Drones Security and Privacy: Services and Protocols

CPS and IoT Security

Alessandro Brighente

Master Degree in Cybersecurity

UNIVERSITÀ DEGLI STUDI DI PADOVA

SPRITZ SECURITY & PRIVACY RESEARCH GROUP

Security Society form Drones

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- We here focus on the threat represented by drones
- Relevant examples include the Gatwick and Erbil attacks
- Three types of target:
  - Individuals: security, safety, and privacy of civilians
  - Companies/Organizations
  - Nations: critical infrastructure, military bases, national facilities

Attacker Profile

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- Attacker is any pilot who uses a drone to violate the security and privacy of individuals, organizations, and nations
- Can act independently, on commission, or be part of a criminal/terrorist organization
- Main threats include
  - Spying: video streaming and tracking jeopardizing privacy
  - Terrorism: physical attack leading to injury or death of people and/or destruction of facilities
  - Cyber-Attacks: attacks against computing devices to break into networks and steal information
  - Smuggling: any type of illegal transfer of goods via drones

Countermeasures

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- Ways to protect against malicious drones
- Based on operational range and TRL
- This envisions a process with specific steps
  - Detection
  - Assessment
  - Interdiction

Remote Drone Identification

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- Remote ID is a standardized feature provided by the Federal Aviation Administration
- "*Remote ID is the ability of a drone in flight to provide identification information that can be received by other parties*"
- Drone authentication and identification can be performed by authenticating the drone or the pilot
- Use an identification platform nearby drones based on dedicated short range communications (e.g., DJI Aeroscope)
- Provide detailed information on operators ID, flight location and altitude

Remote ID Options

- Three ways drone pilots can meet Remote ID requirements
- Operate a standard Remote ID drone: one that is produced with built-in remote ID broadcast capability in accordance with the remote ID rule's requirements
- Operate a drone with Remote ID broadcast module: an additional module that retrofits drones to comply with Remote ID
- Operate without Remote ID equipment: at FAA-recognized identification areas (FRIAs) sponsored by community-based organizations or educational institutions

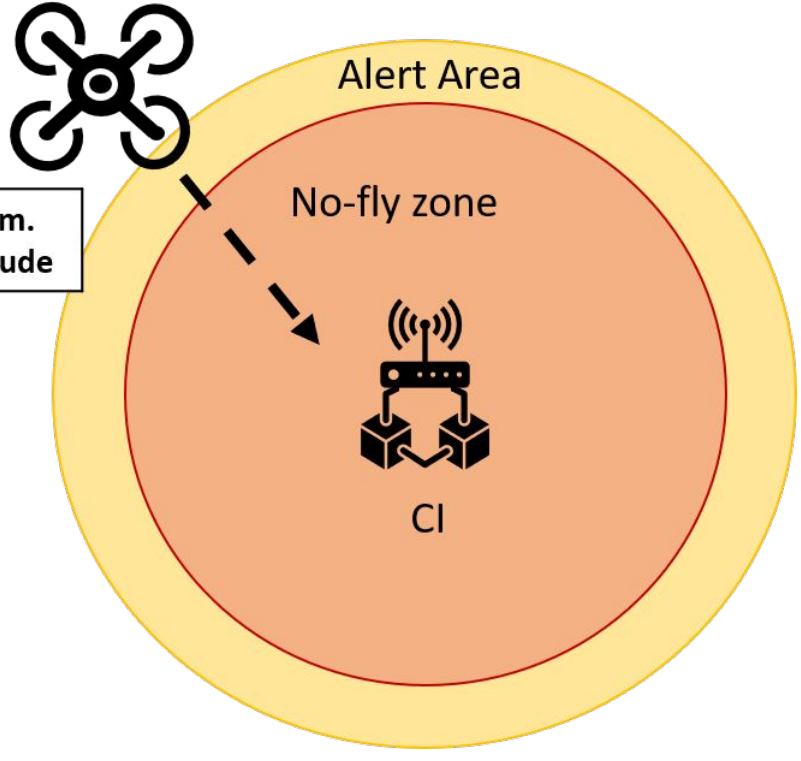
Remote Drone Identification

- Every drone should either natively support RemoteID or be equipped with a RemoteID module
- Messages shall be broadcasted (every max. 1 second) from take-off to shutdown, including
  - Unique identifier
  - Latitude, longitude, geometric altitude, and velocity
  - Latitude, longitude, geometric altitude of control station or take-off location
  - Time mark
  - Emergency status

Remote Drone Identification

Remote ID message

ID	Latitude	Longitude	Geom. altitude



Differentially Private Remote ID

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- We would like drone users to preserve drone's privacy
- Therefore, avoid disclosing precise location information
- At the same time, we want critical infrastructures to be able to detect drone invasions in no-fly zones
- We want to avoid complicated operations at the drone said to save energy
- We do not want to establish secure channels with the CI, as this would require some key exchange or a pre-shared key

Differential Privacy

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- Differential privacy is a methodology to publicly share information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset
- Main idea: if the effect of making an arbitrary single substitution in the database is small enough, the query result cannot be used to infer much about single individuals
- Usually, we publish aggregate statistics without providing means for inferring private sensitive information

epsilon-Differential Privacy

- Mathematical definition for the privacy loss associated with any data release drawn from a statistical database
- Intuition: a person's privacy cannot be compromised by a statistical release if their data are not in the database
- The goal is to give each individual roughly the same privacy that would result from having their data removed
- Inverse relationship between number of contributors and amount of noise

epsilon-Differential Privacy

- Let ε be a positive real number and $\mathcal{A}$ be a randomized algorithm (i.e., an algorithm that employs a degree of randomness as part of its logic)
- The algorithm takes a dataset as input, which is the action of the trusted party holding the data
- Let $\operatorname{im} \mathcal{A}$ denote the image of $\mathcal{A}$
- The algorithm provides ε-DP if, for all datasets D1 and D2 that differ on a single element and all subsets S of $\operatorname{im} \mathcal{A}$

$$\Pr[\mathcal{A}(D_1) \in S] \leq \exp(\varepsilon) \cdot \Pr[\mathcal{A}(D_2) \in S]$$

Properties: Composability

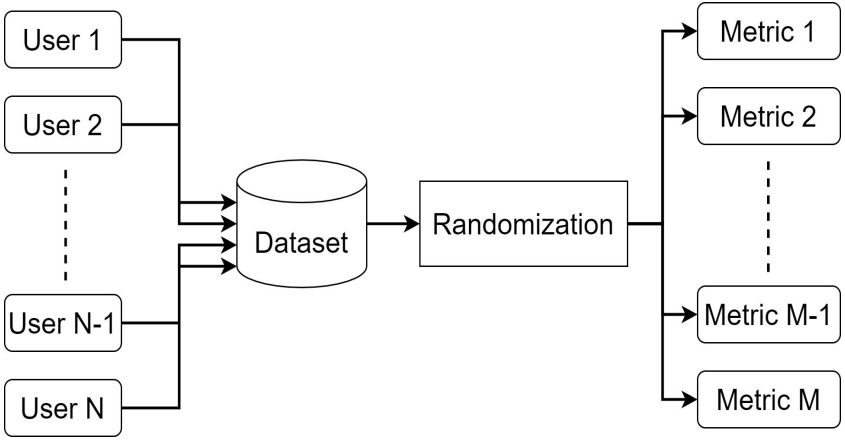
- The joint distribution of the outputs of possibly adaptive chosen differentially private mechanisms satisfies differential privacy
- If we query an epsilon-DP mechanism t times and the outcomes of the randomization algorithm are independent and identically distributed (IID) the result will be εt-DP
- In general, given n independent mechanisms $\mathcal{M}_1, \ldots, \mathcal{M}_n$ with privacy guarantees $\varepsilon_1, \ldots, \varepsilon_n$-DP, then any function g() of them is $\left(\sum_{i=1}^{n} \varepsilon_i\right)$-DP

Properties: Robustness and Group Privacy

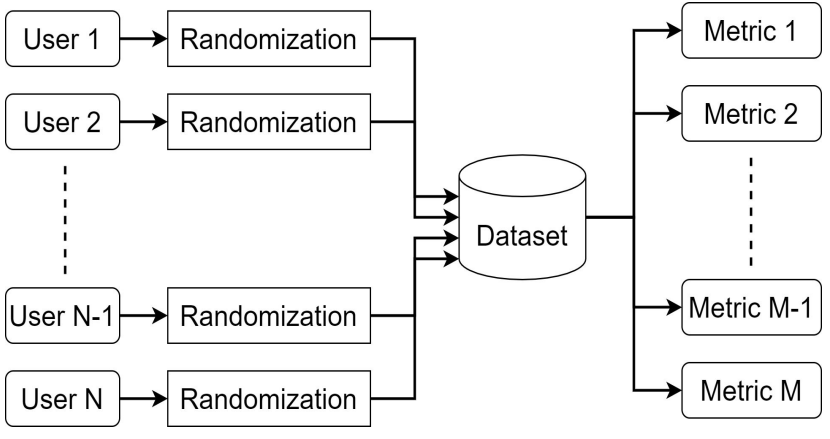
- For any deterministic or randomized function F defined over the image of the randomization algorithm, if $\mathcal{A}$ satisfies epsilon-DP, so does $F(\mathcal{A})$
- We may want to protect the privacy of databases that differ in c > 1 rows, similar to having an adversary with arbitrary auxiliary information knowing if c participants submitted their information
- If c items change, the probability dilatation is bounded by $\exp(\varepsilon c)$

$$\Pr[\mathcal{A}(D_1) \in S] \leq \exp(\varepsilon c) \cdot \Pr[\mathcal{A}(D_2) \in S]$$

DP vs Local DP



Differential Privacy



Local Differential Privacy

Local Differential Privacy

- Opposed to the centralized differential privacy model, in this case we have no central aggregator
- Thus, the adversary cannot have access to the raw data generated by the user
- Let ε be a positive real number and $\mathcal{A}$ be a randomized algorithm
- The algorithm proves local ε-DP if for all pairs of users' possible private data x and x', and S subset of $\mathrm{im}\mathcal{A}$

$$\Pr[\mathcal{A}(x) \in S] \leq e^{\epsilon} \times \Pr[\mathcal{A}(x') \in S]$$

Protecting Users' Locations

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- The concepts of DP and local DP have been introduced for databases
- However, the underlying concept is useful also for other applications
- Location Based Services: based on users' geographic data, offer specific computing services
- Nice, but as user you do not want to provide your exact location to these services
- We need a DP-way to encode such information

Geo-Indistinguishability

- We want to provide users with privacy guarantee within a radius r
- As for DP, we define a parameter ϵ corresponding to the level of privacy
- We define a mechanism called ϵ-Geo-Indistinguishability, where a user at a radius r enjoys ϵr-privacy
- We can specify the user's privacy requirements via a tuple (ℓ, r) where the former is the privacy level aimed for at a radius r
- We can define $\epsilon = \ell / r$

Probabilistic Model

- We define the set $\mathcal{X}$ of points of interest, typically the user's possible locations
- We define the set $\mathcal{Z}$ of possible reported values, i.e., obfuscated locations
- We sometimes consider that $\mathcal{Z}$ may contain spatial points
- The selection of a reported value is probabilistic: z can be obtained by adding noise to the true location

Definition

- We consider a randomization mechanism K, i.e., a function that given the true location reports a value z with a predefined probability
- As in DP, we define the privacy requirements as a constraint on the distance of the distribution between two points x and x'
- Let us consider the euclidean distance d(. , .)
- Enjoying ℓ-privacy within r means that for any x,x' s.t. d(x,x') <= r the distance $d_{\mathcal{P}}(K(x), K(x'))$between the corresponding distributions should be at most ℓ

Formal Definition

- A mechanism K satisfies ϵ-geo-indistinguishability if and only if for all x,x' $d_{\mathcal{P}}(K(x), K(x')) \leq \epsilon d(x, x')$
- Note that fo all points x' within a radius r from x the definition forces the corresponding distributions to be at most ϵr distant
- The definition is very similar to that of differential privacy: we use Euclidean distance instead of Hamming distance
- In our case, Hamming distance would be too strong as we need to partially disclose our location to have access to services

Achieving Geo-Ind.

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

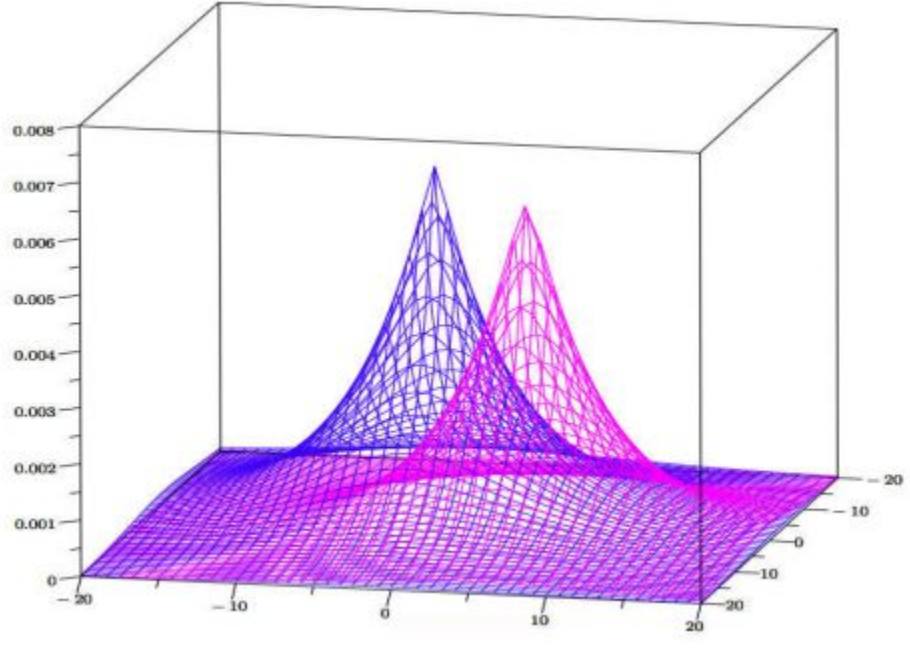
- How do we generate noise to achieve geo-ind.?
- We model the location domain as a discrete Cartesian plane with the standard notion of Euclidean distance
- We proceed as follows:
  - We define a mechanism to achieve geo-ind. In the ideal case of the continuous plane
  - We discretize the mechanism by remapping generated points to the closest point in the discrete domain
  - We truncate the mechanism to report only points within the limits of the area of interest

Mechanism for the Continuous Plane

- Whenever the true location is $x_0 \in \mathbb{R}^2$, we report a point $x \in \mathbb{R}^2$ randomly generated according to a noise function
- Such function shall be such that the probabilities of reporting a point in a certain area around x differs at most by a multiplicative factor $e^{-\epsilon\, d(x_0, x_0')}$
- We can achieve this by requiring an exponential decrease in the probability around the point area
- This is the behavior of the Laplace distribution, with pdf $\epsilon/2\, e^{-\epsilon\,|x-\mu|}$

Mechanism for the Continuous Plane

- Given parameter $\epsilon \in \mathbb{R}^+$ and the actual location $x_0 \in \mathbb{R}^2$, the pdf of our noise mechanism on any other point $x \in \mathbb{R}^2$ is $D_\epsilon(x_0)(x) = \frac{\epsilon^2}{2\pi} e^{-\epsilon\, d(x_0,x)}$



Draw a Random Point

- We notice that the pdf of the Laplacian distribution only depends on the distance from its center point
- We hence use polar coordinates to achieve a convenient representation
- A point is represented as a tuple (r, θ) = distance from center to x and angle between line center-x and horizontal axis of cartesian system

$$D_\epsilon(r, \theta) = \frac{\epsilon^2}{2\pi} r\, e^{-\epsilon r}$$

Draw a Random Point

- Radius and angle are independent
- Therefore, we can draw them independently
- We generate θ uniformly in $[0, 2\pi)$ as the polar pdf is constant with the angle
- We instead use inverse transform sampling to generate r
- Consider the CDF $C_\epsilon(r)$
- Draw a uniform value p and compute $C_\epsilon^{-1}(p)$

Discretization

- We want to approximate the Laplace mechanism on a grid of discrete Cartesian coordinates
- We remap the generate point in (r, θ) to the closest point x on the discrete grid $\mathcal{G}$
- It is not obvious that discretization preserves geo-ind.
- Indeed, noise is generated by machines which are discrete in nature
- However, it is possible to show that discretization provides geo-ind. At the price of a degradation of the privacy parameter epsilon

Truncation

- The Palace mechanism has the potential to generate points everywhere in the plane
- However, this might be a problem due to memory bounds or privacy/usability trade-off
- We still want to be able to use location-based services
- Planar Laplace mechanism: truncate the generated point by remapping it to the closest point in the intersection between the discrete grid and the finite set of admissible locations

Application to Drones

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- The concept of geo-ind. Can be applied to drones to preserve their location privacy while using services
- We consider the case where we want to protect a Critical Infrastructure (CI) from a drone invasion
- We consider three types of adversaries
  - Unaware invader
  - Rookie pilot attacker
  - Eavesdropper

Scenario and Model

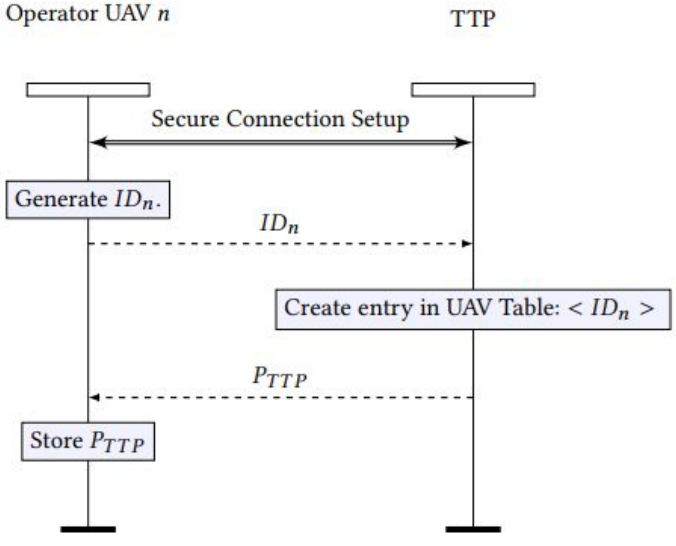
SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

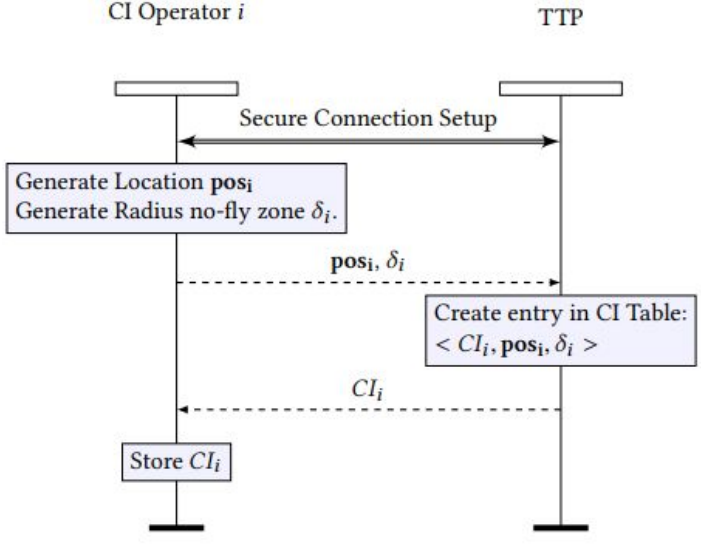
UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- We consider a UAV sending sanitized location reports to a CI
- Using geo-ind. Implies the drone sends statistical information to the CI
- Hence the privacy guarantees come with the possibility of false invasion detections from the CI
- However, once detecting an invasion, we assume the CI wants to check the real location of the drone before deactivating it
- Therefore, we need a Trusted Third Party (TTP)

Registration Phase

- The registration phase is executed before deployment, where both CI and UAV operator register to the TTP
- We assume this connection to be secure





Encrypted Location Reports

- To provide means to the CI operators to verify invasions, we assume the UAV sends encrypted location reports, i.e., ciphertexts encoding the precise location information
- Each UAV stores the public key of the TTP
- Everytime the UAV needs to send a new Remote ID message, it will encode its location via geo-indistinguishability and will generate a new encrypted location report with a one-time key

$$c_{n,0}^{m} = \tilde{E}\left(ID_n || x_{n,0}^{m}, K_n\right)$$

Encrypted Location Reports

- The encrypted location report is generated as $c_{n,0}^{m} = \tilde{E}\left(ID_n || x_{n,0}^{m}, K_n\right)$
- We assume we use a generic symmetric encryption algorithms (e.g., AES)
- Then, using the public key of the TTP, it generates the encrypted key

$$\tilde{K}_n = \tilde{E}_S\left(K_n, P_{TTP}\right)$$

- The encrypted location report is defined as the tuple ($c_{n,0}^{m}$, $\tilde{K}_n$)

Area Invasion Detection

SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- The CI implements a mechanism to detect invasions from UAVs based on the sanitized location
- We define as detection window W the maximum time interval where the detection system needs to provide indications of an invasion of the no-fly zone by any UAV
- To perform the detection, the CI compares its location with the sanitized one reported by the UAV
- If their distance is smaller than the allowable threshold, then report an invasion

Reporting Phase

- As soon as an invasion is detected, the CI reports to the TTP the ID of the CI operator, the remote ID messages emitted by the UAV that caused the invasion detection