

DYNAMIC DIFFERENTIAL LOCATION PRIVACY WITH PERSONALIZED ERROR BOUNDS

LEI YU, LING LIU AND CALTON PU COLLEGE OF COMPUTING GEORGIA INSTITUTE OF TECHNOLOGY CREATING THE NEXT

Location based services and Privacy issues





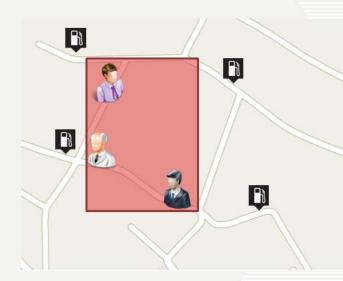


CREATING THE NEXT®

Location Privacy Protection



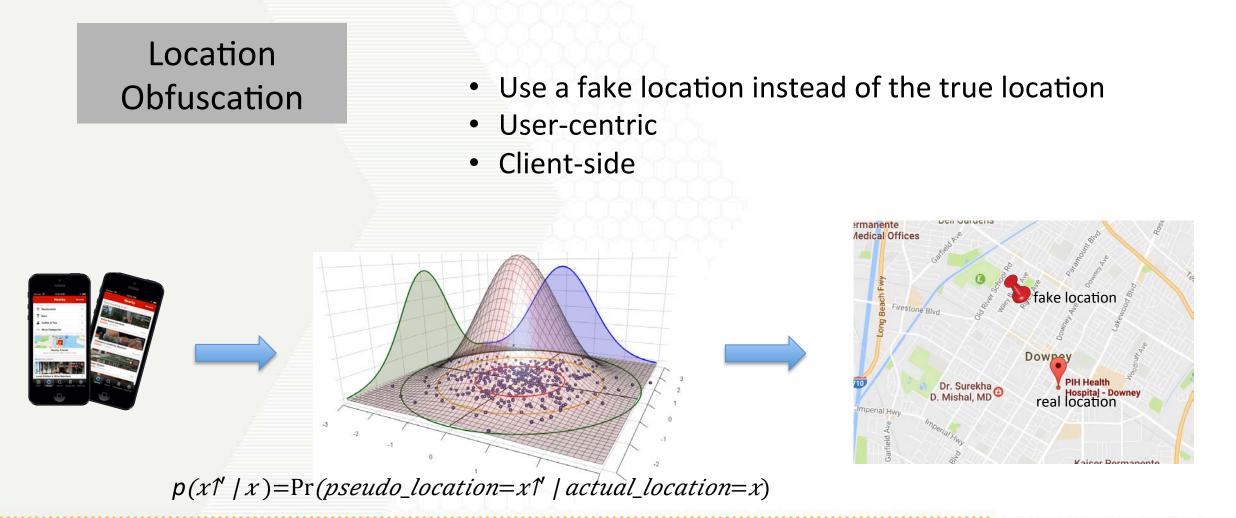
Anonymization



- K-anonymity
- trusted third-party anonymization server

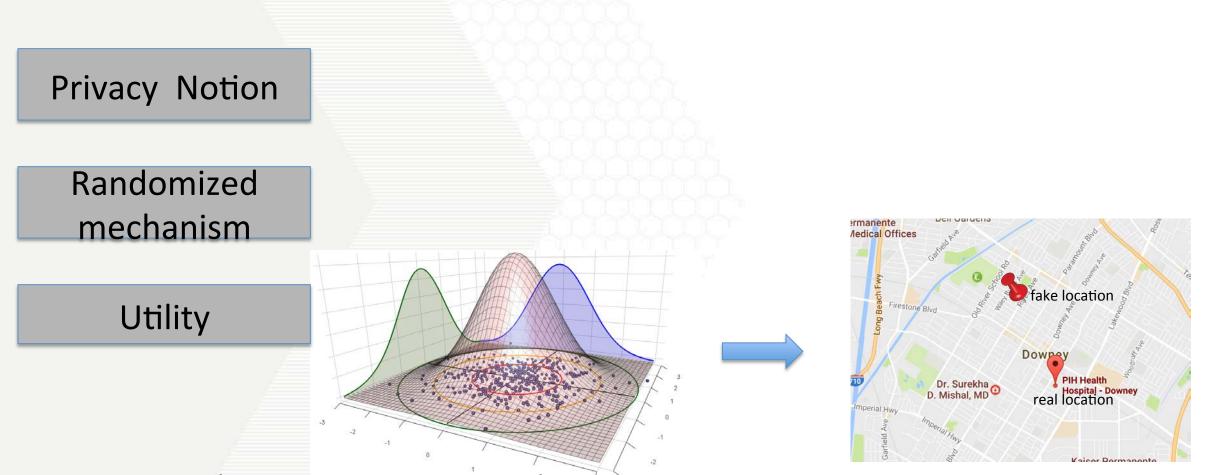
Location Privacy Protection





Location Obfuscation





 $p(x^{\uparrow} | x) = \Pr(pseudo_location = x^{\uparrow} | actual_location = x)$

// CREATING THE NEXT®

Existing Techniques

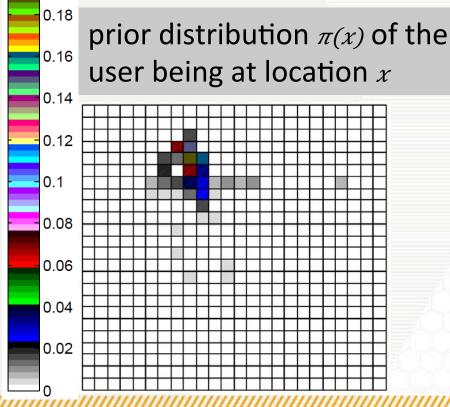
- Privacy Notions:
 - Expected inference error
 - Geo-indistinguishability



Expected inference error



 The expected distance between the user's real location and the location guessed by the adversary.



Given observation x', the probability of actual location being x $\Pr xx \uparrow' = \pi(x)f(x\uparrow'|x)/\sum x \in \chi \uparrow m \pi(x)f(x\uparrow'|x)$

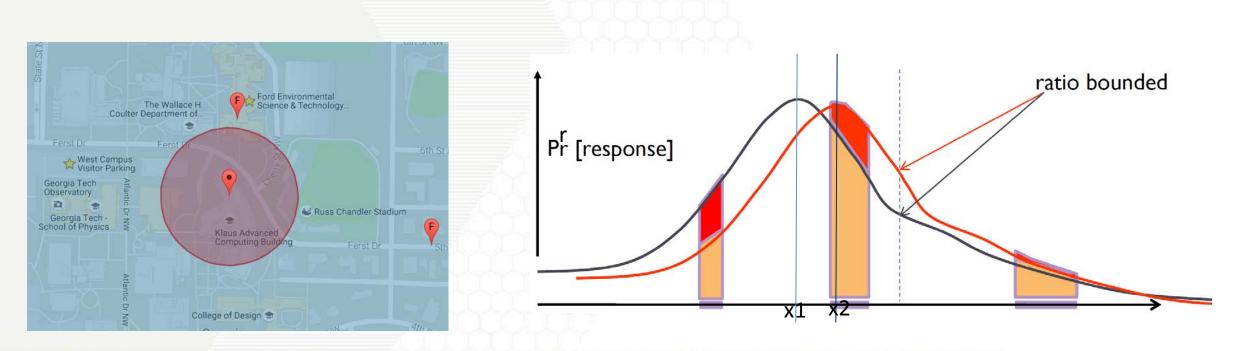
CREATING THE NEXT[®]

Geo-indistinguishability



For any two points x, y in the protection circular area of radius r centered at the actual location, by $\epsilon \downarrow g = \epsilon/2r$

 $f(x\uparrow'|x)/f(x\uparrow'|y) \le e\uparrow\epsilon$



Existing Techniques



• Privacy Notions:

Expected inference error	Geo-indistinguishability							
Bayesian inference	differential privacy							
Rely on a specific prior distribution of user's real location	only depends on the mechanism and does not depend on any prior							
Not robust against any other prior distribution	Adding noise regardless of any prior can be inefficient and insufficient for privacy protection							

Our work



- Limitation of Geo-indistinguishability
- Two-phase location obfuscation framework
 - Adaptive noise level for different locations with guaranteeing a minimum level of inference error
 - Customizability
 - Instantly specify his privacy preference for his current location
 - Existing mechanisms are computed statically once for all, and cannot efficiently support customizability

Experimental Illustration



- Existing mechanisms
 - Optimal Bayesian mechanism [R. Shokri et al., 2012]
 - Optimal geo-indistinguishable mechanism [N. E. Bordenabe et al., 2014]

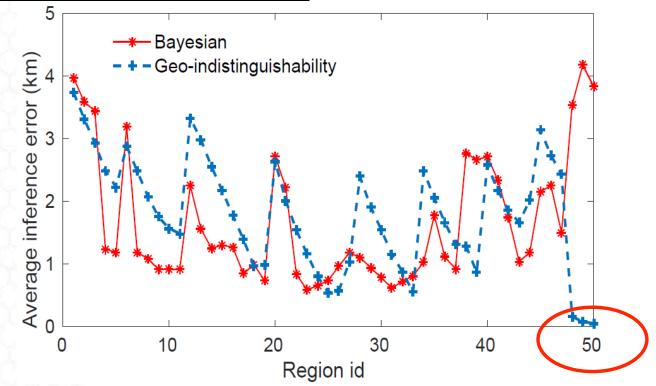
Experimental Illustration



						~ 7												
					10	27		┝─										_
				11	19	20	22	<u> </u>					_					_
				10	10	20	22											_
	 		5	9	17	23	31	37	43	ΔΔ	47				\mathbf{f}	49		_
_			4	9 8	16	22	30	36	42	11	47				\prec	10		
			3	1	115	121	29	35	41		46							
			2	6	14		28	34	40		45							
			1		13													
					12			⊢	00									
						00			<u>39</u> 38									_
	 					20		<u> </u>	აი									_
								-										_
_	 							-		-								-
											T	48	\uparrow			1	50	

50 regions with prior probability >0

Dataset: GeoLife GPS Trajectories dataset Formatted as in [N. E. Bordenabe et al., 2014]



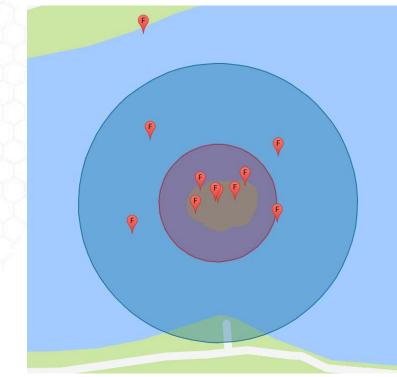
Two mechanisms that achieve the same location privacy in terms of overall expected inference error weighted by prior probability

Experimental Illustration



Geo-indistinguishability





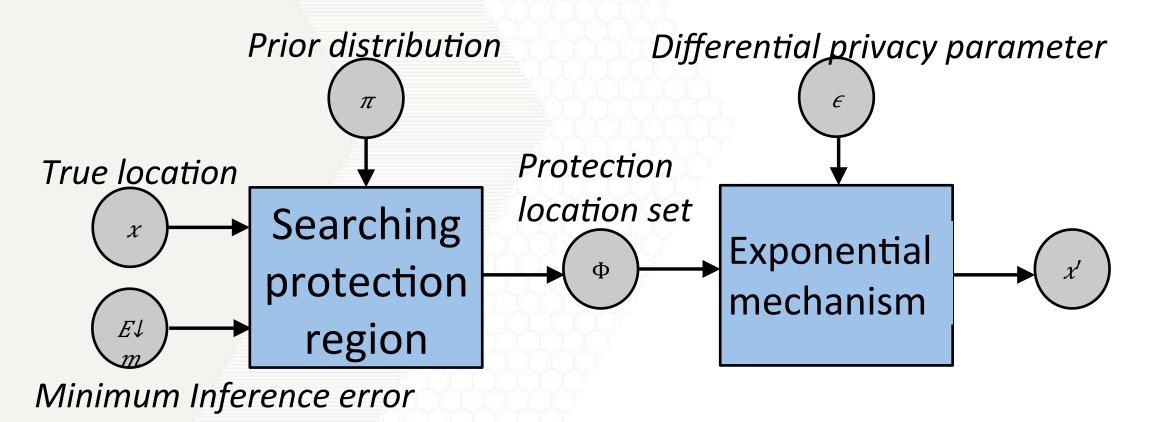
Planar Laplacian Mechanism, Pr(pseudo-location in blue circle) $\geq 95\%$

Not Adaptable: Uniform noise level either insufficient location protection at some skewed locations in terms of prior information or excessive noise for protection at other locations

Two-phase framework



Combine expected inference error and Geo-indistinguishability



Relationship between two privacy notions



 $f(x\uparrow |x)/f(x\uparrow |y) \leq e\uparrow\epsilon$

- Geo-indistinguishability
 - Any two locations x, y in the protection region Φ ,

Lower bound of conditional expected inference error

 $\min_{\tau x} \sum x \in \chi \uparrow \text{ Pr} x x \uparrow d(x,x) \ge e \uparrow -\epsilon \min_{\tau} x \sum x \in \Phi \uparrow \text{ mn}(x) / \sum y \in \Phi \uparrow \text{ mn}(y) d(x,x)$

Protection Location Set

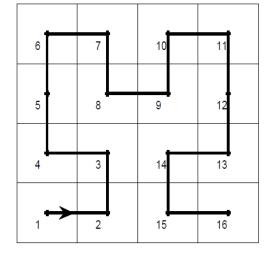


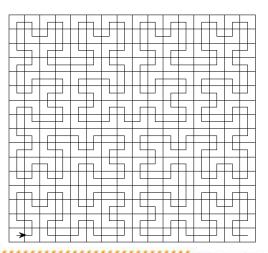
• Theorem: For a location obfuscation mechanism that achieves ϵ -differential privacy on protection location set Φ , if $E(\Phi) \ge e \uparrow \epsilon$ $E \downarrow m$, the optimal inference attack using any observed pseudolocation x', the expected inference error $\ge E \downarrow m$.

 $E(\phi) = \min_{\forall x \in \Phi} f(x) / \sum_{y \in \Phi} f(x) / d(x,x)$

Phase I: Search Protection Region

- $E(\Phi) \ge e \uparrow \epsilon E \downarrow m$
- Hilbert-curve based searching
 - Larger diameter of protection location set indicates higher noise level
 - Improvement with multiple rotated Hilbert curves





CREATING THE NEXT®

Georgia Tech

Phase II: Exponential mechanism

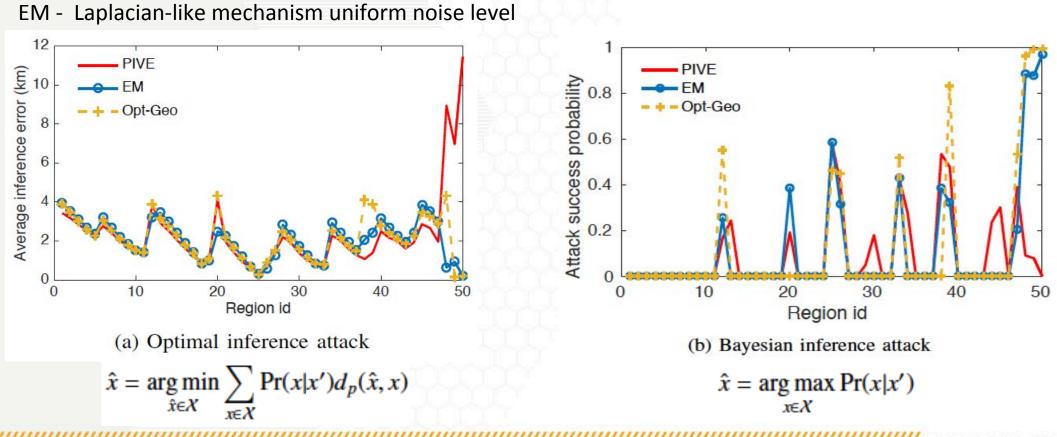


Given the user's location x and location protection set Φ, the exponential mechanism selects and output a pseudo-location x' with probability proportional to exp(-εd(x,x1')/2D), where D is the diameter of Φ.

Evaluation



Comparison with existing mechanisms on location privacy

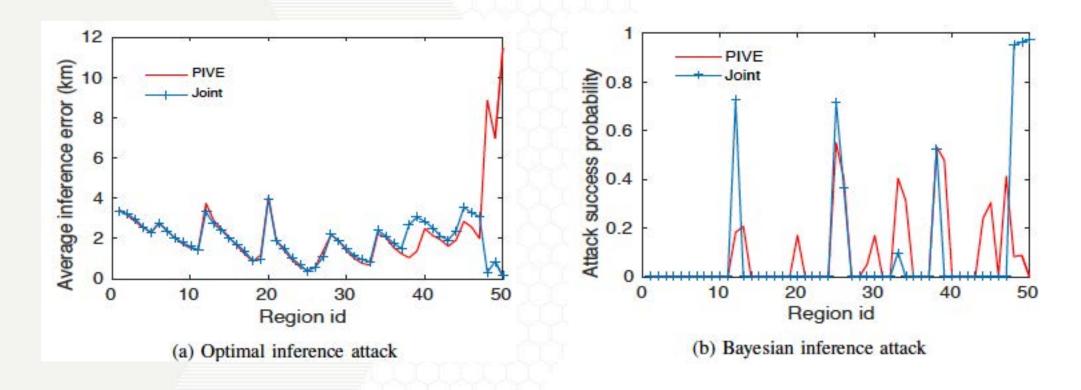


/////// CREATING THE NEXT

Evaluation

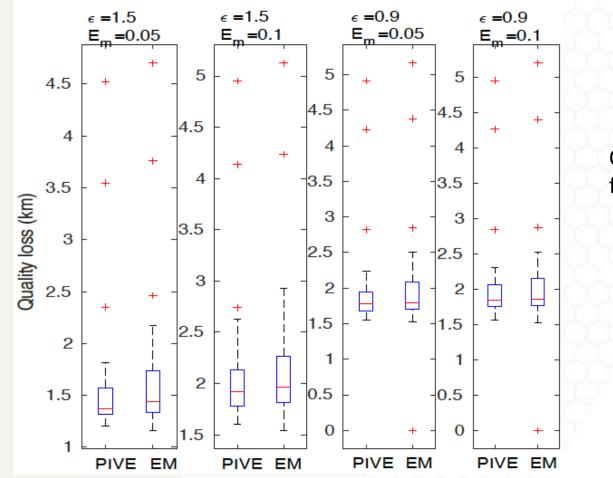


Comparison with joint mechanism on location privacy



Utility





Quality loss: the average distance between the fake location and the real location.

PIVE



- Geo-indistinguishability + prior information
- Adaptively adjust noise level of different privacy according to prior distribution
- Customizability



Thank you! Q&A

CREATING THE NEXT®

Expected inference error



Conditional expected inference error $\sum x, x \in \chi^{\uparrow} = \Pr x x^{\uparrow} h(x x^{\uparrow}) d(x, x)$	Unconditional expected inference error $\sum x, x^{\uparrow}, x \in \chi^{\uparrow} \equiv \pi(x) f x^{\uparrow} x h(x x^{\uparrow}) d(x, x)$						
the distance between the estimation and the actual location							
h(x x t') - Probability of guessing x as the user's actual location, given that x' is observed	Quality loss $\sum x, x \uparrow' \in \chi \uparrow \implies \pi(x) f x \uparrow' x d(x, x')$						
<i>Optimal inference attack</i> : $x = \operatorname{argmin}_{\tau} x \in \chi \sum x \in \chi \uparrow $ Pr $xx \uparrow$	d(x,x)						
Bayesian inference attack: $x = \operatorname{argmax}_{\tau} x \in \chi \operatorname{Pr}(x x\mathcal{T})$							