Problem Setting
Accurate trajectory patterns are necessary for Location-Based Services. A method that can collect exact trajectories in a privacy-preserving manner is needed. A method that uses free and energy-saving short-range P2P communication is desirable. However, during such communication a fixed hardware ID is exposed. Hence it is necessary that when a data item, which contains the private trajectory with possibly secret or embarrassing locations, is communicated, the link between public and private information is broken.

Privacy-Preserving Trajectory Collection in Five Stages

Client Registration (CR)
In the CR stage, the client expresses its privacy requirements \((k, a)\). In response, the server approves a group of \(k\) clients and sends them timing parameters \((\text{start time: } T_r, \text{ reporting period: } r)\). The CR stage ensures the \(k\)-anonymity of clients.

Data Summarization (DS)
In the DS stage the server continuously records the reports, merges trajectory pieces and monitors the number of pdts received for each trajectory piece. For a given trajectory, if after \(T_r + 2\tau\) the majority parity of the number of pdts for the trajectory pieces is even the trajectory is real and is stored in the Trajectory Repository (TR), otherwise the trajectory is synthetic and is discarded. The DS stage ensures the \(k\)-anonymity of the data in TR.

Data Reporting (DR)
After the reporting period has elapsed or the client DB is full, the client enters the DR stage. In the DR stage the client determines a maximal anonymity set of pdts, in which the number of pdts for each ID is statistically equal, and sends this set to the server. The DR stage ensures the \(k\)-anonymity of the reported data.

Location Privacy Definitions
\(k\)-anonymity requires that each data item can be associated with at least \(k\) moving objects and vice versa. \(\alpha\)-diversity requires some spatial or spatio-temporal diversity in a set of locations / trajectories. Finally, \(k\)-\(\alpha\)-anonymity combines the two.

Trajectory Sampling and Anonymization (TSA)
In the TSA stage, the client continuously samples its real trajectory and generates \(k\)-1 realistic and pair-wise \(\alpha\)-diverse synthetic trajectories and cuts the trajectories into pieces at every \(\tilde{T}\)-period. Trajectory pieces of a trajectory are tagged with an ID and form partial data items (pdts). At every \(\tilde{T}\)-period an even number of copies of sampled pdt and odd number of copies of the generated pdts are stored in the trajectory DB of the client. The TSA stage ensures the \(k\)-\(\alpha\)-anonymity of the client trajectory DB.

Trajectory Exchange (TE)
In the TE stage, the client periodically performs a Neighborhoood Discovery (ND) process to find other clients to exchange pdts with. The pdts to be exchanged are randomly selected, but contain at least two sampled or generated-pdts and older pdts are prioritized. The TE stage ensures the \(k\)-\(\alpha\)-anonymity of the exchanged data.

Empirical Evaluation and Results
Realistic simulation shows that the method works under reasonable conditions and anonymity settings (communication range = 10 meters and \(k = 5\) is shown). In particular, most clients can report most of the collected data in a privacy-preserving fashion. The collection is virtually lossless. In summary, the proposed system collects exact trajectories without loss, does not require trusted components, and provides strong privacy guarantees.