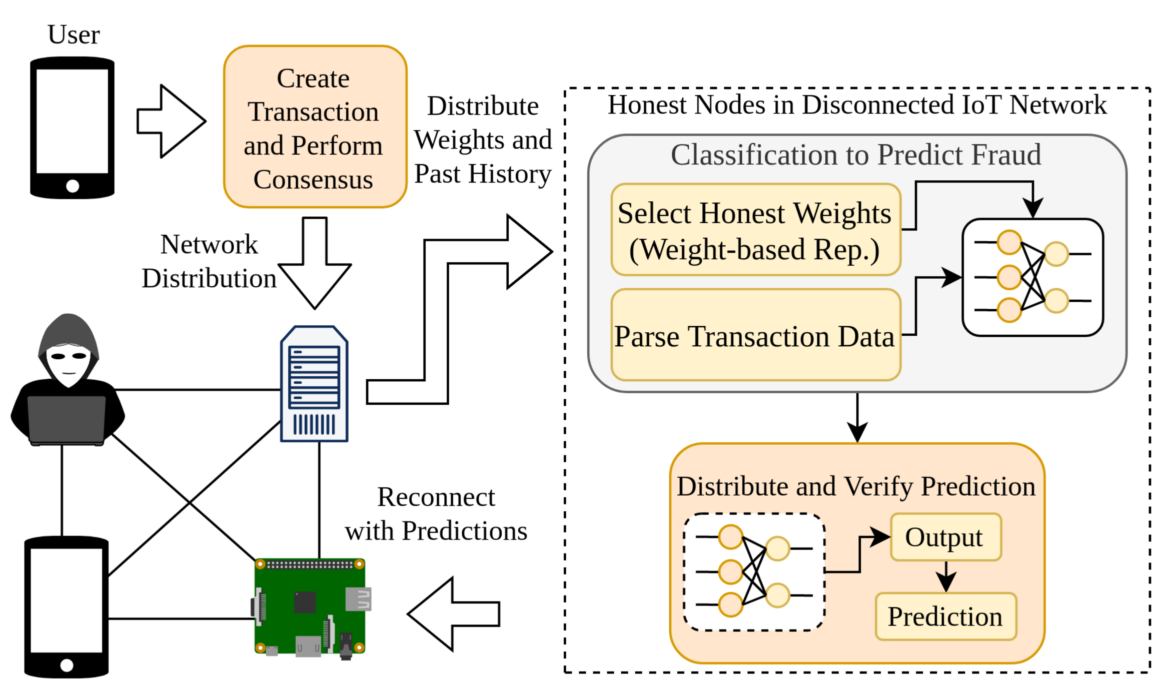
A Secure Reputation-based Consensus Scheme for Robust Decision-making in a Lightweight Machine-learning Framework for IoT Blockchain Networks

Charles Rawlins and Jagannathan Sarangapani

Missouri University of Science and Technology

Introduction and Objectives

- A blockchain is a decentralized network for recording arbitrary data in a transaction format, and cannot be implemented on IoT directly
- Seek to condense blockchain knowledge with machine-learning (ML) classification in distributed training scenario with adversaries
- Proposed effort addresses ML vulnerabilities with data poisoning
- Effort objectives are:
  - Secure ground-truth in distributed ML training setting
  - Develop reward/reputation function to determine node intent



Background and Threat Model

- Classic blockchain consensus like Proof of Work (PoW) [1] has been proven practically robust but incompatible with limited nodes
- Related ML distributed training has seen more attention recently [2], but are not robust in the presence of many adversaries
- Consider distributed parallel SGD (DPSGD) algorithm [2] averaging weight samples ***W*** to update a deep network as a base classifier, a second network uses Q-learning to observe ***W*** as state-space ***S***
- To poison and delay learning, consider two attacks for independent adversaries poisoning ***W*** with noise added to ***k***-th ***W*** values:

$$W_k = W_k + U(0,1) \quad W_k = W_k - KL(W_0, W_k)(W_k(t-1)) + \sqrt{\frac{2}{\beta\eta}}\epsilon$$

with the Uniform and Maximal Action-distance (MAD) attacks [3] adding independent and dependent noise respectively

- To capture both attacks, setup a Bayesian Game which models node knowledge about the network with tuple (*N,A,Θ,p,U,S(t)*):
  - *N* nodes with {N_i = Honest node, N_j = Suspicious Node}
  - *A* actions, where *A* = {A_i = {Accept, Reject}, A_j = {Attack, Not Attack}
  - Node type *Θ* = {$Θ_H$, $Θ_B$}, for honest $Θ_H$ and malicious $Θ_B$
  - Prior *p* at each node developed from other node actions
  - Utility *U*: *A x Θ → R*
  - State space *S(t)*, defined as S_j = {$\boldsymbol{W_j}$, D_j}, contains the weight array and auxiliary data D_j, such as node ID, connection delay, etc.
- In consensus, honest nodes seek global optimal policy, while Byzantine nodes maximize selfish reputation

$$\pi_*^H = argmax_{\pi^H} J^H(\pi^H, \pi^B) \quad \pi^B = max_{\pi^B} \mathbb{E}[\Sigma_{t=0}^{\infty} \gamma(t) Rep_{t+1}^j(S_t, A_t)]$$

Proposed Methodology: WBR and PoH

- *N* nodes execute DPSGD, updating local ***W***, and gossip experience to other nodes, with adversaries perturbing values
- N_i scales $\boldsymbol{W_j}$ with C_{ij}, then conducts SGD update

$$W_i(t) = \sum_{j=1}^{n} C_{ij} W_j(t) \quad W_i(t+1) = W_i(t) - \alpha \nabla f(X_i)$$

- To associate larger C_{ij} values with honest nodes, need to determine which W_j values are closer to ideal weights ***W****

$$\mathcal{W}^* \in \arg\min_{\mathcal{W} \in \mathbb{R}^D} \mathbb{E}(f(\mathcal{W}, X))$$

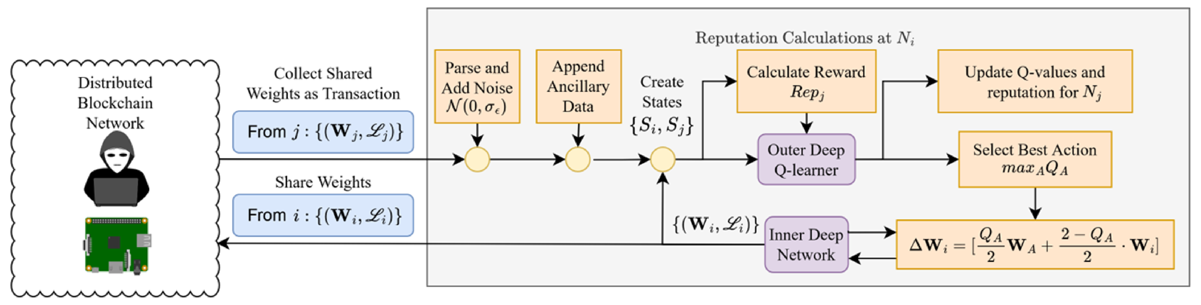
- Honest nodes will maximize U_j by creating values closer to W*. Undisturbed W_j values were proven to be within

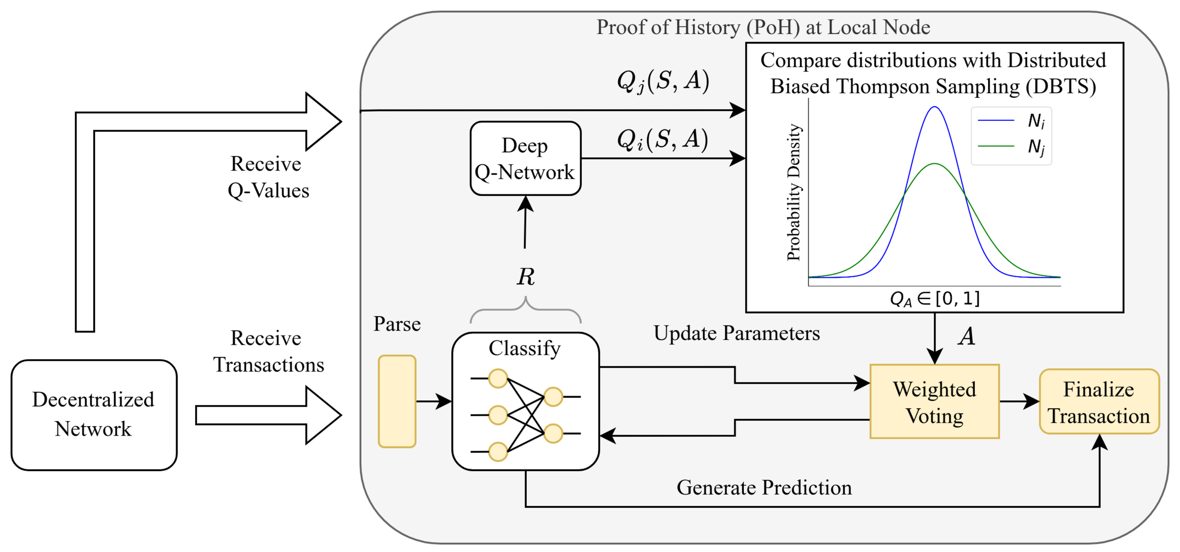
$$\alpha^2 \mathbb{E}(Var(\nabla f_i)) \leq \mathbb{E} \left\| \frac{1}{n} \sum_{i=1}^{n} W_i^* - W_i \right\|^2$$

- Nodes are selected using Weight-based Reputation (WBR):

$$Rep_j = ||S_i - S_j|| + ||W_i - W_j| - \alpha^2 \mathbb{E}(Var(\nabla f_i))||$$

- C_{ij} values are created using min-max-scaled Q-values

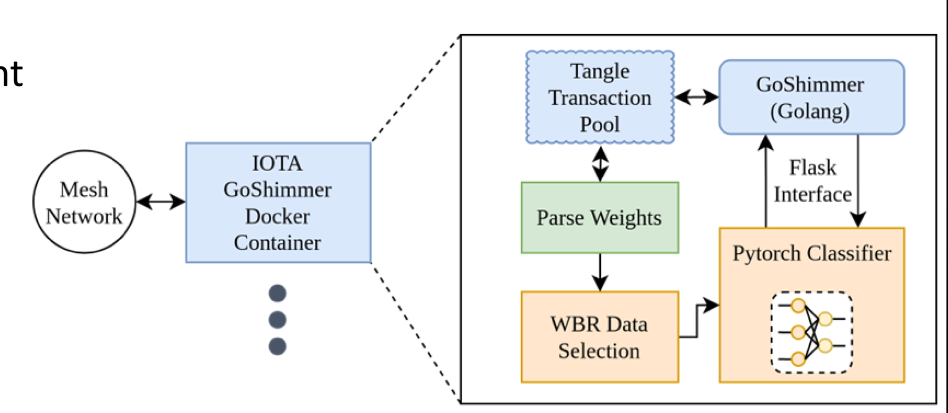


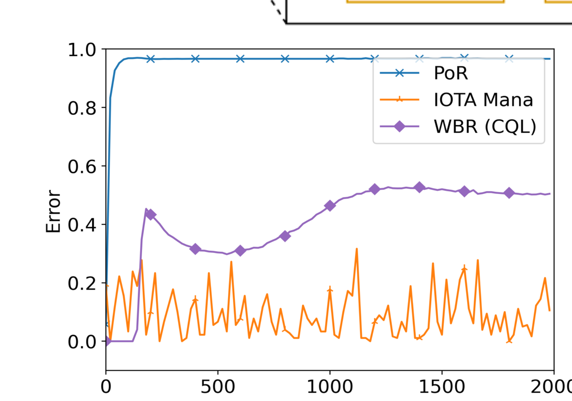


- For consensus and blockchain decisions, array of Q-values are distributed along with ***W*** and compared
- Q-value batches are processed with new action-selection called Distributed Biased Thompson Sampling (DBTS)
- DBTS filters outlier distributions using Levene's Test to create a circle of trust and prevent reputation poisoning from adversaries
- In a blockchain conflict, a weighted threshold voting scheme finalizes transaction based on confidence *C* in a desired label
- Total scheme is called Proof-of-History (PoH), since learning from blockchain history is considered for making decisions

Simulation Results and Discussion

- Tested WBR in both controlled Python and realistic blockchain environment using the IOTA protocol as a base scheme
- Uniform and MAD attacks tested with a variety of Q-learning algorithms and compared against other schemes
- Bottom: Controlled environment shows superior performance for WBR with Deep Q-network (DQN) to both uniform and MAD attacks
- Right: WBR was compared to other blockchain protocols executing similar attack scenarios. WBR has consistent controlled error between nodes compared to the IOTA mana reputation system [4] and Proof-of-Reputation [5]
- Below: WBR is more time/space complex compared to other schemes

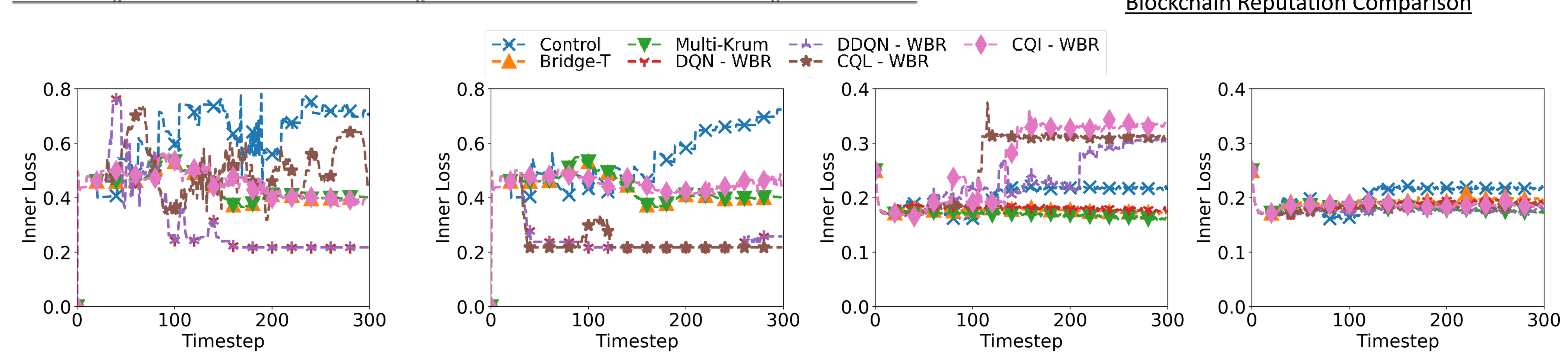




Blockchain Reputation Comparison

Single-step Complexity Comparison

Algorithm	Time	Space	Communication
Multikrum	$\mathcal{O}(mk^2d)$	$\mathcal{O}(md)$	$\mathcal{O}(kd)$
Bridge-T	$\mathcal{O}(bkd)$	$\mathcal{O}(kd)$	$\mathcal{O}(kd)$
WBR	$\mathcal{O}(kQ) \to \mathcal{O}(k(dN + (L-1)N^2))$	$\mathcal{O}(nd+Q) \to \mathcal{O}(nd + Nd + (L-1)N^2)$	$\mathcal{O}(kd)$



Controlled Uniform Attack Comparison

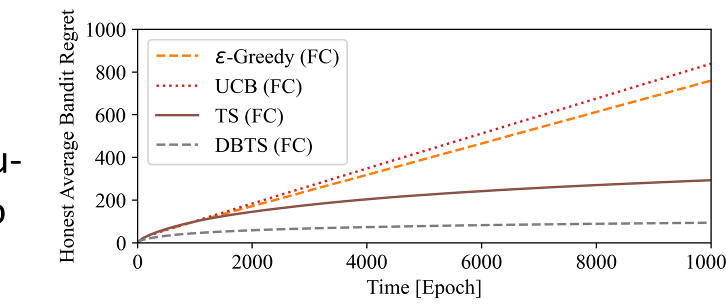
Controlled MAD Attack Comparison

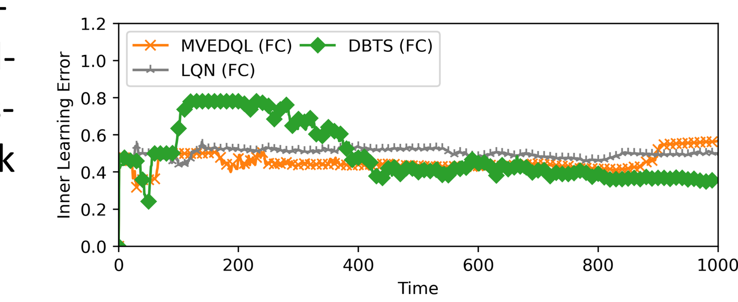
Realistic Uniform Attack Comparison

Realistic MAD Attack Comparison

Simulation Results and Discussion Cont'd.

- DBTS in a controlled setting with the uniform attack shows superior performance compared to other bandits
- Implementing DBTS with a distributed DQN protocol in WBR also improves performance to classifier training with uniform attack





- WBR successfully repels poisoning attacks to distributed training in controlled scenarios for both independent and dependent noise
- Weak averaging consensus with selection of only the 'best' node in WBR alone fails to reduce error in reputation and realistic scenarios
- Resource consumption for WBR, compared to PoW in IOTA, is roughly comparable in memory and CPU consumption, but has better throughput
- PoH can be more robust than other protocols with high *Rep* and *C*

Average Resource Consumption Comparison

Method	IOTA PoW	WBR (Train)	WBR (Test)
Memory [%]	16.4	22.03	**14.11**
CPU [%]	16.0	35.5	21.86
Power [W]	6.32E-6	17.3	8.43E-6
Blocks/Second [bps]	20.93	18.75	**29.3**
Block Delay [ms]	12.4	**0**	**0**

Blockchain Consensus Complexity Comparison

Method	Verification Delay	Message	Fault-tolerance
Nakamoto PoW	$\mathcal{O}(MlogN)$	$\mathcal{O}(M)$	N/2
PoS	$D\Omega(K)$	$MN\Theta(1)$	N/2
IOTA FPC	$\mathcal{O}(KM)$	$\mathcal{O}(N)$	$\sim (N/2)$
PoH (This Work)	$\mathcal{O}(mnk)$	$\mathcal{O}(N)$	$\sum_{i \in N_H} Rep_i C_i$

Conclusions and Future Work

- WBR provides a robust technique for securing ground-truth in an IoT network and effective distributed training for intelligent blockchain
- PoH consolidates opinions and finalizes transactions with a robust bandit update scheme. Threat model for PoH could be expanded in the future
- Future work will explore replacing deep networks with auditable decision trees and reducing computation with SGD-alternatives

Acknowledgements

- This work was supported by the Graduate Assistance in Areas of National Need (GAANN) national fellowship program.

References

[1] Nakamoto, Satoshi, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2009. https://bitcoin.org/bitcoin.pdf (accessed Aug. 22, 2023).

[2] X. Lian, C. Zhang, H. Zhang, C.-J. Hsieh, W. Zhang, and J. Liu, "Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent." arXiv, Sep. 11, 2017.

[3] H. Zhang et al., "Robust Deep Reinforcement Learning against Adversarial Perturbations on State Observations," in Advances in Neural Information Processing Systems, 2020, pp. 21024–21037

[4] IOTA Foundation, "The Coordicide," 2019, [Online]. Available: https://files.iota.org/papers/Coordicide_WP.pdf

[5] F. Gai, B. Wang, W. Deng, and W. Peng, "Proof of Reputation: A Reputation-Based Consensus Protocol for Peer-to-Peer Network," in Database Systems for Advanced Applications, J. Pei, Y. Manolopoulos, S. Sadiq, and J. Li, Eds., Cham: Springer International Publishing, 2018, pp. 666–681.