REVIEWING SCALABILITY AND SECURITY IN LARGE-SCALE PEER-TO-PEER NETWORKS

Supported by tetfund
TERTIARY EDUCATION TRUST FUND

Ikani, Lucy Hassana; Ogwueleka, Francisca Nonyelum
[1,2]Department Of Computer Science, Faculty of Science, University of Abuja
Corresponding author: lucyikani@gmail.com

Received: February 14, 2025, Accepted: April 28, 2025

Abstract: Large-scale peer-to-peer (P2P) networks play a crucial role in various distributed applications, facilitating decentralized data sharing and communication among a vast number of interconnected nodes. However, ensuring scalability and security in these networks remains a significant challenge due to their dynamic and decentralized nature. In this review, we systematically analyze existing literature to explore the primary scalability challenges and security concerns faced by large-scale P2P networks, along with the strategies proposed or implemented to address them. We examine the impact of different architectural models, such as structured and unstructured overlays, on network scalability and assess the effectiveness of various security mechanisms and protocols in mitigating security threats. Furthermore, we discuss the trade-offs between scalability and security measures and provide recommendations for optimizing network performance while ensuring robust security. Our findings highlight the importance of adopting efficient routing algorithms, redundancy mechanisms, and comprehensive security measures to enhance the scalability and security of large-scale P2P networks. By addressing these challenges, stakeholders can unlock the full potential of large-scale P2P networks and pave the way for their widespread adoption in diverse applications.

Keywords: peer-to-peer (P2P) networks; security; Network Optimization; architectural models

Introduction

Peer-to-peer (P2P) networks have emerged as a transformative approach to distributed computing, enabling decentralized resource sharing without reliance on central servers. In a P2P network, each participating node functions as both a client and a server, facilitating direct communication and collaboration among peers. Large-scale P2P networks, such as BitTorrent for file sharing and Bitcoin for decentralized financial transactions, exemplify the scalability and efficiency of this architecture (Lua et al., 2005). These networks have demonstrated their potential in various applications, ranging from content distribution to secure financial transactions and communication. However, as these networks continue to expand, challenges related to scalability and security become increasingly prominent, requiring innovative solutions to maintain their efficiency and reliability (Richa & Scheideler, 2007). Scalability is a critical factor in the performance of large-scale P2P networks. As the number of peers grows, maintaining efficient resource discovery, data management, and network stability becomes more complex. Traditional centralized approaches are insufficient to handle the dynamic nature of these networks, necessitating distributed mechanisms such as Distributed Hash Tables (DHTs) and gossip protocols (Shen, Brodie, & Levine, 2005). These techniques help optimize peer-to-peer interactions by efficiently indexing and retrieving data across large networks. However, scalability improvements must be carefully balanced with security measures, as an increase in network size also introduces greater exposure to potential attacks and vulnerabilities (Ke & Mostafa, 2016).

Security remains a major concern in large-scale P2P networks due to their decentralized and open nature. Unlike traditional client-server models where security can be centrally managed, P2P networks are susceptible to various threats, including data tampering, unauthorized access, and denial-of-service attacks (Nadu, 2012). One of the most significant threats is the Sybil attack, in which malicious entities create multiple fake identities to manipulate the network. Addressing these security issues requires robust cryptographic techniques, access control mechanisms, and trust models that can safeguard the integrity and confidentiality of information exchanged among peers (Hu, Chen, & Chen, 2006). However, security mechanisms should not hinder the scalability and efficiency of the network, highlighting the need for a balanced approach.

The interplay between scalability and security in large-scale P2P networks is an area of ongoing research and development. While previous studies have explored these aspects independently, there is a pressing need for an integrated perspective that examines how scalability strategies can influence security and vice versa. Understanding this relationship is crucial for designing P2P networks that can accommodate large user bases while maintaining strong security protocols. This study aims to provide a comprehensive review of scalability and security in large-scale P2P networks, addressing key challenges, architectural models, security threats, and potential solutions (Richa & Scheideler, 2007).

Methodology

Research Methodology

To undertake a comprehensive and structured investigation into current efforts targeting scalability and security within large-scale peer-to-peer networks, we have adopted the systematic methodology crafted by Kitchenham et al. This methodology serves as our guide for reviewing and analyzing the existing literature in this domain. Our review encompasses six stages: (1) formulating research questions, (2) delineating research procedures, (3) screening relevant articles, (4) extracting pertinent data, and (5) mapping the findings.

Formulating a research question

1. What are the primary scalability challenges faced by large-scale peer-to-peer networks, and what strategies have been proposed or implemented to address them?

This question directly focuses on identifying the scalability challenges in large-scale P2P networks, such as routing efficiency, network overhead, dynamic membership, and content discovery. It also seeks to explore the strategies proposed or implemented to mitigate these challenges, which are crucial for improving the scalability of P2P networks.

2. How do different architectural models (e.g., structured vs. unstructured) impact the scalability of peer-to-peer networks, and what are their respective advantages and limitations?
   This question examines how different architectural models affect the scalability of P2P networks. It explores whether structured or unstructured approaches offer better scalability and discusses their respective advantages and limitations in addressing scalability challenges.
3. What are the most common security threats and vulnerabilities in large-scale peer-to-peer networks, and how do they affect network performance and user privacy?
   This question focuses on understanding the security threats and vulnerabilities inherent in large-scale P2P networks, which is essential for ensuring the security of these networks alongside scalability. It also addresses the impact of security threats on network performance and user privacy, highlighting the interplay between security and scalability concerns.
4. What security mechanisms and protocols are commonly employed to mitigate security risks in large-scale peer-to-peer networks, and how effective are they in practice?
   This question explores the security mechanisms and protocols used to mitigate security risks in large-scale P2P networks, which is crucial for addressing security concerns alongside scalability. It assesses the effectiveness of these measures in practice, highlighting their role in ensuring the security and scalability of P2P networks
5. What are the trade-offs between scalability and security measures in large-scale peer-to-peer networks, and how can these trade-offs be balanced to optimize network performance and security?
   This question examines the trade-offs between scalability and security measures in large-scale P2P networks, acknowledging that implementing stringent security measures may impact scalability and vice versa. It explores strategies for balancing these trade-offs to optimize both network performance and security.

Delineating Research Procedures

To develop the most substantial findings, several approaches need to be implemented including search strategy, inclusion, and exclusion criteria.

Search Strategy

To compile relevant articles in a keyword-based format, the research strategy commenced with the utilization of key terms. This initial approach involved searching terms such as "Scalability and Security of large-scale network", "Scalability and Security of peer-to-peer network", and "problems of peer-to-peer networks", among others, via Google Scholar. Subsequently, all articles were selected and downloaded chronologically, spanning from the earliest to the most recent publications. These sourced materials encompassed a diverse array of origins, including journals, conferences, IEEE, ACM, and SCOPUS.

Exclusion and Inclusion Criteria

After the questions related to the scope were proposed, all the primary studies were considered to identify the suitable information related to this study's systematic review. Meanwhile, 32 prime articles were included and tagged for data extraction. Table 1 below presents the articles with the corresponding ID.

Table 3.1. Exclusion and inclusion criteria.

Criteria	***Details***
Exclusion	Not cited. Published between 2015 to 2024. Duplications of articles from different sources. Titles do not mention or relate to Large-scale peer-to-peer networks
Inclusion	Cited at least once Systematic review related to Large-scale peer-to-peer network between 2015 to 2024. The abstract, introduction, or conclusion is related to the large-scale peer-to-peer network

Screening of Relevant Articles

To ensure the pertinence of each retrieved article to the research inquiries, an iterative methodology has been employed. Initially, duplicate articles sourced from various databases were eliminated during the screening process. Subsequently, the titles of all papers underwent meticulous scrutiny to filter out irrelevant ones—those not pertinent to the research questions. For instance, articles retrieved via the search query related to peer-to-peer networks but failed to address scalability issues were deemed beyond the scope of this Systematic Literature Review (SLR). However, determining relevance based solely on the title of a paper occasionally proved challenging. Thus, a more thorough examination of each paper's abstract was necessary to make a final determination of its inclusion. Our predefined inclusion and exclusion criteria played a crucial role in evaluating each article's relevance to the research questions.

Data extraction

Similar to the data extraction process for blockchain research, a structured approach is crucial for evaluating large-scale peer-to-peer networks. We can leverage a standardized form to consistently capture information relevant to scalability and security. This form could incorporate three key sections:

Network Characteristics: This section would capture details like network type (e.g., Gnutella, BitTorrent), size (estimated number of nodes), and routing protocols employed.

Scalability Analysis: Here, we'd focus on factors impacting scalability, such as message overhead, search efficiency, and

load balancing mechanisms. Data points on throughput, latency, and resource consumption (storage, bandwidth) could be collected.
Security Assessment: This section will delve into security features and potential vulnerabilities. Information on encryption methods, access control mechanisms, and resilience to malicious nodes would be valuable.
Following PRISMA guidelines for quality assessment ensures the validity and reliability of the data collected. Just as the Microsoft Excel form was rigorously tested and iteratively improved, this P2P network analysis form would undergo similar validation processes. This ensures the gathered data accurately reflects the strengths and weaknesses of the network regarding scalability and security.
By employing a structured and well-tested data collection method, we can gain valuable insights into the design and performance of large-scale peer-to-peer networks, paving the way for improvements in both scalability and security.

Analysis and Discussion of Findings
Data Analysis and Presentation
This section delves into the analysis of 32 chosen papers spanning the period from 2015 to 2024. It sheds light on the research trends over the past decade regarding scalability and security issues, as well as the available solutions for large-scale peer-to-peer networks. The discussion primarily revolves around:

1. The temporal distribution of publications related to scalability and security issues in peer-to-peer networks.
2. The categorization of publications on large-scale peer-to-peer networks.
3. The distribution of application domains for peer-to-peer networks.

To adequately address the research questions, the data collected during the extraction process were meticulously compiled, and demographic data were scrutinized for the specified publication years.
Figure 1 provides a graphical representation of the year-wise analysis of the selected papers. The graph indicates a growing interest in academic research concerning the scalability and security of peer-to-peer networks, particularly evident in the increasing number of publications from 2018 to 2021. It is noteworthy that the peak of academic research activity on scalability and security in large-scale peer-to-peer networks occurred in 2019.

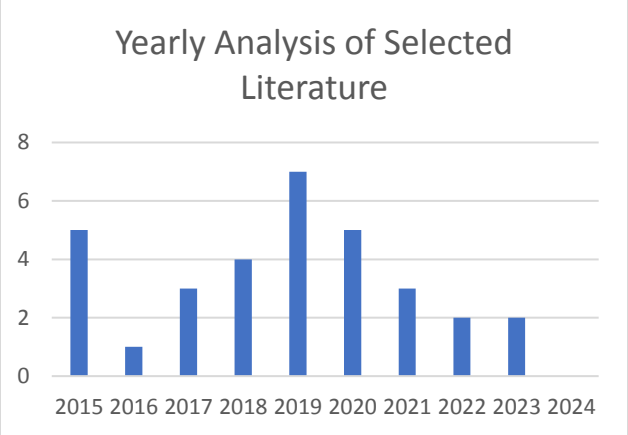

Figure 4.1. The figure illustrates the year-wise distribution of published papers in peer-to-peer network.
Figure 2 provides a breakdown of the publication types for the articles selected in this Systematic Literature Review (SLR). The identified publication types in this study include:
- Journal articles
- Conference Proceedings
- Book chapters
- Workshops

The analysis uncovered that the majority of the publications addressing scalability and security in peer-to-peer networks were disseminated through conferences and journals. Specifically, out of the 32 articles reviewed, 14 were published in conference proceedings, while 16 appeared in journal publications. The remaining articles were distributed across book chapters (1) and workshops (1).

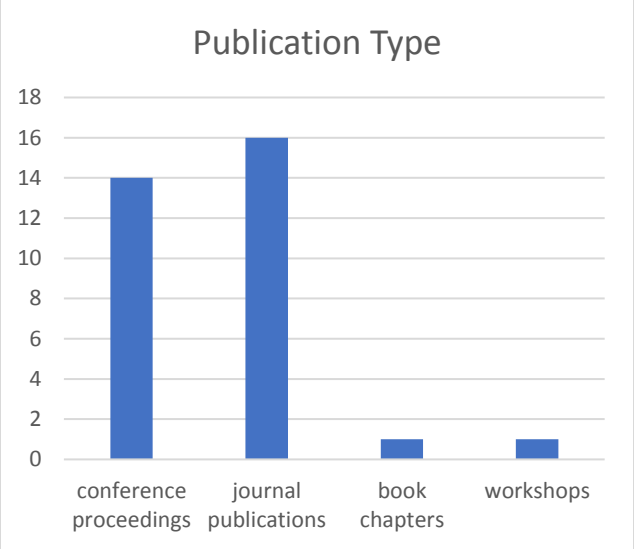

Figure 4.2. The figure illustrates the publication type of the published paper in a peer-to-peer network

Results and Findings
Research Question 1: What are the primary scalability challenges faced by large-scale peer-to-peer networks, and what strategies have been proposed or implemented to address them?
Large-scale peer-to-peer (P2P) networks offer a fascinating yet challenging environment when it comes to scalability

(Vijay, 2023). The cited articles reported some common scalability and security challenges faced by large-scale peer-to-peer networks. These challenges are:

***Resource Discovery and Search**:* The researchers reported that finding specific data or resources within a vast network of constantly changing nodes can be inefficient. Traditional flooding techniques can lead to message overhead and slowdowns.

***Load Balancing and Resource Availability**:* Peers with limited resources (bandwidth, storage) might be overloaded with requests, while others with more resources might be underutilized. Efficiently distributing workload and ensuring resource availability across the network is crucial.

***Scalability of Routing Protocols**:* Routing protocols that determine how data hops between nodes can become cumbersome as the network grows. Maintaining efficient routing tables and minimizing overhead becomes a challenge.

Strategies that can be put in place to eradicate these challenges as reported in the cited articles are:

1. ***Distributed Hash Tables (DHTs):*** These data structures map data to specific nodes based on a hash function, enabling efficient lookup and retrieval of information. Nodes are responsible for specific data that can be easily located, reducing search overhead.
2. ***Super-peers and Hierarchies**:* Introducing super-peers with higher capacities can act as hubs for routing and resource discovery. This creates a hierarchical structure that can handle larger network sizes compared to purely flat P2P models.
3. ***Content-aware Routing**:* Routing decisions are based on the content itself. Nodes with similar content can be clustered, allowing for more targeted searches and reduced network traffic.
4. ***Incentive Mechanisms**:* Encouraging resource sharing and participation is essential. Reputation systems and token-based rewards can incentivize nodes with better resources to contribute more to the network's overall health and scalability.
5. ***Overlay Networks**:* These virtual networks built on top of the existing physical network can implement custom routing protocols and functionalities specifically designed for efficient resource discovery and load balancing within the P2P network.

Research Question 2: How do different architectural models (e.g., structured vs. unstructured) impact the scalability of peer-to-peer networks, and what are their respective advantages and limitations?

Architectural models serve as communication tools for conveying design decisions, requirements, and constraints among project stakeholders, including developers, architects, project managers, and clients (Dubois and Mauger, 2015). They facilitate collaboration, consensus-building, and decision-making throughout the software development lifecycle. In a peer-to-peer network, there are two types of architectural models, structured and unstructured. The architectural model of a peer-to-peer (P2P) network significantly impacts its scalability, with each approach offering distinct advantages and limitations:

i. *Structured P2P Networks:*

Structured peer-to-peer (P2P) networks are decentralized network architectures where peers (nodes or participants) organize themselves in a structured manner to efficiently locate and retrieve resources without relying on centralized servers (Fredrick et.al). Below is the scalability analysis of a structured P2P network.

Scalability:

Structured networks excel in scalability as the network grows. They employ Distributed Hash Tables (DHTs) that efficiently map data to specific nodes based on a hash function. This allows for efficient search and retrieval of information regardless of network size. Additionally, routing protocols are often more efficient compared to unstructured models.

Advantages of structured P2P architectural model:

a. **Fast and efficient search:** Finding specific data is faster due to the organized structure of the network.
b. **Improved load balancing:** DHTs can potentially distribute resources and workload more effectively.

Limitations:

a. **Complexity:** Implementing and maintaining DHTs can be complex.
b. **Single point of failure:** Some DHT implementations have central nodes that could become bottlenecks or single points of failure if compromised.

ii. *Unstructured P2P Networks:*

Unstructured peer-to-peer (P2P) networks are decentralized network architectures where peers (nodes or participants) connect in a more ad-hoc or decentralized manner, without a predefined structure or overlay topology (Xing and Gary, 2010). In unstructured P2P networks, peers typically join and leave the network dynamically, and there is no strict organization or coordination among peers for resource discovery and routing. Below is the analysis of the scalability, advantages, and limitations of an unstructured P2P network.

Scalability:

Unstructured networks face challenges with scalability as the network size increases. Traditional flooding techniques used for search can become inefficient, leading to message overhead and slowdowns. Routing tables can grow unwieldy as the network scales.

Advantages:

i. ***Simplicity:*** Unstructured networks are easier to set up and maintain due to their decentralized nature.
ii. ***High fault tolerance:*** There are no single points of failure as all peers are equal.

Limitations:

i. ***Slow and inefficient search:*** Finding specific data can be slow and resource-intensive due to flooding techniques.

ii. ***Load balancing issues:*** Resource distribution can be uneven as some nodes might be overloaded while others remain underutilized.

Choosing the Right Model:

The optimal model depends on the specific application and its priorities.

- ***Structured P2P*:** Ideal for applications where fast and efficient search for specific data is crucial, such as content-delivery networks (CDNs) or distributed file systems.
- ***Unstructured P2P:*** More suitable for applications where resilience and ease of deployment are prioritized, such as file-sharing networks (e.g., BitTorrent) or instant messaging systems.

In conclusion, understanding the trade-offs between structured and unstructured architectures is crucial for designing efficient and scalable P2P networks. The type of data, search patterns, and desired level of fault tolerance will ultimately guide the choice of the most suitable architectural model.

Research Question 3: What are the most common security threats and vulnerabilities in large-scale peer-to-peer networks, and how do they affect network performance and user privacy?

Security threats and vulnerabilities in large-scale peer-to-peer networks can have significant impacts on network performance, user privacy, and data integrity. Mitigating these threats requires a combination of proactive measures, including robust authentication mechanisms, encryption protocols, access control policies, and intrusion detection systems, to safeguard network assets and protect against malicious activity (Washiyun et.al, 2024). Below are some of the common threats and vulnerabilities.

Threats and Vulnerabilities:

a. Free Riding and Sybil Attacks: Malicious nodes might leech resources (bandwidth, storage) without contributing to the network. In a Sybil attack, a single entity creates multiple fake identities to disrupt searches, manipulate data, or gain undue influence. The impact of free riding and Sybil attacks:

i. ***Performance:*** Free riding reduces available resources, leading to slowdowns and decreased efficiency. Sybil attacks can further disrupt search functionality and manipulate search results.

ii. ***Privacy:*** Sybil attacks can be used to hide the source of malicious activity or manipulate reputation systems.

b. ***Data Integrity Attacks:***

Malicious nodes might spread corrupted data (malware, fake content) or tamper with legitimate data during transfer. Corrupted data can waste network resources and potentially damage user systems. Users might unknowingly download and share malicious content, compromising their systems and potentially exposing personal information.

c. ***Man-in-the-Middle Attacks:***

An attacker intercepts communication between two peers, eavesdropping on data or potentially modifying it in transit. Sensitive information exchanged between peers could be intercepted, exposing user credentials or private data. Depending on the attack strategy, additional processing by the attacker could introduce delays.

d. Denial-of-Service (DoS) Attacks:

Attackers overload the network with traffic, making it unavailable for legitimate users. The network becomes unusable for legitimate purposes, hindering resource sharing and communication. DoS attacks can potentially be used to mask other malicious activities.

Impact on User Privacy:

- ***Lack of Centralized Control:*** P2P networks often lack a central authority, making it harder to enforce privacy policies or track down malicious actors. This can leave users vulnerable to data breaches and unauthorized access.
- ***File Sharing Risks:*** The very nature of P2P networks involves sharing files, which can expose sensitive information if not done cautiously. Downloaded malware or data breaches on individual nodes can compromise user privacy.

Overall Impact on Network Performance:

- ***Resource Consumption:*** Security threats like DoS attacks and excessive free riding can deplete network resources, leading to slowdowns and reduced efficiency.
- ***Overhead from Security Measures:*** Implementing encryption or other security measures can add overhead to communication, potentially impacting network performance.
- ***Disrupted Communication:*** Attacks like Man-in-the-Middle attacks or data poisoning can disrupt communication channels between peers, hindering resource sharing and file transfer processes.

Mitigating these threats requires a multi-pronged approach:

- ***Reputation Systems:*** Encourage responsible behavior by rewarding nodes that contribute and penalizing those that free-ride or engage in malicious activities.
- ***Encryption:*** Implement encryption protocols to protect data in transit and at rest, safeguarding user privacy and preventing data tampering.
- ***Decentralized Security Mechanisms:*** Leverage the decentralized nature of P2P networks to distribute security tasks and avoid single points of failure.
- ***User Education:*** Educate users on safe P2P practices like file verification and avoiding suspicious content to minimize the risk of malware and data breaches.

By acknowledging these security vulnerabilities and implementing appropriate measures, developers and users can work together to create a more secure and efficient environment for large-scale peer-to-peer networks.

Research Question 4: What security mechanisms and protocols are commonly employed to mitigate security

risks in large-scale peer-to-peer networks, and how effective are they in practice?

Large-scale peer-to-peer (P2P) networks offer advantages like decentralization and resource sharing but also face unique security challenges. This research question delves into the security mechanisms and protocols commonly employed to mitigate these risks, along with their effectiveness in practice.

Common Security Mechanisms and Protocols:

1. ***Cryptography and Encryption:***

Encryption scrambles data using a key, making it unreadable without decryption. This protects data confidentiality during transfer and storage on nodes.

- ***Effectiveness:*** Encryption is highly effective in safeguarding data privacy. However, it adds computational overhead and requires proper key management to prevent unauthorized access.

2. ***Digital Signatures:***

Digital signatures allow verification of data origin and integrity. A node "signs" the data with its private key, and other nodes can verify its authenticity using the corresponding public key.

- ***Effectiveness:*** Digital signatures help prevent data tampering and ensure authenticity. However, they require robust public key infrastructure (PKI) and can add processing overhead.

3. ***Hashing and Integrity Verification:***

Hashing functions generate a unique fingerprint (hash) for a data block. Any changes to the data will result in a different hash. This allows verification of data integrity during transfer.

- ***Effectiveness:*** Hashing is computationally efficient and helps ensure data hasn't been corrupted. However, it doesn't guarantee confidentiality and doesn't identify the source of the data.

4. ***Reputation Systems:***

These systems track the behavior of nodes within the network. Nodes with positive contributions (sharing resources, verifying data) gain good reputations, while malicious nodes (free riding, spreading malware) are penalized.

- ***Effectiveness:*** Reputation systems can incentivize cooperation and deter malicious behavior. However, they can be susceptible to manipulation by Sybil attacks (creating fake identities) and require careful design to be fair and effective.

5. ***Decentralized Access Control (DAC):***

DAC allows nodes to define access permissions for their shared resources. This can restrict access to authorized users and prevent unauthorized downloads.

- ***Effectiveness:*** DAC offers granular control over resource access. However, it requires efficient mechanisms for managing access control policies across a large-scale network.

6. ***Cooperative Intrusion Detection Systems (C-IDS):***

Nodes collaborate to detect and report suspicious activity within the network. This distributed approach leverages the collective intelligence of the network to identify potential threats.

- ***Effectiveness:*** C-IDS can offer broader threat detection capabilities. However, it requires efficient communication protocols for sharing information and mitigating false positives.

Effectiveness in Practice:

The effectiveness of these mechanisms depends on several factors:

- ***Network Design and Implementation*****:** Security protocols need to be well-integrated into the P2P architecture for optimal results.
- ***Scalability of Solutions:*** Security mechanisms should be efficient and scalable to handle large numbers of nodes and data transfers.
- ***User Education and Behavior:*** User awareness of security best practices is crucial for minimizing risks associated with file sharing and node interactions.

Research Question 5: What are the trade-offs between scalability and security measures in large-scale peer-to-peer networks, and how can these trade-offs be balanced to optimize network performance and security?

Large-scale peer-to-peer (P2P) networks face a constant struggle to balance **scalability** (efficiently handling a growing number of nodes and data) with robust **security** measures. Here's a breakdown of the key trade-offs and strategies for achieving a balance:

Trade-offs:

- ***Complexity vs. Efficiency:*** Implementing robust security measures often involves complex protocols and computations. This can add overhead to communication and resource consumption, potentially hindering scalability as the network grows.
- ***Centralization vs. Decentralization*****:** Certain security mechanisms, like centralized reputation systems, might offer better control but introduce a single point of failure, which contradicts the decentralized nature of P2P networks.
- ***Transparency vs. Privacy:*** Enhanced security might require some level of transparency in user activity to identify malicious behavior. This can create tension with user privacy concerns.

Strategies for Balancing Trade-offs:

- ***Lightweight Security Protocols:*** Develop and implement security protocols that are efficient and scalable, minimizing overhead without compromising effectiveness. Techniques like lightweight cryptography and hashing can be explored.
- ***Decentralized Security Mechanisms:*** Leverage the distributed nature of P2P networks by employing decentralized reputation systems, intrusion detection, and access control mechanisms that don't rely on central authorities.
- ***Incentive-based Systems:*** Encourage secure and responsible behavior by rewarding nodes that contribute positively to network security and

penalizing malicious activity. This can be integrated with reputation systems.

- ***Privacy-Preserving Security:*** Develop security techniques that maintain user privacy while achieving adequate levels of protection. Techniques like homomorphic encryption or anonymous authentication can be explored.
- ***Layered Security Approach:*** Implement a layered security architecture with different mechanisms at various levels (e.g., encryption at the data layer, reputation systems at the network layer). This provides a flexible and adaptable defense against diverse threats.
- ***Scalable Infrastructure:*** Utilize efficient and scalable network infrastructure that can handle the increasing demands of security measures in conjunction with growing network size.

Finding the Optimal Balance:

The ideal balance between scalability and security depends on the specific application and its priorities.

- ***Applications prioritizing fast file sharing (e.g., BitTorrent)*** might prioritize scalability with lighter security measures, relying on user discretion and basic encryption.
- ***Applications handling sensitive data (e.g., secure communication platforms)*** might prioritize security with more robust protocols, even if it impacts performance slightly.

Balancing scalability and security in P2P networks is an ongoing challenge. By continuously researching and implementing new techniques like those mentioned above, we can create a more secure and efficient environment for these valuable networks.

Summary and Conclusion

Large-scale peer-to-peer (P2P) networks offer a unique paradigm for resource sharing and distributed computing. However, ensuring their scalability to accommodate a growing number of nodes and data, while simultaneously maintaining robust security, presents a significant challenge. This review explored the key issues surrounding scalability and security in P2P networks. We discussed the limitations of both structured and unstructured architectures, highlighting the trade-offs between search efficiency and overall network efficiency. We also delved into the common security threats faced by P2P networks, such as free-riding, data integrity attacks, and Denial-of-Service attacks, along with their impact on user privacy and network performance. Furthermore, we examined the various security mechanisms and protocols employed to mitigate these risks. Encryption, digital signatures, and hashing functions provide essential tools for data protection and integrity verification. Reputation systems and decentralized access control offer ways to incentivize cooperation and deter malicious behavior. However, the effectiveness of these mechanisms depends heavily on factors like network design, scalability, and user awareness.

Finally, we explored the crucial trade-offs between security measures and scalability. Implementing complex security protocols can introduce overhead and hinder network efficiency. Conversely, prioritizing scalability might leave the network vulnerable to security threats.

The path forward lies in achieving a balanced approach. Lightweight and scalable security protocols, combined with decentralized mechanisms and incentive-based systems, offer promising avenues for improvement. Privacy-preserving security techniques and a layered security architecture further enhance network resilience without compromising user privacy.

By continuously researching and implementing innovative solutions, we can create a more secure and efficient environment for large-scale P2P networks. This will ultimately pave the way for wider adoption and unlock the full potential of this powerful technology.

Recommendation

After a thorough analysis and review of existing literature on scalability and security in large-scale peer-to-peer networks, several key recommendations emerge to improve the performance, reliability, and resilience of these networks:

1. Adopt Structured Overlay Networks: Implementing structured overlay networks, such as Distributed Hash Tables (DHTs), can significantly enhance routing efficiency and scalability in large-scale peer-to-peer networks. By organizing nodes into a structured topology and employing efficient routing algorithms, structured overlays can reduce latency and network overhead, thereby improving overall performance.
2. Utilize Redundancy and Replication: Leveraging redundancy and replication mechanisms is essential to ensure data availability and reliability in large-scale peer-to-peer networks. By storing multiple copies of data across different nodes and employing data replication strategies, networks can withstand node failures and maintain high levels of data availability.
3. Implement Security Mechanisms: Integrating robust security mechanisms and protocols is crucial to protecting large-scale peer-to-peer networks against various security threats and vulnerabilities. Techniques such as encryption, authentication, and intrusion detection can help safeguard network communication and data integrity, enhancing overall security.
4. Optimize Resource Discovery: Efficient content discovery mechanisms are essential for large-scale peer-to-peer networks to locate and retrieve content distributed across multiple nodes. By implementing distributed indexing and search mechanisms, networks can reduce search latency and improve scalability, enhancing the user experience.
5. Balance Scalability and Security: Striking a balance between scalability and security measures is paramount to ensure optimal network performance and resilience. While stringent security measures may impact scalability, it is essential to implement measures that mitigate security risks without compromising network scalability.

References

Alotibi, B., Alarifi, N., Abdulghani, M., & Altoaimy, L. (2019). Overcoming free-riding behavior in peer-to-peer networks using a points system approach. *Procedia Computer Science*, *151*, 1060-1065.

Angrish, A., Craver, B., Hasan, M., & Starly, B. (2018). A case study for Blockchain in manufacturing: "FabRec": A prototype for a peer-to-peer network of manufacturing nodes. Procedia Manufacturing, 26, 1180-1192.

Augustine, J, Pandurangan,G. Robinson, P, Roche, S & Upfal, E.(2015) "Enabling Robust and Efficient Distributed Computation in Dynamic Peer-to-Peer Networks, IEEE 56th Annual Symposium on Foundations of Computer Science, Berkeley, CA, USA, 2015, pp. 350-369, doi: 10.1109/FOCS.2015.29.

Balani, N., Chavan, P. & Ghonghe, M. (2022). Design of high-speed blockchain-based sidechaining peer-to-peer communication protocol over 5G networks. *Multimed Tools Appl* **81**, 36699–36713 https://doi.org/10.1007/s11042-021-11604-6

Chergui, N., Kechadi, T., & Chikhi, S. (2017). Scalability-aware mechanism based on workload

Dennis, R., & Owenson, G. (2016). Rep on the roll: a peer-to-peer reputation system based on a rolling blockchain. *International Journal for Digital Society*, *7*(1), 1123-1134.

Dubois, Eric & Mauger, Cyril. (2015). On the Role of Architectural Models: What Can We Learn from Information System and from Construction Projects?. Proceedings - International Conference on Research Challenges in Information Science. 2015. 10.1109/RCIS.2015.7128858.

Ehiagwina, Frederick & Iromini, Nurudeen & Olatinwo, Ikeola Suhurat & Raheem, Kabirat & Anifowose Nee Mustapha, Khadijat. (2022). A State-of-the-Art Survey of Peer-to-Peer Networks: Research Directions, Applications and Challenges. 1. 19-38. 10.55708/js0101003.

Eltamaly, A. M., & Ahmed, M. A. (2023). Performance Evaluation of Communication Infrastructure for Peer-to-Peer Energy Trading in Community Microgrids. *Energies*, *16*(13), 5116.
environments. *IEEE Network, 20*(4), 22-31. The authors propose a Voronoi-based overlay network (VON) for scalable P2P-based virtual environments.

Gai, F., Wang, B., Deng, W., & Peng, W. (2018). Proof of reputation: A reputation-based consensus protocol for peer-to-peer network. In *Database Systems for Advanced Applications: 23rd International Conference, DASFAA 2018, Gold Coast, QLD, Australia, May 21-24, 2018, Proceedings, Part II 23* (pp. 666-681). Springer International Publishing.

Hao, W., Zen, J., Dai, X., Xiao, J., Hua, Q., Chen, H., Li, K., & Hai, J. (2019). BlockP2P: Enabling Fast Blockchain Broadcast with Scalable Peer-to-Peer Network Topology. In: Miani, R., Camargos, L., Zarpelão, B., Rosas, E., Pasquini, R. (eds) Green, Pervasive, and Cloud Computing. GPC 2019. Lecture Notes in Computer Science(), vol 11484. Springer, Cham. https://doi.org/10.1007/978-3-030-19223-5_16

Heilman, E., Kendler, A., Zohar, A., & Goldberg, S. (2015). Eclipse attacks on {Bitcoin's}{peer-to-peer} network. In *24th USENIX security symposium (USENIX security 15)* (pp. 129-144).
highlights the role of overlay networks in scalable P2P systems.

Hou, W., Jiang, Y., Lei, W. (2020). A P2P network-based edge computing smart grid model for efficient resource coordination. *Peer-to-Peer Netw. Appl.* **13**, 1026–1037. https://doi.org/10.1007/s12083-019-00870-9

Hu, S.-Y., Chen, J.-F., & Chen, T.-H. (2006). VON: A scalable peer-to-peer network for virtual

Jafari Navimipour, N. & Sharifi Milani, F (2015).. A comprehensive study of the resource discovery techniques in Peer-to-Peer networks. *Peer-to-Peer Netw. Appl.* **8**, 474–492 https://doi.org/10.1007/s12083-014-0271-5

Javadpour, A., Wang, G. & Rezaei, S. (2020). Resource Management in a Peer-to-Peer Cloud Network for IoT. *Wireless Pers Commun* **115**, 2471–2488. https://doi.org/10.1007/s11277-020-07691-7

Jin, Xing & Chan, S.-H. (2010). Unstructured Peer-to-Peer Network Architectures. 10.1007/978-0-387-09751-0_5.

Kavalionak, H., Carlini, E., Ricci, L. *et al* (2015). Integrating peer-to-peer and cloud computing for massively multiuser online games. *Peer-to-Peer Netw. Appl.* **8**, 301–319. https://doi.org/10.1007/s12083-013-0232-4

Ke, W., & Mostafa, J. (2016). Scalability analysis of distributed search in large peer-to-peer

Li, Z., Barenji, A. V., & Huang, G. Q. (2018). Toward a blockchain cloud manufacturing system as a peer-to-peer distributed network platform. *Robotics and computer-integrated manufacturing*, *54*, 133-144.

M. Khorasany, Y. Mishra, B. Babaki, and G. Ledwich,(2019). "Enhancing scalability of peer-to-peer energy markets using adaptive segmentation method," in Journal of Modern Power Systems and Clean Energy, vol. 7, no. 4, pp. 791-801, doi: 10.1007/s40565-019-0510-0.

Masinde, N &, Graffi, K. (2020).. Peer-to-Peer-Based Social Networks: A Comprehensive Survey. *SN COMPUT. SCI.* **1**, 299 https://doi.org/10.1007/s42979-020-00315-8

Mohammadi, B., & Navimipour, N. J. (2019). Data replication mechanisms in the peer-to-peer networks. *International Journal of Communication Systems*, *32*(14), e3996. https://doi.org/10.1002/dac.3996.

Musa, A., Abubakar,A., Gimba, U. A., & Rasheed, R. A.. (*2019*). "An Investigation into Peer-to-Peer Network Security Using Wireshark," *15th International Conference on Electronics, Computer and Computation (ICECCO)*, Abuja,

Nigeria, 2019, pp. 1-6, doi: 10.1109/ICECCO48375.2019.9043236.

Nadu, T. (2012). Performance Analysis of Controlled Scalability in Unstructured Peer-to-Peer

Naik, A.R., & Keshavamurthy, B.N. (2020). Next level peer-to-peer overlay networks under high churns: a survey. *Peer-to-Peer Netw. Appl.* **13**, 905–931. https://doi.org/10.1007/s12083-019-00839-8 networks. *2016 IEEE International Conference on Big Data (Big Data)*, 909-914.

Neudecker, T., Andelfinger, P., & Hartenstein, H. (2015, May). A simulation model for analysis of attacks on the Bitcoin peer-to-peer network. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)* (pp. 1327-1332). IEEE.

Pouya, I., Pronk, S., Lundborg, M., & Lindahl, E. (2017). Copernicus, a hybrid dataflow and peer-to-peer scientific computing platform for efficient large-scale ensemble sampling. *Future Generation Computer Systems*, *71*, 18-31. prediction in ultra-peer networks. *Peer-to-Peer Networking and Applications*, *11*, 431-449. This paper discusses workload prediction mechanisms to improve scalability in ultra-peer networks.

Richa, A. W., & Scheideler, C. (2007). Overlay Networks for Peer-to-Peer Networks. This work

Risson, J., & Moors, T. (2006). "Survey of research towards robust peer-to-peer networks: Search methods." Journal of Network and Computer Application.

Sankar, S.P., Subash, T.D., & Vishwanath, N. *et al.* (2021). Security improvement in blockchain technique enabled peer-to-peer networks beyond 5G and the Internet of Things. *Peer-to-Peer Netw. Appl.* **14**, 392–402. https://doi.org/10.1007/s12083-020-00971-w

Shen, H., Brodie, A., Xu, C., & Shi, W. (2005). Scalable and Secure Peer-to-Peer Overlay

Soto, E. A., Bosman, L. B., Wollega, E., & Leon-Salas, W. D. (2021). Peer-to-peer energy trading: A review of the literature. *Applied Energy*, *283*, 116268.

Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., & Balakrishnan, H. (2021). "Chord: A scalable peer-to-peer lookup service for internet applications." ACM SIGCOMM Computer Communication Review.

Stoykov, L., Zhang, K., & Jacobsen, H. A. (2017, December). Vibes: fast blockchain simulations for large-scale peer-to-peer networks. In *Proceedings of the 18th ACM/IFIP/USENIX Middleware Conference: Posters and Demos* (pp. 19-20).

Thantharate P, & Thantharate A. (2023). ZeroTrustBlock: Enhancing Security, Privacy, and Interoperability of Sensitive Data through ZeroTrust Permissioned Blockchain. *Big Data and Cognitive Computing*. 7(4):165. https://doi.org/10.3390/bdcc7040165

Tian, Y. (2007). Performance and security issues in peer-to-peer based content distribution

Verma, P., O'Regan, B.,& Hayes, B. *et al*(2018). EnerPort: Irish Blockchain project for peer-to-peer energy trading. *Energy Inform* **1**, 14. https://doi.org/10.1186/s42162-018-0057-8

Vijay, K. (2023). What Is Peer-To-Peer? Meaning, Features, Pros, and Cons. https://www.spiceworks.com/tech/networking/articles/what-is-peer-to-peer/. (Accessed on 21st February, 2024)

Vijay, K.(2023). What Is Peer-To-Peer? Meaning, Features, Pros, and Cons. Retrieved from (https://www.spiceworks.com/tech/networking/articles/what-is-peer-to-peer/)

Wasyihun Sema Admass, Yirga Yayeh Munaye, Abebe Abeshu Diro, Cyber security: State of the art, challenges and future directions, Cyber Security and Applications, Volume 2, 2024, 100031, ISSN 2772-9184, https://doi.org/10.1016/j.csa.2023.100031. (https://www.sciencedirect.com/science/article/pii/S2772918423000188)

Wongthongtham, P., Marrable, D., Abu-Salih, B., Liu, X., & Morrison, G. (2021). Blockchain-enabled Peer-to-Peer energy trading. *Computers & Electrical Engineering*, *94*, 107299.

Yang, J., & Garcia-Molina, H. (2002) "Improving search in peer-to-peer networks." Proceedings of the International Conference on Very Large Data Bases (VLDB).

Ye, Y., Tang, Y., Wang, H., Zhang, X.P., & Strbac, G. (2021) "A Scalable Privacy-Preserving Multi-Agent Deep Reinforcement Learning Approach for Large-Scale Peer-to-Peer Transactive Energy Trading," in *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5185-5200, doi: 10.1109/TSG.2021.3103917

Yifan, M., Soubhik, D., Shaileshh, B.V., Sreeram, K., & Kannan, S.(2020). Perigee: Efficient Peer-to-Peer Network Design for Blockchains. In Proceedings of the 39th Symposium on Principles of Distributed Computing (PODC '20). Association for Computing Machinery, New York, NY, USA, 428–437. https://doi.org/10.1145/3382734.3405704

Zhao, B. Y., Huang, L., Stribling, J., Rhea, S. C., Joseph, A. D., & Kubiatowicz, J. D. (2004) "Tapestry: A resilient global-scale overlay for service deployment." IEEE Journal on Selected Areas in Communications.

Zhuang and J. M. Chang,(2019) "Enhanced PeerHunter: Detecting Peer-to-Peer Botnets Through Network-Flow Level Community Behavior Analysis," in *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1485-1500, doi: 10.1109/TIFS.2018.2881657.