush, D. Bacon, and et al., “Quantum supremacy using a programmable superconducting processor,” Nature, vol. 574, pp. 505–510, Oct. 2019.

18. J. Preskill, “Quantum computing in the nisq era and beyond,” Quantum, vol. 2, p. 79, Aug. 2018.

19. Google Quantum AI and Collaborators, “Quantum error correction below the surface code threshold,” Nature, vol. 638, pp. 920–926, Feb. 2025

20. M. AbuGhanem, “Ibm quantum computers: Evolution, performance, and future directions,” 2024.

21. T. M. Fernández-Caramès and P. Fraga-Lamas, “Towards post-quantum blockchain: A review on blockchain cryptography resistant to quantum computing attacks,” IEEE Access, vol. 8, pp. 21091–21116, 2020.

22. D. A. Bard, J. J. Kearney, and C. A. Perez-Delgado, “Quantum advantage on proof of work,” Array, vol. 15, p. 100225, Sept. 2022

23. A. Aydeger, E. Zeydan, A. K. Yadav, K. T. Hemachandra, and M. Liyanage, “Towards a quantum-resilient future: Strategies for transitioning to post-quantum cryptography,” in 2024 15th International Conference on Network of the Future (NoF), pp. 195–203, 2024.

24. M. Webber, V. Elfving, S. Weidt, and W. K. Hensinger, “The impact of hardware specifications on reaching quantum advantage in the fault tolerant regime,” AVS Quantum Science, vol. 4, p. 013801, Mar. 2022.

25. L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, and D. Stehlé, “Crystals-dilithium: A lattice-based digital signature scheme,” IACR Transactions on Cryptographic Hardware and Embedded Systems, vol. 2018, p. 238–268, Feb. 2018.

26. D. Soni, K. Basu, M. Nabeel, N. Aaraj, M. Manzano, and R. Karri, SPHINCS+, pp. 141–162. 10 2020

27. J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. Schanck, P. Schwabe, G. Seiler, and D. Stehle, “Crystals - kyber: A cca-secure module-lattice-based kem,” pp. 353–367, 04 2018.

28. J. Hoffstein, J. Pipher, and J. H. Silverman, “Ntru: A ring-based public key cryptosystem,” in International Workshop on Ant Colony Optimization and Swarm Intelligence, 1998.

29. A. Bansal, Arju, Esha, and P. Mehra, "A post-quantum consortium blockchain based secure ehr framework," pp. 1–6, 06 2023.

30. J. Señor, J. Portilla, and G. Mujica, "Analysis of the ntru post-quantum cryptographic scheme in constrained iot edge devices," IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18778–18790, 2022.

31. N. Chandran and E. Manuel, "Performance analysis of modified sha-3," Procedia Technology, vol. 24, pp. 904–910, 12 2016.

32. C. C. San Jose, "Comparative and security performance analysis of sha3," 11 2019.

33. K. F. Hasan, L. Simpson, M. A. R. Baee, C. Islam, Z. Rahman, and et al., "A framework for migrating to post-quantum cryptography: Security dependency analysis and case studies," IEEE Access, vol. 12, pp. 23427– 23450, 2024.

34. A. Amadori, T. Attema, M. Bombar, J. D. Duarte, V. Dunning, S. Etinski, and et al., The PQC Migration Handbook: Guidelines for Migrating to Post-Quantum Cryptography. revised and extended second edition ed., December 2024.

Curriculum vitae

Miljenko Švarcmajer was born on 28 March 1987 in Osijek, Croatia. He received the B.Sc. and M.Sc. degrees in computer science from the Josip Juraj Strossmayer University of Osijek in 2010 and 2013, respectively. His engagement with computing networks dates to his secondary school education, establishing an early and sustained interest that would define his academic and professional trajectory.

Following his graduate studies, he pursued a career in the private sector as ICT Manager for a group of companies, overseeing enterprise computing infrastructure. After seven years in industry, he returned to academia with a research focus on blockchain technology and cybersecurity, bringing a practitioner's perspective on the real-world challenges of distributed computing systems.

He is currently pursuing the Ph.D. degree in computer science at the Faculty of Electrical Engineering, Computer Science and Information Technology (FERIT), Josip Juraj Strossmayer University of Osijek, where he is also employed as a Research and Teaching Assistant. He is a collaborator on a NextGenerationEU-funded research project investigating advanced algorithms and cloud-based business intelligence solutions (NPOO.C3.2.R3-I1.04.0128, 2024–2026), within the scope of which he conducted research visits to companies in the United States and participated in an international summit in London. He has presented his research at conferences and symposia across Croatia and the region, and participated in academic study visits to Hungary and Serbia.

He is the author or co-author of 14 scientific publications, including 6 papers in Q1 journals. His work has been recognised with the HIKS Award for Outstanding Scientific Work (2025), the FERIT Dean's Recognition for Outstanding Achievement in Doctoral Studies (2026), and a finalist nomination at the SC Awards, London (2026), an international award programme in cybersecurity.

U Osijeku, 2026.

Miljenko Švarcmajer