Using Weighted Voting to Accelerate Blockchain Consensus

How to make sure that the latency that the nodes report prior to AWARE's algorithm is realistic?

Filip Błaszczyk[1]

Supervisor(s): Jérémie Decouchant[1]**, Rowdy Chotkan**[1]

[1]**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 23, 2024

Name of the student: Filip Błaszczyk
Final project course: CSE3000 Research Project
Thesis committee: Jérémie Decouchant, Rowdy Chotkan, Kaitai Liang

Abstract

This research addresses the challenge of managing latency in distributed computer systems. Maintaining correct delays in data transmission across various network conditions is crucial for system efficiency and security. We focus on improving the Adaptive Wide-Area Replication (AWARE) algorithm, a method used to coordinate data across different locations in a way that minimizes delays. To enhance AWARE, we incorporate two concepts: Vivaldi network coordinates and Newtonian invariants. Vivaldi network coordinates help the system better understand and calculate the physical layout of the network by embedding network members in Euclidean space, where the distance in this space represents latency between them. Newton invariants are rules based on physics that help the system detect and adjust for any unusual changes in network latency that might be caused by technical issues or security threats. We evaluated the original and enhanced AWARE algorithms by simulating typical network operations and various attack scenarios designed to slow down the process of reaching consensus. Our findings show that the enhanced AWARE algorithm provides more accurate and robust management of network latency, especially under attack conditions, leading to a more reliable and secure distributed system. This study confirms that integrating correction techniques into latency management processes significantly improves the resilience and accuracy of distributed systems.

1 Introduction

State machine replication (SMR) is a well-established methodology for achieving fault tolerance in distributed systems. It replicates the state of a machine across multiple nodes to ensure continuity and consistency of service despite failures. With the evolution of distributed technologies, particularly in blockchain and Byzantine fault-tolerant (BFT) systems, the efficiency of reaching consensus in geographically dispersed nodes has become important. This is because the latency of messages between nodes can significantly affect the speed and reliability of consensus mechanisms, which are fundamental to the security and robustness of distributed ledgers and other decentralized applications. The Adaptive Wide-Area Replication (AWARE [1]) algorithm addresses this by adjusting voting weights and dynamically selecting leaders based on latency measurements, thereby optimizing the consensus process. However, a significant challenge arises when these latency measurements are not reliable—due to faults in transmission, network congestion, or deliberate manipulation by malicious nodes aiming to disrupt the consensus or degrade the system's performance. This thesis focuses on enhancing the AWARE algorithm by introducing mechanisms to detect and mitigate the impact of faulty or malicious nodes on latency reporting. These nodes can skew the algorithm's perception of network conditions, leading to suboptimal leader selection and weighting decisions that can severely affect the overall system performance and reliability. By embedding detection mechanisms, we aim to preserve the integrity of the latency measurements that are critical for the adaptive functionalities of AWARE. To validate and refine our approach, we will simulate the enhanced AWARE algorithm using real-life latency datasets such as the King dataset for DNS-based latency measurements and the WonderNetwork global ping statistics. These datasets will provide a robust foundation for testing our hypothesis that movements within the Vivaldi [3] coordinates can effectively signal potential security risks or faults in node behavior.

2 Background

This paper addresses the challenges of stability and accuracy in latency prediction within AWARE, leveraging the Vivaldi coordinate system. By incorporating the Newton enhancement, which enforces physical laws within the algorithm, we can disregard updates that arise from network failures or malicious behavior, thereby enhancing robustness.

2.1 AWARE

AWARE advances the WHEAT [1] algorithm by dynamically reconfiguring weights assigned to replicas, with the primary goal of optimizing consensus times in distributed systems. Unlike WHEAT, which utilizes static weight assignments based on initial conditions, AWARE adapts in real-time to latency changes, making it more responsive to the evolving state of the network.

The core of AWARE's operational framework lies in its utilization of latency matrices that each replica maintains. Two principal mechanisms enhance the reliability of these latency measurements: matrix sanitization and the injection of randomness.

Matrix sanitization addresses asymmetrical latency readings between nodes. If the recorded latency from node i to node j (denoted as i, j) is greater than from j to i (j, i), the algorithm conservatively adopts the higher value for both i, j and j, i. This ensures that the system's operational parameters always consider potential maximum delays, thereby guarding against underestimation of latency, which could jeopardize consensus accuracy.

The second mechanism, the injection of randomness into latency measurement requests, serves as a safeguard against manipulative behaviors by replicas. By requiring that a random number be added to each latency response, AWARE mitigates the risk of replicas artificially appearing more responsive by prematurely acknowledging requests. This randomness helps in masking the exact response time, thus complicating any attempt by a malicious node to consistently manipulate latency measurements for strategic advantage.

Despite these mechanisms, AWARE faces challenges in environments where nodes may exhibit random latency fluctuations due to network instability or where a group of nodes attempts to skew the consensus process. These issues highlight the need for continuous refinement of the algorithm to effectively handle diverse threats in large-scale, dynamic networks.

2.2 Vivaldi

Vivaldi [3] utilizes an n-dimensional Euclidean space to model the latencies between nodes in distributed systems. Each node within the network is assigned coordinates in this space, with the distances between nodes representing the latency estimates. The model is modeled as a dynamic spring

system, where each connection between nodes acts like a spring that can either be compressed or stretched based on real-time latency measurements.

The primary operation of Vivaldi lies in its ability to adjust these 'spring tensions', which corresponds to recalibrating the estimated distances between nodes. This adjustment is triggered when a difference between the actual measured latencies and those predicted by the model arises. Such discrepancies could be due to changes in network traffic, routing alterations, or hardware issues, which affect the latency between nodes.

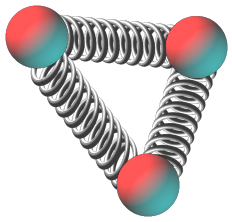


Figure 1: Vivaldi's spring system visualization

To update the node positions, Vivaldi applies a force-based algorithm. When a node i receives a latency measurement from node j, it calculates the force exerted by this latency difference on the connecting spring. If the measured latency is greater than the estimated distance, the spring is considered to be too compressed and needs to be stretched. Conversely, if the measured latency is less, the spring is deemed to be overly stretched and needs to be compressed. The node then adjusts its position in the Euclidean space to either increase or decrease the distance to node j accordingly, aiming to bring the estimated latency closer to the measured value.

Over time, the system tends to stabilize as the positions of the nodes adjust to reflect the true latency landscape of the network. This stabilization is crucial for maintaining the accuracy of the latency model, particularly in highly volatile environments where network conditions can change rapidly.

Moreover, Vivaldi's design is inherently scalable, as each node operates independently, making localized decisions based on its interactions. This decentralized nature ensures that the system can scale to large networks without a significant increase in computational overhead or communication complexity. Each node only needs to communicate with a subset of other nodes to adjust its position, rather than requiring a global view of the network, which enhances the efficiency and scalability of the algorithm.

2.3 Newton

The Newton [8] enhancement significantly refines the Vivaldi model by integrating principles from Newton's laws of motion, which introduces physics-based invariants into the system. These invariants are designed to detect and mitigate anomalies that could distort the accuracy of network latency estimations.

IN1: Centroid Stability: The first invariant concerns the centroid, or geometric center, of a node i and its neighboring nodes. Under ideal conditions, where no external perturbations affect the network, this centroid should align at the origin of the coordinate system. Any displacement from this point indicates the presence of an unbalanced force, which could be the result of malicious activities or asymmetric network behaviors. By monitoring the centroid's location, Newton can identify and counteract these influences.

IN2: Force Consistency: The second invariant examines the consistency of force vectors between closely situated nodes. If node i is influenced by a force $\vec{f_{ij}}$ due to its interaction with node j, then any other node k that is close to i should experience a similar force from j, scaled by the vector projection onto the line connecting j and k. This invariant leverages the physical principle that forces in a stable system should behave predictably across similar distances. Anomaly here can indicate coordinated attempts to manipulate latency measurements.

IN3: Damping Force Reduction: The third invariant focuses on the dynamic behavior of the system as nodes adjust towards their equilibrium positions, or 'rest positions'. According to Newton's laws, as the system nears equilibrium, the magnitudes of the forces acting on the nodes should decrease, manifesting a damping behavior. Persistent high levels of force or sudden spikes can suggest external disruptions or anomalies in network performance.

These physics-based invariants allow the Newton enhancement to act as a sophisticated filter that continuously assesses the reliability of data fed into the AWARE algorithm. By ensuring that only consistent, equilibrium-conforming measurements influence the latency predictions, Newton boosts the accuracy and stability of the entire system. Moreover, the application of these invariants provides an additional layer of security against manipulation and attacks, as any anomalous forces can be quickly identified and neutralized.

3 Enhancing AWARE with Coordinate-Based Verification

The refinement implemented in AWARE replaces the traditional latency matrix within the algorithm with a model where each replica holds the coordinates of the nodes. This spatial approach allows for a more intuitive and direct method of assessing the relative positions and distances among nodes, which is essential for managing latency effectively. During updates, these coordinates are evaluated against Newton's invariants, a set of physics-based rules designed to ensure that changes in network conditions are reflected accurately and malicious manipulations are quickly identified and mitigated. If an update does not satisfy these criteria, it is discarded across the network.

This method not only simplifies the verification process at each node, thereby enhancing transparency and consistency across the system but also effectively manages subtle network adjustments and filters out anomalous or malicious changes in reported latencies. We begin by applying the first invariant, which asserts that node coordinates should be proximal to the origin of Euclidean space. A significant deviation from this

norm suggests that a node might be under the influence of a malicious node that falsifies data. In the experiments, we set the threshold of 20 units for the deviation from the origin. This parameter should be adjusted for the scale and magnitude of changes in the system.

Following insights from Mercury [9], we have also incorporated a maximum allowable force post-system stabilization. We define system stabilization as the state where the error between two nodes is less than twenty percent. This threshold ensures that only minor adjustments are needed to maintain equilibrium, thus stabilizing the system more rapidly and reducing the likelihood of large-scale disruptions.

The final verification involves examining the forces applied to a node over the last ten rounds. Each node maintains historical data regarding the magnitudes of forces imposed by its peers. This historical analysis allows the system to understand typical force patterns and identify outliers effectively. If the magnitude of a new force substantially exceeds the median force, $\tilde{F}$, i.e.,

$$|\tilde{f}_{\text{new}}| > \tilde{F} + k \times D$$

it is disregarded. Here, D is the median absolute deviation, and based on empirical data, it has been found that a value of $k = 8$ is effective [9]. This criterion helps to ensure that only legitimate and reasonable updates are processed, which is particularly important in dynamic and potentially hostile network environments.

By implementing these enhanced verification processes, the AWARE algorithm can more effectively detect and respond to both subtle and significant changes in the network, ensuring robust and reliable operation even under adverse conditions. This approach helps with verifying the positions between all the nodes. It patches the gap from the original approach with the latency matrix, where a group of nodes may lie about their latencies between each other. In the enhanced version of the system, this behavior should now be detected by the validation of the invariants.

4 Experimental Setup and Datasets

For all experiments, we will use $n = 3 \times f + 1 + \delta$ where n represents the number of nodes, f the number of faulty nodes, and δ the number of additional replicas. We will evaluate the performance of the standard AWARE algorithm against the enhanced version that incorporates Newtonian principles. Each dataset simulation will mimic all network behaviors as f grows, with $\delta = f$. Simulations will run over 10 rounds, maintaining consistent latencies but with different randomly selected malicious nodes each time. For Vivaldi, we will initially stabilize the coordinates, similar to using landmarks in the real world to quickly establish initial coordinates when a new node joins. Each round will involve simulated annealing to optimize the weights and manage malicious behaviors, such as falsifying coordinates or reporting incorrect latencies.

4.1 Behavior in the Network

As documented in Newton [8] and other studies on threats to Vivaldi [4], we can pinpoint several attacks on coordinate systems that are analogous to attacks in a matrix format:

- **Inflation:** Attackers exaggerate their coordinates significantly, misdirecting benign nodes from accurate coordinates—an attack on precision. In a matrix context, this would be akin to reporting high latencies to degrade the perceived value of well-positioned nodes.
- **Deflation:** Attackers underreport their coordinates near the origin, preventing benign nodes from updating accurately, thus compromising precision. For the matrix, this involves reporting unusually low latencies to appear more favorably positioned in the network.
- **Oscillation:** Attackers randomly alter their coordinates and delay measurements, affecting both accuracy and stability. In matrix terms, this means entering random values.
- **Frog-boiling:** Attackers gradually distort their coordinates, slowly increasing deviations. Over time, this leads to significant inaccuracies. For the matrix, it involves incrementally increasing latency reports to eventually manipulate the network's favor.

These behaviors were simulated and their impacts measured against the enhanced algorithm's performance in a controlled environment with all nodes behaving normally.

4.2 Data Sources for Latency Measurements

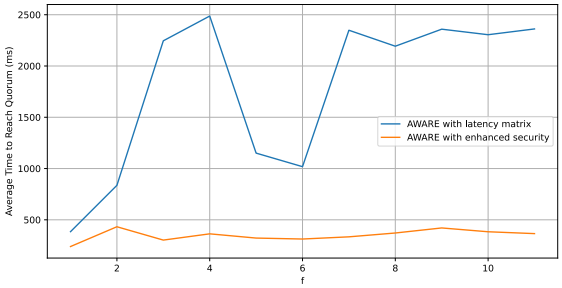
To see how the enhanced algorithm will perform in real networks and in different environments, we have selected three datasets with real-life latency measurements from around the world. The datasets are:

- **Wonder Network Ping Table:** Provides a global view of internet latencies, ideal for simulating a broad network [7].
- **PlanetLab Dataset:** A reliable source of network performance data from a vast array of nodes worldwide [10].
- **King Dataset:** Contains precise measurements between various internet hosts, commonly utilized in research on internet distance estimation and topology [5].

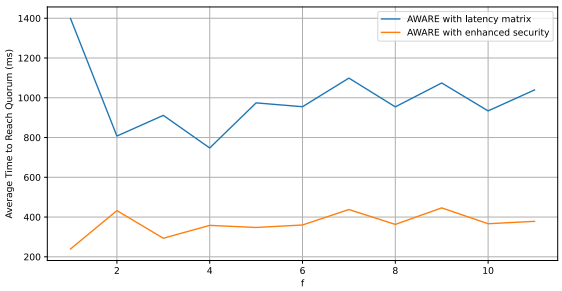
4.3 Results

The first comparison, as illustrated in Figure 5, was conducted by simulating normal behavior using both the AWARE algorithm and its enhanced version employing Vivaldi coordinates instead of the traditional latency matrix. The latency of quorum reaching time, averaged over 10 rounds of simulation, demonstrates that the quorums proposed by the enhanced algorithm, based on Vivaldi coordinates, perform comparably to those suggested by the original latency matrix-based method.

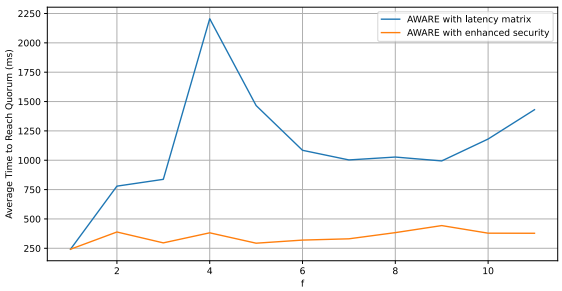
We observe that the enhanced algorithm performs exceptionally well on the King and Wonder Network datasets (Figures 2, 4). The algorithm effectively detects and disregards malicious nodes, thereby maintaining a lower average quorum collection time. However, the PlanetLab dataset shows some irregularities in performance (Figure 3). On average, the enhanced algorithm still handles attack scenarios better, though there are instances where it is also affected.



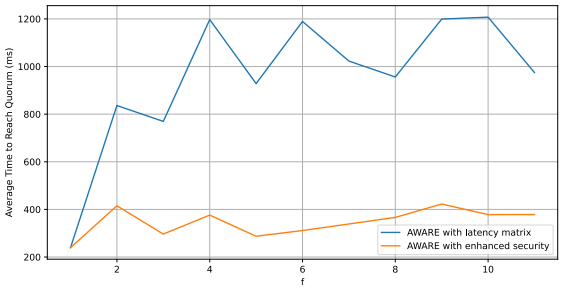
(a) Deflation Attack



(b) Inflation Attack

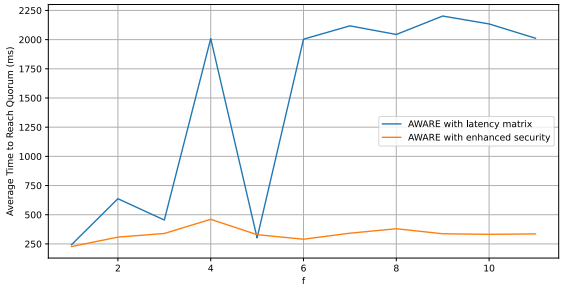


(c) Frog-Boiling Attack

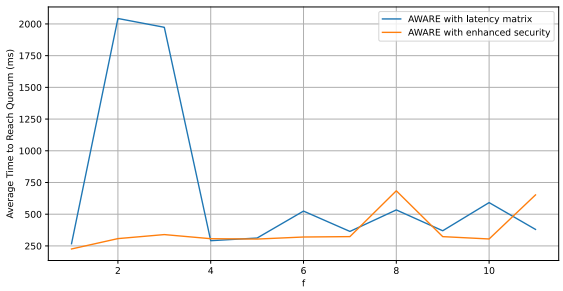


(d) Oscillation Attack

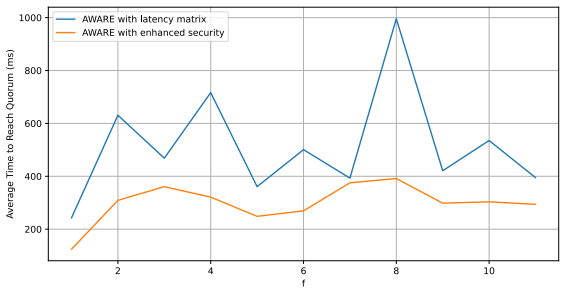
Figure 2: Network simulations performed on the latencies from the Wonder Network ping table



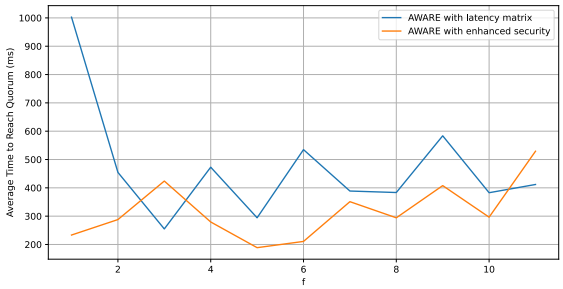
(a) Deflation Attack



(b) Inflation Attack



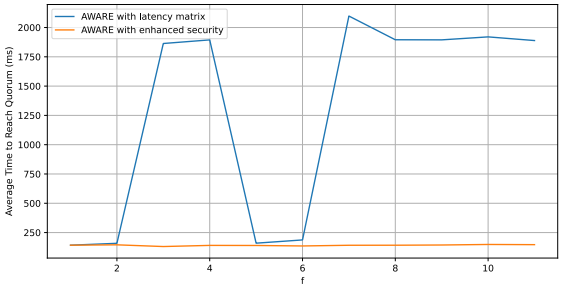
(c) Frog-Boiling Attack



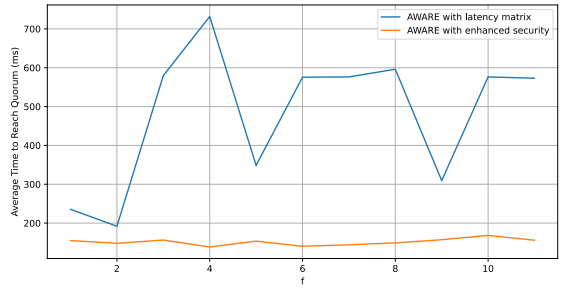
(d) Oscillation Attack

Figure 3: Network simulations performed on the latencies from the PlanetLab dataset

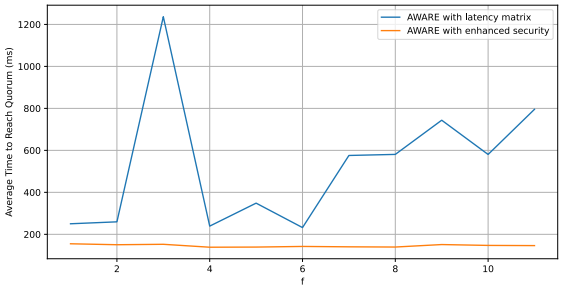
Further investigation into these irregularities revealed that the PlanetLab dataset contains several outliers in terms of latency. These outliers cause nodes to be more spread out in Vivaldi space, providing more room for movement. While this



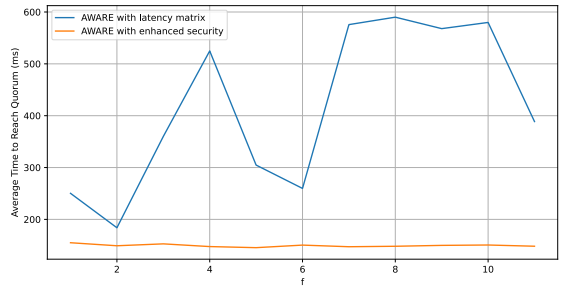
(a) Deflation Attack



(b) Inflation Attack



(c) Frog-Boiling Attack



(d) Oscillation Attack

Figure 4: Network simulations performed on the latencies from the King dataset

does not affect the first invariant, it impacts the system's stabilization, which could be the cause of the anomalies. Since

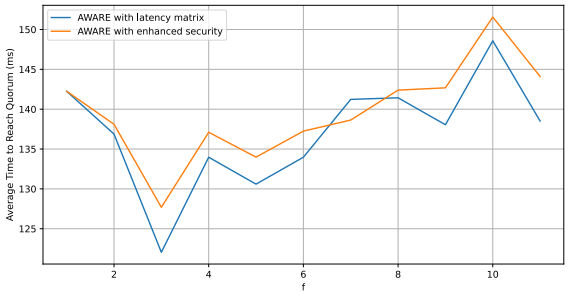


Figure 5: Comparing AWARE to the enhanced algorithm during normal network behavior on the Wonder Network Dataset

the simulation is limited to f, larger-scale simulations with more nodes need to be performed to ensure the enhanced algorithm's performance trend continues.

5 Discussion

The results obtained not only allow us to draw conclusions about the performance of enhanced AWARE protocol [1] but also help us understand potential causes for the observed behaviors. The performance of the enhanced algorithm mirrors that of the latency matrix-based approach under normal conditions. This shows that our added layer of security does not affect overall performance. In our simulations, we gave Vivaldi coordinates forty updates to stabilize.

In scenarios involving network attacks, the enhanced algorithm significantly outperforms the standard version by more effectively selecting non-malicious nodes, thanks to the additional security measures integrated into the system. Notably, the original AWARE protocol is most vulnerable to inflation attacks, where groups of malicious nodes report deceivingly low latencies among themselves, suggesting this is a particularly effective strategy against the standard protocol.

Furthermore, the effectiveness of the Frog Boiling attack in smaller groups is notable. The variation in node positions results in a cascading effect with fewer nodes present, underscoring the vulnerability of systems to strategic manipulations in small network clusters. Due to time constraints, we were unable to implement more dedicated attacks against AWARE, such as reporting minimal latencies among a group of f malicious nodes while reporting greater latencies between members of f and every other node in the network. These attacks warrant deeper exploration. This study has demonstrated that the Enhanced AWARE [1] algorithm, through its integration of network coordinate systems and enforcement of Newton's invariants [8], exhibits resilience to network variations compared to the original AWARE algorithm. Enhanced AWARE maintains robustness while still being responsive to legitimate network changes, effectively mitigating the impact of significant latency discrepancies caused by anomalous or malicious activities.

Looking forward, several areas have been identified for further development and research. The next step involves deploying Enhanced AWARE in real-world distributed systems to assess its practical efficacy and efficiency. This phase will

also help in understanding the operational challenges and resource requirements in a live environment. A comparative analysis is crucial, and it should compare this enhanced algorithm with other latency protection methods used in distributed systems, such as the matrix factorization techniques employed in Pharos [2]. Given the potential of machine learning techniques in predictive analytics, future research could also explore the integration of ML algorithms to further refine latency predictions [6] and network anomaly detection within the Enhanced AWARE framework. Additionally, more experiments should be designed to test the algorithm under a wider range of network conditions and configurations, particularly focusing on scalability and the handling of highly mobile network environments.

6 Responsible Research

The improvements to the AWARE algorithm, particularly with the incorporation of Vivaldi coordinates and Newtonian invariants, aim to enhance the reliability and security of distributed systems. By mitigating the risk of latency manipulation, our enhancements contribute to the broader goal of creating more secure and robust networks, which is essential in critical applications such as finance, healthcare, and public administration.

Additionally, our work respects the ethical principle of transparency. The simulations and datasets used in this research are based on publicly available data, and we have ensured that no private or sensitive information is utilized. To ensure the reproducibility of our research, we have documented the processes and methodologies used. The datasets used for simulating network behaviors, including the Wonder Network Ping Table, PlanetLab Dataset, and King Dataset, are publicly accessible. Detailed references to these datasets are provided, enabling other researchers to obtain and use the same data for replication studies. The protocols for running simulations, including the setup of malicious nodes and the application of various attack strategies, are clearly described. Parameters such as the number of nodes, fault thresholds, and the specific conditions under which the simulations were conducted are explicitly stated. The modifications to the AWARE algorithm, incorporating Vivaldi coordinates and Newtonian invariants, are detailed in the paper. We provide comprehensive explanations of the mathematical models and the logic behind the invariant checks. Software and Tools: The tools and software libraries used for simulations, data analysis, and plotting results are posted on the open source repository. By adhering to these principles, we aim to contribute to the field of distributed systems research in a manner that is both ethically sound and scientifically robust. Our commitment to transparency and reproducibility ensures that our findings can be validated, extended, and applied by other researchers in the community.

7 Conclusion

The promising results from this study pave the way for a transformative improvement in network latency management, positioning Enhanced AWARE as a pivotal advancement in the field of distributed systems.

References

[1] Christian Berger, Hans P. Reiser, Joao Sousa, and Alysson Bessani. AWARE: Adaptive Wide-Area Replication for Fast and Resilient Byzantine Consensus. *IEEE Transactions on Dependable and Secure Computing*, 19(3):1605–1620, May 2022.

[2] Yang Chen, Yongqiang Xiong, Xiaohui Shi, Beixing Deng, and Xing Li. Pharos: A Decentralized and Hierarchical Network Coordinate System for Internet Distance Prediction. In *IEEE GLOBECOM 2007 - IEEE Global Telecommunications Conference*, pages 421–426, November 2007. ISSN: 1930-529X.

[3] Frank Dabek, Russ Cox, Frans Kaashoek, and Robert Morris. Vivaldi: a decentralized network coordinate system. *ACM SIGCOMM Computer Communication Review*, 34(4):15–26, August 2004.

[4] Mohamed Ali Kaafar, Laurent Mathy, Thierry Turletti, and Walid Dabbous. Real attacks on virtual networks: Vivaldi out of tune. In *Proceedings of the 2006 SIGCOMM workshop on Large-scale attack defense*, LSAD '06, pages 139–146, New York, NY, USA, September 2006. Association for Computing Machinery.

[5] Steven D. Gribble Krishna P. Gummadi, Stefan Saroiu. Index of /archive/p2psim/kingdata. https://pdos.csail.mit.edu/archive/p2psim/kingdata/, 2005. Accessed: 2024-06-16.

[6] Shady A. Mohammed, Shervin Shirmohammadi, and Sa'di Altamimi. A Multimodal Deep Learning-Based Distributed Network Latency Measurement System. *IEEE Transactions on Instrumentation and Measurement*, 69(5):2487–2494, May 2020. Conference Name: IEEE Transactions on Instrumentation and Measurement.

[7] Wonder Netowrk. Global ping statistics. https://wondernetwork.com/pings/, 2024. Accessed: 2024-06-16.

[8] Jeffrey Seibert, Sheila Becker, Cristina Nita-Rotaru, and Radu State. Securing Virtual Coordinates by Enforcing Physical Laws. In *2012 IEEE 32nd International Conference on Distributed Computing Systems*, pages 315–324, Macau, China, June 2012. IEEE.

[9] Mingxun Zhou, Liyi Zeng, Yilin Han, Peilun Li, Fan Long, Dong Zhou, Ivan Beschastnikh, and Ming Wu. Mercury: Fast Transaction Broadcast in High Performance Blockchain Systems. In *IEEE INFOCOM 2023 - IEEE Conference on Computer Communications*, pages 1–10, New York City, NY, USA, May 2023. IEEE.

[10] Rui Zhu. Netlatency-data. https://github.com/uofa-rzhu3/NetLatency-Data, 2017. Accessed: 2024-06-16.