Attacks on PoW systems

Yujin Kwon

KAIST

Various Attacks

- Double Spending
  - Generate forks intentionally
- Selfish mining
  - Generate forks intentionally
    - "Majority Is Not Enough: Bitcoin Mining Is Vulnerable", FC 2014
- Block withholding (BWH) attack
  - Exploit the pools' protocol
  - It is possible to launch the BWH attack each other.
    - "The Miner's Dilemma", SP 2016
    - "On Power Splitting Games in Distributed Computation: The Case of Bitcoin Pooled Mining", CSF 2016
- Fork after withholding (FAW) attack
  - Generate forks intentionally through pools

Various Attacks

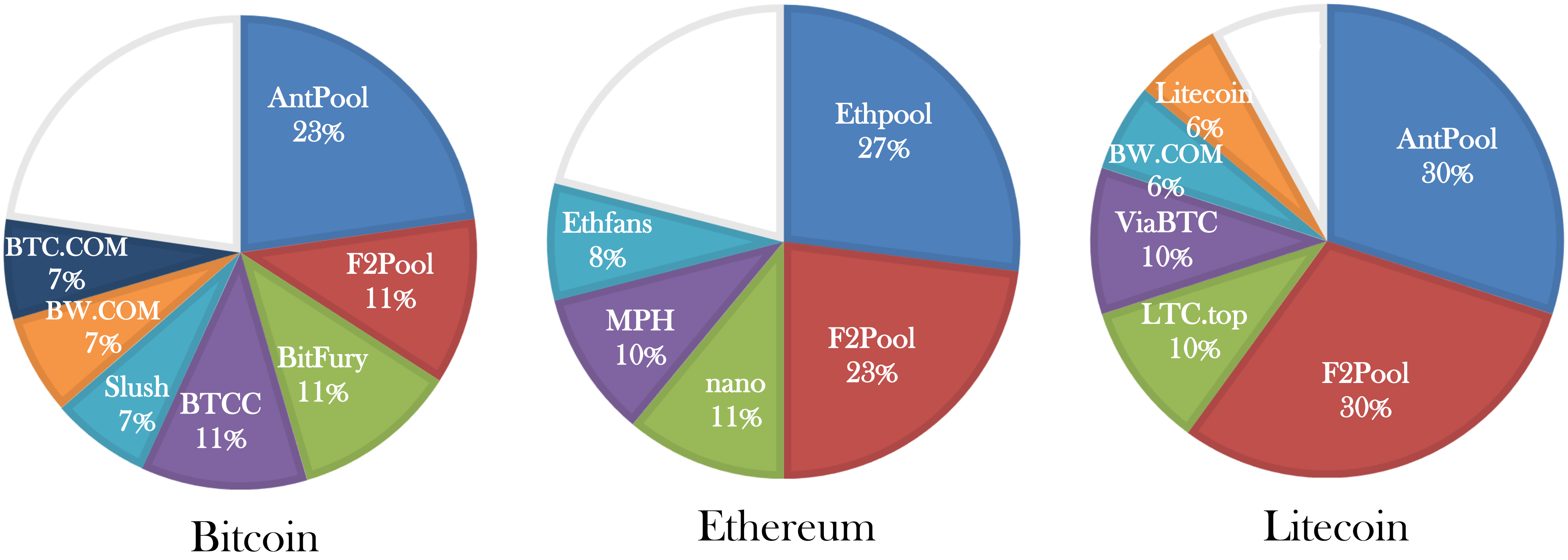
- Double Spending
  - Generate forks intentionally
- Selfish mining
  - Generate forks intentionally
    - "Majority Is Not Enough: Bitcoin Mining Is Vulnerable", FC 2014
- Block withholding (BWH) attack
  - Exploit the pools' protocol
  - It is possible to launch the BWH attack each other.
    - "The Miner's Dilemma", SP 2016
    - "On Power Splitting Games in Distributed Computation: The Case of Bitcoin Pooled Mining", CSF 2016
- Fork after withholding (FAW) attack
  - Generate forks intentionally through pools

The Miner's Dilemma

Ittay Eyal
Cornell University

2015 IEEE Symposium on Security and Privacy

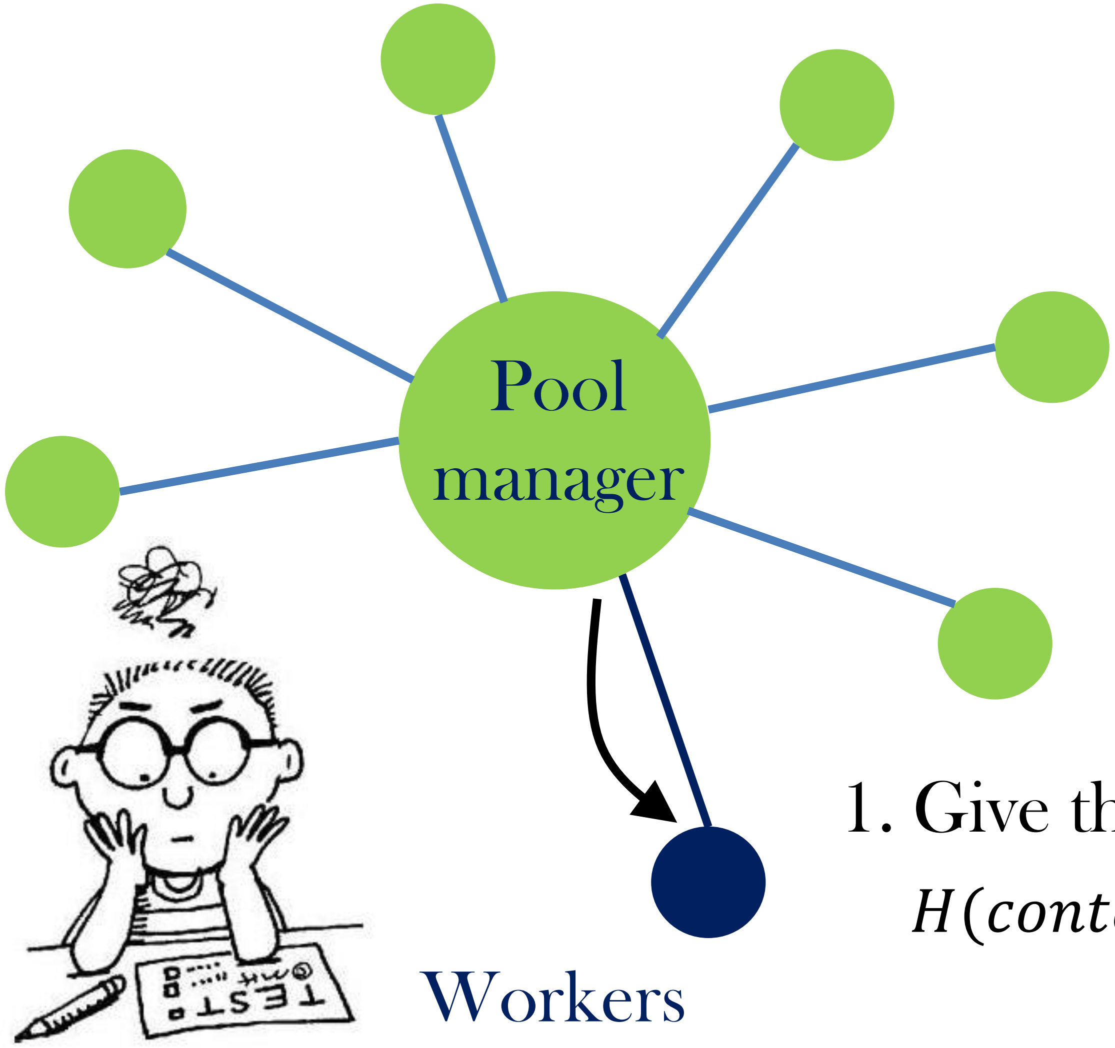
Mining Pool

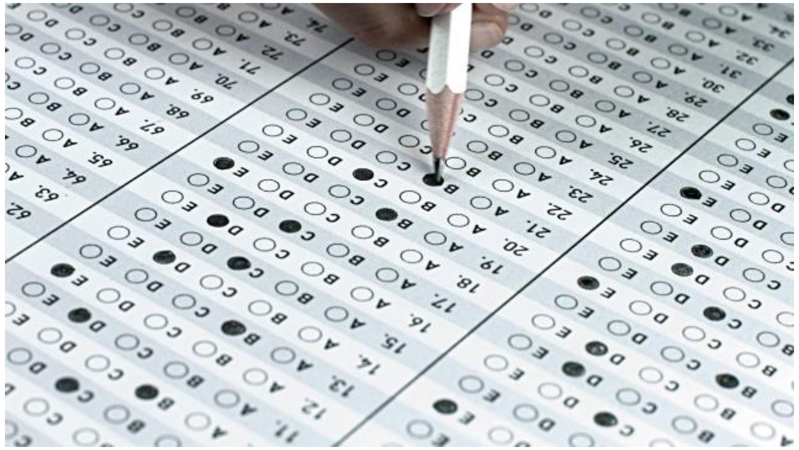


❖ Miners can organize pools and mine together to reduce the variance of reward.

❖ Currently, major players are pools.

Mining Pool

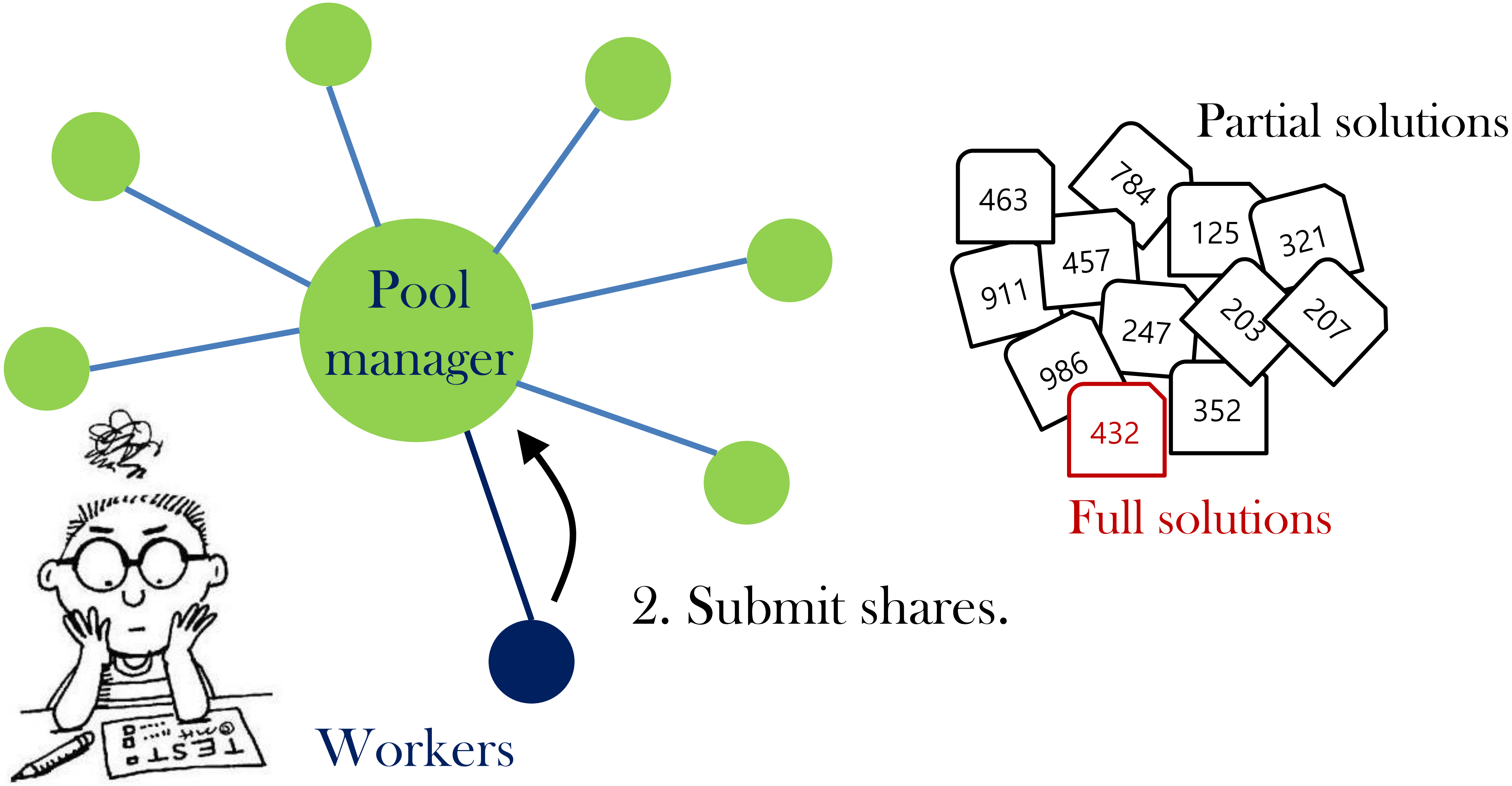




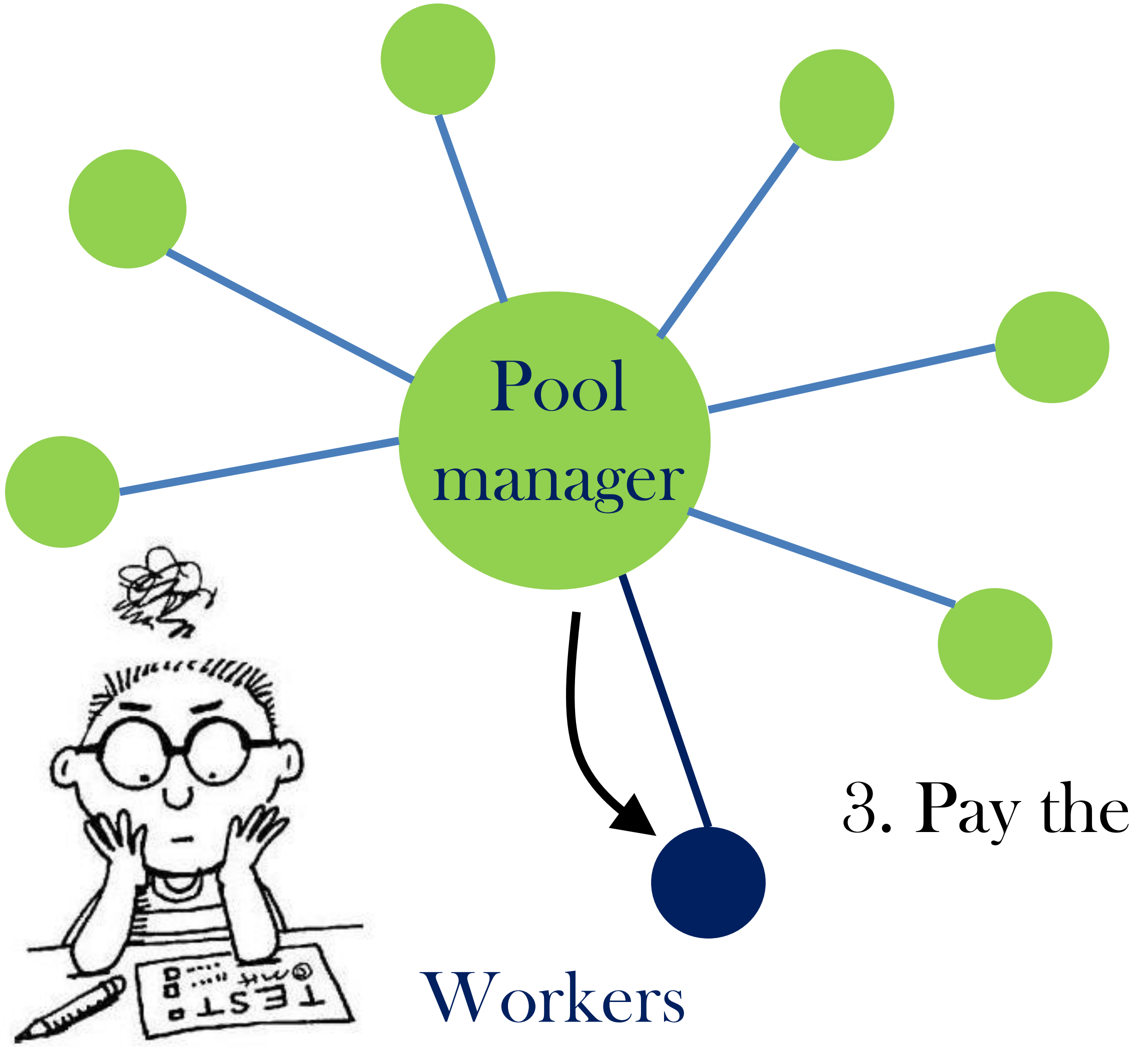
1. Give the problem.

$$H(contents||nonce) < target \ ?$$

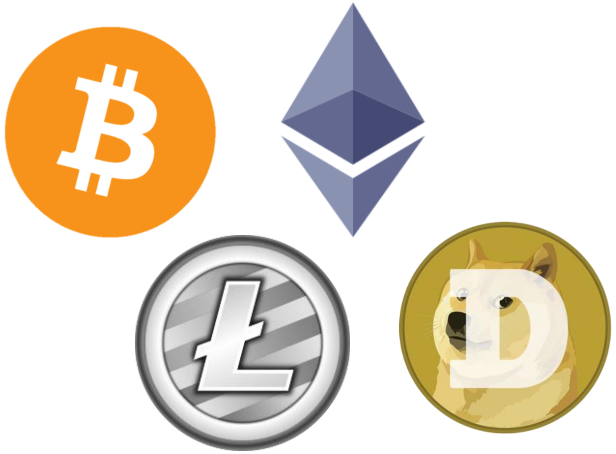
Mining Pool


Pool manager
Workers
2. Submit shares.
Partial solutions
463
784
125
321
457
911
247
203
207
986
432
352
Full solutions

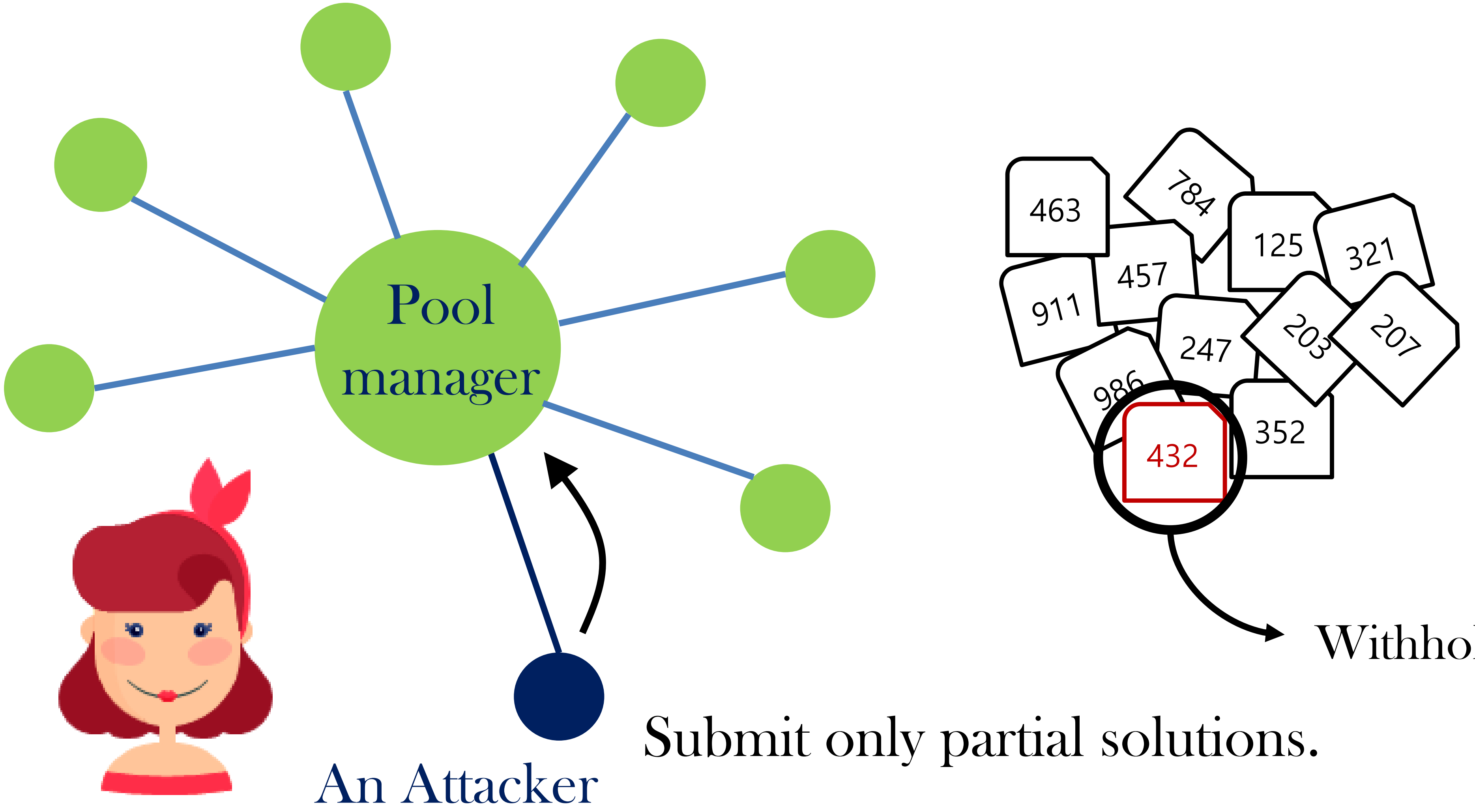
Mining Pool



3. Pay the reward.



Block Withholding (BWH) Attack

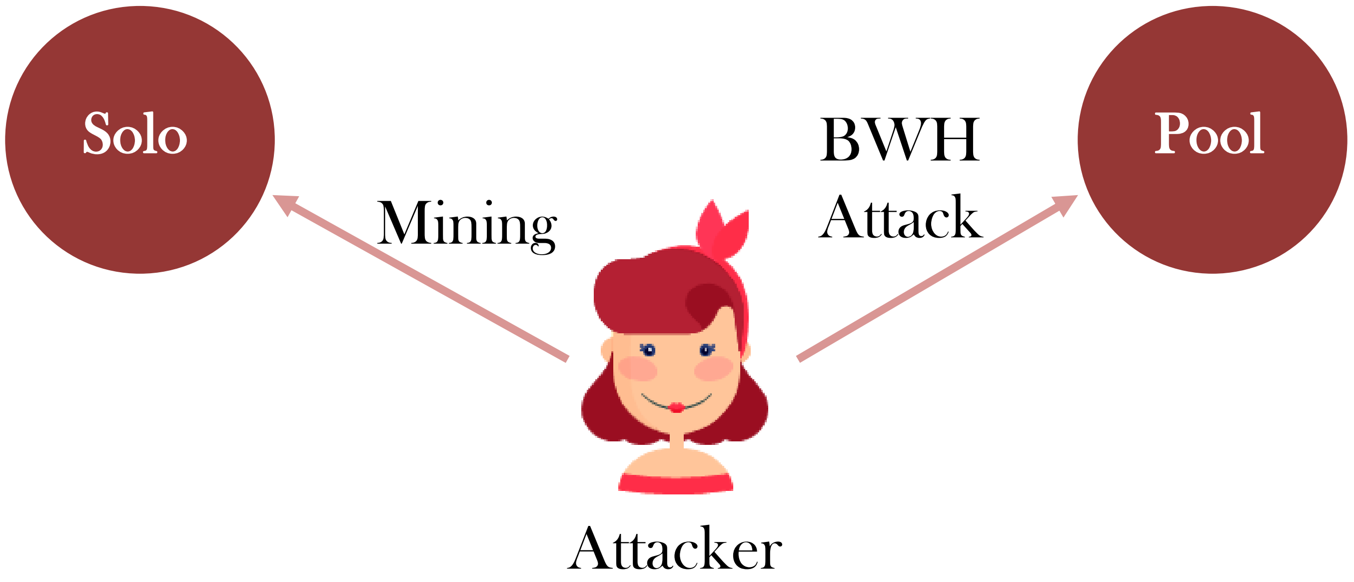


History

- 2011 : Analysis of Bitcoin Pooled Mining Reward Systems (by Meni Rosenfeld)
  - "This has no direct benefit for the attacker, only causing harm to the pool operator or participants. "
- 2014 : On Subversive Miner Strategies and Block Withholding Attack in Bitcoin Digital Currency
  - "They showed that an attacker can earn profit by this attack"
- 2015 : The miner's dilemma
- On Power Splitting Games in Distributed Computation: The Case of Bitcoin Pooled Mining
  - "Attack strategy && game theory"

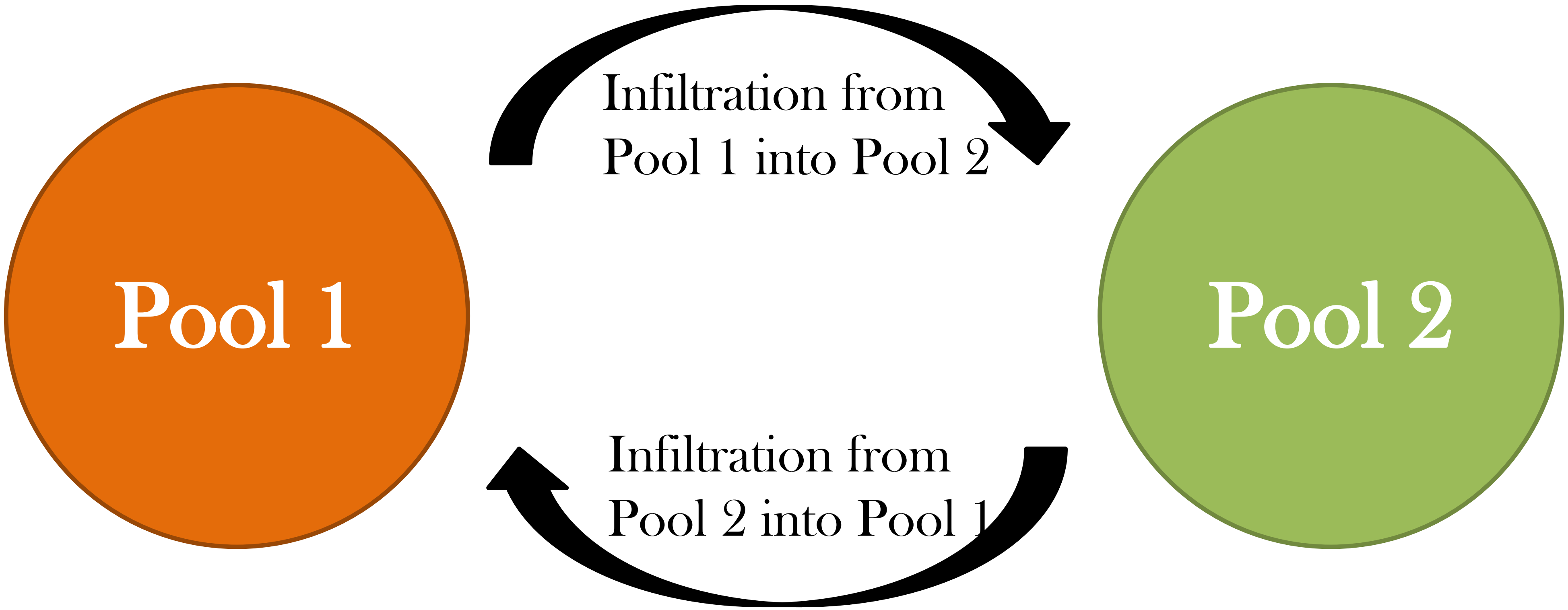
Block Withholding (BWH) Attack

- An attacker joins the victim pool.
- She should split her computational power into solo mining and malicious pool mining (BWH attack).
- She receives unearned wages while only pretending to contribute work to the pool.

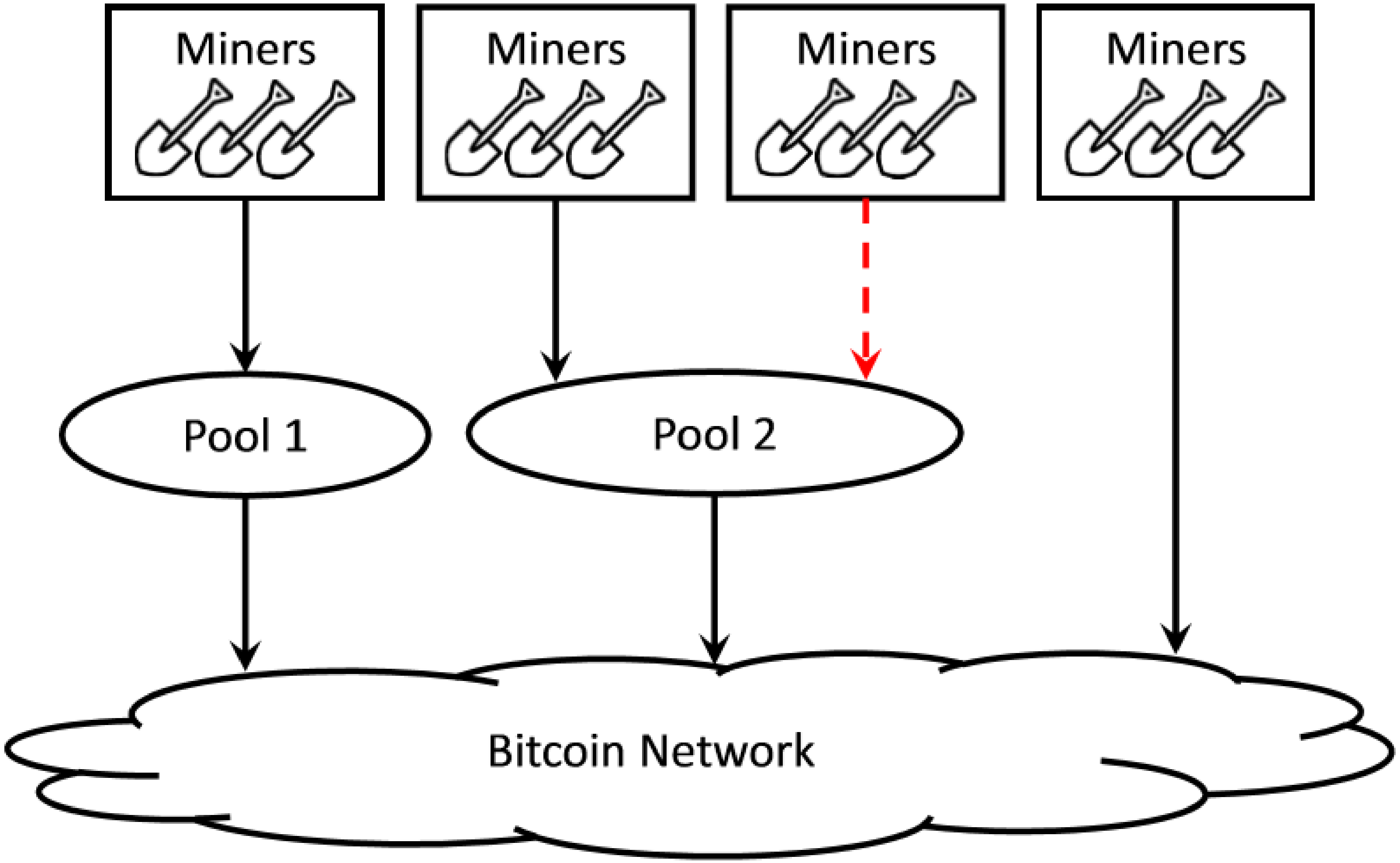


Pool game

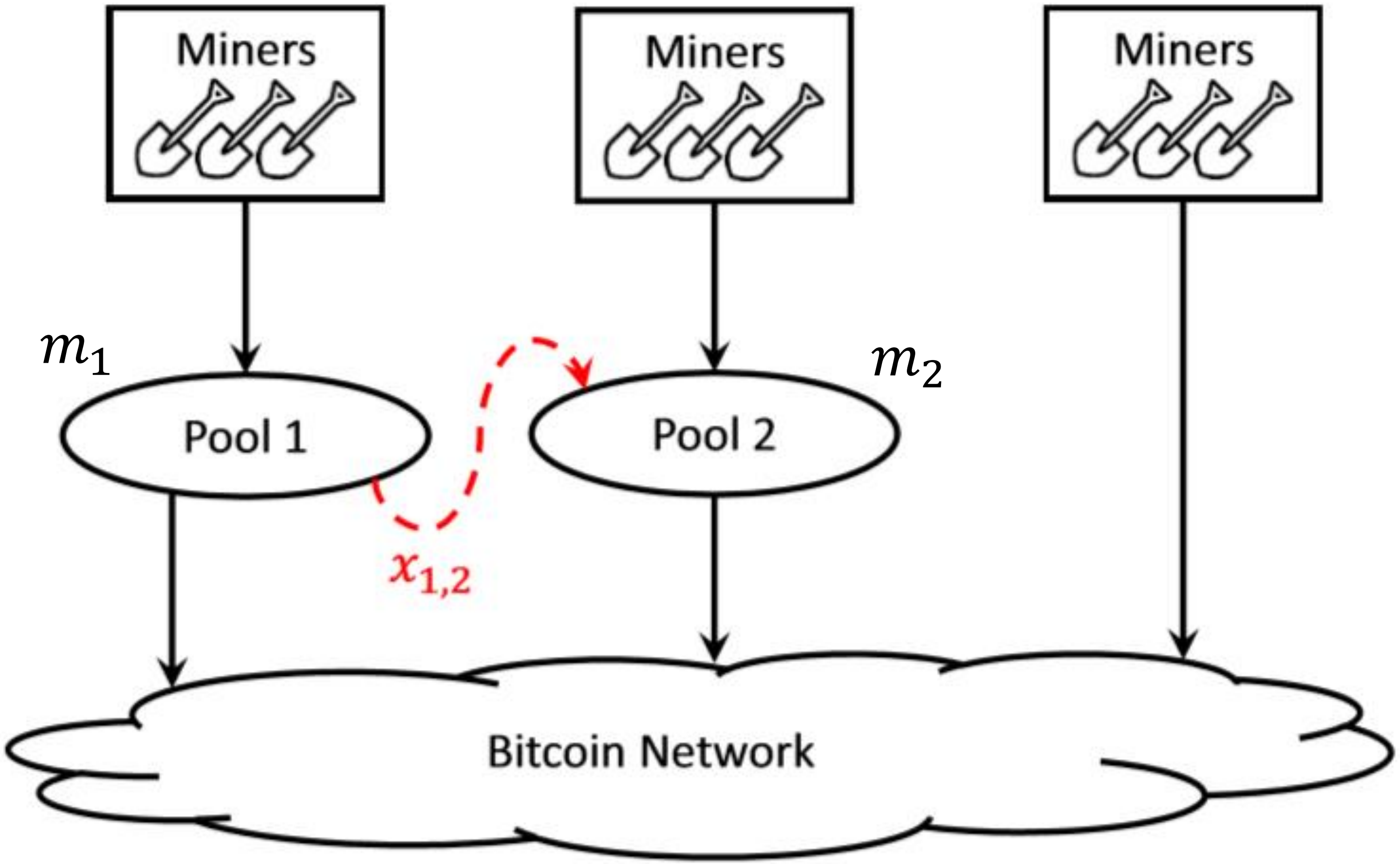
- Pools can launch the BWH attack each other through infiltration.



Classical BWH attack



BWH attack among pools



Analysis

$$R_1 = \frac{m_1 - x_{1,2}}{m - x_{1,2}}$$

$$R_2 = \frac{m_2}{m - x_{1,2}} \ .$$

$$r_1 = \frac{R_1 + x_{1,2} \cdot r_2}{m_1} \ .$$

$$r_2 = \frac{R_2}{m_2 + x_{1,2}} \ .$$

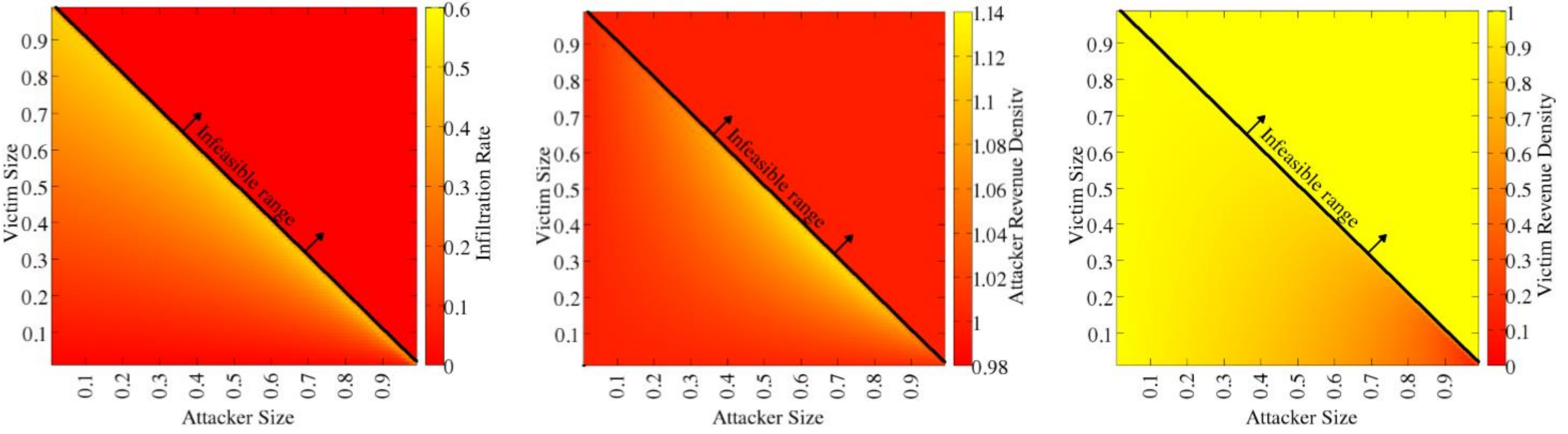
$$r_1 = \frac{m_1(m_2 + x_{1,2}) - x_{1,2}^2}{m_1(m - x_{1,2})(m_2 + x_{1,2})}$$

↓

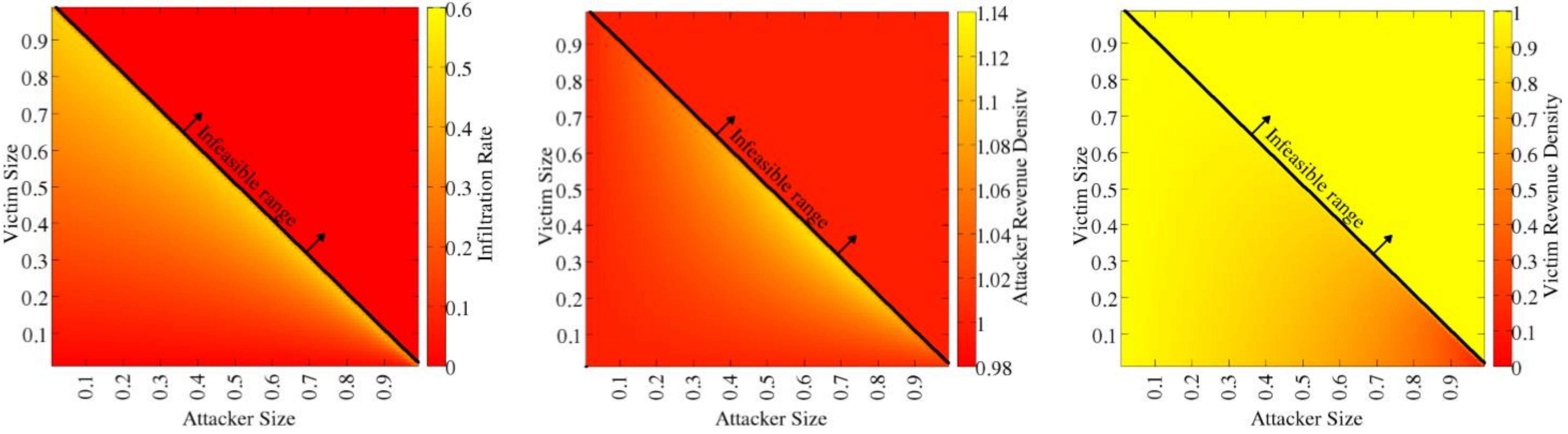
$$\bar{x}_{1,2} = \frac{m_2 - m_1 m_2 - \sqrt{-m_2^2(-1 + m_1 + m_1 m_2)}}{-1 + m_1 + m_2}$$

$$\bar{r}_1 = \frac{m_1 + (2 + m_1)m_2 - 2\sqrt{-m_2^2(-1 + m_1 + m_1 m_2)}}{m_1(1 + m_2)^2}$$

$$\bar{r}_2 = \frac{m_2(-1 + m_1 + m_2)^2}{\left(m_2^2 - \sqrt{-m_2^2(-1 + m_1 + m_1 m_2)}\right)\left(1 - m_1(1 + m_2) - \sqrt{-m_2^2(-1 + m_1 + m_1 m_2)}\right)}$$


Victim Size
Attacker Size
Infiltration Rate
Infeasible range
Attacker Revenue Density
Victim Revenue Density

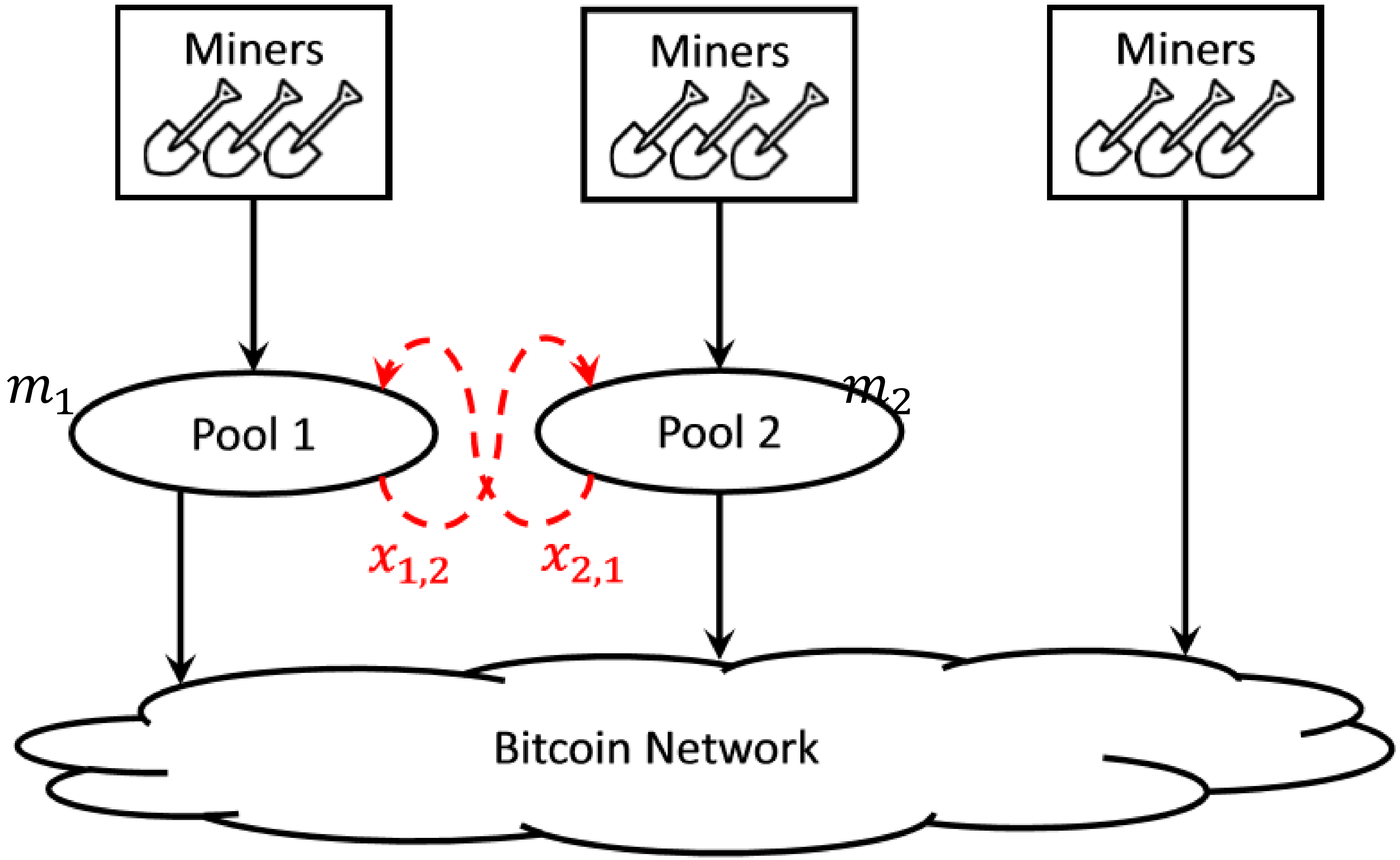
(a) $x_{1,2}$ (b) r_1 (c) r_2



(a) $x_{1,2}$ (b) r_1 (c) r_2

Therefore, the case for no attack is not an equilibrium.

Two Pools



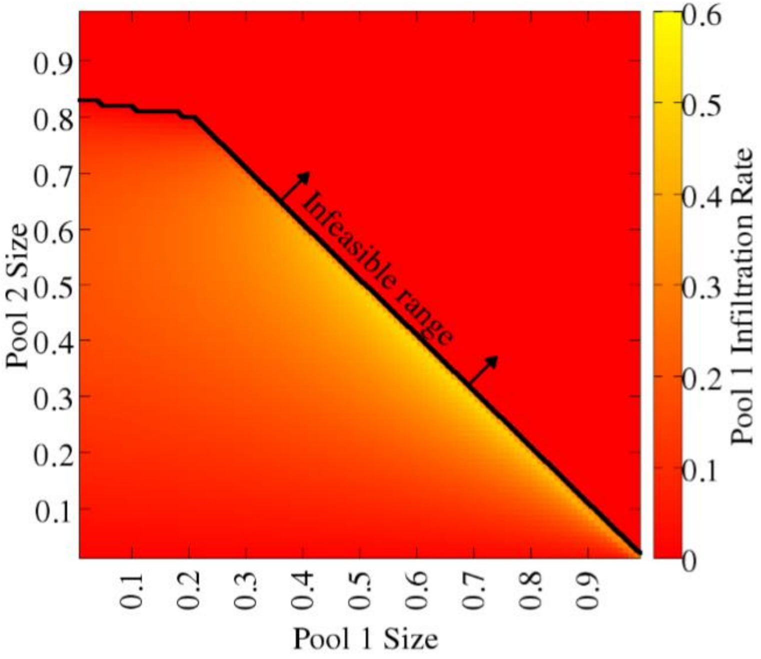
Miners
Miners
Miners
m_1
Pool 1
Pool 2
m_2
$x_{1,2}$
$x_{2,1}$
Bitcoin Network

Analysis

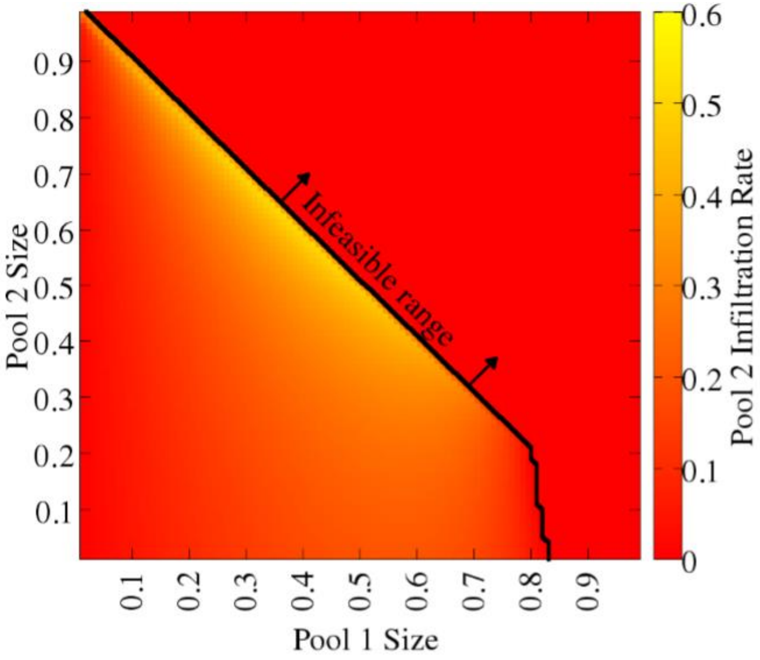
$$R_1 = \frac{m_1 - x_{1,2}}{m - x_{1,2} - x_{2,1}}$$
$$R_2 = \frac{m_2 - x_{2,1}}{m - x_{1,2} - x_{2,1}}$$

$$r_1 = \frac{R_1 + x_{1,2} r_2}{m_1 + x_{2,1}}$$
$$r_2 = \frac{R_2 + x_{2,1} r_1}{m_2 + x_{1,2}} .$$

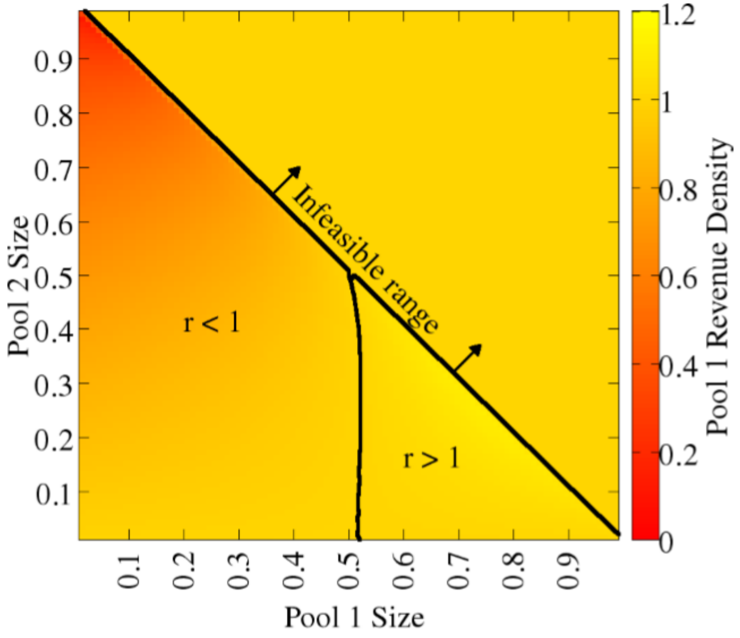
$$r_1(x_{1,2}, x_{2,1}) = \frac{m_2 R_1 + x_{1,2}(R_1 + R_2)}{m_1 m_2 + m_1 x_{1,2} + m_2 x_{2,1}}$$
$$r_2(x_{2,1}, x_{1,2}) = \frac{m_1 R_2 + x_{2,1}(R_1 + R_2)}{m_1 m_2 + m_1 x_{1,2} + m_2 x_{2,1}}$$



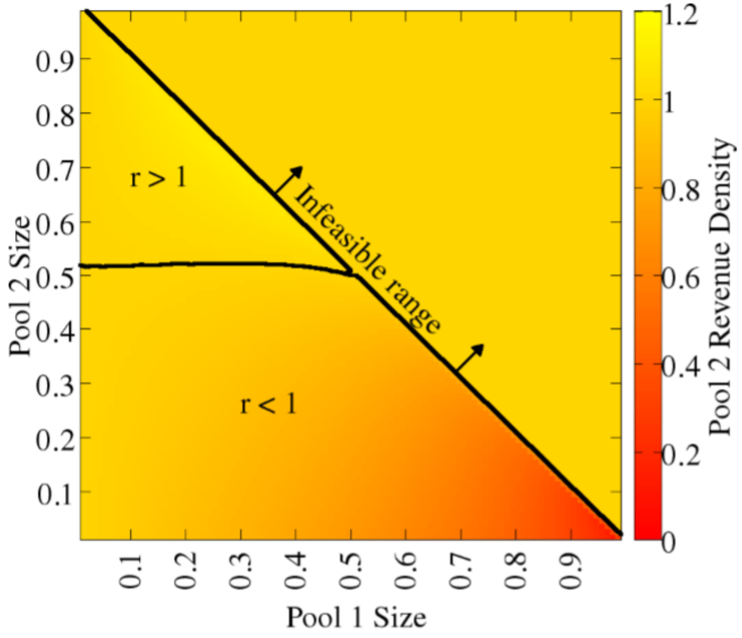
(a) $x_{1,2}$



(b) $x_{2,1}$



(c) r_1



(d) r_2

The prisoner's dilemma

- The **prisoner's dilemma** is a standard example of a game analyzed in game theory
- Two prisoners are separated into individual rooms and cannot communicate with each other.

Prisoner B / Prisoner A	Prisoner B stays silent (*cooperates*)	Prisoner B betrays (*defects*)
Prisoner A stays silent (*cooperates*)	Each serves 1 year	Prisoner A: 3 years Prisoner B: goes free
Prisoner A betrays (*defects*)	Prisoner A: goes free Prisoner B: 3 years	Each serves 2 years

The Miners' dilemma

Pool 2 \ Pool 1	no attack	attack
no attack	$(r_1 = 1, r_2 = 1)$	$(r_1 > 1, r_2 = \tilde{r}_2 < 1)$
attack	$(r_1 = \tilde{r}_1 < 1, r_2 > 1)$	$(\tilde{r}_1 < r_1 < 1, \tilde{r}_2 < r_2 < 1)$

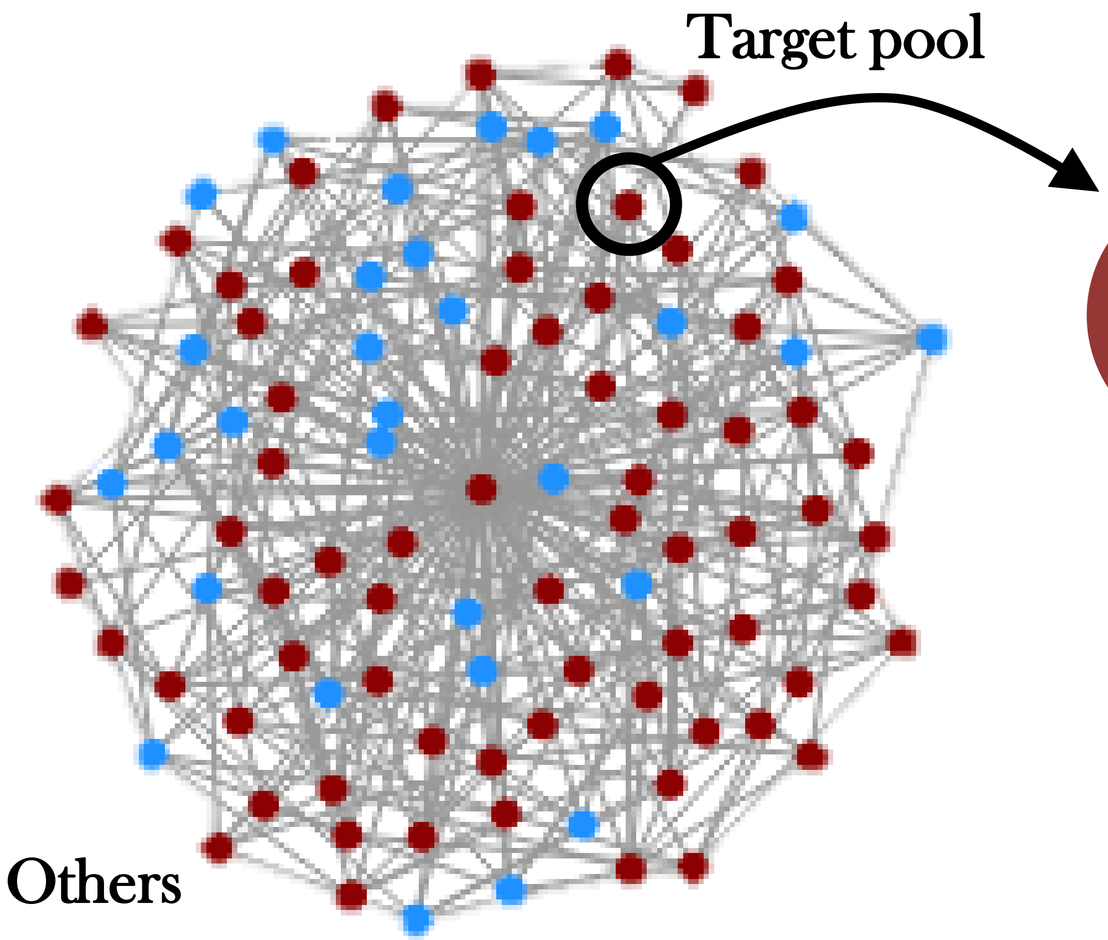
From "The Miner's Dilemma"

❖ The equilibrium reward of the pool is **inferior** compared to the no-attack scenario.

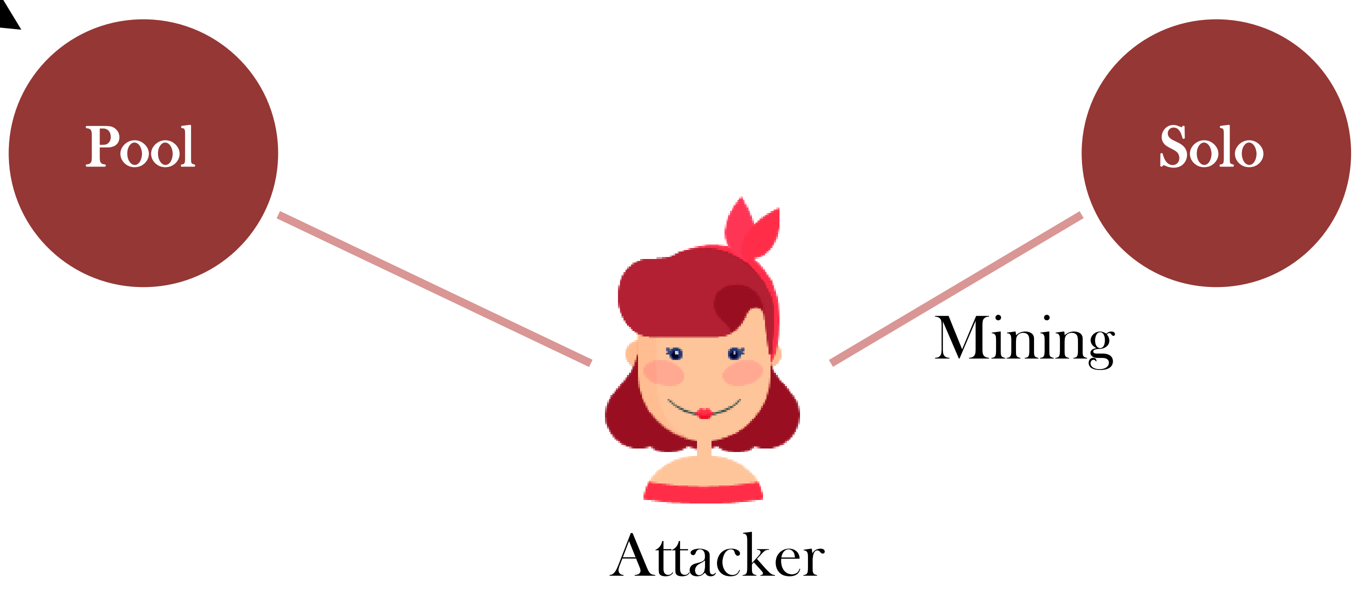
❖ The fact that the BWH attack is **not common** may be explained.

The FAW Attack

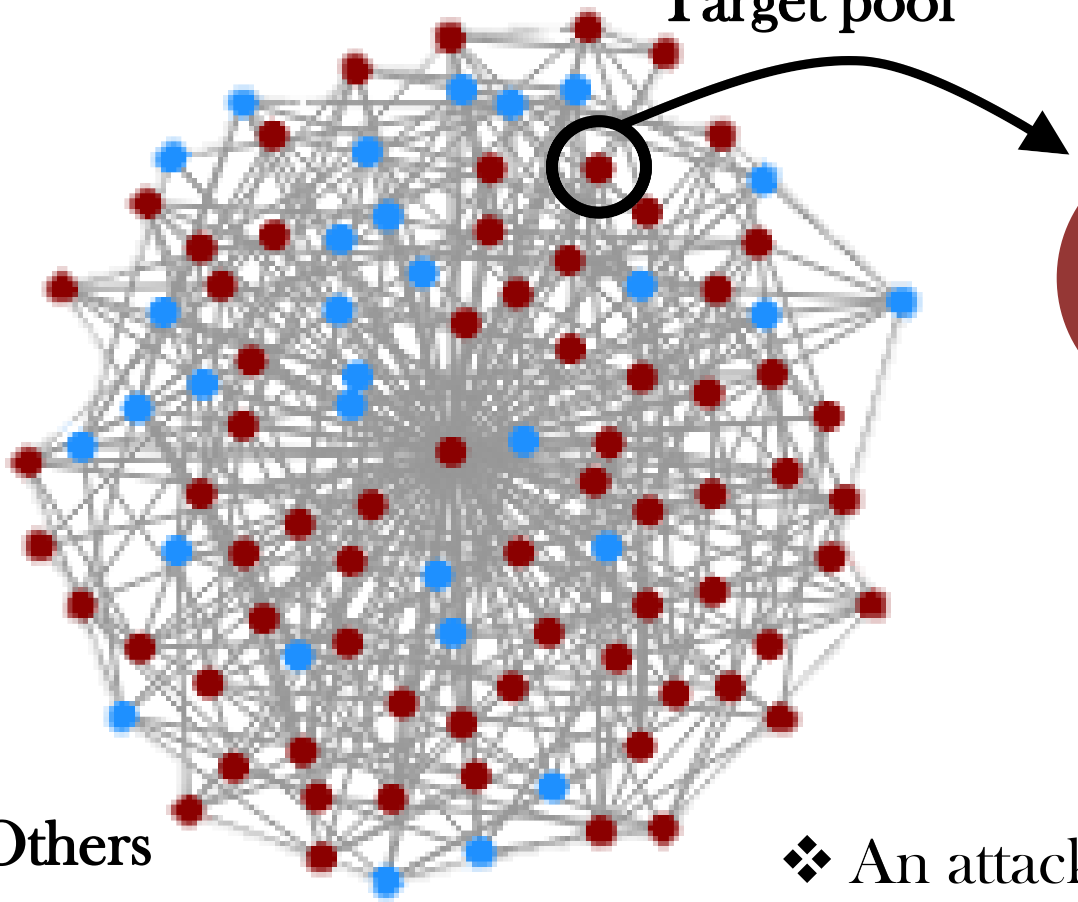
FAW Attack Against One Pool



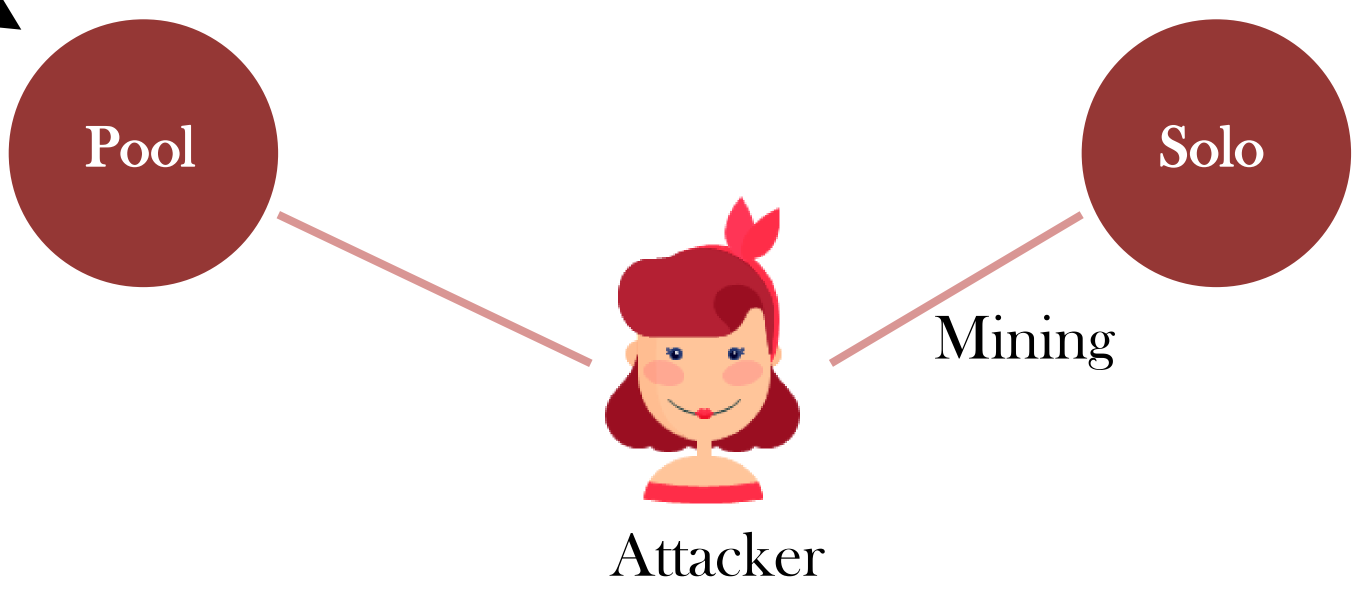
Submit an FPoW to the pool only if others generate another block. Otherwise, throw her FPoW.



FAW Attack Against One Pool



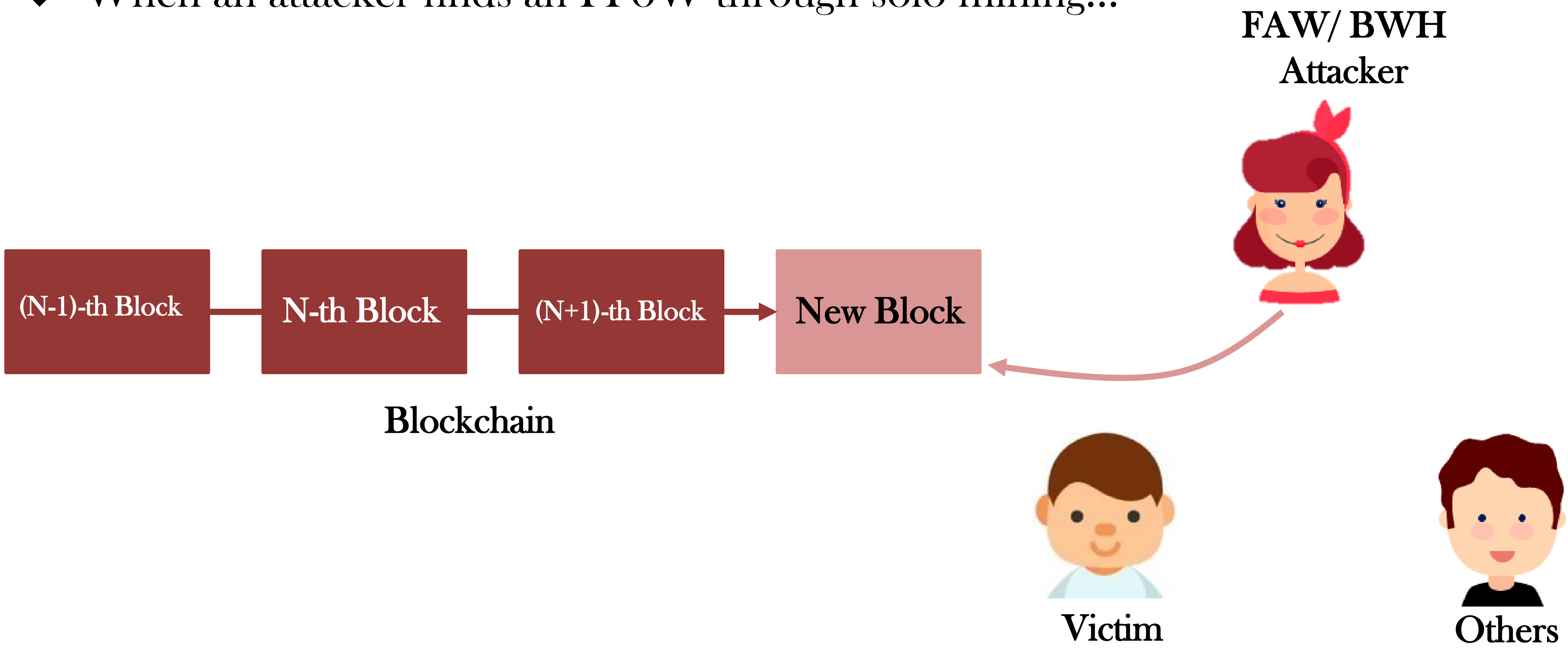
Submit an FPoW to the pool only if others generate another block. Otherwise, throw her FPoW.



❖ An attacker generates forks intentionally through a pool!

FAW vs BWH

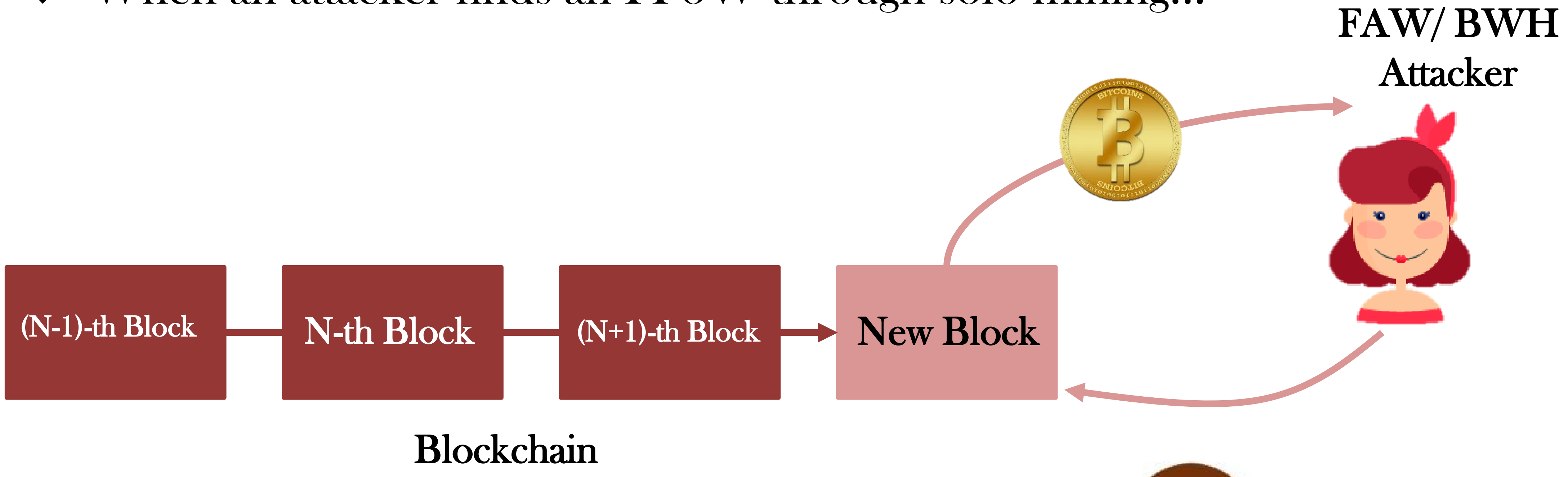
❖ When an attacker finds an FPoW through solo mining…



SysSec
System Security Lab

FAW vs BWH

- When an attacker finds an FPoW through solo mining…

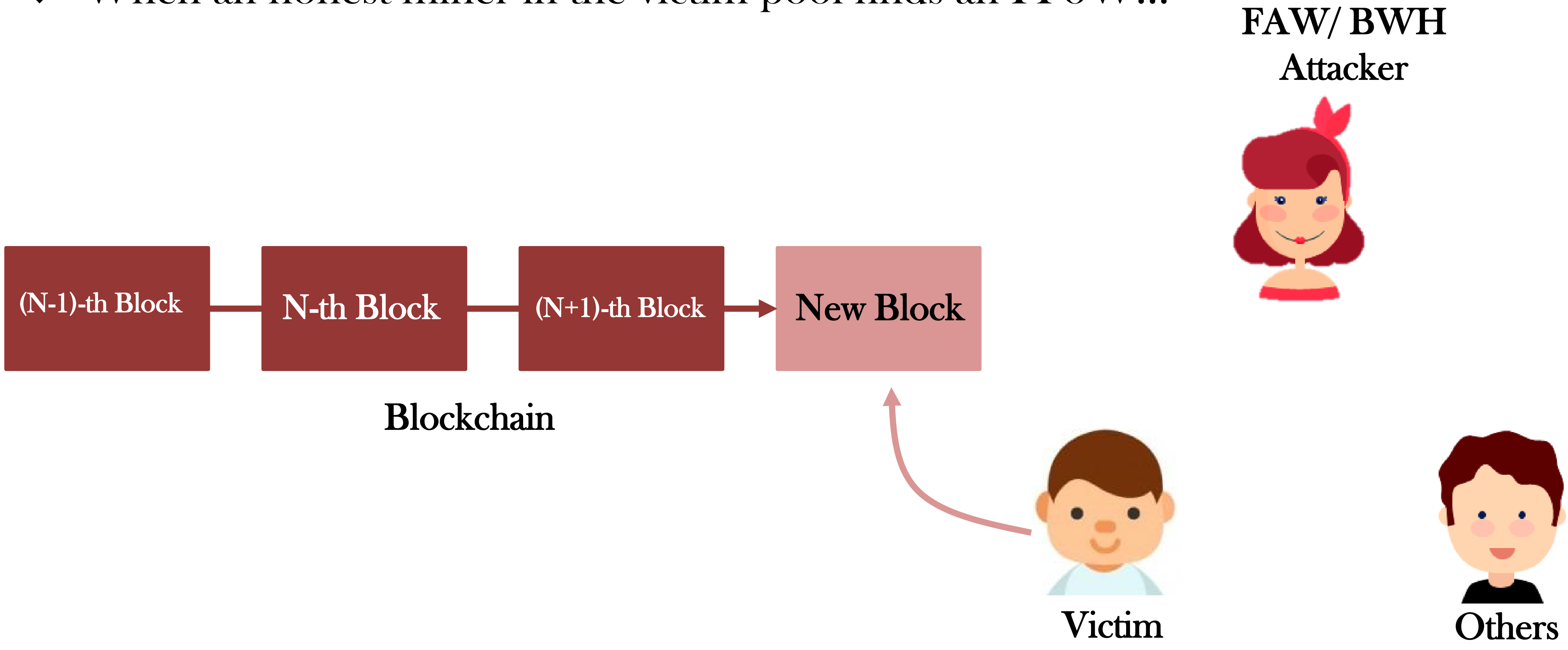


The attacker earns the block reward.

SysSec System Security Lab

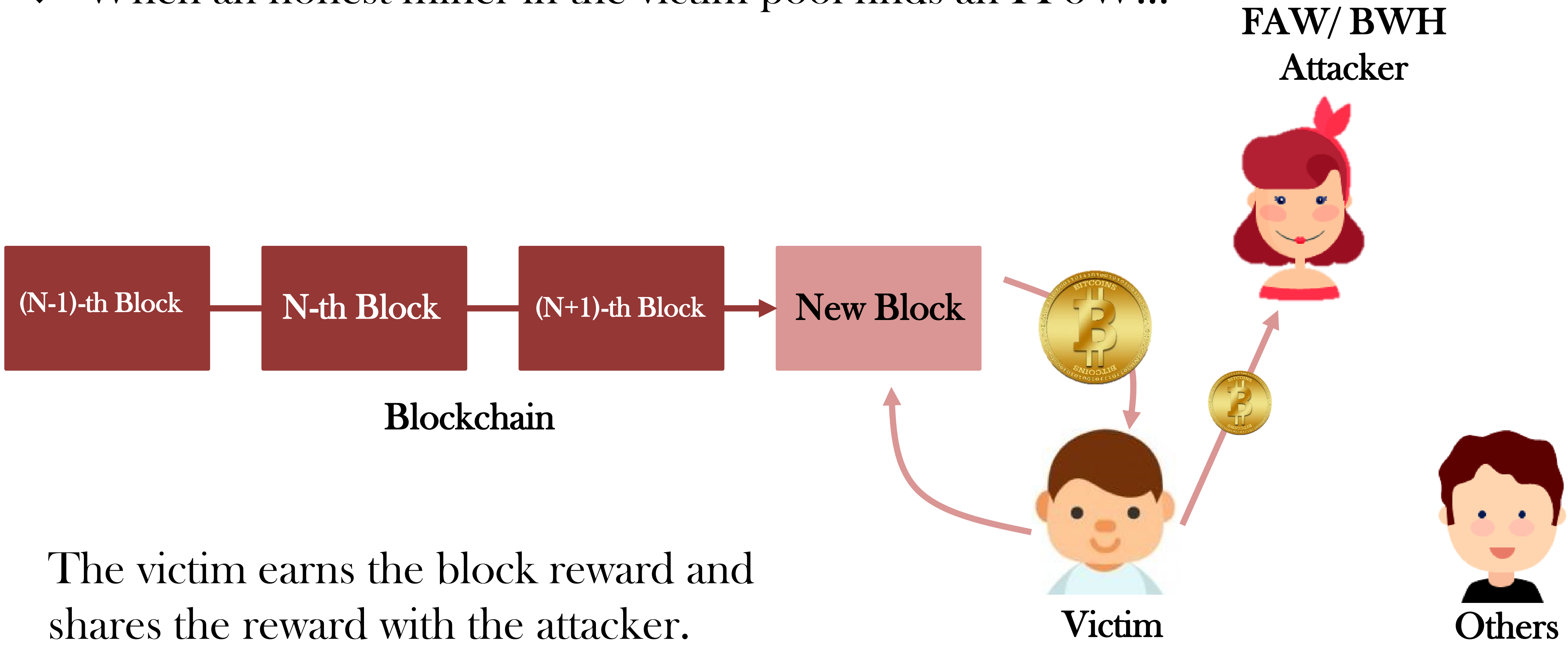
FAW vs BWH

❖ When an honest miner in the victim pool finds an FPoW…



FAW vs BWH

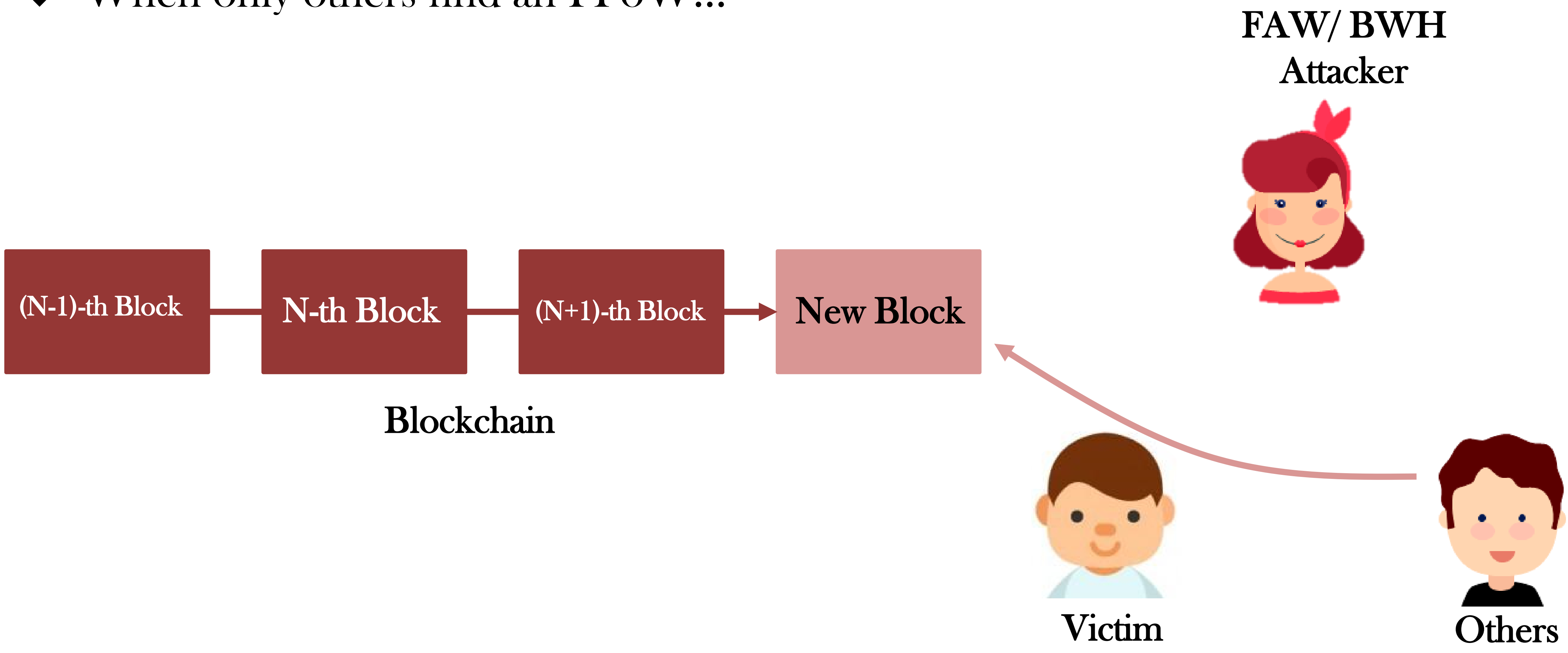
- When an honest miner in the victim pool finds an FPoW…



The victim earns the block reward and shares the reward with the attacker.

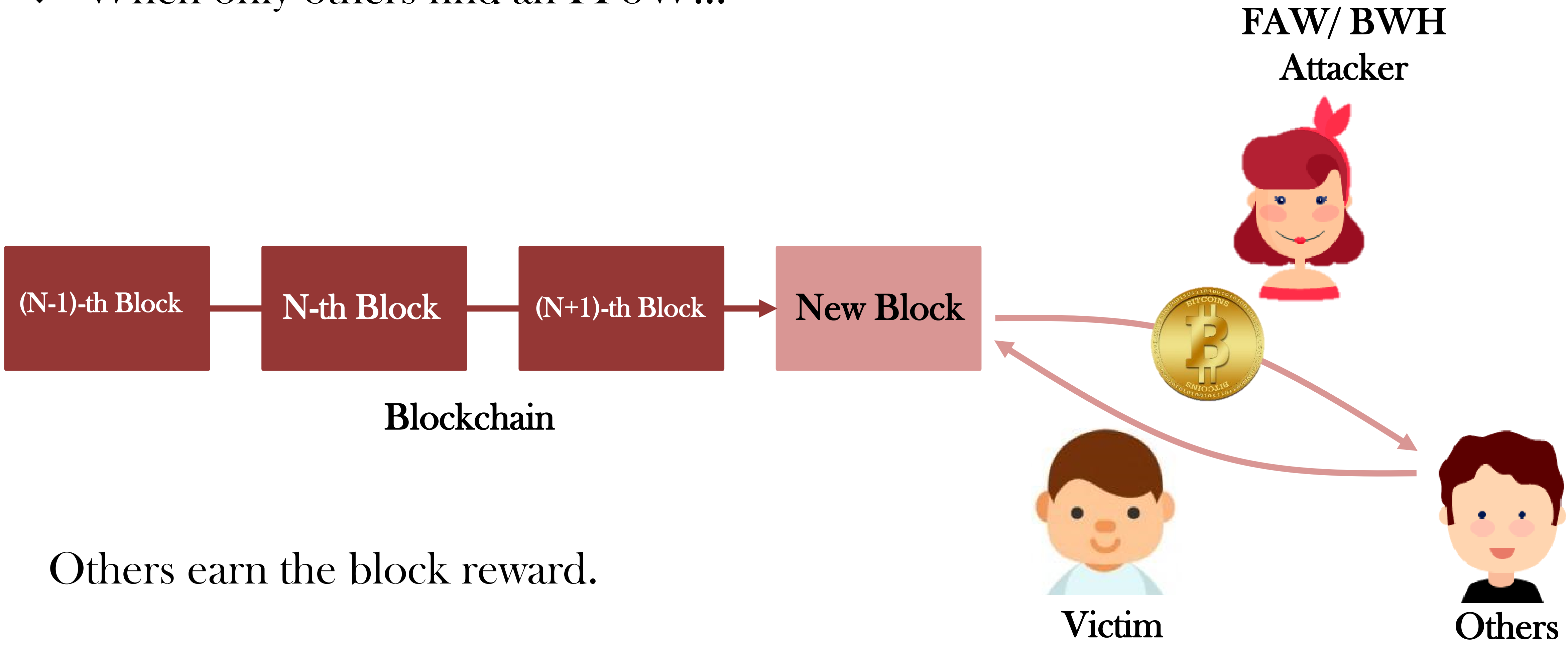
FAW vs BWH

- When only others find an FPoW…



FAW vs BWH

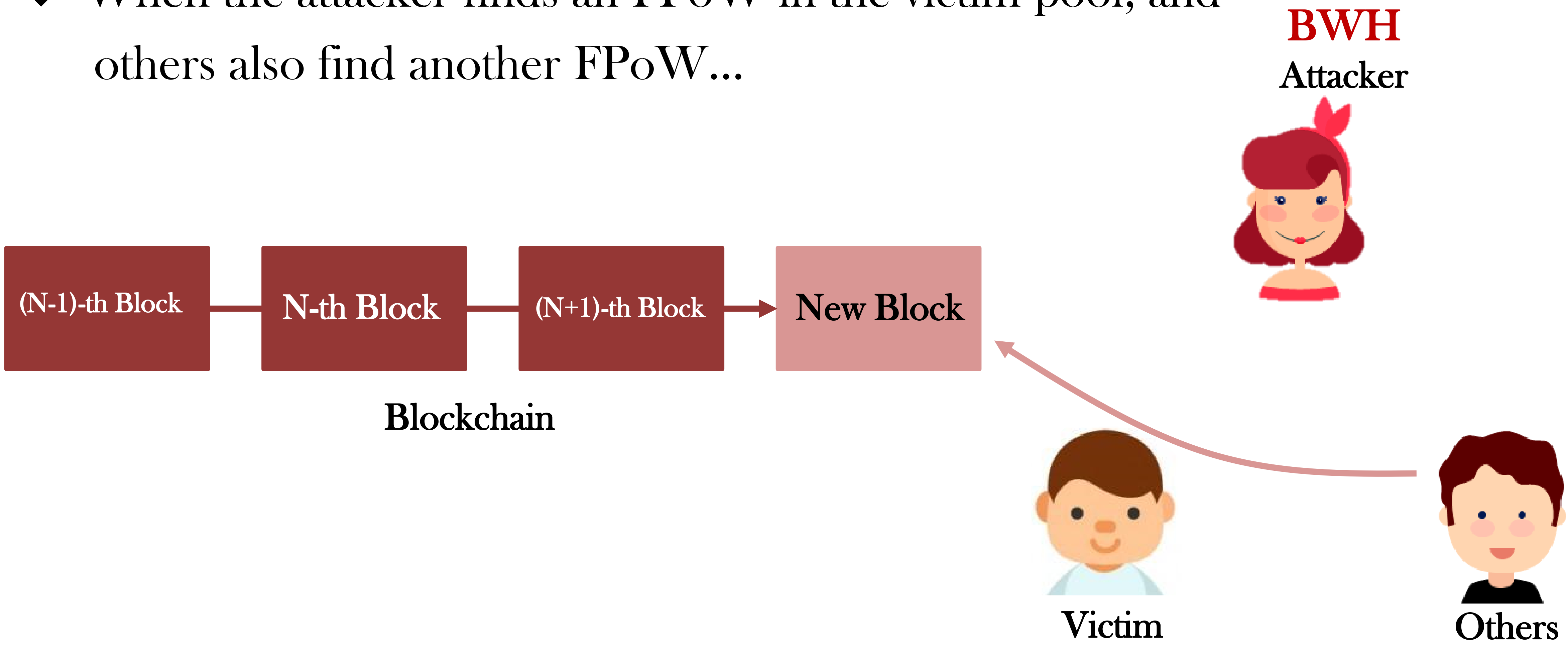
- When only others find an FPoW…



Others earn the block reward.

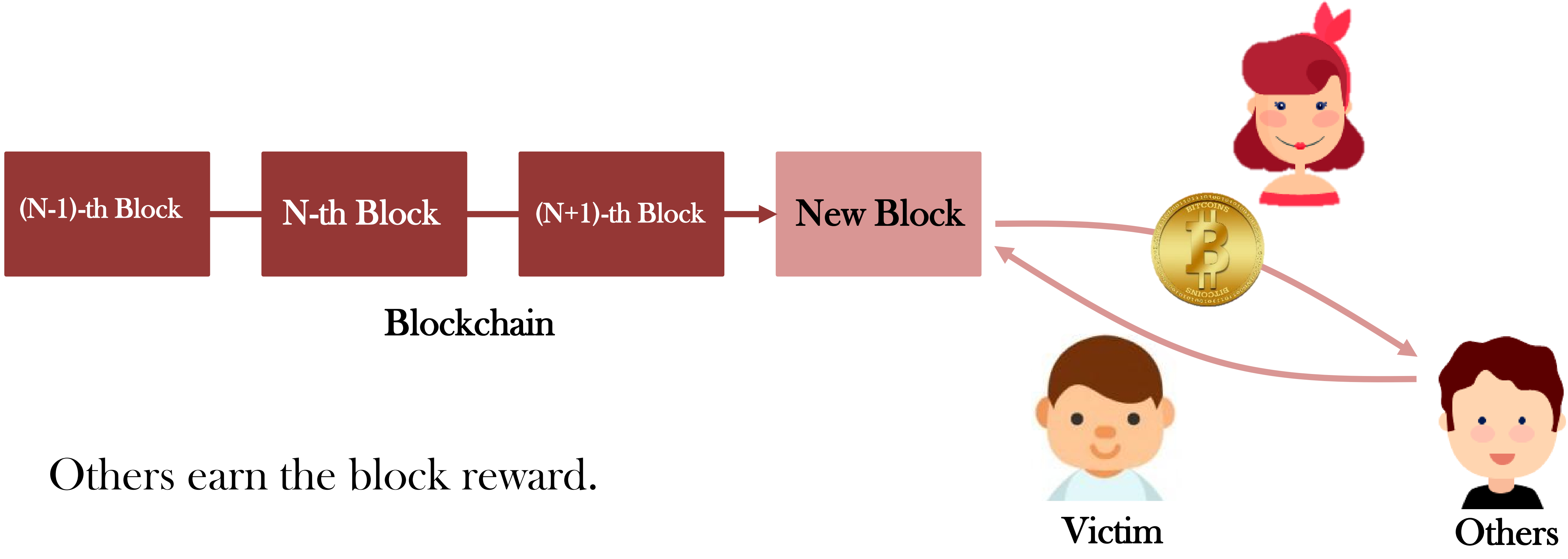
FAW vs BWH

❖ When the attacker finds an FPoW in the victim pool, and others also find another FPoW…



FAW vs BWH

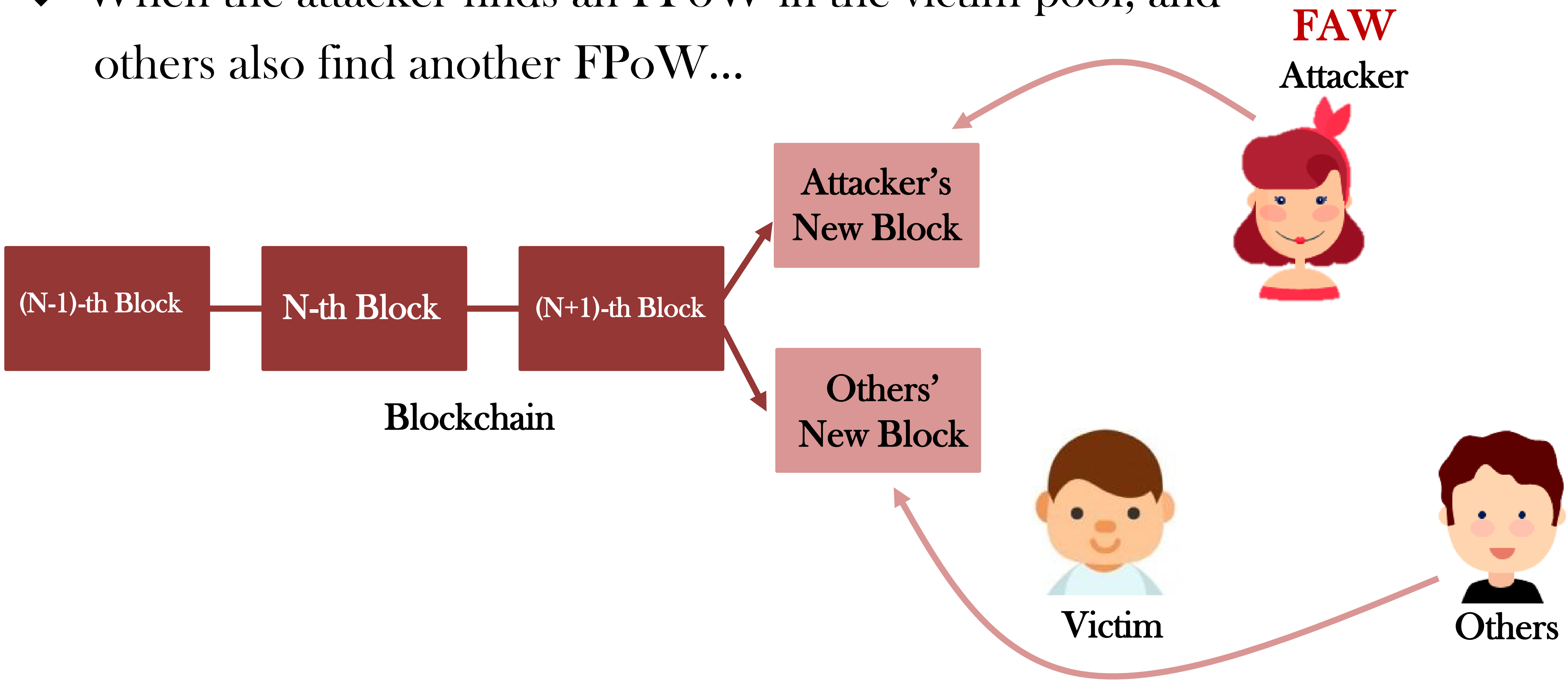
- When the attacker finds an FPoW in the victim pool, and others also find another FPoW…



Others earn the block reward.

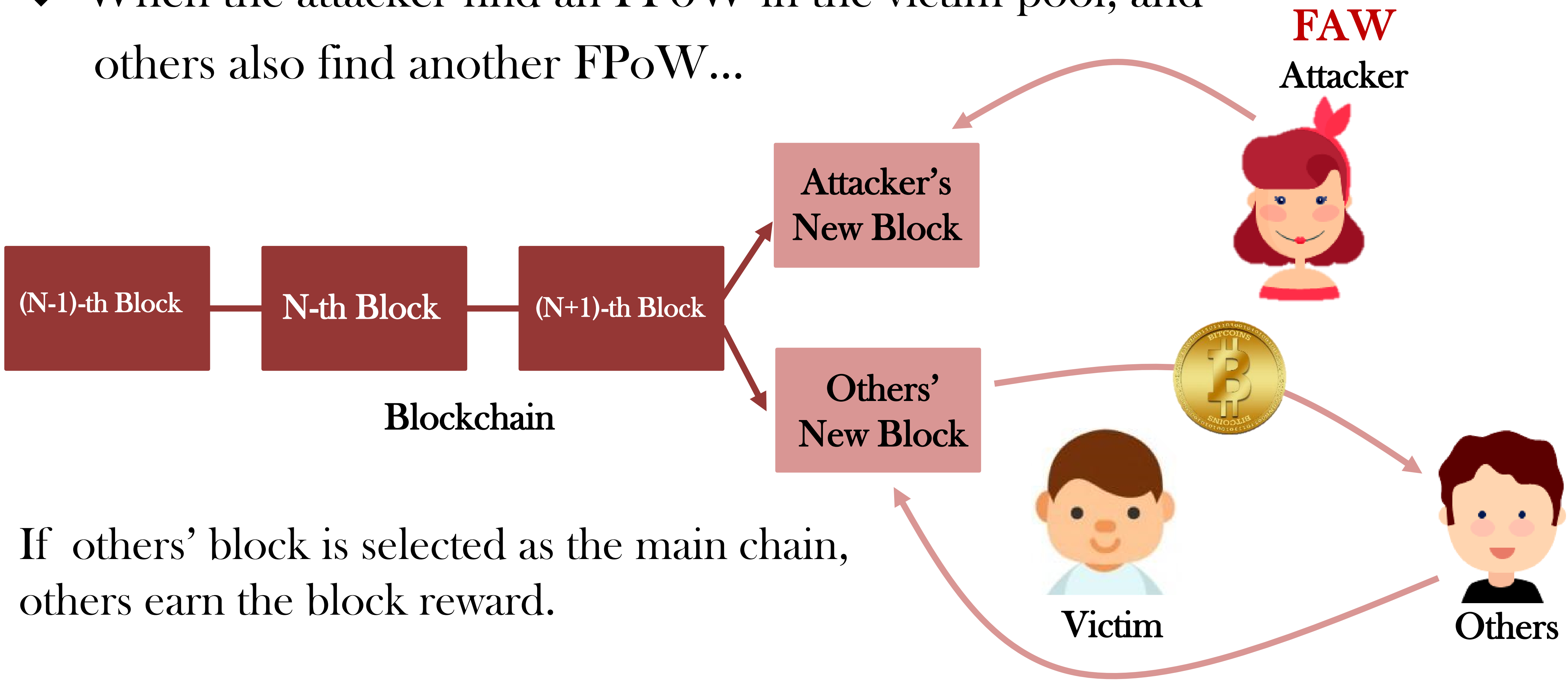
FAW vs BWH

- ❖ When the attacker finds an FPoW in the victim pool, and others also find another FPoW…



FAW vs BWH

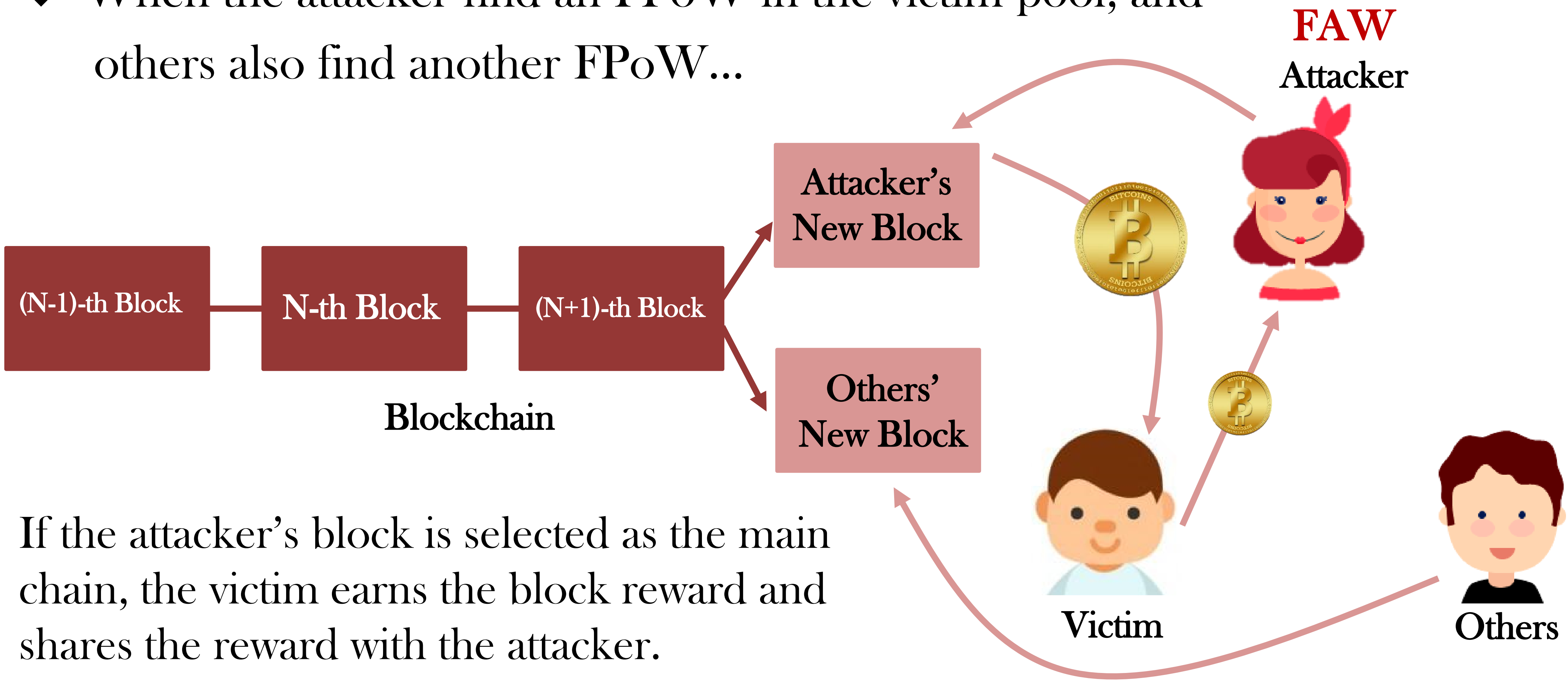
❖ When the attacker find an FPoW in the victim pool, and others also find another FPoW…



If others' block is selected as the main chain, others earn the block reward.

FAW vs BWH

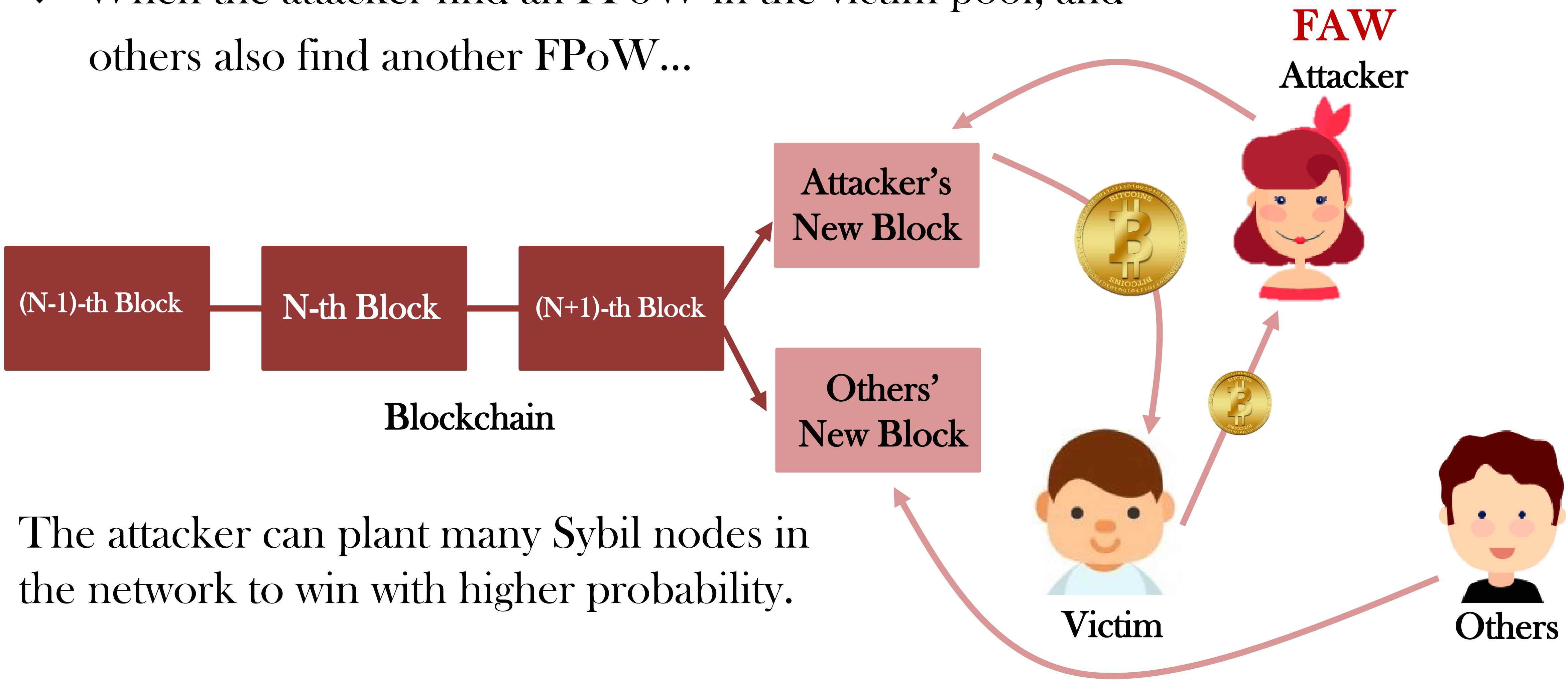
❖ When the attacker find an FPoW in the victim pool, and others also find another FPoW…



If the attacker's block is selected as the main chain, the victim earns the block reward and shares the reward with the attacker.

FAW vs BWH

❖ When the attacker find an FPoW in the victim pool, and others also find another FPoW…



The attacker can plant many Sybil nodes in the network to win with higher probability.

FAW Attack Against One Pool

- Notation
  - α: Computational power of the attacker
  - β: Total computational power of a victim pool
  - γ: The infiltration mining power divided by α
  - c: Attacker′s network capability
  - R_a (R_p): An attacker’s (The victim′s) reward

Analysis

THEOREM 5.1. *An FAW attacker can earn*

$$R_a(\tau) = \frac{(1-\tau)\alpha}{1-\tau\alpha} + \left(\frac{\beta}{1-\tau\alpha} + c\tau\alpha \cdot \frac{1-\alpha-\beta}{1-\tau\alpha}\right) \cdot \frac{\tau\alpha}{\beta+\tau\alpha}. \quad (1)$$

The reward is maximized when the optimal τ value, $\overline{\tau}$, is

$$\frac{(1-\alpha)(1-c)\beta+\beta^2 c-\beta\sqrt{(1-\alpha-\beta)^2 c^2+((1-\alpha-\beta)(\alpha\beta+\alpha-2))c-\alpha(1+\beta)+1}}{\alpha(1-\alpha-\beta)(c(1-\beta)-1)} \quad (2)$$

THEOREM 7.1. *In the FAW attack game between two pools, the rewards R_1 of $Pool_1$ and R_2 of $Pool_2$ are:*

$$R_1 = \frac{\alpha_1-f_1}{1-f_1-f_2} + c_2 f_2 \frac{1-\alpha_1-\alpha_2}{1-f_2} + c_2' f_1 f_2 \left(\frac{1}{1-f_1} + \frac{1}{1-f_2}\right)\frac{1-\alpha_1-\alpha_2}{1-f_1-f_2} + R_2 \frac{f_1}{\alpha_2+f_1} \quad (6)$$

$$R_2 = \frac{\alpha_2-f_2}{1-f_1-f_2} + c_1 f_1 \frac{1-\alpha_1-\alpha_2}{1-f_1} + c_1' f_1 f_2 \left(\frac{1}{1-f_1} + \frac{1}{1-f_2}\right)\frac{1-\alpha_1-\alpha_2}{1-f_1-f_2} + R_1 \frac{f_2}{\alpha_1+f_2} \quad (7)$$

FAW vs BWH

	Attacker	Victim	Others
FAW			
BWH			

Numerical Analysis

An attacker's power — Increasing →

c \ α	0.1	0.2	0.3	0.4
0	0.53 (0.53)	1.14 (1.14)	1.85 (1.85)	2.70 (2.70)
0.25	0.65 (0.67)	1.38 (1.38)	2.20 (2.20)	3.1 (3.13)
0.5	0.85 (0.85)	1.74 (1.74)	2.70 (2.70)	3.75 (3.75)
0.75	1.21 (1.22)	2.37 (2.37)	3.52 (3.52)	4.69 (4.70)
1	2.12 (2.12)	3.75 (3.75)	5.13 (5.13)	6.37 (6.36)

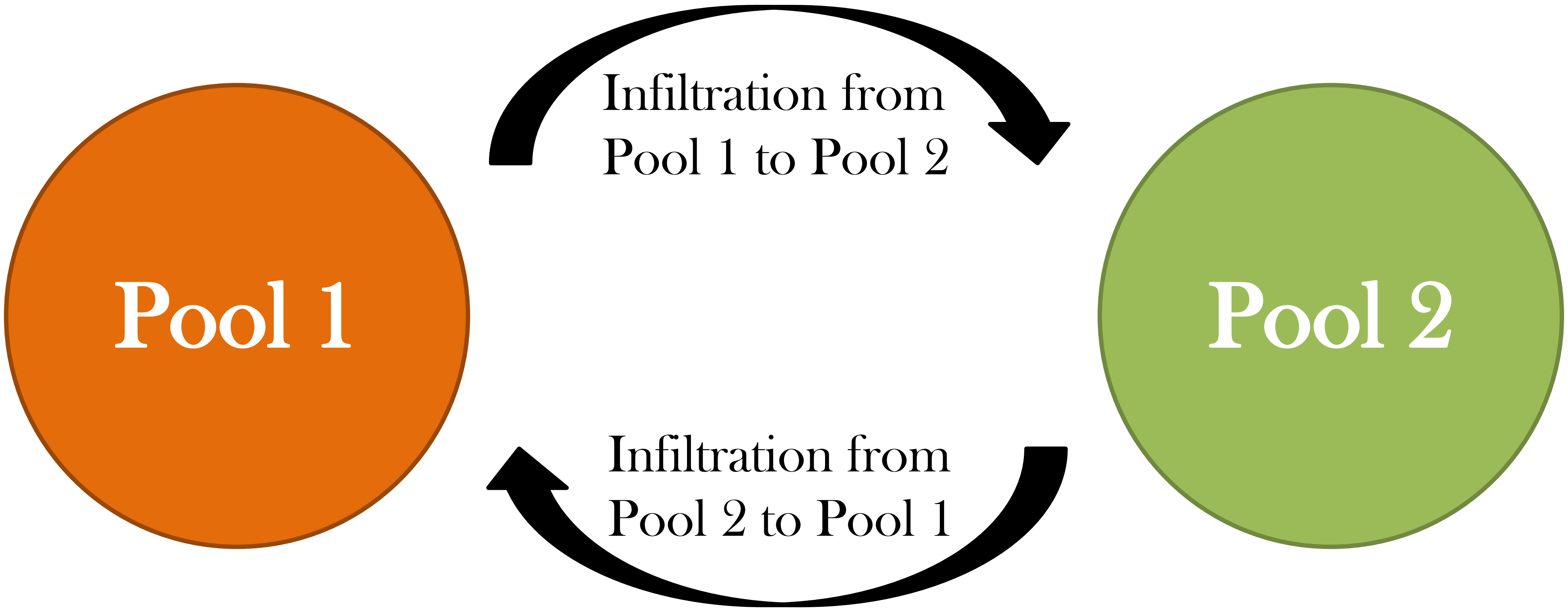
Increasing ↓ (c)

The case is equivalent to the case of the BWH attack. → (row c = 0)

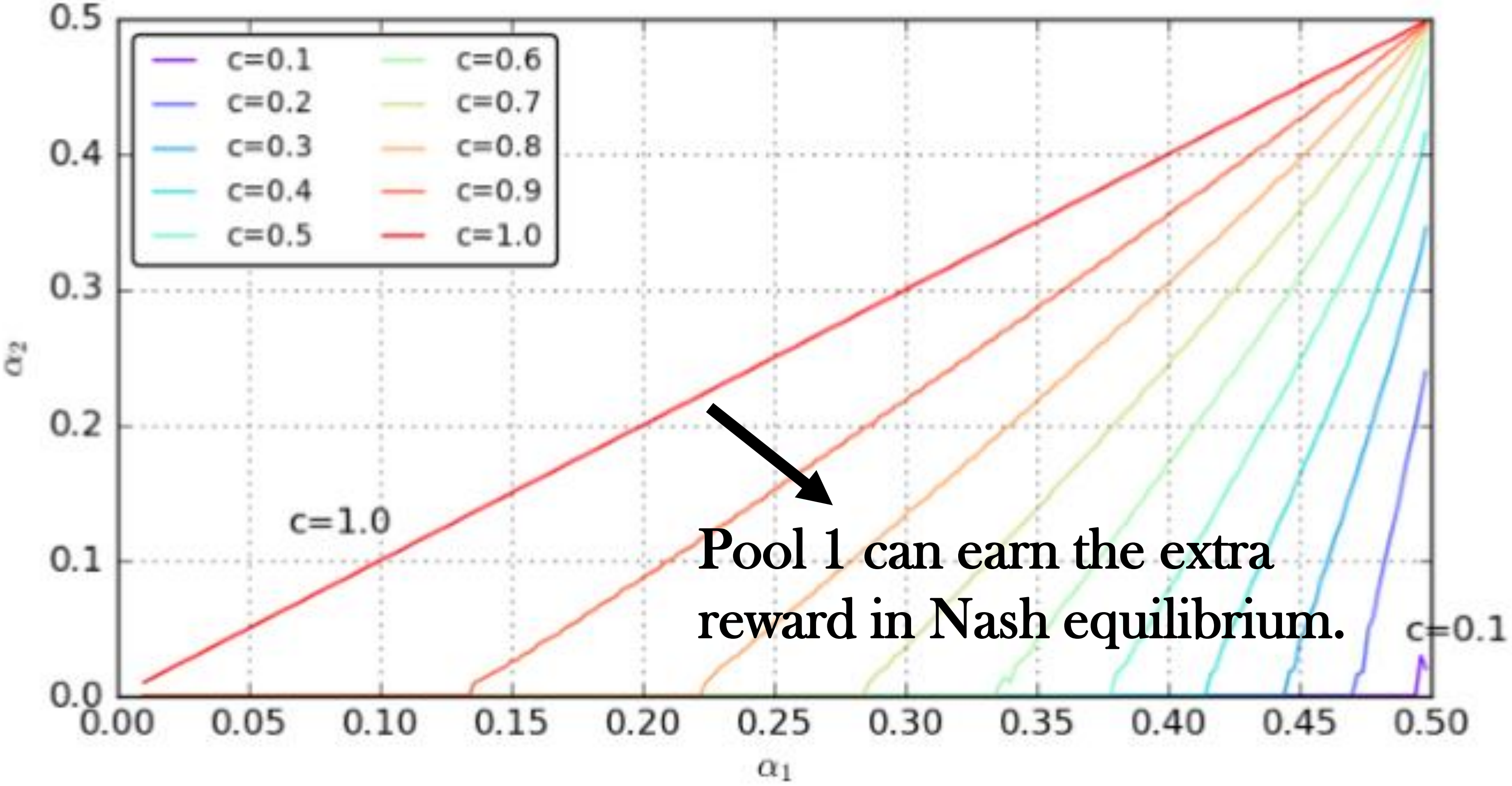
❖ We can see that the FAW attack is more profitable than the BWH attack numerically.

FAW Attack Game

❖ Pools can launch the FAW attack each other through infiltration.



Break Dilemma



❖ FAW attacks between two pools lead to a pool size game: the larger pool can always earn the extra reward.

Identification

- The FAW attack causes high fork rate.
- The FAW attacker leaves a trace of the only victim pools' identities but not the attacker's identity.
- The manager can suspect a miner who submits FPoWs used for forks.
- The attacker may easily launch the FAW attack using many **Sybil nodes** in the victim pool.
- The attacker's behavior makes the detection **useless**.

No Silver Bullet

- New reward system
  - High variance of rewards
- Change Bitcoin protocol
  - Two-phase proof-of-work
  - Not backward compability
- **There is no one silver bullet.** 

Thank You!

dbwls8724@kaist.ac.kr