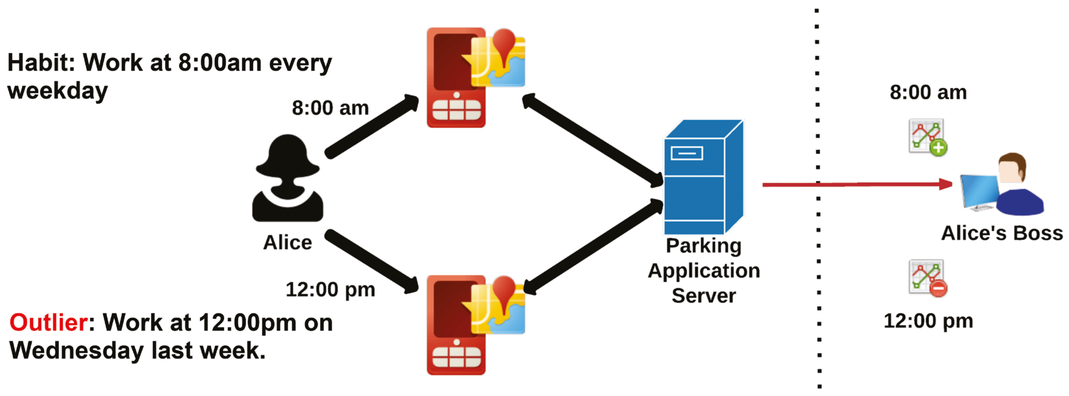
Privacy Zones:
Privacy Aware Sharing of Sensitive Data

Dr. Fayola Peters, Supervised by: Prof. Bashar Nuseibeh

1 Motivation

Privacy is the ability to understand, choose, and control what personal information an individual shares, with whom, and for how long.

» Users have the opportunity to set privacy preferences but **do not act on them in practice.**
» Data collected by mobile applications without regard for user privacy, are open to **sensitive attribute disclosure**, which occurs when a user is associated with sensitive data such as an outlier.

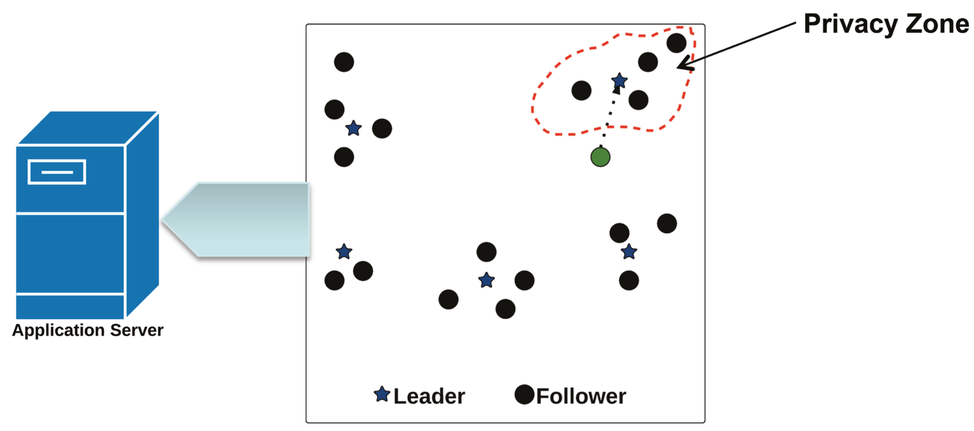


Our Idea

Privacy Zones (PZs), is an approach that allows developers to provide users with a default privacy setting based on their habits which assures their privacy and allows them to obfuscate data that are prone to sensitive attribute disclosure.

2 Privacy Zones
Identifies user habits using location and time

» PZs approach have three core features
1. Algorithm that clusters user data, clusters represent users' habits (privacy zones).
2. Privacy threat detection which identifies deviations from the norm (non-privacy zones).
3. Obfuscation to change data from non-privacy zones to their nearest privacy zones so the user can continue to use an application even when data are privacy sensitive.



PZs approach operate in the application server

The Parking Space Finder Example

Alice, passes location and time to the application server.
» Past data has been clustered.
» New data finds the nearest cluster.
» If it falls within the cluster, the parking application alerts Alice about available parking spots near her current location, otherwise;
» Alice is alerted about being in a non-privacy zone.

3 Experimental Setup

Check-in Data from Alice, Bob and Eve

Compare 3 Data Sharing Strategies

→ Always sharing
→ Probabilistic sharing
→ Privacy Zone sharing

Comparison Metrics

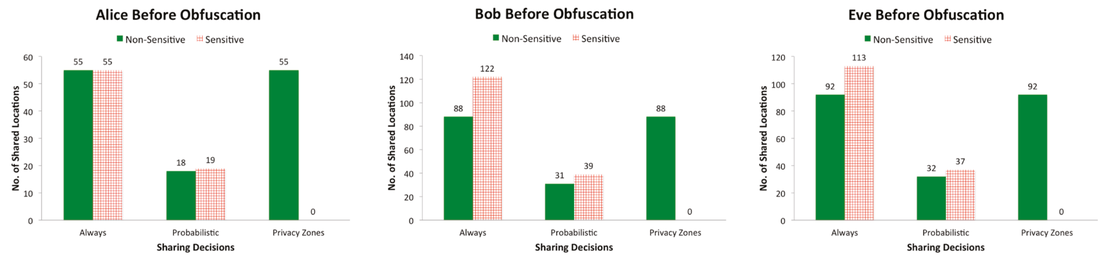
» Effectiveness of Privacy Zones:
› Trained on past (90%) – Cluster;
› Tested on future (10%) – For each check-in, find out if in the zone or not?

» Adaptability of Privacy Zones
› Trained week by week on past (90%);
› Tested on future (10%);
› Benefit = number of PZ check-ins shared / total number of check-ins

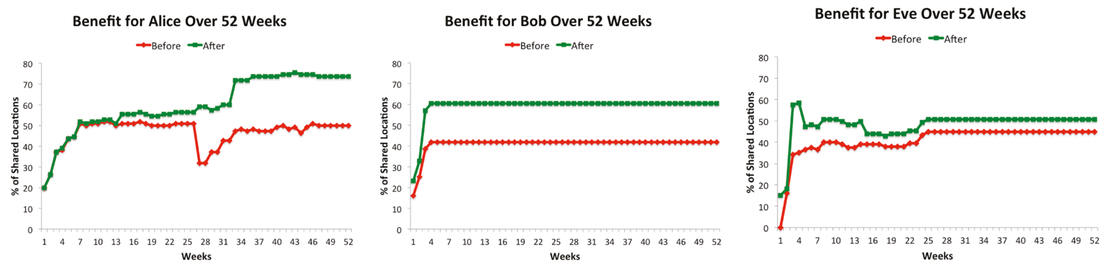
» Cost of Privacy Zones:
› Recorded time taken to return results for 10% (6 runs);
› Report execution times for mean, lower (2.5%) and upper (97.5%) quantiles.

4 Results

Effectiveness of Privacy Zones



Adaptability of Privacy Zones



Cost of Privacy Zones

Table I: Execution time (seconds) before obfuscation.

Subjects	lower quantile (2.5%)	mean	upper quantile (97.5%)
Alice	0.098	0.100	0.101
Bob	0.147	0.148	0.148
Eve	0.128	0.131	0.134

Table II: Execution time (seconds) after obfuscation.

Subjects	lower quantile (2.5%)	mean	upper quantile (97.5%)
Alice	0.868	0.874	0.885
Bob	1.580	1.590	1.600
Eve	6.974	6.999	7.023

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
United Technologies Research Center
SOUTHERN Regional Assembly
Promoting our Region