CAG – Collector vs. Adversary

1. Model Overview

This case study recreates the Collector–Adversary Game (CAG) as described in Wu et al. (2020), modeled as a hide-and-seek game between a trusted data collector and a strategic adversary. In the CAG setting, the data collector aggregates user requests to preserve privacy, while the adversary attempts to track or deanonymize a target user. The central mechanism examined is the p-destination threshold used by the data collector, which determines how many perturbed user requests must be observed before releasing an aggregated result.

The model implemented here follows the structure used in location-privacy systems:

- The data collector chooses a privacy-protection mode--either Protected (P) with threshold ppp, or Transparent (T) with minimal protection.
- The adversary chooses whether to Exploit (E) and attempt deanonymization, or Tolerate (T) and not attack.
- The system studies the strategic interaction between these two players using a 2x2 normal-form game and computes the resulting mixed-strategy Nash equilibrium.

This game fits the Collector vs. Adversary category because the data collector acts as a privacy-preserving intermediary, while the adversary aims to extract information about users. The collector's choice of protection level directly shapes the adversary's incentives to attack.

2. Players

- Data Collector (DC)
    - A trusted third party responsible for aggregating user requests.
    - Seeks to reduce privacy leakage while minimizing processing cost.
    - Chooses whether to use protection (P) with user-defined threshold ppp, or operate transparently (T).
- Adversary (ADV)
    - Attempts to track or deanonymize a target user.
    - Gains value from successful attacks but pays a cost to launch them.
    - Chooses to attack (E) or refrain (T).

3. Strategies

3.1 Data Collector

- P (Protect): Aggregate requests only after observing ppp perturbed reports. Reduces adversary success probability but increases computation cost.
- T (Transparent): Minimal protection and lower cost but higher probability of successful attack.

3.2 Adversary

- E (Exploit): Attack a user, attempting to infer true location.
- T (Tolerate): Do not attack; avoid cost.

4. Payoff Functions

- The payoff structure in the CAG is based on that stronger protection helps the data collector but costs more, and the adversary only attacks when it is worth the effort.
- When the data collector (DC) uses protection (P), the adversary's chance of success goes down, but the DC pays higher processing cost.
- When the DC is transparent (T), the system is cheaper to run, but the adversary has a higher chance of succeeding.
  The adversary attacks (E) only if the expected value of breaking privacy exceeds the attack cost.
- Because each player's best choice depends on the other's action, there is no pure strategy equilibrium--they end up mixing between their strategies.

To describe these outcomes, we use:

- advValueSuccess – value adversary gains if deanonymization succeeds
- advAttackCost – cost of attempting an attack
- successProbTransparent -- attack success chance when DC plays T
- successProbProtected(p) -- attack success chance when DC plays P
- dcLossOnBreach -- loss to DC if an attack succeeds
- dcPrivacyBenefitTransparent / dcPrivacyBenefitProtected(p) – privacy benefits under T or P
- dcCostTransparent / dcCostProtected(p) -- system costs under T or P

4.1 Outcomes

(P, E) – DC protects, adversary attacks

- Adversary gets expected success value minus attack cost
- DC gets privacy benefit minus protection cost minus expected breach loss

(P, T) – DC protects, adversary does not attack

- Adversary gets 0
- DC gets benefit of P minus cost of P

(T, E) – DC transparent, adversary attacks

- Adversary gains more (higher success chance)
- DC suffers breach loss along with normal transparent-mode cost

(T, T) – DC transparent, adversary does nothing

- Both get simple baseline payoffs; no attack-related costs or losses

5. Equilibrium Concept

5.1 Mixed-Strategy Nash Equilibrium

Wu et al. shows that the CAG's payoff structure prevents any pure strategy Nash equilibrium from forming. Instead the data collector mixes between P and T and the adversary mixes between E and T with.

5.2 Why this fits:

- The CAG is not a mechanism-design problem; instead, it is a strategic two-player zero-sum-like game.
- Each player reacts to the other's incentives, leading to mixed behavior that balances privacy risk and attack cost.
- Mixed equilibrium is the solution concept used in Wu et al. and fits the payoff geometry.

6. Implementation Details

The play the game, computation are set in order. The following steps are how a game is performed.

- params.py: defines economic and privacy parameters and how ppp affects payoff functions.
- mechanisms.py: builds the 2x2 payoff matrix and solves for mixed-strategy Nash equilibrium.
- simulate.py: sweeps over different values of ppp and records leakage, DC payoff, and ADV payoff.
  players.py: defines strategy types and optional simulation actors.
- main.py: runs the ppp-sweep and prints results.

7. Results

The output is:

p=1: x*=0.646, y*=0.323, leakage=0.194
p=4: x*=0.226, y*=0.452, leakage=0.271
p=8: x*=0.165, y*=0.661, leakage=0.397

7.1 Key Observations

- Increasing p means that DC chooses P less often. Protection becomes too costly, so the collector only occasionally uses it in equilibrium.
- Increasing p → adversary attacks more frequently. As DC plays P less, attacks become more attractive.

- Leakage probability increases with p. Larger thresholds reduce protection effectiveness, increasing risk.
- DC payoff decreases with p. High protection cost and growing adversary exploitation drive down utility.

8. Interpretation

The CAG simulation demonstrates how the data collector's protection choice interacts with adversarial incentives. Unlike the OCG, where mechanism design enforces truthfulness and optimizes welfare, the CAG reveals a fundamentally adversarial dynamic:

- Stronger thresholds do not necessarily improve privacy.
- Higher costs make the DC avoid protection, encouraging more attacks.
- Mixed strategies capture the stochastic nature of privacy–attack interactions in real systems.

The takeaway from Wu et al. and from our reproduction is that privacy protection is not only a technical problem, it is strategic. Raising thresholds without considering attacker incentives can worsen outcomes. The CAG thus highlights the need for systems where privacy mechanisms must account for rational attackers rather than relying on static protection heuristics.