# Location Obfuscation For Location Data Privacy

Vaibhav Ankush Kachore, J. Lakshmi, S. K. Nandy

Indian Institute of Science, Bangalore

*vaibhav_kachore@ssl.serc.iisc,.in, {jlakshmi, nandy}@serc.iisc.ernet.in*

# Overview

- Motivation
- Encryption v/s Obfuscation
- Attacker Model
- Challenges
- Related Work
- User Obfuscation Functions
  - Ellipsoidal Random Obfuscation Function
  - Modified Random Obfuscation Function
  - Grid Obfuscation Function
- Results
- Conclusion and Future Work
- References

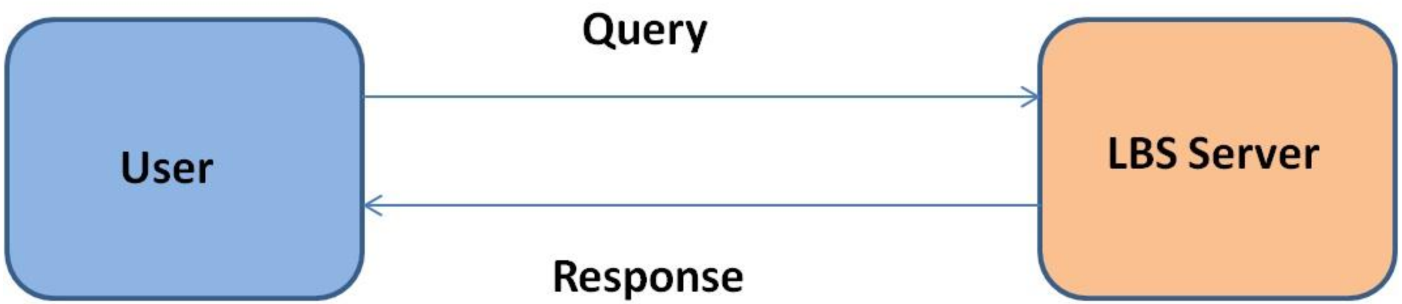
# Introduction and Motivation

- Advancement of wireless internet, sensing and mobile positioning technologies.
- Popularity of location based services (LBSs) among mobile users.
- According to PRIP(Pew Research Internet Project),74% of adult smartphone user's use their phone to get directions or other information based on their current location.
- Many enterprises are willing to purchase geo-location data, and use them to analyze potential customer preferences.
- They can better understand customer requirements and expectations, they can analyze market trends and customize the content of their applications.
- But, while doing this, the user privacy needs to be maintained.

# Why obfuscation and not encryption?

- Encryption can give protection against third party attacks.
- Encryption cannot provide protection against privacy threats from server side.
- Location obfuscation is a technique to protect user privacy by altering the location of the users.
- Obfuscation preserves capability of server to compute few mathematical functions over the obfuscated location information.
- So, this study tries to bridge the gap between user privacy and accurate query results of LBSs without much overhead.
- This study mainly concentrates on giving user privacy in LBSs which wants to know the distance travelled by user for providing their services. e.g. RunKeeper, SportsTracker, Runtastic, etc.

# Attacker Model



- Attackers can spoof a link between user and LBS server.
    - Solution : Use Encrypted Service.
- LBS server itself can be an attacker.
    - Solution : Encryption is not useful in this case. Use Obfuscated Service.

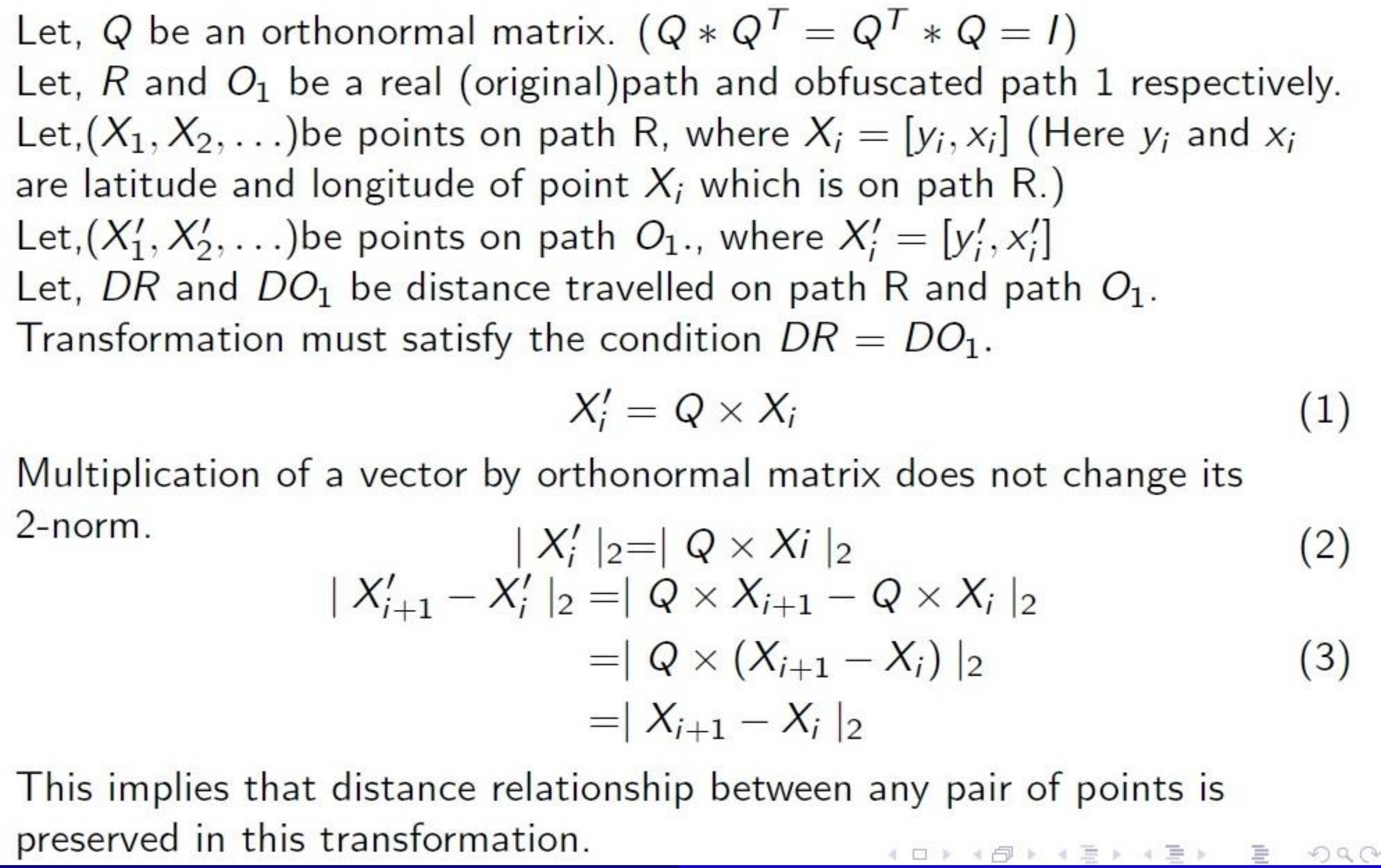
# Challenges

- Ensuring privacy of user without much obfuscation technique overhead.
- Inverse relationship between functionality of application and user privacy.
- Accuracy.

# Related Work

- Use of dummies.
    - Problem: Huge Overhead.
- Addition of noise.
    - Problem: Accuracy of results.
- Use of pseudonym: technique in which the real identity of user is replaced by fake identity.
    - Problem: Chances of revealing actual identity due to attack by intruder.
- Use of accelerometer and gyroscope sensors.
    - Problem: Accuracy of results.
- Matlock.
    - Problem: Need of irreversible layer for complete privacy.
- Obfuscation functions : Random Obfuscation Function (ROF), Linear Obfuscation Function (LOF)

# Proposed Approach for EROF



Let, $Q$ be an orthonormal matrix. $(Q * Q^T = Q^T * Q = I)$
Let, $R$ and $O_1$ be a real (original)path and obfuscated path 1 respectively.
Let,$(X_1, X_2, \ldots)$be points on path R, where $X_i = [y_i, x_i]$ (Here $y_i$ and $x_i$ are latitude and longitude of point $X_i$ which is on path R.)
Let,$(X_1', X_2', \ldots)$be points on path $O_1$., where $X_i' = [y_i', x_i']$
Let, $DR$ and $DO_1$ be distance travelled on path R and path $O_1$.
Transformation must satisfy the condition $DR = DO_1$.

$$X_i' = Q \times X_i \tag{1}$$

Multiplication of a vector by orthonormal matrix does not change its 2-norm.

$$\mid X_i' \mid_2 = \mid Q \times Xi \mid_2 \tag{2}$$

$$\begin{aligned} \mid X_{i+1}' - X_i' \mid_2 &= \mid Q \times X_{i+1} - Q \times X_i \mid_2 \\ &= \mid Q \times (X_{i+1} - X_i) \mid_2 \qquad (3) \\ &= \mid X_{i+1} - X_i \mid_2 \end{aligned}$$

This implies that distance relationship between any pair of points is preserved in this transformation.

## Continue. . .

Let, "$O_2$" be an obfuscated path 2.
Let,$(X_1'', X_2'', \ldots)$ be points on path $O_2$, where $Xi = [y_i'', x_i'']$
Let, "$DO_2$" be distance travelled on path $O_2$.
Applying ellipsoidal transformation on path $O_1$,
Ellipse has property that sum of distances of any point on the ellipse from it's foci is constant.

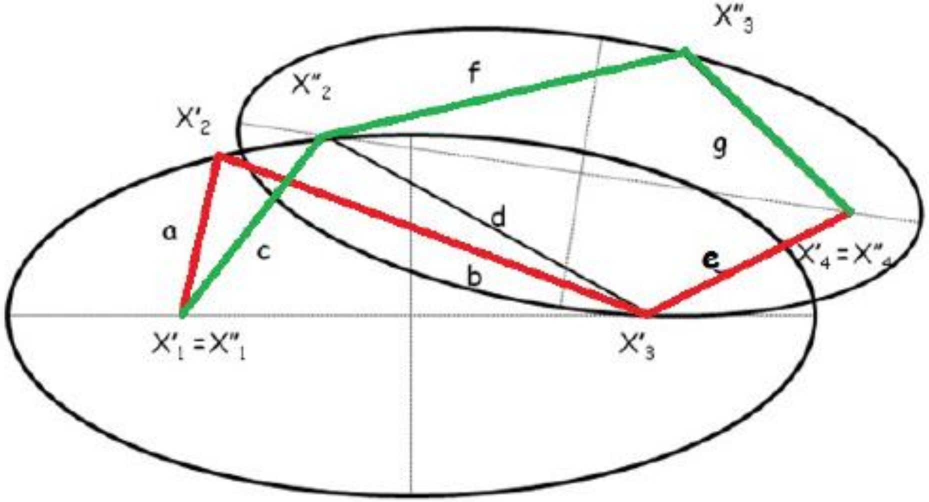


Figure: Ellipsoidal Random Obfuscation Function Example

# Continue. . .

In example of fig.2, path travelled by $O_1$ is [ $X'_1 \rightarrow X'_2 \rightarrow X'_3 \rightarrow X'_4$ ] and path travelled by $O_2$ is $[X''_1 \rightarrow X''_2 \rightarrow X''_3 \rightarrow X''_4]$.

$$DO_1 = a + b + e \tag{4}$$

$$DO_2 = c + f + g \tag{5}$$

Consider an ellipse with $X'_1$ and $X'_3$ as foci.
Using property of ellipse,

$$a + b = c + d \tag{6}$$

Similarly, for an ellipse with $X''_2$ and $X'_4$ as foci,

$$d + e = f + g \tag{7}$$

From (4), (5), (6) and (7),

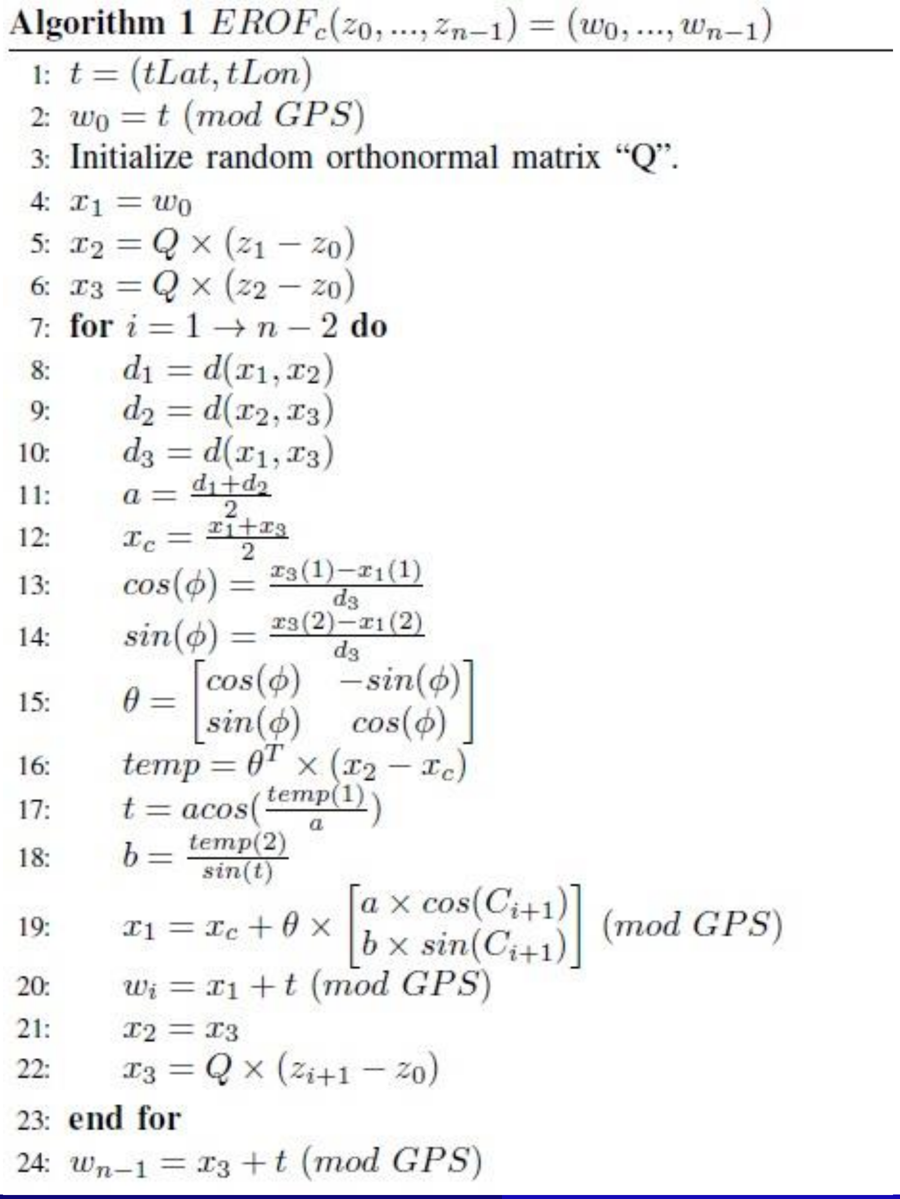
$$DO_1 = c + d + e = c + f + g = DO_2 \tag{8}$$

$\therefore$ Distance travelled on path $O_1$ and $O_2$ is same.

# Proof of irreversibility of EROF

- For calculating equation of ellipse whose major axis is at some angle with respect to X - axis , 3 points are needed because 3 parameters of the ellipse i.e. semi major axis a, semi minor axis b and angle which its major axis is making with X - axis are unknowns.
- Now, to find path O1 from O2, if any 3 consecutive points are chosen on path O2, then equation of ellipse can be found but it is not possible to know which point on this ellipse was there in path O1.
- Because all points will satisfy distance criteria. Hence, EROF is irreversible.

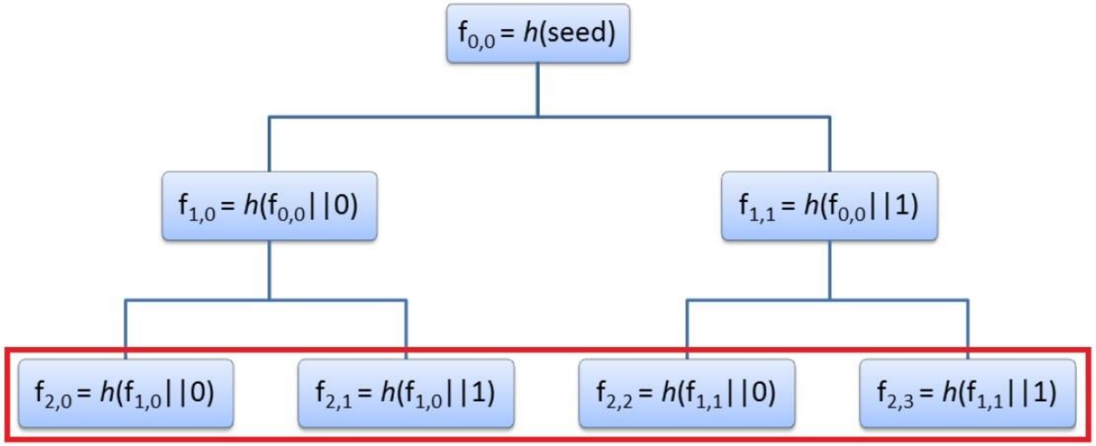
# Ellipsoidal Random Obfuscation Function



**Algorithm 1** $EROF_c(z_0, ..., z_{n-1}) = (w_0, ..., w_{n-1})$

```
$t = (tLat, tLon)$
$w_0 = t \ (mod \ GPS)$
Initialize random orthonormal matrix "Q".
$x_1 = w_0$
$x_2 = Q \times (z_1 - z_0)$
$x_3 = Q \times (z_2 - z_0)$
for $i = 1 \rightarrow n - 2$ do
    $d_1 = d(x_1, x_2)$
    $d_2 = d(x_2, x_3)$
    $d_3 = d(x_1, x_3)$
    $a = \frac{d_1 + d_2}{2}$
    $x_c = \frac{x_1 + x_3}{2}$
    $cos(\phi) = \frac{x_3(1) - x_1(1)}{d_3}$
    $sin(\phi) = \frac{x_3(2) - x_1(2)}{d_3}$
    $\theta = \begin{bmatrix} cos(\phi) & -sin(\phi) \\ sin(\phi) & cos(\phi) \end{bmatrix}$
    $temp = \theta^T \times (x_2 - x_c)$
    $t = acos(\frac{temp(1)}{a})$
    $b = \frac{temp(2)}{sin(t)}$
    $x_1 = x_c + \theta \times \begin{bmatrix} a \times cos(C_{i+1}) \\ b \times sin(C_{i+1}) \end{bmatrix} \ (mod \ GPS)$
    $w_i = x_1 + t \ (mod \ GPS)$
    $x_2 = x_3$
    $x_3 = Q \times (z_{i+1} - z_0)$
end for
$w_{n-1} = x_3 + t \ (mod \ GPS)$
```

# Merkel Tree Based Random Number Chains


$f_{0,0} = h(\text{seed})$
$f_{1,0} = h(f_{0,0} || 0)$
$f_{1,1} = h(f_{0,0} || 1)$
$f_{2,0} = h(f_{1,0} || 0)$
$f_{2,1} = h(f_{1,0} || 1)$
$f_{2,2} = h(f_{1,1} || 0)$
$f_{2,3} = h(f_{1,1} || 1)$

- It is computationally infeasible to calculate a hash value which is on the level l of the tree from another hash value which is on the level l+k of the tree with k > 0.
- It is not possible to calculate any number random number in that chain from the knowledge of any single random number.
- This ensures backward security, forward security, and the impossibility of collusion.

# Proposed Approach for Modified Random Obfuscation Function

•Problem with Random Obfuscation Function: Path can get initialized from the point which not feasible.

•Proposed solution Linear Obfuscation Function.

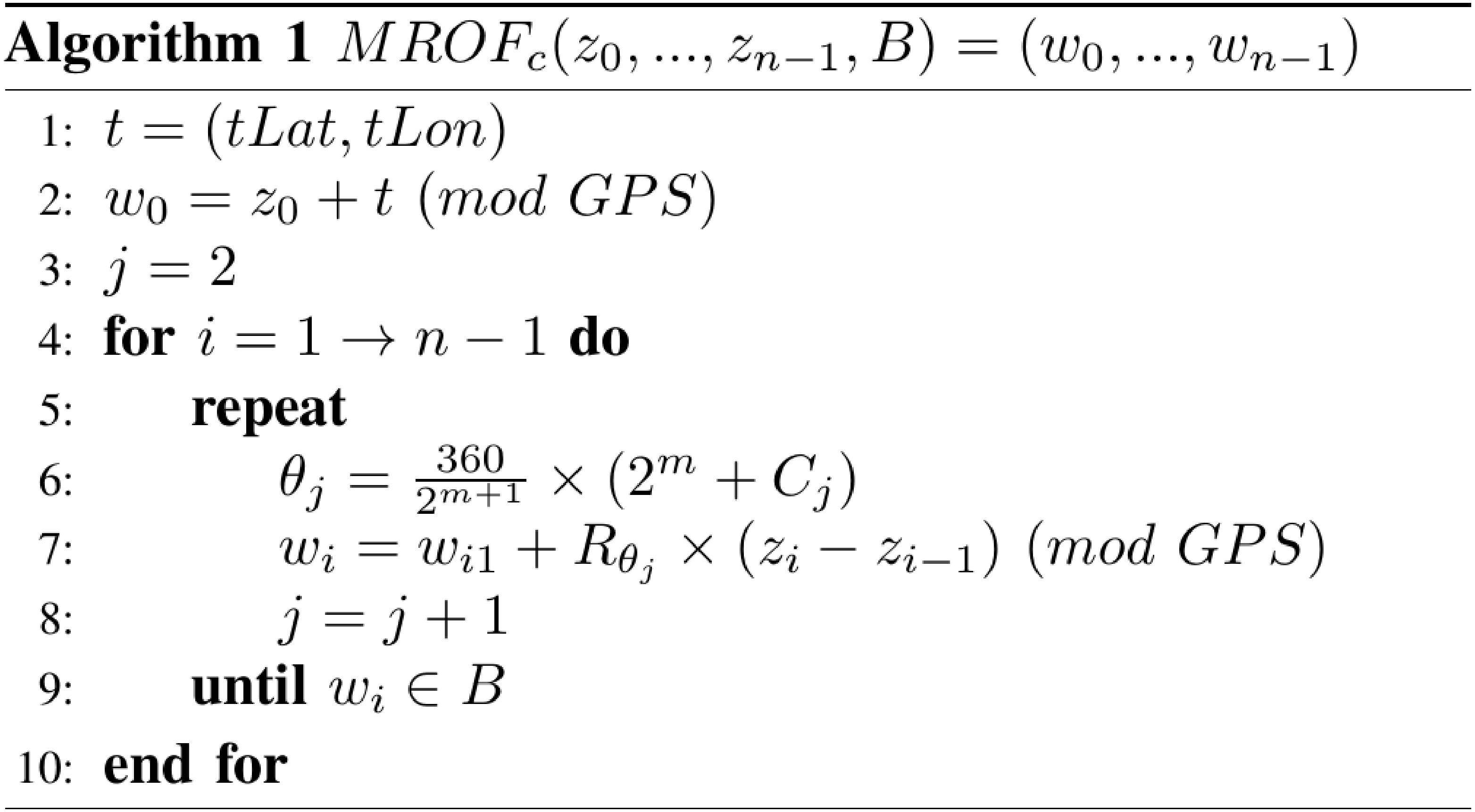
•Main aim of any LBSs (server) is to give its customers good service and in turn get monetary benefits from it.

•There can be LBSs which check authenticity of location information prior to processing user's request. Such LBSs server will stop giving its services, if server is sure that user is doing something from its side (like obfuscating actual location of user) for its privacy. MROF can also handle such situation.

•GPS has inaccuracies of around 5-10 meters in many cases.

•Consider a region having very high density of roads. If sometimes user goes out of road(on obfuscated path), still server cannot be sure of the fact that user is changing its actual coordinates and sending obfuscated coordinates.

# Modifed Random Obfuscation Function



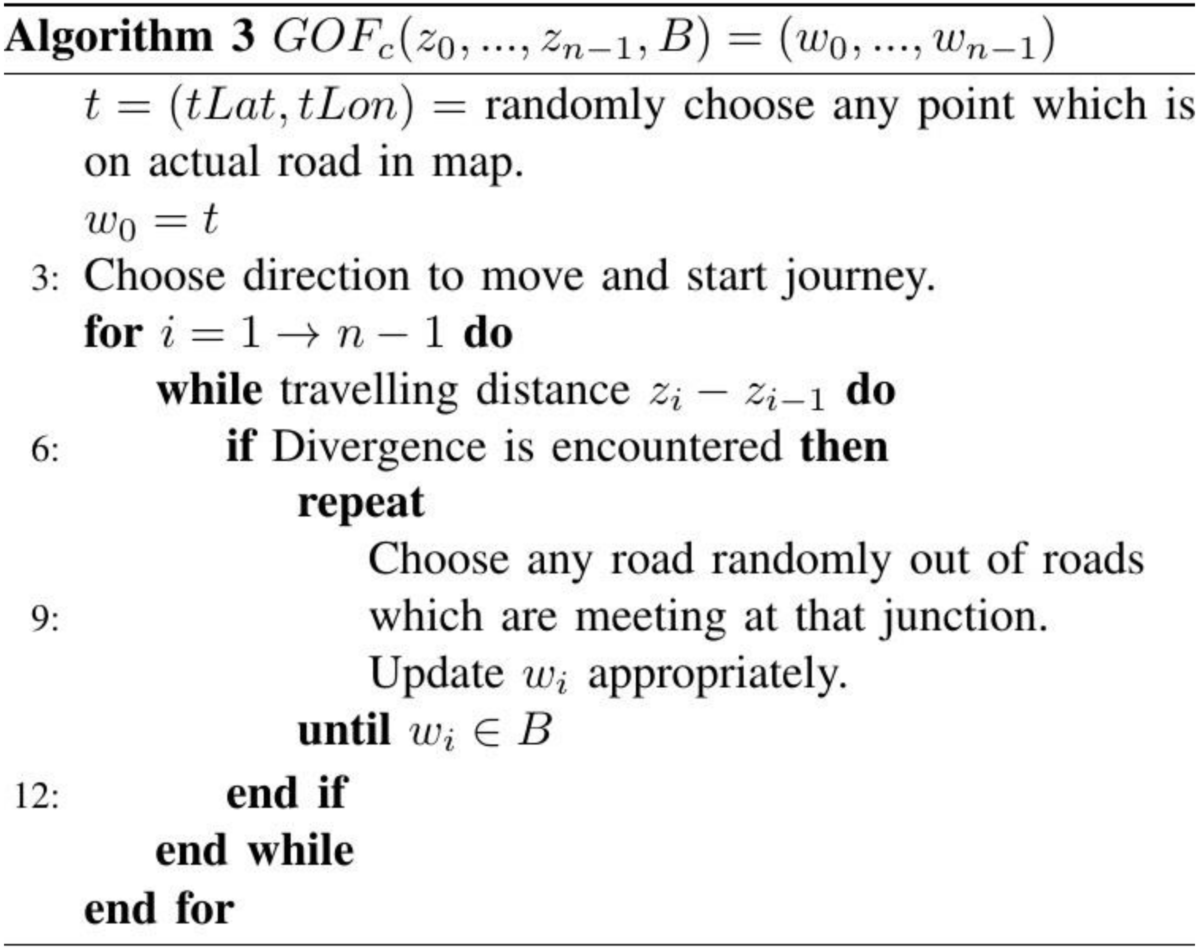
**Algorithm 1** $MROF_c(z_0, ..., z_{n-1}, B) = (w_0, ..., w_{n-1})$

```
$t = (tLat, tLon)$
$w_0 = z_0 + t \ (mod\ GPS)$
$j = 2$
for $i = 1 \rightarrow n - 1$ do
    repeat
        $\theta_j = \frac{360}{2^{m+1}} \times (2^m + C_j)$
        $w_i = w_{i1} + R_{\theta_j} \times (z_i - z_{i-1}) \ (mod\ GPS)$
        $j = j + 1$
    until $w_i \in B$
end for
```

# Proposed Approach for Grid Obfuscation Function

- Problem with Linear Obfuscation Function: It moves a point in obfuscated path on same straight line again and again.
- In Grid Obfuscation Function, off-line map is used for deciding the feasible direction of travel.
- Obfuscated path will start on any point which is on the road. It will move along the road and as soon as it reaches junction, it will randomly choose any road which is meeting at the junction.
- Algorithm make sure that obfuscated path should not go beyond certain region. If obfuscated path is not restricted, then processing very large spatial data (map) will be required.
- Algorithm can ensure security and avoid processing huge amount of map data.
- By increasing size of bounded region, probability of detecting the fact that user is obfuscating its original coordinates by server can be reduced.

# Grid Obfuscation Function



**Algorithm 3** $GOF_c(z_0, ..., z_{n-1}, B) = (w_0, ..., w_{n-1})$

```
$t = (tLat, tLon)$ = randomly choose any point which is
on actual road in map.
$w_0 = t$
Choose direction to move and start journey.
for $i = 1 \rightarrow n - 1$ do
    while travelling distance $z_i - z_{i-1}$ do
        if Divergence is encountered then
            repeat
                Choose any road randomly out of roads
                which are meeting at that junction.
                Update $w_i$ appropriately.
            until $w_i \in B$
        end if
    end while
end for
```

# Results

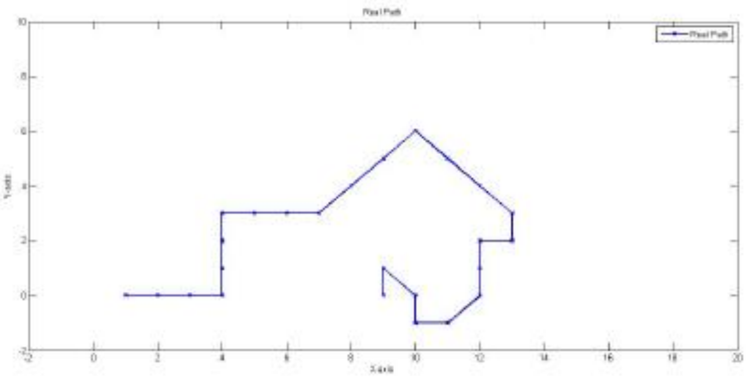


Figure: Real path for EROF

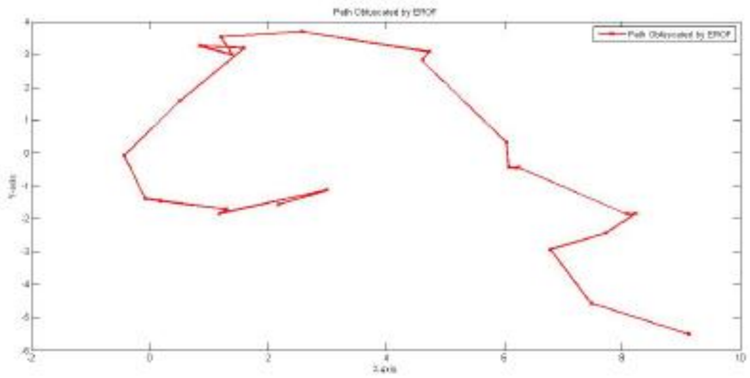


Figure: EROF obfuscated path

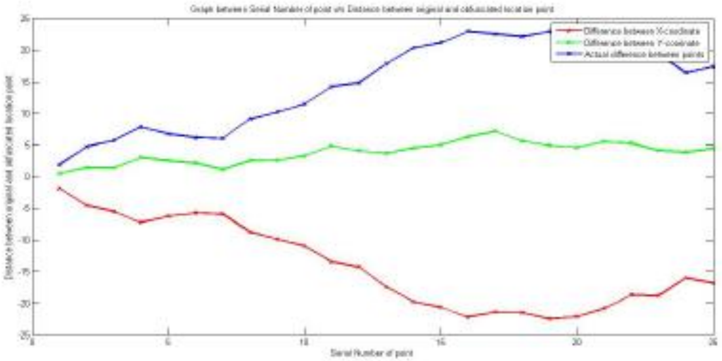


Figure: Difference between Real Path and Obfuscated Path in EROF.

# Results

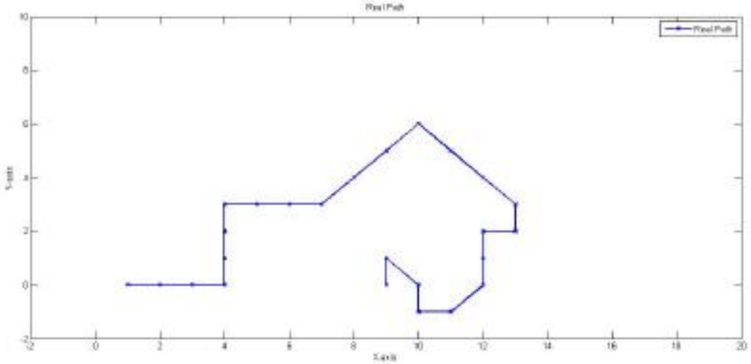


Figure: Real path for MROF

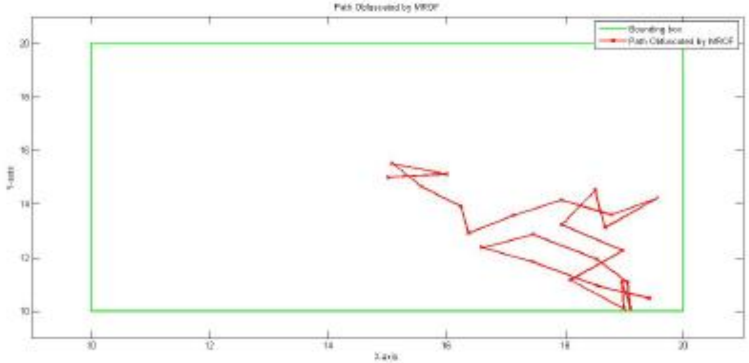


Figure: MROF obfuscated path

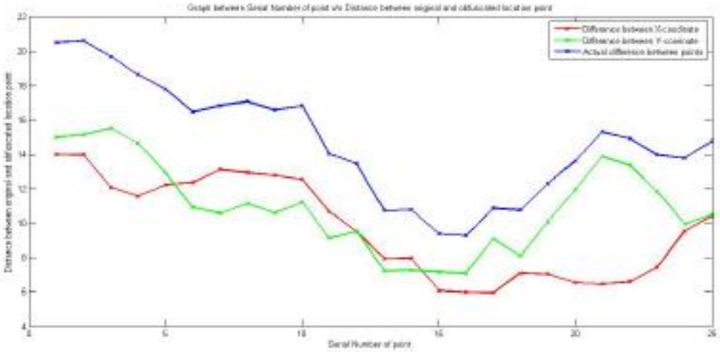


Figure: Difference between Real Path and Obfuscated Path in MROF.

# Results

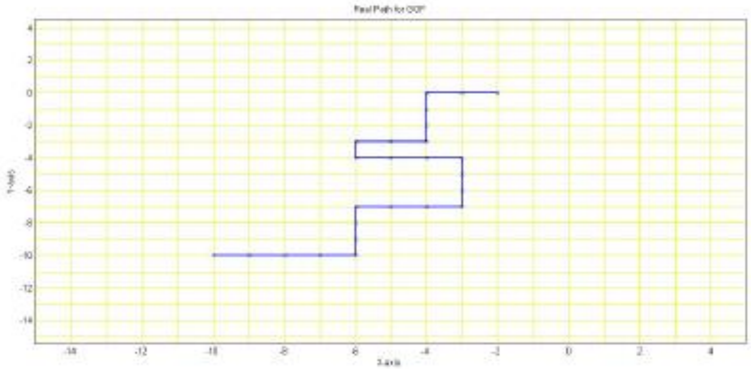


Figure: Real path for GOF

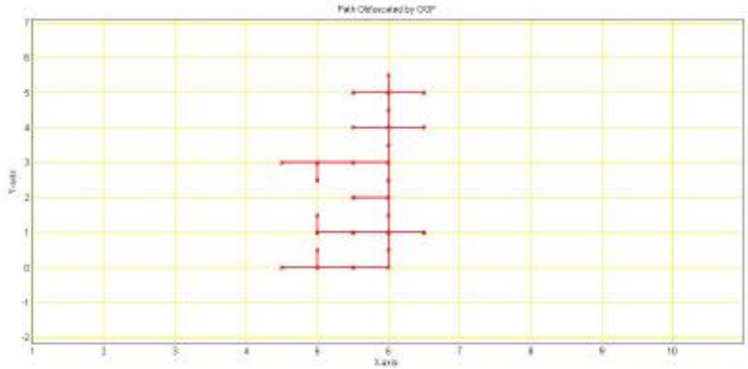


Figure: GOF obfuscated path

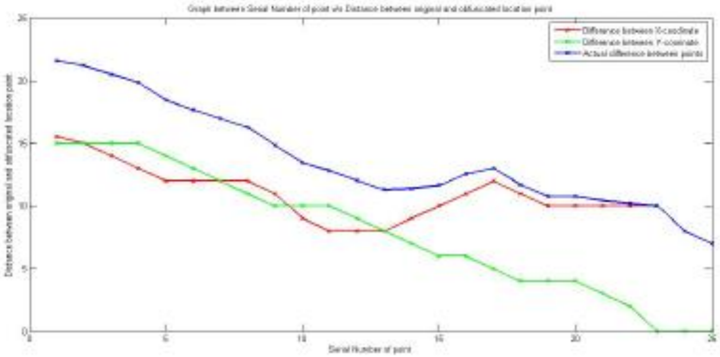


Figure: Difference between Real Path and Obfuscated Path in GOF.

# Results

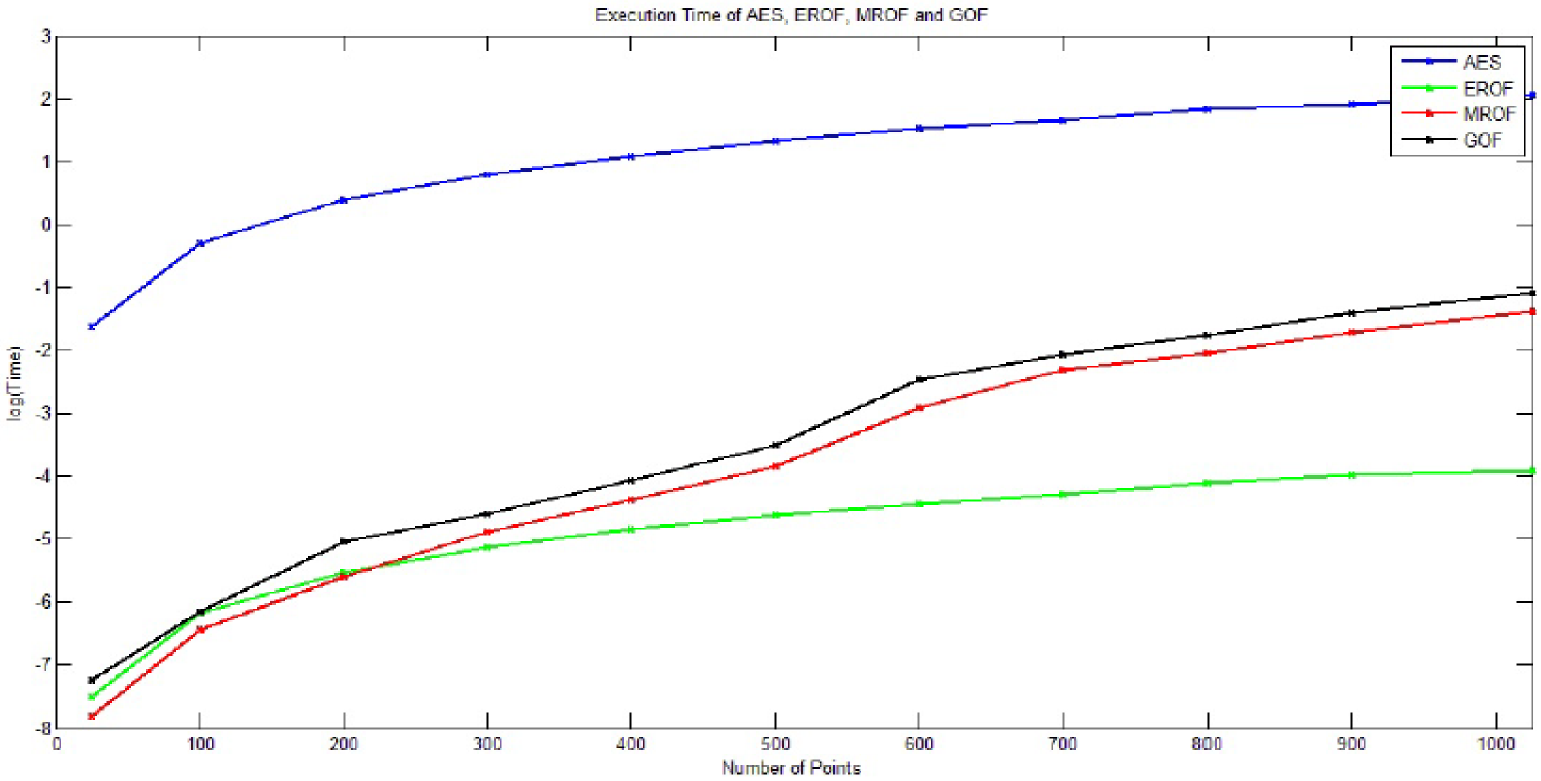


Figure: Execution Time of AES, EROF, MROF and GOF

# Conclusion and Future Work

- This solution provide user privacy without sacrificing service accuracy of LBS and comparison of obfuscation functions.
- Suitable for those services that need to evaluate distance travelled by user.
- Experimental evaluation shows that original and obfuscated path using our approach are quite different.
- Obfuscation techniques are highly application dependent and hence choice of appropriate obfuscation technique is closely related to the application requirement.
- Presently, privacy protection in navigation application is being explored and suitability of obfuscation techniques for such applications is being evaluated.

# References

- K. P. Puttaswamy and B. Y. Zhao (2010)
  Preserving privacy in location-based mobile social applications.
  *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications. ACM, 2010*

- C.-Y. Chow, M. F. Mokbel, and W. G. Aref (2009)
  Casper*: Query processing for location services without compromising privacy
  *ACM Transactions on Database Systems (TODS),vol. 34, no. 4, p. 24, 2009.*

- R. Di Pietro, R. Mandati, and N. V. Verde (2013)
  Track me if you can: Transparent obfuscation for location based services
  *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a. IEEE, 2013, pp. 19.*

- M. Gruteser and D. Grunwald(2003)
  Anonymous usage of location-based services through spatial and temporal cloaking
  *Proceedings of the 1st international conference on Mobile systems, applications and services. ACM, 2003, pp. 3142.*

# References

- Runtastic. [Online].
  Available: https://www.runtastic.com/

- Endomondo. [Online].
  Available: https://www.endomondo.com/

- Runkeeper. [Online].
  Available: http://runkeeper.com/

- Sportstracker. [Online].
  Available: http://www.sports-tracker.com/

- K. P. Puttaswamy, S. Wang, T. Steinbauer, D. Agrawal, A. El Abbadi, C. Kruegel, and B. Y. Zhao (2014)
  Preserving location privacy in geosocial applications.
  *Mobile Computing, IEEE Transactions on, vol. 13, no. 1, pp. 159173, 2014.*

# Results

- H. Kido, Y. Yanagisawa, and T. Satoh, (2005)
  Protection of location privacy using dummies for location-based services.
  *Data Engineering Workshops, 2005. 21st International Conference on. IEEE, 2005, pp. 1248 1248.*

- P. M. Wightman, M. A. Jimeno, D. Jabba, and M. Labrador (2012)
  Matlock: A location obfuscation technique for accuracy-restricted applications.
  *Wireless Communications and Networking Conference (WCNC), 2012 IEEE. IEEE, 2012, pp. 18291834.*

- D. Riboni, L. Pareschi, C. Bettini, and S. Jajodia (2009)
  Preserving anonymity of recurrent location-based queries.
  *16th International Symposium on. IEEE, 2009, pp. 6269.*

# Results

Thank you