T²-SCORE: A REPUTATION SCORE WITH AN ADAPTIVE Q-LEARNING BASED TRUST MODEL TO AVOID MALICIOUS MOBILITY DATA ENDORSEMENTS IN BLOCKCHAIN CONSENSUS

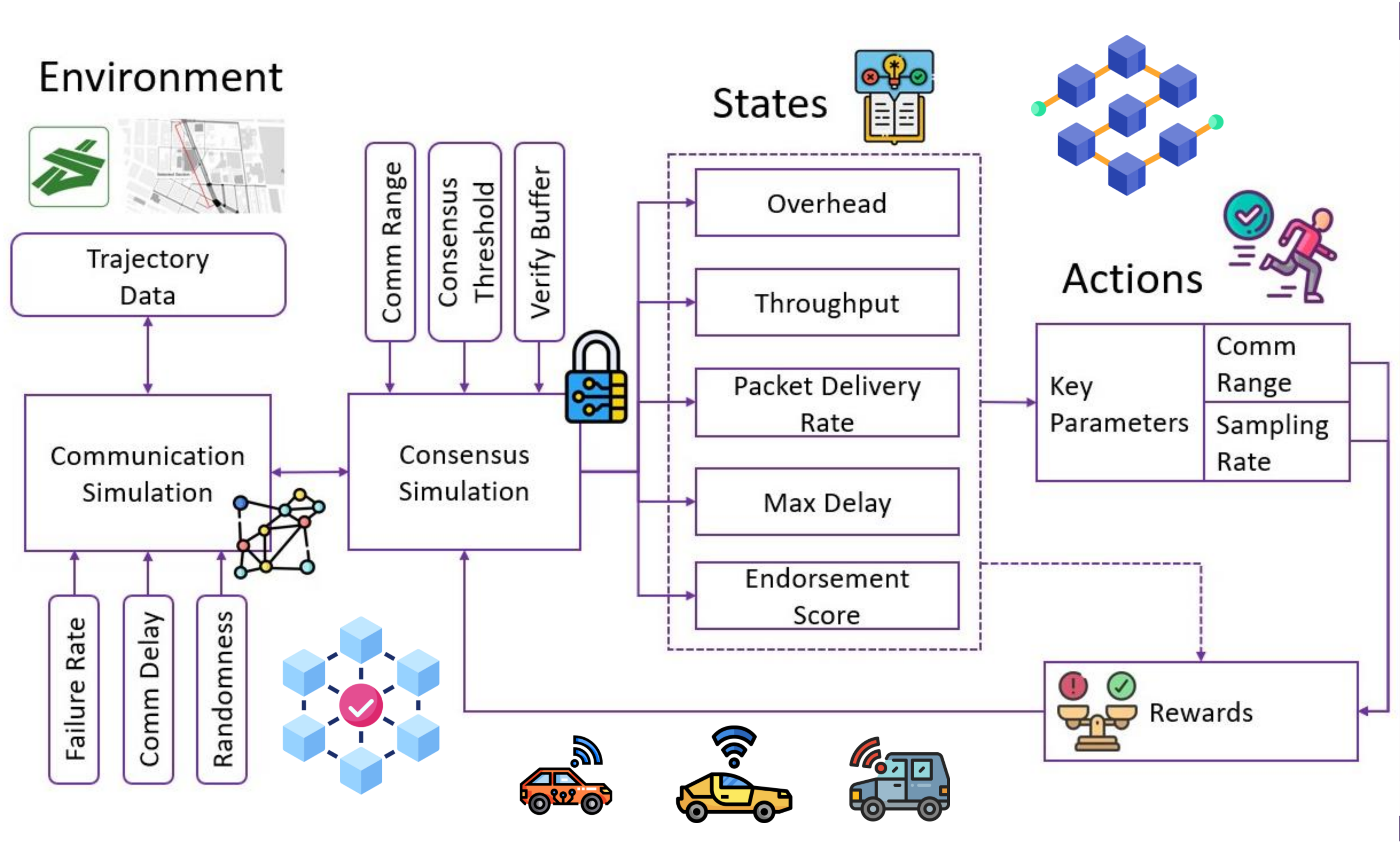
Junaid Ahmed Khan[a], Fan Zuo[b], Kaan Ozbay[b]
[a]Western Washington University, [b]New York University

Paper 24-04759 | Contact: Junaid.khan@nyu.edu

Abstract

GPS mobility data is vital for Intelligent Transportation Systems (ITS). Blockchain technology, used for storing data from devices like smartphones and autonomous vehicles, faces efficiency issues due to its consensus-based processing. This paper leverage reinforcement learning to optimize the consensus process to reduce devices' resource utilization and increase overall endorsements over time. Nearby mobility generating devices or nodes participating in a consensus process learn the optimal configuration and thresholds for the validation of each other's mobility data and adapt strategies to reduce their communication and computing resources consumption. Additionally, we define a novel reputation scoring mechanism, namely, Trajectory trustworthiness score or T2-Score for nodes to gain trust over time based on the frequency and freshness of peer endorsed mobility data information on blockchain, where other nodes use the score to determine whether the mobility data regarding a node is trustworthy for navigation and other safety-related mission critical decisions. Simulation results from New York City data show the efficiency of the proposed scheme as it achieves lower overhead with higher network throughput.

Framework



Data collection:

- SUMO generate trajectory based on the Flatbush Avenue Scenario.
- Full trajectory file contains 6484 Vehicles in 2 hours. Test set contains 139 vehicles in 100 seconds.

Control Variables:

- Communication Range: 20~210 meters.
- Sampling Rate: The frequency of launching a consensus process, 0.1~9.5seconds
- # of Random Seeds: 20

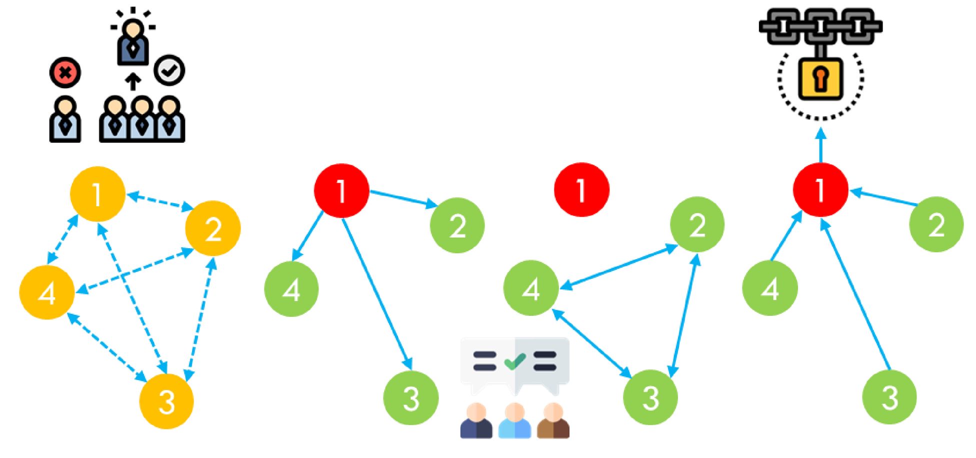
Methods, Data, and Settings

1. Consensus for proof of mobility

Nodes identify a consensus initiator based on the highest T2-Score. They predict, observe, and validate each other's trajectories against a deviation threshold. Validated trajectories are efficiently recorded in a distributed ledger, prioritizing privacy and resource conservation.

T²-Score: Trajectory Trustworthiness Score

$$\Omega_v(t+1) = (1-\alpha)\left(\frac{\xi_v(t+1)}{\Delta t(t+1)}\right) + \alpha\Omega_v(t)$$



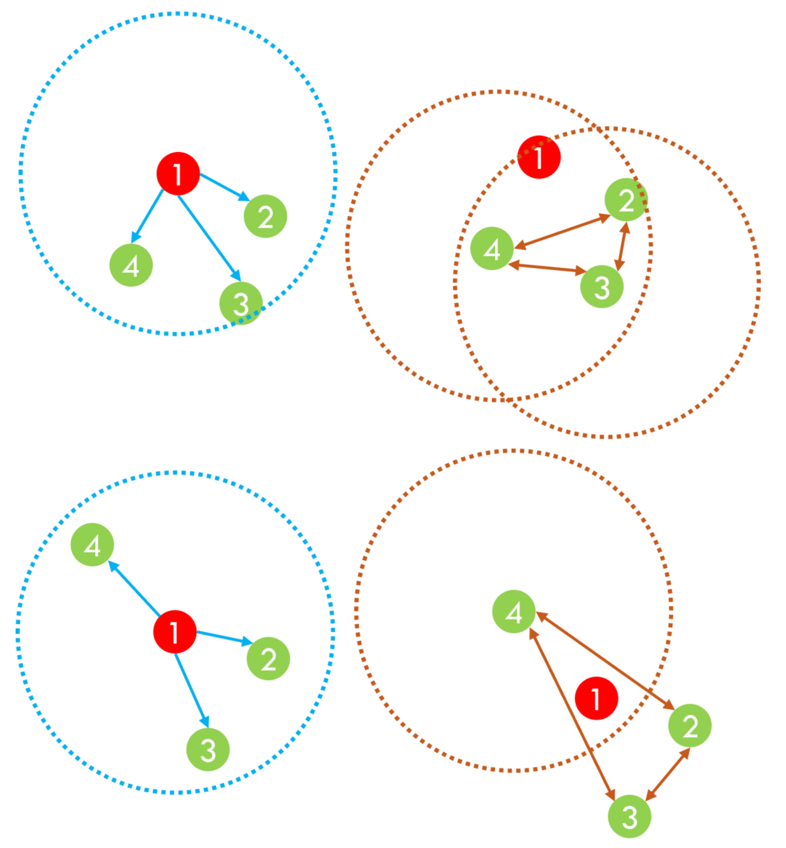
Proof of mobility consensus

A node can be in one of three states:

- Follower: Default state for nodes, where they participate passively in the network.
- Candidate: Nodes transition to the candidate state in the absence of communications from the current leader, initiating an election process.
- Leader: Upon securing a majority of votes from the participating nodes(> 50%) a candidate ascends to the leader state, acquiring the authority to orchestrate system modifications.

Algorithm 1 T^2-Score with Q-learning

```
INPUT: S(l,t), A(l,t) for v ∈ V,
OUTPUT: U(S(l,t)) for v ∈ V,
for each node v ∈ V do
    repeat
        Initial State S_0 = (v_0, g_0), Q = 0
        for i = t_1, t_2, ..T do
            Scan for existing initiator node for next time slot
            if No initiator node detected then
                Broadcast initiator announcement after t_b
                Select action A based on ε-greedy policy
            end if
            Observe S, predict r̂
            Send Threshold λ and deadline and other configuration
            Q(S,A) = Q(S,A) + β[r̂ + γ max_{A'} Q(S',A') − Q(S,A)]
            S = S'
        end for
    until Q(S,A) converge
    g(l,t) = arg max_A Q(g(l,t−1),A)
else
    g(l,t) = g(l,t−1)
end for
```



Two different scenarios of the topology structure of the nodes and the impact on the consensus process.

2. Q-learning based block creation

Consensus process modeled as a Markov Decision Process, the initiator node sets each round's parameters, aiming to validate maximum nodes' mobility data on the blockchain while efficiently managing resources and participation limits. The nodes learns and adapts in real time without sharing raw data.

$$U_\pi(S) = E\left[\sum_l \sum_t \gamma(l,t) r_S^{A\pi} | S(0,0) = S\right]$$

- **State**: The state is defined by the tuple $S(l,t) = \{v(l,t), g(l,t)\}$ with the initiator node $v(l,t)$ and the consensus configuration function $g(l,t)$ to decide the time requirements.
- **Action**: The action set is defined for the node responsible for the consensus round, where it learns the configuration function to be used for the next round based on the feedback from the current consensus round.
- **Reward**: The reward of the action $A(S(l,t))$ is defined as $r_S^A(l,t) = f(S(l,t), A(l,t), \Omega_v(t), \forall v \in V)$. Thus, an increase in T²-Score positively contributes to an individual node reward.

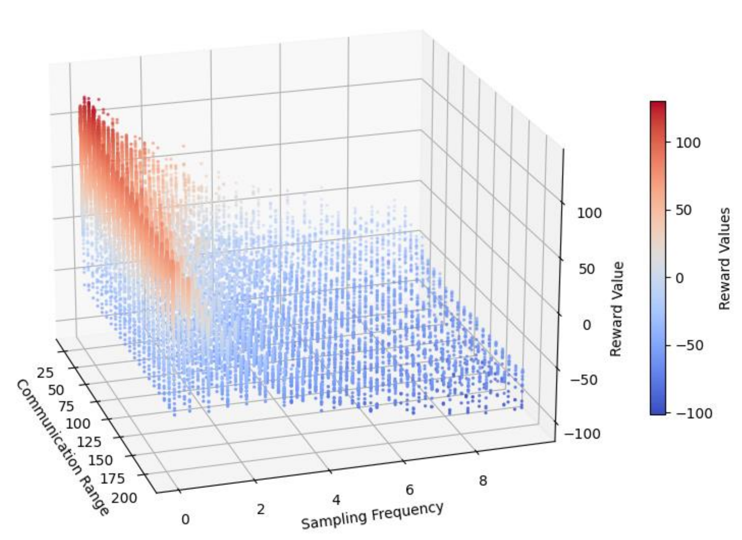
Results and Takeaways

Objective:

- Design a secure and privacy-preserving consensus mechanism for connected vehicles.
- Examine state transitions during the learning process.
- Exhaustively explored all possible combinations (discrete) of key parameters within their feasible intervals to capture the parameter space adequately.

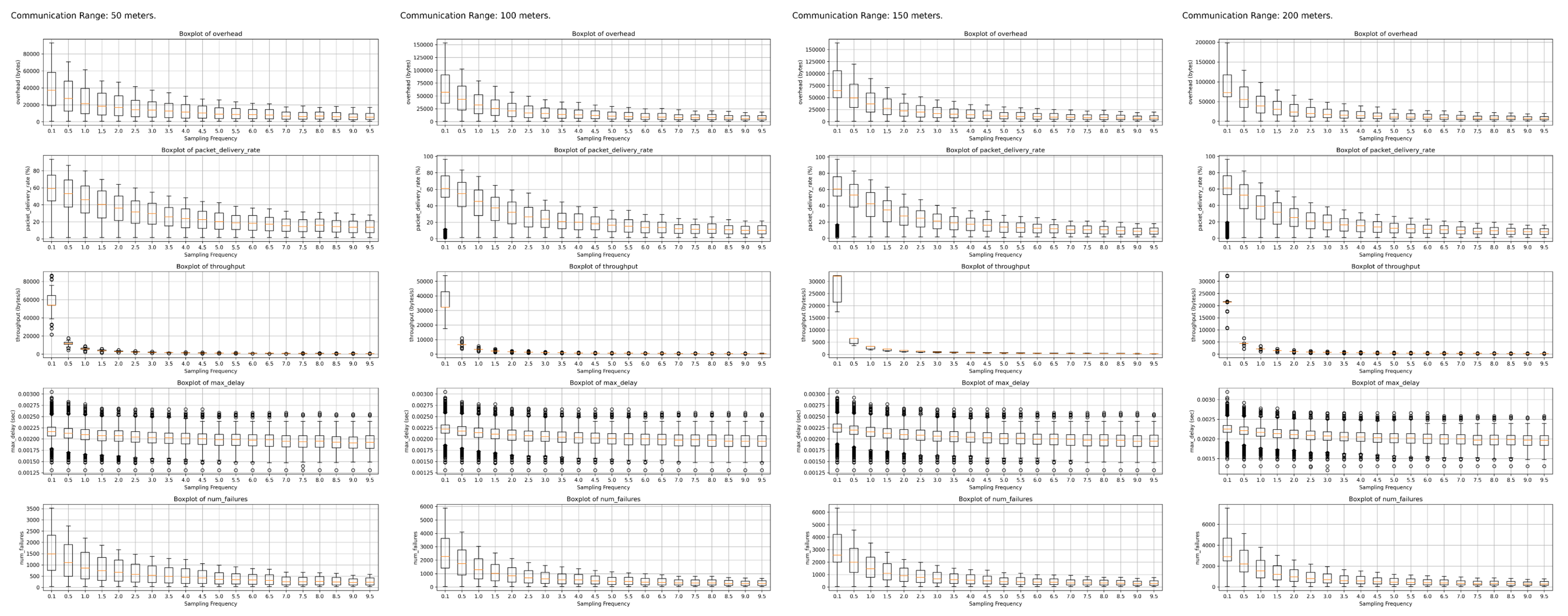
Key Performance Metrics

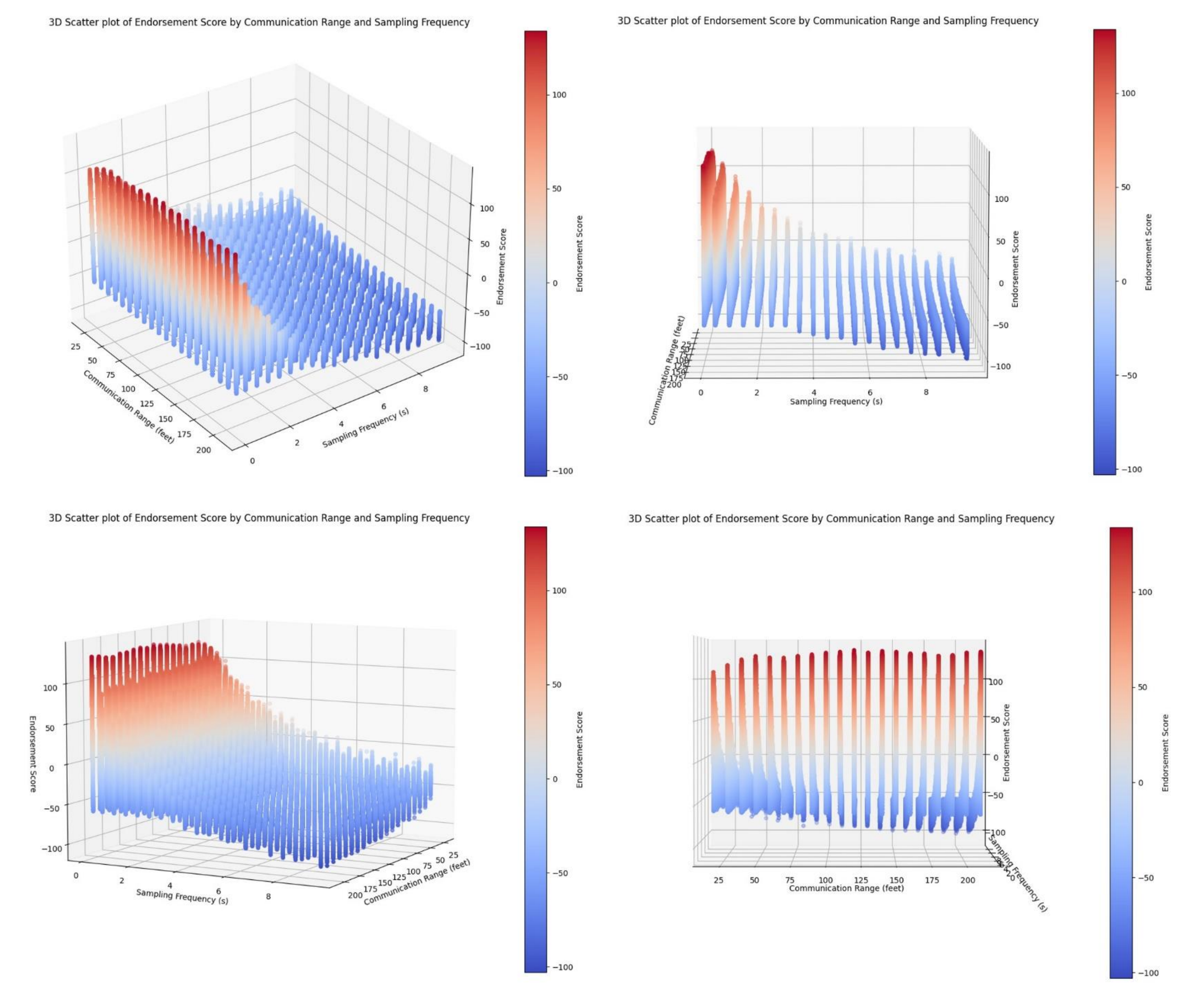
- **Overhead (bytes):** the amount of data exchanged between nodes participating in the consensus process. Lower values are desirable.
- **Packet Delivery rate (%):** the percentage of vehicles endorsed out of the total number of vehicles in the study area. Higher values are preferable.
- **Throughput (bytes/s):** a measure of network utilization, where lower throughput would indicate low resource utilization by the consensus process. Lower values are preferable.
- **Max delay (s):** the maximum delay during the consensus process. Lower values are preferable.



Reward Distribution

The boxplots of essential performance measurement outputs over different communication ranges.





The distribution of the total T² Score along with all the ranges and sampling frequency rate combination.

Takeaways

- The total T² score distribution is particularly sensitive to the sampling frequency.
- Throughput was somewhat deprioritized or "sacrificed" to elevate the importance of other key parameters.
- The group with a 0.1s sampling rate exhibits significantly higher values than other groups (up to five times).
- Beyond a certain threshold, any reduction in overhead becomes negligible.

Future Work

Future directions include scaling the proposed solution to a large number of nodes and further optimizing the delays and overhead. Additionally, we plan to develop a prototype to demonstrate the practical aspect of the blockchain-based reputation and trust model.

The work in this paper is sponsored by C2SMARTER Tier 1 University Transportation Center Led by New York University.