Evaluating Risks in Unfair Gameplay on Pokémon GO with Bayesian Network

Zhengxiao Wei

Risk Modeling and Assessment

June 15, 2019

Identified Risks

Uncertainty: punishment for unfair gameplay

Focus area: Pokémon GO spoofing, the behavior of falsifying location data

Risk areas:

- bans
  - first strike: warning
  - second strike: suspension
  - third strike: termination

- If a player violates the Terms of Service (TOS) by simulating GPS on iOS, then any strike will possibly be issued.

- If a player has received the first two strikes, then a third strike will possibly be issued.

- If a player deceptively appeals against a second strike to the support team, then a third strike will possibly be issued.

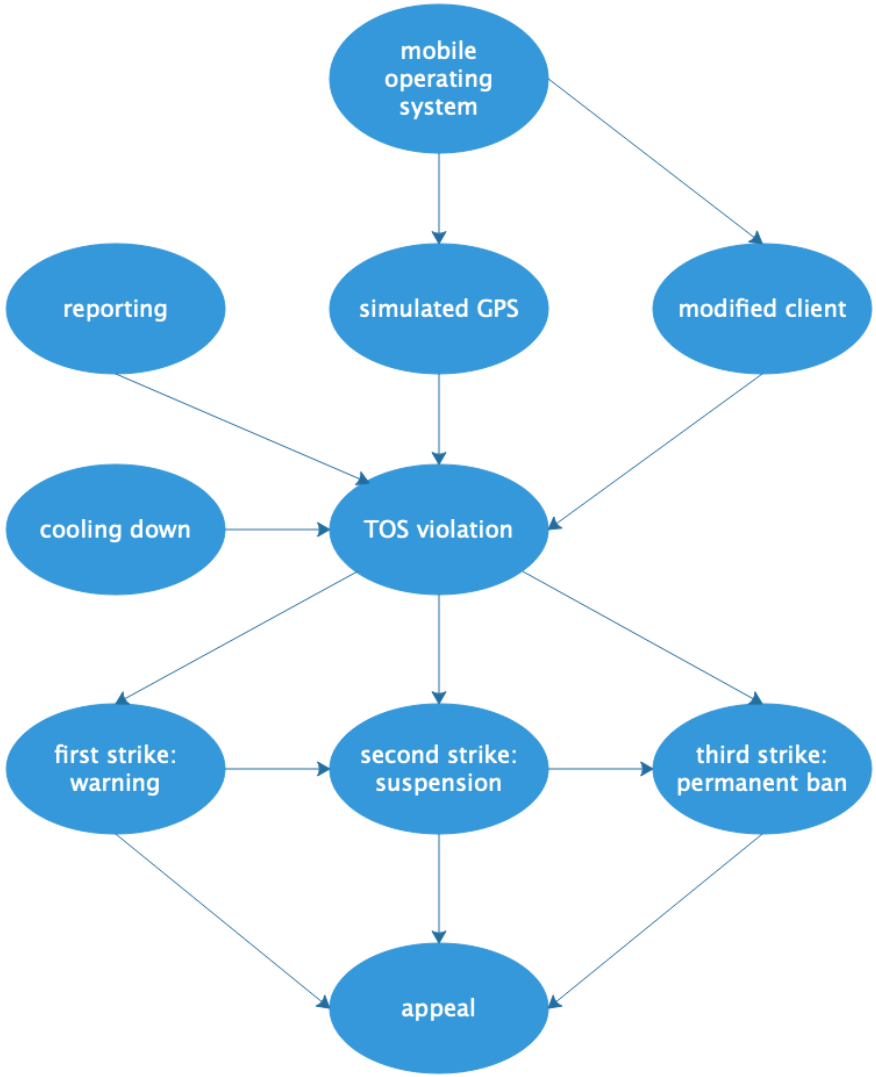
Modeling

Assumptions:

- 10 Bernoulli random variables
- the subject is the **GPS cheaters** instead of active players
- a directed acyclic graph

Sub-models:

- a chain: $X \to Y \to Z$
- a fork: $X \leftarrow Y \to Z$, a structure of common cause
- a collider: $X \to Y \leftarrow Z$, a structure of common effect
- detection $\to$ warning $\to$ decision-making



Data source: Statista

Node	NPT		Justification
mobile operating system (os)	iOS	0.38	62% of spoofers are Android users.
	Android	0.62	

simulated GPS	mobile os	iOS	Andriod
	Simulated	0.1	1
	Not	0.9	0

All Android spoofers simulate GPS.
10% of iOS spoofers simulate GPS.

modified client	mobile os	iOS	Andriod
	Modified	0.8	0.05
	Not	0.2	0.95

5% of Android spoofers use a modified client.
80% of iOS spoofers use a modified client.

cooling down		
	Respected	0.96
	Not	0.04

96% of cheaters respect the cool-down time.

reporting		
	Reported	0.01
	Not	0.99

1% of cheaters get reported.

TOS violation	simulated GPS	Simulated							
	reporting	Reported				Not			
	modified client	Modified		Not		Modified		Not	
	cooling down	Respected	Not	Respected	Not	Respected	Not	Respected	Not
	Detected	0.6	0.8	0.5	0.75	0.58	0.79	0.4	0.74
	Not	0.4	0.2	0.5	0.25	0.42	0.21	0.6	0.26

The probability of being detected given the four conditions. For example, 80% are detected violation given simulated, reported, modified, and not respected.

TOS violation (cont'd)	simulated GPS	Not							
	reporting	Reported				Not			
	modified client	Modified		Not		Modified		Not	
	cooling down	Respected	Not	Respected	Not	Respected	Not	Respected	Not
	Detected	0.3	0.7	0.01	0.65	0.25	0.72	0	0.55
	Not	0.7	0.3	0.99	0.35	0.75	0.28	1	0.45

1% are detected violation when reported, not simulated, not modified, and respected.
It is totally safe when not simulated, not report, not modified, and respected.

first strike: warning	TOS violation	Detected	Not
	Issued 1st	0.9	0.01
	Not	0.1	0.99

90% of players that have been detected violation are warned. For those not detected, still 1% are warned.

second strike: suspension	TOS violation	Detected		Not	
	first strike	Issued 1st	Not	Issued 1st	Not
	Issued 2nd	0.95	0.25	0.15	0.01
	Not	0.05	0.75	0.85	0.99

Given being detected and warned, 95% are suspended. Given not detected or warned, still 1% are suspended.

third strike: permanent ban	TOS violation	Detected		Not	
	second strike	Issued 2nd	Not	Issued 2nd	Not
	Issued 3rd	0.98	0.05	0.1	0
	Not	0.02	0.95	0.9	1

98% are permanently banned after being detected and suspended. If not detectd or suspended, no one is permanently banned.

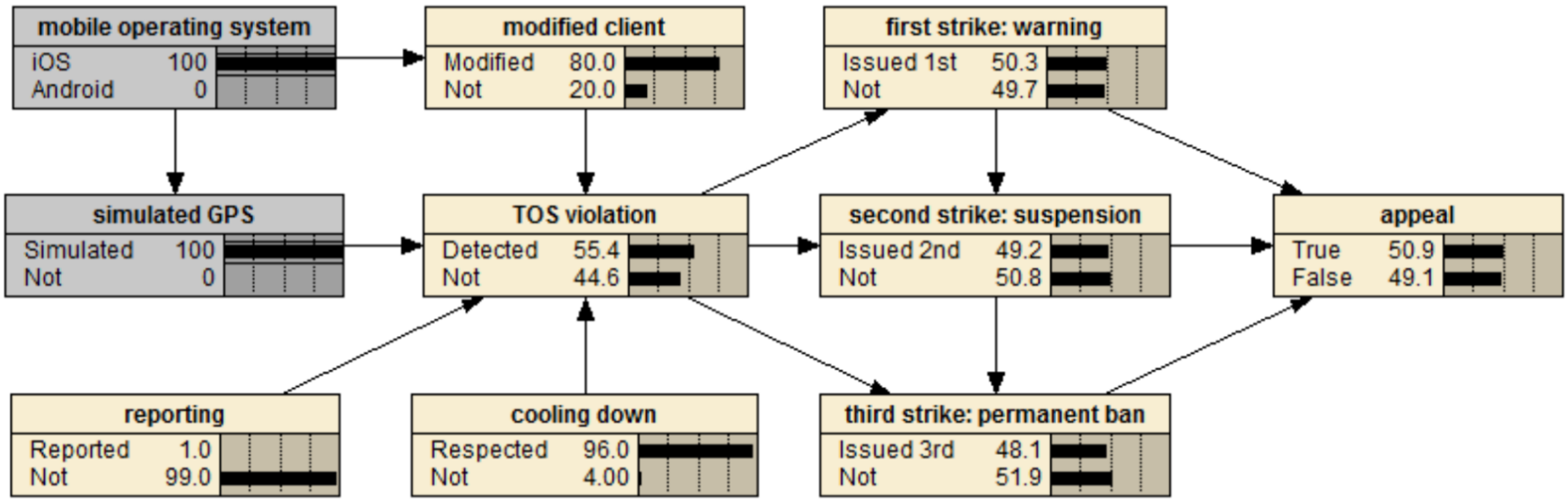
appeal	first strike	Issued 1st				Not			
	second strike	Issued 2nd		Not		Issued 2nd		Not	
	third strike	Issued 3rd	Not	Issued 3rd	Not	Issued 3rd	Not	Issued 3rd	Not
	True	1	0.95	0.75	0.6	0.99	0.8	0.5	0
	False	0	0.05	0.25	0.4	0.01	0.2	0.5	1

The probability of remaining banned after appeal given the three conditions. For example, the status is unchanged if all three strikes have been issued.

Risks Assessment

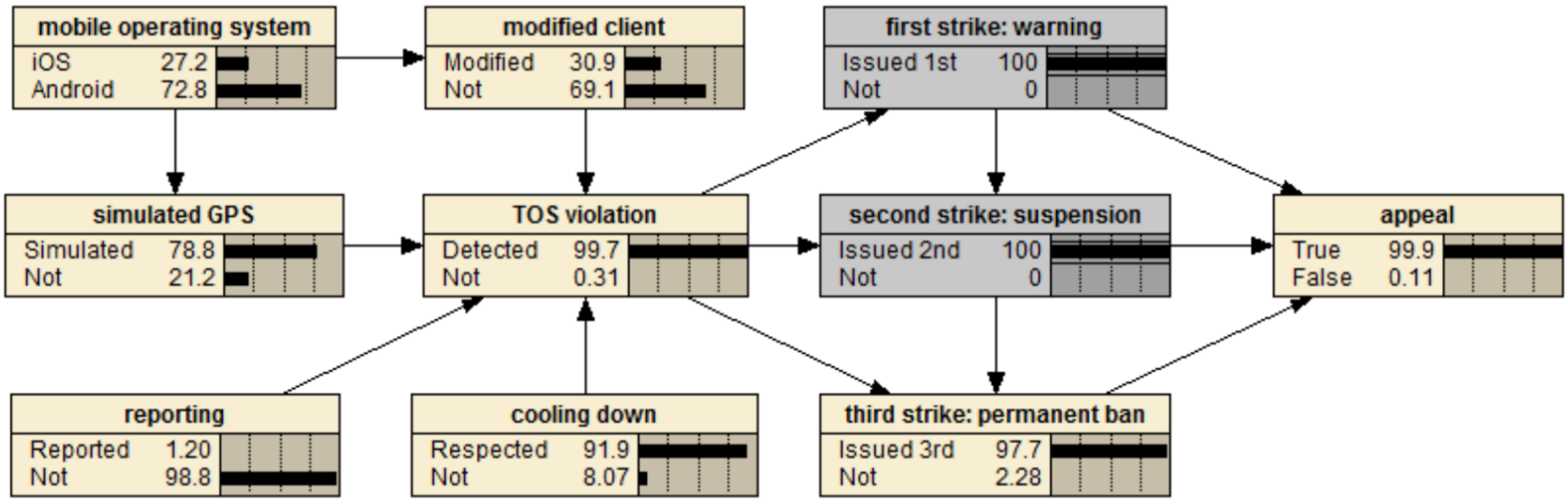
The probability of *at least* one strike issued given iOS and simulated GPS is

$$1 - (1 - 0.497) * (1 - 0.508) * (1 - 0.519) \approx 0.8810$$



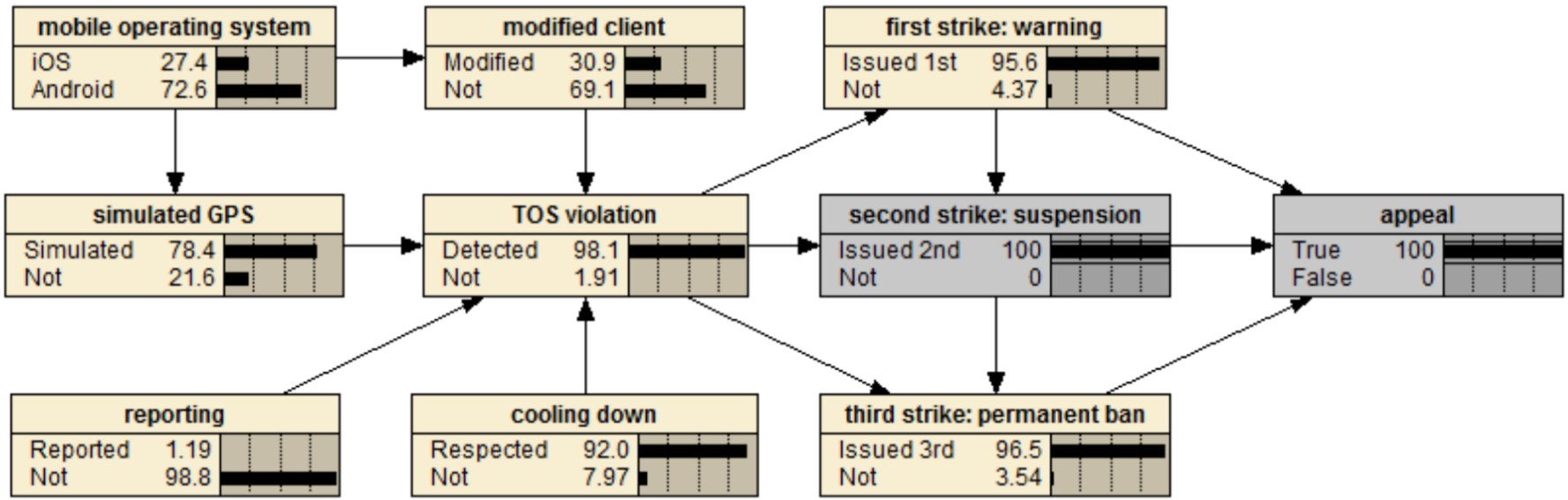
Risks Assessment

The probability of a third strike issued given both of the first two strikes already issued is 97.7%.



Risks Assessment

The probability of a third strike issued given the second strike issued and the true appeal is 96.5%.



Modification and Conclusion

Other variables:

- login from multiple devices, walking on water or through buildings, GPS drifting, and repeating routes, *etc*.

Game cheaters take the risk of their accounts being banned, while the company takes the risk of losing customers. Similar to the *credit card issuing model*, a binary classification can assist with this simulation.

- assure that the cost of conducting risk mitigation is not higher than the cost of risk outcome
- the sunk cost

If the risk is violating the integrity or the laws, one should never take the first step.