

Charting the Uncharted: The Landscape of Monero Peer-to-Peer Network

1st Yu Gao* 2nd Matija Piškorec 3rd Yu Zhang 4th Nicolò Vallarano 5th Claudio J. Tessone
UZH Blockchain Center University of Zurich UZH Blockchain Center UZH Blockchain Center UZH Blockchain Center
University of Zurich Rudjer Boskovic Institute University of Zurich University of Zurich University of Zurich
yugao@ifi.uzh.ch matija.piskorec@irb.hr zhangyu@ifi.uzh.ch nicolo.vallarano@uzh.ch claudio.tessone@uzh.ch

Abstract—The Monero blockchain enables anonymous transactions through advanced cryptography in its peer-to-peer network, which underpins decentralization, security, and trustless interactions. However, privacy measures obscure peer connections, complicating network analysis. This study proposes a method to infer peer connections in Monero’s latest protocol version, where timestamp data is unavailable. We collect peerlist data from TCP flows, validate our inference algorithm, and map the network structure. Our results show high accuracy, improving with longer observation periods. This work is the first to reveal connectivity patterns in Monero’s updated protocol, providing visualizations and insights into its topology. Our findings enhance the understanding of Monero’s P2P network, including the role of supernodes, and highlight potential protocol and security improvements.

Index Terms—Monero P2P Protocol, Monero P2P Network Mapping, Robustness in Monero P2P network

I. INTRODUCTION

The peer-to-peer (P2P) network is the foundational layer of decentralized ledger technology (DLT) systems, ensuring decentralization, consensus, and network security through broadcast of transactions and blocks [1], [2]. Although different DLT systems vary in design, all rely on the P2P network as the backbone for decentralization, security, and trustless interactions [3]–[5]. The P2P network consists of independent nodes (peers) that connect through an overlay structure to share resources. Each peer plays an equal role in sending messages and sharing resources [6]. The introduction of consensus and synchronization mechanisms adds complexity, affecting the network’s security, consensus, and smart contract vulnerabilities.

A key feature of P2P networks in DLT is dynamic discovery and interaction between nodes, forming a decentralized network structure governed by the P2P protocol [7], [8]. This eliminates the need for central authority management, making the network resilient and scalable, driven by autonomous nodes. However, this self-organizing capability complicates the study of the network, requiring experimental measurements and tools such as “traceroute” to map the structure of the system [9], [10].

Monero, an open-source blockchain, uses advanced cryptographic techniques such as ring signatures, stealth addresses, and ring confidential transactions to ensure transaction anonymity on its permissionless P2P network. Monero

employs Proof of Work (PoW) as its consensus mechanism but distinguishes itself by ensuring untraceable transactions. Although DLT research spans smart contracts [11], [12], consensus mechanisms [13], and decentralized finance [14], research on Monero’s P2P network remains scarce, with studies focusing more on transaction traceability [15], [16]. However, the decentralized nature of Monero’s network has also made it a target for cryptojacking [9], [10], an illicit practice in which attackers exploit the computing power of users to mine Monero without consent. This issue not only challenges user privacy, but also poses security risks to the integrity of the network, highlighting the need for stronger safeguards in Monero’s P2P ecosystem.

Few studies investigate Monero’s P2P network, the most notable being Cao et al. [17], which notes an update of the protocol to hide the *time_stamp* field in peer lists. This update removes the timestamp information, making it harder to infer peer connections. Existing studies on Monero focus mainly on traceability [15], highlighting the need for more research on its P2P network.

In this paper, we analyze Monero’s P2P network by collecting peer lists extracted from the TCP data exchanged between our Monero node and its peers. To overcome the absence of timestamps, we propose an algorithm that infers peer connections using timestamp-free TCP packet analysis. We validate this approach by comparing the inferred connections with the actual neighbors, achieving high accuracy. Our analysis of Monero’s network topology provides insights into the structure of the system, potential vulnerabilities, and areas for future improvements in protocol design.

II. MONERO PEER-TO-PEER PROTOCOL

By default, a Monero node establishes 8 outgoing connections. At least one outgoing connection is required to join the network (no such requirement for incoming connections). When a new node joins, it connects to the public seed nodes to discover other active nodes. Through a handshake protocol, the node establishes connections, validates them, and synchronizes data such as blocks and peer lists, while maintaining and updating its peer list, managing connection requests, and ensuring network stability.

Unlike Bitcoin-based P2P protocols, Monero uses its unique protocol. Each peer has two lists: *white_list* (1000 slots) for

recent handshakes and *grey_list* (5000 slots) for unresponsive peers or those with lower timestamp rankings. Under the new Monero protocol, a peer returns 120% of peers from its top 300 *last_seen* peers in the white list during the TCP flow process¹.

III. PEER LIST COLLECTION AND NEIGHBOR INFERENCE

In P2P networks within DLT systems, it is important to differentiate between “active neighbors” and “potential peers”.

A. Definition of “Neighbor” and “Interaction”

To formalize this distinction, we define a “connection” and an “interaction” as follows:

Neighbor/Connection: A connection is a persistent, active communication channel between peers, allowing the exchange of data. Active neighbors are peers with ongoing, direct connections, essential for maintaining network integrity and performance as peers join and leave over time.

Interaction: An interaction refers to a temporary relationship between peers, including: (1) Temporary connections, such as during initial handshakes, without forming a persistent channel. (2) Potential peers listed in a node peer list, based on discovery protocols, without established communication or validation. This distinction helps to accurately represent and manage the network topology, ensuring that only meaningful interactions are considered active neighbors.

To infer P2P connections between Monero nodes, we first set up a peer list data collection pipeline, as outlined in [17], and then develop a new inference algorithm. Due to changes in Monero’s peer list sharing protocol, which no longer includes the *last_seen* timestamp, we propose a modified algorithm that works without these data.

B. Peer list data collection

Our data collection pipeline (Figure 1) consists of three machines deployed in the United States, Europe, and Singapore. Geographical dispersion is crucial to capturing a broad view of the Monero P2P network, while the controlled nodes allow for validation of the P2P inference. Peer list data is collected by running a Monero node and monitoring TCP traffic with the *tcpflow*² program, listening on port 18080, through which communication with other nodes occurs. Peer lists are extracted from the binary dump of the traffic data, as they are not exposed through the node’s RPC interface.

The connection data from our three nodes is used to validate the P2P network inference. Data for this paper were collected over three weeks, from 21.12.2024 to 10.1.2025.

1) *Identifying Real Neighbors:* To identify real neighbors, we analyze peer list data based on the P2P protocol, where each peer shares up to 250 entries from its top 300 whitelisted peers. Real neighbors, which maintain frequent handshakes, appear more often in shared *peer_lists* than non-neighbors. The nodes belong to one of two categories:

- **Highly connected nodes:** for example, public seed nodes with numerous connections.

¹Monero GitHub Repository

²<https://github.com/simsong/tcpflow>

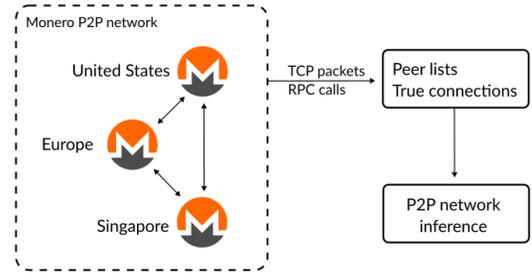


Fig. 1: Data collection pipeline.

- **Typical nodes:** default 8 outgoing neighbors.

The challenge is to set a frequency threshold to distinguish real neighbors. Since peer selection is bidirectional and is governed by the `IDLE_HANDSHAKE` protocol, real neighbors are expected to appear consistently in updated whitelists.

For a node i , we define the relative presence of an observed address a as:

$$p_i(a) = \frac{\sum_{j=1}^{n_i} \chi(a \in P_j)}{n_i}$$

where $P^{(i)}$ represents i ’s received TCP packets and χ is an indicator function.

A node with 8 neighbors, receiving 250 peer addresses per handshake, selects a neighbor with probability $P_{\text{neighbour}} \approx 0.833$. For non-neighbors, selection follows two stages:

- 1) Entering the top 300 with probability $P_{\text{enter}} \approx 0.302$.
- 2) Being chosen within the top 300 with $P_{\text{selected}} \approx 0.833$.

Thus, the overall probability for random nodes is: $P_{\text{random}} = P_{\text{enter}} \cdot P_{\text{selected}} \approx 0.252$.

This structural bias results in real neighbors appearing three times more frequently than random nodes, making them distinguishable through cumulative frequency analysis.

C. Neighbor identification by *k*-means

The simple method above intuitively identifies high-frequency peer interactions compared to low-frequency ones. This observation led us to apply *k*-means clustering to separate high- and low-frequency links.

To identify real neighbors, we first collect the frequency of each peer’s connections in shared *peer_lists* or TCP flows, then use the *k*-means clustering algorithm to differentiate between high-frequency and low-frequency links.

The dataset consists of triplets $(ip1, ip2, \text{count})$, where $ip1$ and $ip2$ are IP addresses, and count is the interaction frequency between them. We excluded rows where $ip1$ corresponds to Singapore, the US, or the EU, since our nodes perform frequent handshakes. To identify relevant connections and infer real neighbors, we applied the following steps:

First, we remove the rows where $\text{count} \leq 1$, as these likely represent noise, leaving the dataset D_{filtered} with meaningful interactions.

Second, we grouped the data by $ip1$, creating groups G_s for each source IP. For each group:

Algorithm 1 Data Filtering and Neighbor Inference Process

Require: Raw dataset D with columns $ip1, ip2, count$.

- 1: $C_{\min} \leftarrow 2$, threshold of minimum connection count.
 - 2: $N_{\min} \leftarrow 8$, threshold of minimum number of out-going neighbors.
 - 3: $D_{\text{filtered}} = \{row \in D : row[count] \geq C_{\min}\}$
 - 4: **for** each unique source IP s in D_{filtered} **do**
 - 5: $G_s = \{row \in D_{\text{filtered}} : row[ip1] = s\}$
 - 6: **if** $|G_s| \geq N_{\min}$ **then**
 - 7: Perform k -means clustering on the unique count values in G_s with $k = 2$. Greater count values' labels are set as 1 and as 0 otherwise.
 - 8: **end if**
 - 9: **end for**
 - 10: $D' = \{row \in D_{\text{filtered}} : row[label] = 1\}$
 - 11: **return** D'
-

- We discarded groups with fewer than two interactions.
- We applied k -means clustering with $k = 2$ on the unique count values, which distinguishes true connections from non-true connections based on frequency.

Finally, we retained all connections labeled as true neighbors (for example, those clustered as 1), forming the dataset D' , which includes the most relevant inferred connections.

By clustering the frequency data, we identified high-frequency peers as likely true neighbors, with the k -means method offering computational efficiency and revealing meaningful patterns in large-scale interaction datasets.

D. Neighbor inference validation

We compare the inferred neighbors of our Singapore, EU, and US nodes with the real connection lists obtained using Monero's `monero-daemon-rpc`. These real connection lists serve as a benchmark for validating the neighbor inference.

Classifier quality measures like precision and recall help assess our method's performance. Precision is the ratio of true positives to all detected positives (true + false positives), representing the proportion of real neighbors correctly identified. Recall is the ratio of true positives to all actual neighbors, indicating the proportion of real neighbors successfully detected.

Since peers dynamically connect and disconnect over time, even real connection lists capture only a transient snapshot of the network. To validate our approach, we first identify neighbors of our node in the inferred network and then check how many of these inferred neighbors appear in the real connection lists during data collection.

We measure precision as the ratio of "our node's neighbors in the inferred network" to "neighbors identified from real connections over one week of data." The following tables present the validation results.

Without incorporating the `last_seen` timestamps, the prediction accuracy of our inference method is already relatively high. This result in Table I is based on only one week of listening data, excluding the `last_seen` timestamps in the peer lists under the current protocol. When the data collection period was extended, the accuracy of the EU node improved by approximately 10% (check Table II), highlighting the

TABLE I: One-week data neighbor inference accuracy

Benchmark comparison for one week data			
Location	Singapore	US	EU
Inferred network neighbors	489	534	534
Neighbors in connection list	388	367	371
Neighbor identification Precision	79.35%	68.73%	69.48%

TABLE II: Three weeks data neighbor inference accuracy

Benchmark comparison for three weeks data			
Location	Singapore	US	EU
Inferred network neighbors	394	437	713
Neighbors in connection list	327	319	572
Neighbor identification Precision	82.99%	73.00%	80.22%

significant impact of a larger dataset collected over a longer period on the prediction accuracy. This improvement aligns with the theoretical expectation that longer TCP listening times enhance accuracy.

IV. MONERO P2P NETWORK TOPOLOGY ANALYSIS

After applying k -means clustering to filter the links, we obtained the "high-frequency clusters" edge list and mapped it into a network.

1) *Network Topology Analysis:* P2P networks in DLT systems, defined by their decentralized architecture, present both challenges and opportunities that require a thorough analysis. Network Science, grounded in graph theory and statistical physics, offers a solid framework for studying the structural and functional properties of networks [18], [19]. By representing a complex system as a graph, we can analyze topological features, develop network models, and evaluate robustness.

Although visualizing the network provides an initial insight, this approach becomes limited for larger networks. Therefore, a qualitative analysis of the network features using statistical data is essential. We define centrality and betweenness centrality as follows [18]:

Degree Centrality: Measures the number of direct connections (edges) a node has. The degree k_i of the node i is defined as: $k_i = \sum_{j=1}^n A_{ij}$, where A_{ij} is the adjacency matrix of the network. Degree centrality is useful for analyzing network centralization and structure.

Betweenness Centrality: Quantifies how often a node lies on the shortest paths between pairs of nodes, indicating its influence on network interactions. The betweenness centrality b_i of node i is given by: $b_i = \sum_{st} \frac{n_{st}^i}{g_{st}}$, where n_{st}^i is the number of shortest paths from s to t passing through i , and g_{st} is the total number of shortest paths between s and t .

Fig. 2 visualizes the LCC of the Monero P2P network, where node color represents betweenness centrality and size reflects degree centrality. Warmer-colored nodes indicate higher betweenness, highlighting their role in connecting different parts of the network, while larger nodes signify high-degree peers. The network consists of 4,837 nodes, with 3,153 directly linked to the 14 top-degree nodes, collectively

covering 82.1% of the network. These core nodes—primarily public seed nodes or supernodes—maintain multiple connections, facilitating efficient data exchange and ensuring network integrity.

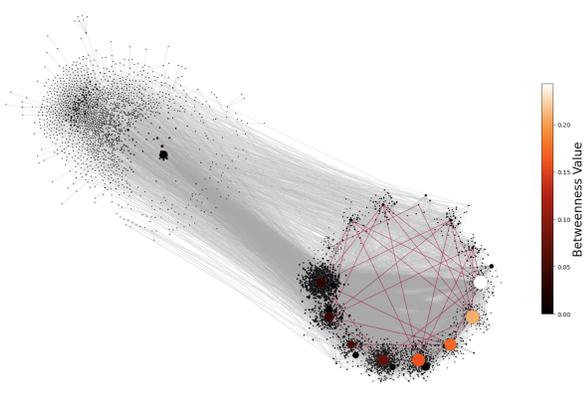


Fig. 2: Visualization of the LCC, with node colors representing betweenness centrality and node sizes proportional to their degree centrality.

2) *Interconnection Analysis Among Supernodes*: To explore the interconnections among these supernodes, we examine the overlap of their direct neighbors. The heatmap in Fig. 3 shows the overlap rate of direct neighbors among the 14 top-degree nodes. The value in each cell represents the overlap rate of the one-hop neighbors between the nodes indexed by the X and Y ticks, calculated as: $r = \frac{n_{overlap}}{\min(n_x, n_y)}$, where n_x and n_y are the number of neighbors of nodes x and y , and $n_{overlap}$ is the number of shared neighbors. The heatmap reveals that over 91% of each top-degree node’s direct neighbors are connected to other top-degree nodes, with this proportion nearing 100% for 9 out of the 14 nodes. This suggests a strong level of interconnectivity and resilience within the core of the network.

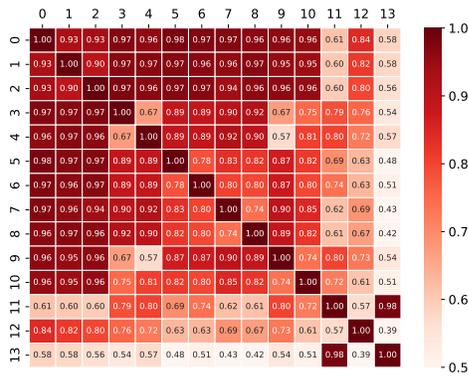


Fig. 3: One-hop neighbor overlap rate among the 14 top-degree nodes. The value in each cell represents the overlap rate of the one-hop neighbors between the nodes indexed by the X and Y ticks.

3) *Robustness of the Monero P2P Network*: The absence of a central node is a defining feature of DLT systems. However, this does not preclude the existence of supernodes within the network. In an unstructured DLT P2P network, if small-world characteristics emerge from user social relationships, nodes with significantly higher degrees can exert disproportionate influence on the network structure, introducing potential risks such as Eclipse attacks [20], [21], Sybil attacks [22], [23], and DoS attacks [24].

To assess the network robustness, we focus on high-degree and high-betweenness centrality nodes and evaluate how quickly the largest connected component (LCC) disintegrates. As shown in Fig. 4, the network demonstrates strong resilience. After removing the first 14 supernodes, the LCC size decreases by 20%. In particular, the LCC collapses to nearly zero when around 9.4% and 12% of nodes are removed based on betweenness and degree centrality, respectively. This suggests that compromising these nodes or their neighbors would significantly disrupt the network, highlighting its inherent robustness.

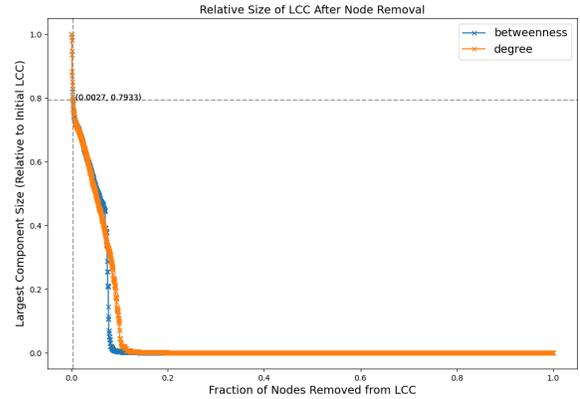


Fig. 4: LCC attack by removing high betweenness and degree centrality nodes, where the turning point of LCC ≈ 0 are 0.094 and 0.12 for betweenness centrality and degree centrality.

V. CONCLUSION

In conclusion, this study introduces a novel method for mapping the Monero peer-to-peer network by identifying real neighbors between nodes. Our approach accurately uncovers the network’s structure and provides an intuitive visualization of its topology. The findings offer the first comprehensive insights into connectivity patterns within the Monero P2P network under the new peer protocol.

Our results show that the network is highly centralized around several super nodes with significant betweenness centrality and high degrees. While this centralization strengthens security and robustness, it also introduces potential vulnerabilities. This study provides a foundation for future improvements to the Monero peer protocol, enhancing security and decentralization.

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of lipschitz-hankel type involving products of bessel functions," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 247, no. 935, pp. 529–551, 1955.
- [2] G. Wood *et al.*, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, no. 2014, pp. 1–32, 2014.
- [3] R. Böhme, N. Christin, B. Edelman, and T. Moore, "Bitcoin: Economics, technology, and governance," *Journal of economic Perspectives*, vol. 29, no. 2, pp. 213–238, 2015.
- [4] A. M. Antonopoulos and D. A. Harding, *Mastering bitcoin*. " O'Reilly Media, Inc.", 2023.
- [5] N. El Ioini and C. Pahl, "A review of distributed ledger technologies," in *On the Move to Meaningful Internet Systems. OTM 2018 Conferences: Confederated International Conferences: CoopIS, C&TC, and ODBASE 2018, Valletta, Malta, October 22-26, 2018, Proceedings, Part II*. Springer, 2018, pp. 277–288.
- [6] J. Buford, H. Yu, and E. K. Lua, *P2P networking and applications*. Morgan Kaufmann, 2009.
- [7] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Satoshi Nakamoto*, 2008.
- [8] Y. Marcus, E. Heilman, and S. Goldberg, "Low-resource eclipse attacks on ethereum's peer-to-peer network." *IACR Cryptology ePrint Archive*, vol. 2018, no. 236, 2018.
- [9] S. Eskandari, A. Leoutsarakos, T. Mursch, and J. Clark, "A first look at browser-based cryptojacking," in *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2018, pp. 58–66.
- [10] E. Tekiner, A. Acar, A. S. Uluagac, E. Kirda, and A. A. Selcuk, "Sok: cryptojacking malware," in *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2021, pp. 120–139.
- [11] W. Zou, D. Lo, P. S. Kochhar, X.-B. D. Le, X. Xia, Y. Feng, Z. Chen, and B. Xu, "Smart contract development: Challenges and opportunities," *IEEE transactions on software engineering*, vol. 47, no. 10, pp. 2084–2106, 2019.
- [12] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, and M. Imran, "An overview on smart contracts: Challenges, advances and platforms," *Future Generation Computer Systems*, vol. 105, pp. 475–491, 2020.
- [13] L. M. Bach, B. Mihaljevic, and M. Zagar, "Comparative analysis of blockchain consensus algorithms," in *2018 41st international convention on information and communication technology, electronics and micro-electronics (MIPRO)*. Ieee, 2018, pp. 1545–1550.
- [14] S. Werner, D. Perez, L. Gudgeon, A. Klages-Mundt, D. Harz, and W. Knottenbelt, "Sok: Decentralized finance (defi)," in *Proceedings of the 4th ACM Conference on Advances in Financial Technologies, 2022*, pp. 30–46.
- [15] M. Möser, K. Soska, E. Heilman, K. Lee, H. Heffan, S. Srivastava, K. Hogan, J. Hennessey, A. Miller, A. Narayanan *et al.*, "An empirical analysis of traceability in the monero blockchain," *arXiv preprint arXiv:1704.04299*, 2017.
- [16] Y. Li, G. Yang, W. Susilo, Y. Yu, M. H. Au, and D. Liu, "Traceable monero: Anonymous cryptocurrency with enhanced accountability," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 679–691, 2019.
- [17] T. Cao, J. Yu, J. Decouchant, X. Luo, and P. Verissimo, "Exploring the monero peer-to-peer network," in *Financial Cryptography and Data Security: 24th International Conference, FC 2020, Kota Kinabalu, Malaysia, February 10–14, 2020 Revised Selected Papers 24*. Springer, 2020, pp. 578–594.
- [18] M. Newman, *Networks*. Oxford university press, 2018.
- [19] D. J. Watts, "Networks, dynamics, and the small-world phenomenon," *American Journal of sociology*, vol. 105, no. 2, pp. 493–527, 1999.
- [20] S. Henningsen, D. Teunis, M. Florian, and B. Scheuermann, "Eclipsing ethereum peers with false friends," *arXiv preprint arXiv:1908.10141*, 2019.
- [21] K. Delmolino, M. Arnett, A. Kosba, A. Miller, and E. Shi, "Step by step towards creating a safe smart contract: Lessons and insights from a cryptocurrency lab," in *Financial Cryptography and Data Security: FC 2016 International Workshops, BITCOIN, VOTING, and WAHC, Christ Church, Barbados, February 26, 2016, Revised Selected Papers 20*. Springer, 2016, pp. 79–94.
- [22] G. Bissias, A. P. Ozisik, B. N. Levine, and M. Liberatore, "Sybil-resistant mixing for bitcoin," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, 2014, pp. 149–158.
- [23] U. Asfia, V. Kamuni, S. Sutavani, A. Sheikh, S. Wagh, and N. M. Singh, "A blockchain construct for energy trading against sybil attacks," in *2019 27th Mediterranean Conference on Control and Automation (MED)*. IEEE, 2019, pp. 422–427.
- [24] L. Garber, "Denial-of-service attacks rip the internet," *computer*, vol. 33, no. 04, pp. 12–17, 2000.