Estimating the Node Degree of Public Peers and Detecting Sybil Peers Based on Address Messages in the Bitcoin P2P Network

Matthias Grundmann Max Baumstark

Institute of Information Security and Dependability (KASTEL)
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Abstract

Some peers in the Bitcoin P2P network distributed a huge amount of spam IP addresses during July 2021. These spam IP addresses did not belong to actual Bitcoin peers. We found that the behavior of the spamming peers can be used to determine the number of neighbors of public peers and to find Sybil peers (peers that have multiple addresses). We evaluate the method by running an analysis based on data collected by our monitor nodes and compare the data to a ground-truth based on few peers that we run ourselves. The node degree of public peers is found with high precision and Sybil peers are correctly classified with very high precision and high recall if the spamming peers and the monitor are connected to all Sybil addresses.

1 Observations

During July 2021, many spam IP addresses were distributed in the Bitcoin [3] P2P network. In this report, we describe our observations and ways we found to analyze them so we could estimate the node degree of public peers and detect Sybil peers. We noticed the event because we run monitor nodes [2] that connect to all public peers in the Bitcoin network and the spam of IP addresses affected our monitor nodes because they try to connect to the spam IP addresses and the huge amount of addresses depleted their resources. Consequently, we increased the number of threads in our monitor nodes to handle the increased load. When looking at the number of unique addresses received in small messages per day, we can see that this number has increased enormously: While the monitor nodes received about 40'000 unique IP addresses per day before the event started, they now receive about 6'000'000 unique IP addresses per day and even received more than 11'000'000 unique IP addresses on July 12, 2021. On July 12, 2021, user piotr_n reported this event in the BitcoinTalk Forum [1]. piotr_n found that the behavior of the spamming peers is to connect to public peers, send them 500 ADDR messages with 10 addresses each, and then disconnect.

We found that the spam IP addresses are distributed with a timestamp in the ADDR entries that is set to about 9 minutes into the future. This increases the

number of peers that the IP address is potentially propagated to.[1] We have already observed this behavior of sending many IP addresses with timestamps into the future a couple days before on June 12, 2021.

On July 22, 2021, we set up a public peer that logs all incoming ADDR messages and installed logging of ADDR messages in two public peers that were already running. On these peers, we observed the behavior described by piotr_n: A peer sent us 5000 unique IPv4 addresses within few seconds. The IP addresses were associated with the same timestamp that was 531-532 seconds into the future when the messages were received. We analyzed the distribution of these IPv4 addresses and found that they were distributed uniformly over the IPv4 space (including IPs from reserved IPv4 address blocks like 192.168.0.0/16 and 10.0.0.0/8).

2 Extractable Information from Observations

We do not know for what purpose the spam IP addresses are sent into the network. However, we can look at the effects that the spamming has and what information can be extracted from observing the effects. When a peer[2] receives the 500 ADDR messages with 10 addresses each, the addresses are propagated because they are contained in small ADDR messages (a message is small if it has up to 10 entries). For each address, a peer chooses two connected peers to forward the address to. These peers are chosen based on, among others, the hash value of each peer's address and the address that is to be forwarded. The peer that the address was received from is excluded from the possible peers to forward the address to. If a peer has n neighbors, the peer forwards $c = 5000 \cdot 2/(n-1)$ addresses to each neighbor. The neighbor will forward these addresses only if the neighbor received 10 or fewer addresses. Thus, if c is greater than 10, we probably cannot deduce the connection between the peer receiving the spam and its neighbor. As the spamming peers connect to public peers only (see post by piotr_n) and our monitor nodes connect to all public peers, our monitor nodes are neighbors of (nearly) all peers that receive the spam addresses. Thus, our monitor nodes receive $5000 \cdot 2/(n-1)$ addresses from each peer that receives 5000 spam addresses where n is the number of connections of that peer. As we can use the future timestamps to distinguish the spam traffic from the usual traffic in the network, we can use this data to estimate the number of neighbors of each public peer.

As far as we observed, the 5000 IP addresses that are sent to one peer with one timestamp are not sent to another peer with the same timestamp. Thus, we can deduce pairs of IP addresses that belong to the same peer using the following heuristic: Say two peers with IP addresses a and b send sets of IP addresses S_a and S_b to the monitor. Continue if both sets S_a and S_b contain IP addresses with the same timestamp and are received only a couple of seconds apart. If both sets S_a and S_b are of about the same size and the intersection of both sets is not empty, then a and b belong to the same peer.

[1] Background: Peers propagate an IP address if its associated timestamp is not older than 10 minutes. Thus, an IP address is usually propagated for 10 minutes. As a peer forwards an address on average after 30 seconds, this limits the number of peers an address is propagated to. Peers accept timestamps that are up to 10 minutes into the future. Thus, if the timestamp is initially 9 minutes into the future, the address is propagated for 19 minutes and reaches more peers.

[2] We assume that a peer runs Bitcoin Core.

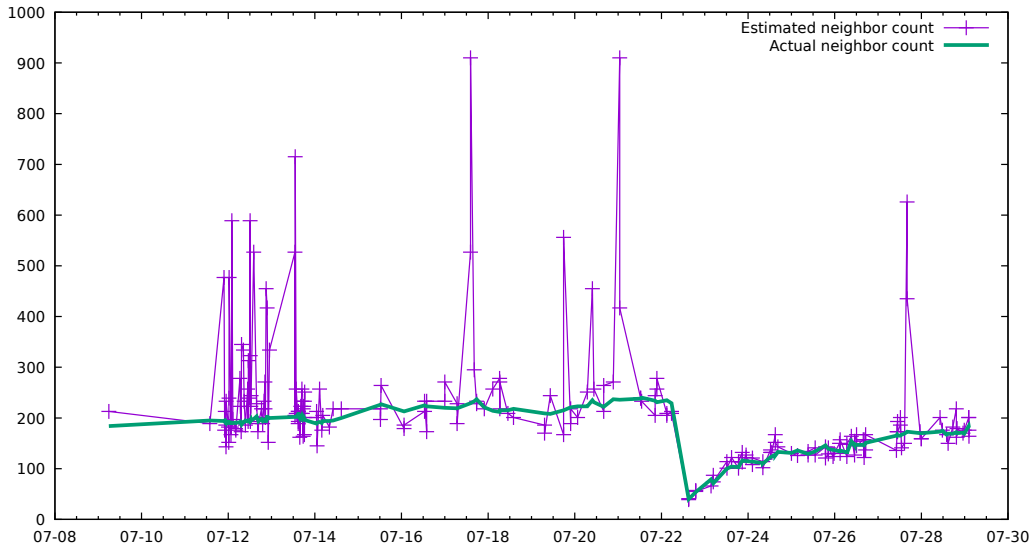

Figure 1: Estimated neighbor count and actual neighbor count of one of our peers.

3 Applying the Heuristic

3.1 Estimating the Number of Neighbors of Public Peers

We estimate the number of neighbors of each public peer and validate the estimate against one peer that we run. We analyze all ADDR messages that were received by one of our monitor nodes and have more than 10 entries. For each of these, we filter all address entries that have a timestamp that is 8 to 10 minutes into the future from the point when the ADDR message was received. Let c be the number of addresses we received with the same timestamp from the peer with IP address a. If $c > 10$, we output c and the IP address a that we received the c addresses from. The estimation of the number of neighbors of the peer with address a is $n = 1 + 5000/c \cdot 2$.

Figure 1 shows the results over the observed timespan for one of our peers. At this peer, we logged the number of neighbors of p every three minutes. The estimated neighbor count (purple) is calculated using the formula for n above and the actual neighbor count (green) is taken from the nearest logging point of our logs. While the estimated count has some outliers especially during the first days, its median fits the actual neighbor count mostly well. An explanation for the outliers during the first days would be that during this time our assumption that the spamming peers send exactly 5000 addresses was not always met. In general, it seems that for a low number of neighbors (after a restart of the peer on July 22, 2021), the precision is very high, while for higher number of neighbors the estimates tend to have more outliers.

3.2 Finding Sybil Peers

We define a Sybil peer as a peer that has multiple addresses. Common reasons for a peer being a Sybil peer are that a peer has an IPv4 and an IPv6 address or that a peer has multiple IPv6 addresses.

We analyze all ADDR messages that were received at the monitor nodes and have more than 10 entries[3]. In an ADDR message, we only consider addresses that have a timestamp that was between 0 and 10 minutes into the future when it was received by the monitor. This should reduce the processed addresses to the addresses that were originally sent by the spamming peers. We find peers p_i and p_j that have sent us the same tuple of address and timestamp and count how many such tuples we received from both p_i and p_j. The more same (address, timestamp) tuples we received from a pair of peers, the higher is our confidence that these peers are actually the same peer.

For three of our peers that are using an IPv4 and IPv6 address, we used this method to evaluate the detection of Sybil peers. For two peers that were running for the whole timespan of the event, we received 749 and 707 pairs for the correct combination of IPv4 and IPv6 address and up to four pairs for false positives. For a peer that did not forward ADDR messages at the beginning of the event until it was modified, we received 388 pairs for the correct combination of IPv4 and IPv6 address and up to two pairs for false positives. We conclude that the detection of Sybil peers works with very high precision if the spamming peers and the monitor are connected to all Sybil addresses.

The whole number of pairs of Sybil peers found with more than ten pairs over the whole timespan is 5964. The majority of these pairs is made up by a group of 284 IPv6 addresses of the same /118 subnet that seem to belong to the same peer. Without these IPv6 addresses, there are 1818 pairs of Sybil peers found with more than ten pairs. Most of these pairs of IP addresses are an IPv4 and IPv6 address, however, there are some pairs of addresses that are both IPv4 or IPv6 addresses. Such pairs of IPv6 addresses could be by peers using the IPv6 privacy extensions that created a new IPv6 address and are reachable via the new IPv6 address but are still reachable via the old IPv6 address. We manually inspected some pairs of IPv4 addresses and found that they were located in the same region which indicates that they might actually be assigned to the same machine.

4 Conclusion

An unknown party distributed a huge amount of spam IP addresses in the Bitcoin P2P network during July 2021. We do not know the party's intentions, however, we have shown that the effects of the spam IP addresses can be used to find the degree of public peers and to find Sybil peers. While the number of neighbors might not seem to be very sensitive information at the first glance, it might reveal peers that are well-connected and make them preferred targets for DoS attacks. The attribution of multiple addresses to the same peer could lead to a reduction of user privacy especially if a peer uses an anonymity network such as Tor and IP at the same time. We encourage a discussion of whether methods to reveal such information should be prevented and how countermeasures could be implemented.

[3]Then, the sending peer has received the sent addresses in more than one ADDR message (if the peer is Bitcoin Core) and it is, thus, more likely that the peer has directly received the addresses from the spamming peers.

References

[1] BitcoinTalk Forum (2021), https://bitcointalk.org/index.php?topic=5348856.msg57469495

[2] Decentralized Systems and Network Services Research Group: Bitcoin network monitoring (2021), https://dsn.kastel.kit.edu/bitcoin/

[3] Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System. Tech. rep. (2008)