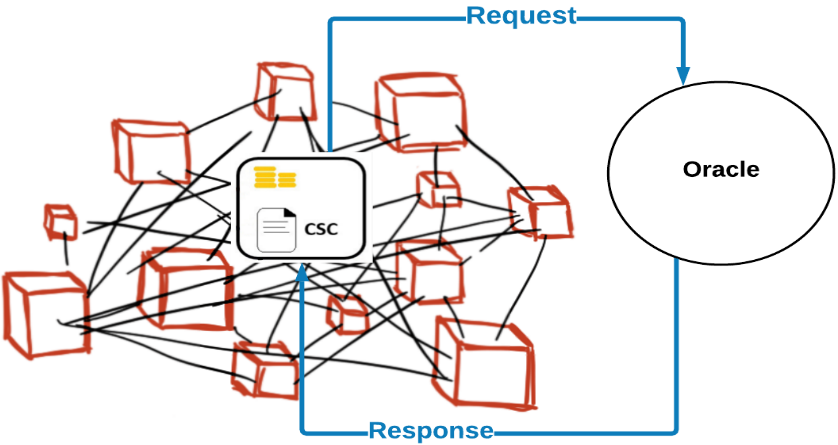
Bet and Attack: Incentive Compatible Collaborative Attacks Using Smart Contracts

Zahra Motaqy[1] Ghada Almashaqbeh[1] Behnam Bahrak[2] Naser Yazdani[2]

[1]UConn, [2]University of Tehran

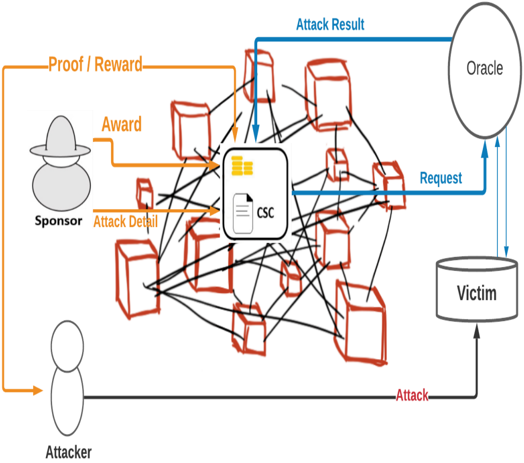
GameSec 2021

Blockchain, Smart Contract, and Oracle



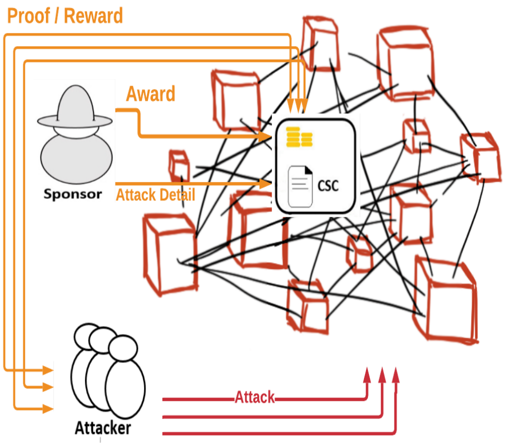
Criminal Smart Contract I

Solo Attacker on Real-World Target



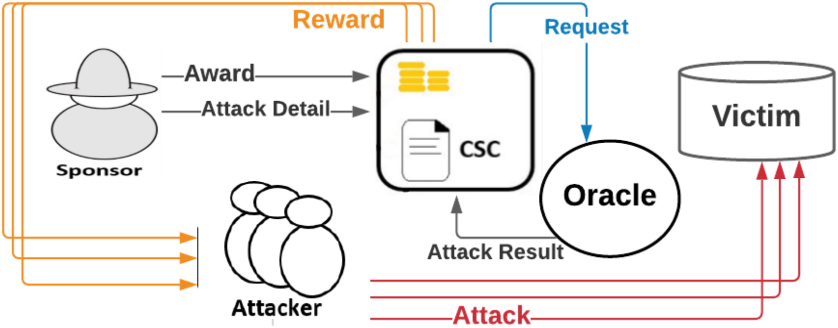
Criminal Smart Contract II

Collaborative Attack on Blockchain / Cryptocurrency



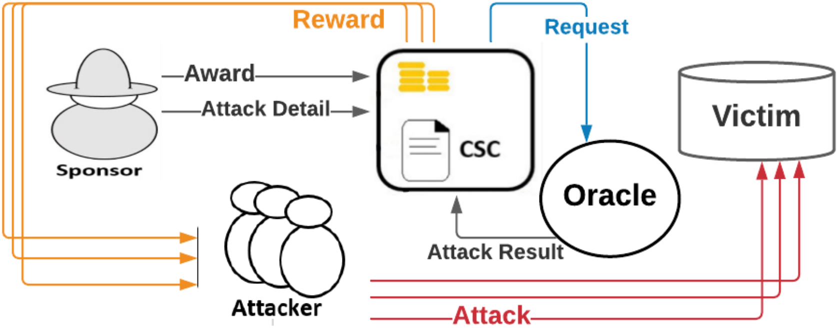
Criminal Smart Contract III

Collaborative Attack on Real-World Target



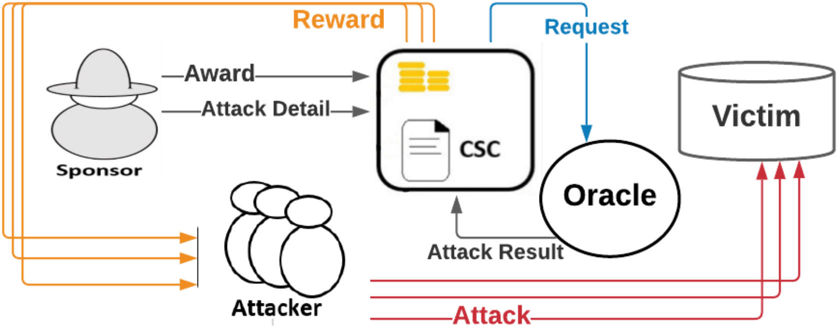
Criminal Smart Contract III

Collaborative Attack on Real-World Target



- *How to measure each attacker's contribution?*

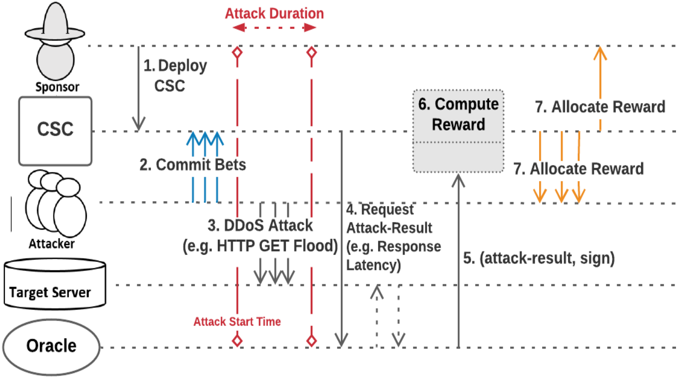
Collaborative Attack on Real-World Target



- *How to measure each attacker's contribution?*
- *When the Attack is successful?*

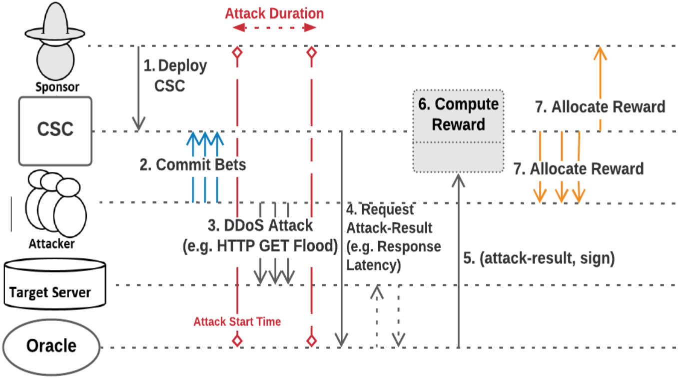
Attack Model
Use case: Distributed Denial of Service attacks

- **Phase I**: Design and Deployment of CSC



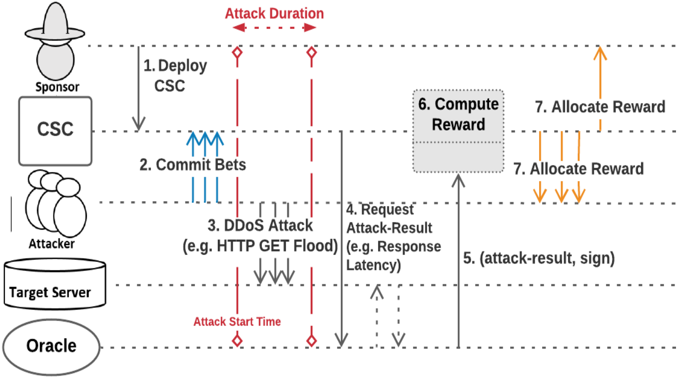
Attack Model
Use case: Distributed Denial of Service attacks

- **Phase I**: Design and Deployment of CSC
- **Phase II**: The Attack



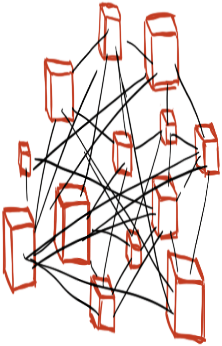
Attack Model
Use case: Distributed Denial of Service attacks

- **Phase I**: Design and Deployment of CSC
- **Phase II**: The Attack
- **Phase III**: Reward Allocation



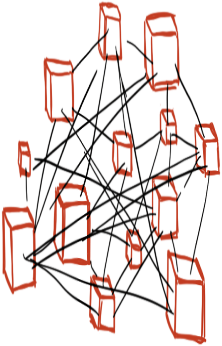
Blockchain is (pseudo) anonymous I

- Attackers will create multiple address (bets) if it get them more reward



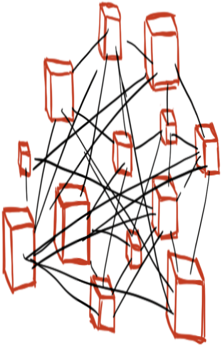
Blockchain is (pseudo) anonymous I

- Attackers will create multiple address (bets) if it get them more reward
- **Private information**
- The number of attackers n
- The amount of their individual bets bet_i



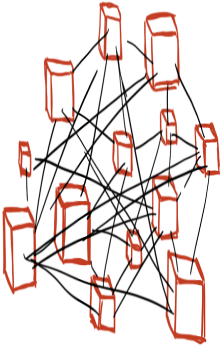
Blockchain is (pseudo) anonymous II

- First, in game model we assume they have bet honestly (under one address (one bet)



Blockchain is (pseudo) anonymous II

- First, in game model we assume they have bet honestly (under one address (one bet)
- Then, in incentive mechanism model we show why they will bet honestly



Game Model I

Interdependent Attackers Game (IAG)

- e_{th}: the total traffic needed for a successful attack
- ω_S: the award of the sponsor
- bet_i: the bet value of the i^{th} attacker $bet_{tot} = \sum_{i \in N} bet_i$
- $t_i = \frac{bet_i}{\omega_S}$: the private information that i^{th} attacker has and it represents his type

Game Model II

Interdependent Attackers Game (IAG)

- **Choice Variable**

  - e_i: the relative contribution of the i^{th} attacker in e_{th}

  $e_{tot} = \sum_{i \in N} e_i$

Game Model II

Interdependent Attackers Game (IAG)

- **Choice Variable**

  - e_i: the relative contribution of the i^{th} attacker in e_{th}

  $e_{tot} = \sum_{i \in N} e_i$

- **Model Parameters**

  - N: the attackers set ($|N| = n$)
  - E: the set of all action profile $\hat{e} = (e_1, \ldots, e_n)$
  - T: the set of all type profiles $\hat{t} = (t_1, \ldots, t_n)$

Game Model III

- Reward Allocation Function (linear with respect to e_i)

$$R(bet_i, e_{tot}) = M \cdot e_{tot} \cdot \frac{bet_i}{bet_{tot}} \tag{1}$$

$$R(t_i, e_{tot}) = M \cdot t_i \cdot e_{tot} \cdot \left(\frac{bet_{tot}}{\omega_S}\right)^{-1}$$

Game Model III

- Reward Allocation Function (linear with respect to e_i)

$$R(bet_i, e_{tot}) = M \cdot e_{tot} \cdot \frac{bet_i}{bet_{tot}} \quad (1)$$

$$R(t_i, e_{tot}) = M \cdot t_i \cdot e_{tot} \cdot \left(\frac{bet_{tot}}{\omega_S}\right)^{-1}$$

- Cost Function (convex with respect to e_i)

$$C(e_i) = \alpha \cdot \frac{\exp(e_i) - 1}{e_{max} - e_i} \qquad \forall i = 1, \ldots, n \quad (2)$$

Game Model III

- Reward Allocation Function (linear with respect to e_i)

$$R(bet_i, e_{tot}) = M \cdot e_{tot} \cdot \frac{bet_i}{bet_{tot}} \tag{1}$$

$$R(t_i, e_{tot}) = M \cdot t_i \cdot e_{tot} \cdot \left(\frac{bet_{tot}}{\omega_S}\right)^{-1}$$

- Cost Function (convex with respect to e_i)

$$C(e_i) = \alpha \cdot \frac{\exp(e_i) - 1}{e_{max} - e_i} \qquad \forall i = 1, \ldots, n \tag{2}$$

- Utility Function (concave with a unique maximum)

$$U(t_i, e_i, e_{tot}) = R(t_i, e_{tot}) - C(e_i) - t_i \cdot \omega_S \tag{3}$$

Equilibrium Analysis I

- Best-response strategy of a rational player in IAG

$$S^*(t_i, \hat{e}_{-i}) = \underset{e_i \in [0,1]}{\arg\max}\ U\left(t_i, e_i,\ \hat{e}_{-i}\right) \tag{4}$$

- $$-\alpha \cdot \frac{\exp\left(e_i\right)}{e_{max} - e_i} - c \cdot \frac{\exp\left(e_i\right) - 1}{\left(e_{max} - e_i\right)^2} + \frac{\omega_S \cdot t_i \cdot \left(\omega_S + bet_{tot}\right)}{bet_{tot}} = 0 \tag{5}$$

Equilibrium Analysis I

- Best-response strategy of a rational player in IAG

$$S^*(t_i, \hat{e}_{-i}) = \arg\max_{e_i \in [0,1]} U(t_i, e_i, \hat{e}_{-i}) \quad (4)$$

- $$-\alpha \cdot \frac{\exp(e_i)}{e_{max} - e_i} - c \cdot \frac{\exp(e_i) - 1}{(e_{max} - e_i)^2} + \frac{\omega_S \cdot t_i \cdot (\omega_S + bet_{tot})}{bet_{tot}} = 0 \quad (5)$$

The only parameters (other than t_i) that determine $S^*(t_i) = e_i^*$ are the cost of the required attack traffic α and the quantity $\frac{bet_{tot}}{\omega_S}$

Equilibrium Analysis II

$S^*(t_i)$ is a strongly dominant strategy that is the best response regardless of $\hat{e}_{-i}$

Theorem

IAG has a Strong Dominant Strategy Equilibrium

Equilibrium Analysis III

Now we know

The contribution of each attacker with type t_i: $S^*(t_i) = e_i^*$

Equilibrium Analysis III

Now we know

The contribution of each attacker with type t_i: $S^*(t_i) = e_i^*$

We want to know

The attack result and the payments in the equilibrium of the game

- $\sum_{i \in N} S^*(t_i) = e_{tot}^*$
- $p_i(\hat{t}) = R(t_i, AR(\hat{t})) - t_i \cdot \omega_S$

Equilibrium Analysis III

Now we know

The contribution of each attacker with type t_i: $S^*(t_i) = e_i^*$

We want to know

The attack result and the payments in the equilibrium of the game

- $\sum_{i \in N} S^*(t_i) = e_{tot}^*$
- $p_i(\hat{t}) = R(t_i, AR(\hat{t})) - t_i \cdot \omega_S$

We need to know

Attacker's true bets

Equilibrium Analysis III

Now we know

The contribution of each attacker with type t_i: $S^*(t_i) = e_i^*$

We want to know

The attack result and the payments in the equilibrium of the game

- $\sum_{i \in N} S^*(t_i) = e_{tot}^*$
- $p_i(\hat{t}) = R(t_i, AR(\hat{t})) - t_i \cdot \omega_S$

We need to know

Attacker's true bets

Will attackers bet honestly?

Incentive Mechanism Model I

Mechanism Formulation

- $AR(\hat{t}) = \sum_{i \in N} S^*(t_i) = \sum_{i \in N} e_i^* = e_{tot}^*$: Attack Result Function

Mechanism Formulation

- $AR(\hat{t}) = \sum_{i \in N} S^*(t_i) = \sum_{i \in N} e_i^* = e_{tot}^*$: Attack Result Function
- $G : T \to O$: Outcome Function, $o = (e_{tot}^*, \hat{p})$

  non-monetary part

Mechanism Formulation

- $AR(\hat{t}) = \sum_{i \in N} S^*(t_i) = \sum_{i \in N} e_i^* = e_{tot}^*$: Attack Result Function
- $G : T \rightarrow O$: Outcome Function, $o = (e_{tot}^*, \hat{p})$
  non-monetary part
- $V(e_{tot}^\star, t_i) = V(t_i) = -(C(S^\star(t_i)) + k \cdot \delta)$: Valuation Function

Mechanism Formulation

- $AR(\hat{t}) = \sum_{i \in N} S^*(t_i) = \sum_{i \in N} e_i^* = e_{tot}^*$: Attack Result Function
- $G : T \rightarrow O$: Outcome Function, $o = (e_{tot}^*, \hat{p})$
  non-monetary part
- $V(e_{tot}^\star, t_i) = V(t_i) = -(C(S^\star(t_i)) + k \cdot \delta)$: Valuation Function
- $U(t_i, o) = V(t_i) + p_i$

Incentive Mechanism Model II

Theorem

The proposed direct mechanism modeling our CSC-based collaborative attacks is Dominant Strategy Incentive Compatible.

Numerical Simulation

Under some mild conditions on the attack cost and total amount of bets, the proposed incentive mechanism provides *individual rationality* and *fair allocation of rewards*

Conclusion

Main Result - CSC-based Collaborative Attack

The attack sponsor can design a **cheat-proof** and **budget-balanced** mechanism to encourage collaboration of selfish rational attackers.

Side Result

The sponsor can predict and adapt the attack result, i.e., determine under what conditions attackers will participate in the attack.

Thank you!

Questions?