Location Based Services

Shubhranshu
The University of Auckland
23, Symonds Street, Auckland
snot422@auckland.ac.nz

ABSTRACT

General population first heard about location based services in late 90's, partly as a response to E911 and partly due to recent technical developments. Since then the flood of ever cheaper, more powerful hardware had pushed these services in everyday devices. With over hundred million location aware devices in operation this phenomenan had shaped an entirely differetnt kind of market model. These devices can help target specific users with very specific details about their daily activities. Such information can be easily misused to track real world identities. As a well known fact, location information present a significant risk to user privacy, still there is no common concensus on privacy control. Most users underestimate value of their location information and security risks posed by such information. Even some of the most trusted obfuscation techniques can not completely mitigate security risks and hence individual perception plays an importand role in privacy models. The fine balance between obfuscation and quality of location data is hard to achieve. We explore various security risks, technical challanges, user perception and some real world examples to discuss implications of location data and it's sustained growth.

Author Keywords

Privacy; Location; Obfuscation; Anonymization

ACM Classification Keywords

K4.2 COMPUTERS AND SOCIETY: Social Issues

General Terms

Human Factors; Security

INTRODUCTION

With introduction of mobile phones in 1984 personal communication became mobile and it was envisioned that a broader range of fixed line like cervices can be made available on this new found mobile infrastructure. Military technologies have always emphesised triangulation of signal source from an early age and as mobile systems came to prominance in early 90's it became possible to apply these location detection technologies in conusmer devices. When the Federal Communications Commission (FCC) mandated E911 in 1996 it was percieved that the technology can easily provide the required degree of accuracy easily. This era marked the begining of first phase of location based services(LBS). Initially these services suffered mostly due to lack of required infrastructure to perform optimally. The accuracy requirement for E911 were 50-100 meters for 67% of all calls and within 150-300 meters for 95% of all calls. This kind of location accuracy using mobile cell triangulation can only be provided in dense urban setups. Apart from that it was not until early 2000's that accoumpnying technologies like 3g and GPS were widely adopted.[6]

THE FIRST PHASE

This first phase of LBS has already proved it's usefulness and it was realised that location services must be delivered considering evolving technologies like 3G. First phase of LBS is identifed by passive LBS where users have to initiate a location aware service which can be served with a context aware response; one such example is Points of Interest(POI) enquiry. POI enquiry allowed users to send a SMS enquiring about post office, banks, atm in nearby locations. However two major technical difficulties prevented it from becoming a phenomenon. Firstly, cell triangulation is a crude method to approximate user location as accuracy degrades rapidly in urban settings due to presence of numerous structures, extensive sampling is required to compensate for man made structures while calculating. Secondly, due to uneven and relatively sparse distribution of transmission towers made it difficult to achieve a respentable error margin. Since these passive sevices were easier to implement (no handset/hardware change for consumer) it was economically viable to implement while maintaining a userbase. Surprisingly, even these early attempts of location detection raised privacy concerns among few users. This technology still being used with niche users allowed these privacy concerns to be

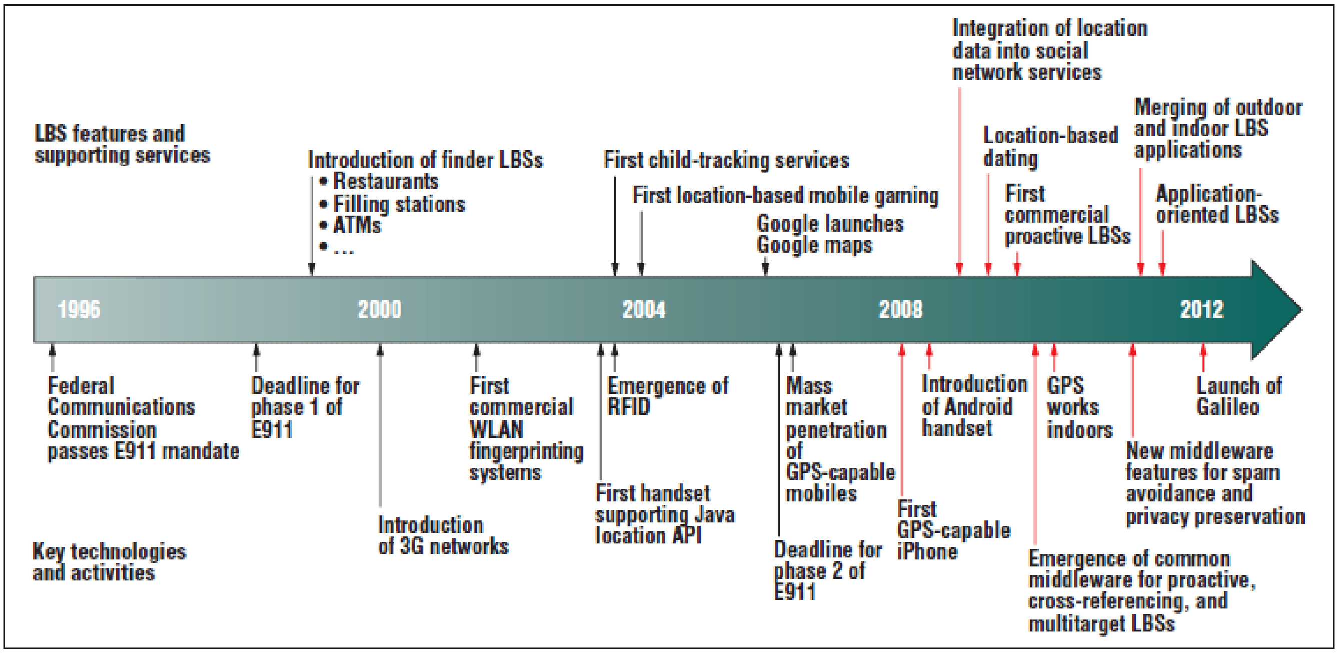

LBS features and supporting services
Introduction of finder LBSs
• Restaurants
• Filling stations
• ATMs
• ...
First child-tracking services
First location-based mobile gaming
Google launches Google maps
Integration of location data into social network services
Location-based dating
First commercial proactive LBSs
Merging of outdoor and indoor LBS applications
Application-oriented LBSs
1996
2000
2004
2008
2012
Federal Communications Commission passes E911 mandate
Deadline for phase 1 of E911
Introduction of 3G networks
First commercial WLAN fingerprinting systems
Emergence of RFID
First handset supporting Java location API
Mass market penetration of GPS-capable mobiles
Deadline for phase 2 of E911
Introduction of Android handset
First GPS-capable iPhone
GPS works indoors
Emergence of common middleware for proactive, cross-referencing, and multitarget LBSs
New middleware features for spam avoidance and privacy preservation
Launch of Galileo
Key technologies and activities

Figure 1: Timeline of LBS evolution [8]

ignored until technology becomes cheaper and more available.[6][8]

THE SECOND PHASE

Early age LBS sustained a steady growth till early 2000's due to relatively higher hardware prices. Higher cost of acquisition for a mobile broadband ready device, for which services were still not widely available proved a roadblock for consumers. The situation migigated in coming years as hardware become more powerful and cheaper than ever. Second phase of LBS is marked by introduction of numerous technologies in rapid succession. Withing past 10 years consumers have witnessed the rise of 3G, GPS, A-GPS and ever increasing device computing power. As GPS system become publically available in early 2000's it paved way for more accurate location detection using GPS staellite. Unlike cell based triangulation gps can provide services without any installed infrastructure. GPS satellites take a few minutes to synchronize and "lock" on the device before they can detect user location. In urban settings a combination of cell and GPS based triangulation can reduce this lock time significantly. More recently Assisted GPS (aGPS) use similar technology to reduce lock times and provide faster location detection. Second phase introduces numerous technologies to market. Mobile devices are more powerful, more processing, ability to transmit and recieve data, integrated GPS and so on. A modern day mobile device can track user position and provide LBS results in a matter of seconds. Figure 1 displays the growth of LBS sice it's inception in 1996.

LBS PROVIDERS! WHO ARE THEY ?

Initial implementation of LBS were carried out in a time when mobiles were a novelty items and a userbase of niche customer who percieved their carrier (LBS provider) as a trusted entity. This soon changed with introduction of GPS and mobile data; as users can choose to send/recieve location information without carrier concent. One such example is Google Maps, initially developed on J2ME platform in 2004 it had the ability to transmit location data on device to third party applications. Like internet, mobile data and subsequently location data becomes independent entities; without a central control. Today most of the applications utilize location information to provide theor services, let it be maps, photo sharing or even communication. Rise of internet and mobile data technologies like 3g have accelerated the growth of location aware services. Today an LBS provider can be any one, it may be a news site or application providing bus schedules. Since there is no central control and location data is easily available, privacy concerns are very obvious.[11]

In an experiment by Iris A. Junglas and Richard T. Watson, 58 subjects in their early 20's were asked to explore the surrounding areas and track eack other using wifi and GPS enables/disabled devices. The satisfaction percentage for the study was very high(88%). However number of participants were amazed and thrilled to know how easily they can track eack other usng ordinary devices. For some being trackable was a real concern.[6] In another study a team of microsoft researchers asked 32 participants from 12 households to carry a GPS enabled device to record their location data over a period of two months. Before the study it was made clear that the location data will be made public for other users. After data collection users were presented with location trace of their devices, it alarmed them as their real world identities can be easily esteblished using simple calculations of their home and work locations. Subjects were given a choice to obfuscate the data so to identity detection can be avoided. Interestingly, given a faily sufficient amount of location history most of the obfuscation techniques(with one exception) will fail and user identity will be reaveled. So far only k-anonymization in conjucation with other techniques is the only technique that can prevent identification of user. Researchers were surprised to know that privacy perception is very different for different groups, even within same household; a mutual concensus on level and type of obfuscation was never realized. Female users and users with owned homes shows significant alerts regarding privacy of data and applied multiple techniques to dismiss finer details.[1] Figure 2 demostrates the nature of techniques used in this experiment.

LOCATIONS AND SOCIAL MEDIA

Leighton Evans takes a philosophical view on emergence of technology like LBS. He quotes Heidegger (1977, p. 11), considering emerging technologies as *poesis*(bringing foth), in a manner that technology evolve due to it’s interaction with human nature. In process of harnessing the true potential of these services humans themselves have become resources. The equivelence is more striking if we consider the fact that, for a better location aware system to exist there must be more human generated location data. Social media has been around for a decade now, hi5, myspace, orkut and other early social spaces have become largely defunct. The type of social interaction provided by these systems never considered a moving user as a prime asset.[10] With web2.0 new technologies have emerged and have esteblished themselves in most of the daily objects. A fridge ordering a pizza was a science fiction a few years back, but not anymore. Phenomenal growth of platforms like Facebook, Four Square, Twitter and others can be attributed to two main factors. One, constant social pressure to evolve and comminicate has pushed even the shy and reserved types on face of public networks. Second, emergence of new technologies enabled user to stay connected even while on move. Check-in is a term well known to most youths today. Four Square, the original check-in website allowed users to check-in to place to show their presence. Four Square was designed and engineered with mobile devices in focus, GPS, aGPS and data services all are available on relatively inexpensive mobile platforms today. This trend was soon replicated by Facebook and Twitter.[11]

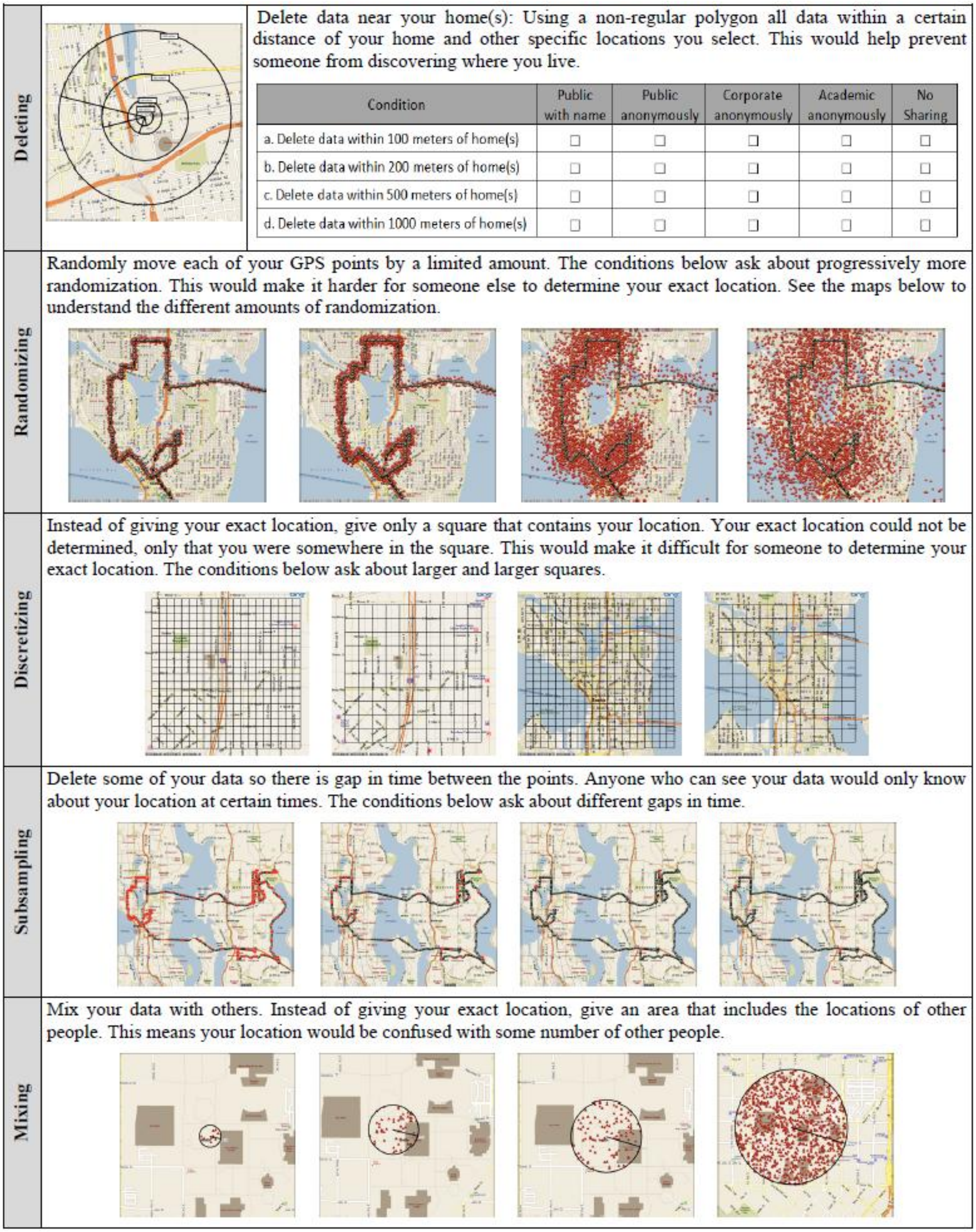
It’s not just social meyham that has evolved in location aware world, some real world commecrial applications are also making their way into our daily lives. Location aware advertisement is one such applications. Today wifi systems and even the fixed line data connections have been upgraded to provide location data with user concent. Some browsers were able to foresee this shift(in turn they motivated location aware fixed networks) and implemented location sharing features in desktops and laptops. Today, economics of advertisement is largely governed by localized advertisements. Google, once a goliath is trying new startegies to tackle content and user rich facebook. Yuan et.al. elaborates a very efficient and productive location taxi system. By constantly monitoring the passanger behaviour it was possible to achieve less than 10% of average idle time for the system. It’s not the appllcaitions that take away our data, today we are distributing it freely; at will.[12]

PRIVACY

Consider a mobile device with GPS and a suitable mobile data connection, how this data will leave the device and reach external users ? The answer is very simple; with express permission of the user ! 40% of users using facebook on mobile update their status regularly, most of them provide a check-in which is enabled by default.[10] Similarly Four Square encourage it’s users to tag places they are currently in(sole purpose of the application), this location data is meant to travel to all your connections and to public if configured so.

Surprisingly mobile applications are not the only things that can raise a serious location threat. Flicker for example is an image sharing weebsites, lack of personal data makes it more confortable for people to share their photographs publically. However, with more and more location aware phones and digital camera lots of geotagged photographs can be found on such websites. As mentioned before, long term location data can help determine real identities faily easily. Similary, given sufficient data real identities of users can be esteblished much faster and easily than anticipated. A dated, geotagged digital photo album can easteblish user patters in a certain vicinity with relative ease. Such data is deliberately uploaded by users. Like early age LBS, initial social media applications were passive and required user interactions to initiate a communication. Today most of the application act proactively; google latitude can locate and transmit realtime user location on network. Instead of collecting data, today applications rely on user submitted location data which miost of us do willingly. 40% of users are confortable disclosing their location data and they are already doing so using convenient applications.

As mentioned, there is no mutual concensus on ownership of location data, for the most part it is user’s responsibility to keep check on such data. Anonymization techniques results in s loss of location quality and hence a loss in LBS. While most of the techniques can not twart risks from long term location trace, only k-anonymization can provide sufficiently usable data for an LBS to work with acceptable quality. K-anonymization can provide sufficiently usable data for a LBS rpovider since it aggregates k-user



Deleting

Delete data near your home(s): Using a non-regular polygon all data within a certain distance of your home and other specific locations you select. This would help prevent someone from discovering where you live.

Condition	Public with name	Public anonymously	Corporate anonymously	Academic anonymously	No Sharing
a. Delete data within 100 meters of home(s)	□	□	□	□	□
b. Delete data within 200 meters of home(s)	□	□	□	□	□
c. Delete data within 500 meters of home(s)	□	□	□	□	□
d. Delete data within 1000 meters of home(s)	□	□	□	□	□

Randomizing

Randomly move each of your GPS points by a limited amount. The conditions below ask about progressively more randomization. This would make it harder for someone else to determine your exact location. See the maps below to understand the different amounts of randomization.

Discretizing

Instead of giving your exact location, give only a square that contains your location. Your exact location could not be determined, only that you were somewhere in the square. This would make it difficult for someone to determine your exact location. The conditions below ask about larger and larger squares.

Subsampling

Delete some of your data so there is gap in time between the points. Anyone who can see your data would only know about your location at certain times. The conditions below ask about different gaps in time.

Mixing

Mix your data with others. Instead of giving your exact location, give an area that includes the locations of other people. This means your location would be confused with some number of other people.

Figure 2: An illustration of various obfuscation techniques[1]

bahaviour on input, however from a user perspective it will also degrade LBS qulaity.

Another type of privacy risk is inferred risk. Even if user have never used a location service he can still be targeted and identified using the social network they are using. User identification can be done indirectly using feeds, check-ins, tags, tweets made by user's friends in the social application. Websites like Pleaserobme an Icanstalkyou shows the real potential of oversharing. For an inferred privacy exploit to work a social graph of target can be made for a specific Online Social Network(OSN), this is a relatively easy and feasible task to achieve. Later specific filters like node closeness and trustworthiness can be applied to discover potential targets. Once a paticular target had been acquired in a specified vicinity of the target user, various hypothesis and analysis can provably produce a very good approximation of target locations.[4][13]

LBS SECURITY

Security measures in LBS are directly proportional to difficulty of identifying the real user. This requirement conflicts with the LBS quality directly. Various methods of providing obfuscated data to prevent identification includes, noise, interval sampling, randomization and mixing. Threats of reidentification is increased many folds if user can be identified using a home-work pair. Longitudinal Employer-Household Dynamics (LEHD) program collected household and work data for nearly 95% of American working population, any LBS provider with a subset of such information can identify location trace to an individual. In real world, if home and work locations can be tracked to a specified location grid, we can prepare a candidate shortlist just by referring to a white book or business directory.[2]

K-anonymization is a good approach for obfuscation, but with combination of other techniques and social skills even it can be compromised. For anonymization, device must contact an anonymization server(AS) in order to mix information with at least k-1 users. If k is very low it is ineffective, if k is too high it suffers from overcrowding in anonymization space. Alternate approaches are being explored to address security as perceived entity instead of measured entity. Super-Ego is once such framework providing security as a perceived entity. Security requirements can't be perceived by any algorithm, but algorithm can learn from user's context aware security response. For example, with training it can be taught to switch off location data near specified locations like home, office e.t.c. If super-ego is not confident about disclosing or hiding the information, it will ask for user input to determine if location should be shared for a particular location. With time, this algorithm can learn user perception and provide required degree of home/work anonymization.[5]

A different variation of precieved security measure employs a dynamic anonymization matrix to prvide a varying degree of privacy. For obfuscating between k-1 other users, obfuscation factor should be increased near location pairs like [work, home]. For a constant k factor in a k-anonymous system, it is highly likely that it might not be able to provide sufficient resolution on some point or may become too noisy. For a constant k, the number of users to be mixed are also fixed, however unlikely, but it is possible that user might be mapped to unique location pairs and hencec reidentified. Feeling based privacy protection keeps a dynamic obfuscation factor. For sensitive locations like home and office, this factor k is adjusted so that there are atleast k more individual with similar traces. While on a highway It can adjust the granularity of location to provide necessary navogation and traffic feedback.[9]

Instead of relying on an AS for informations retrieval, devices can issue private queries and retrieve information using Personal Information Retrieval(PIR). PIR technique allows user to download data from other entities wthout the server itself realizing the data being downloaded. This approach requires the device to hide it's network identity as well as location information to make PIR queries safely. This is a considerable restriction and most mobile platforms are not capable to running multiple cryptographic sessions required by PIR simultaneously. Another approach; caching, counter the identity issues more gracefully. For caching, the location request is translated to a spatial grid on the device itself, and content request for this grid is issued. Once spatial data is obtained it is cached on the device and no further requests are made for this particular grid. Data is usually downloaded on higher scales to maintain high anonymity factor. Practically it is possible for the device to run for weeks before another location request to be made for the same lcoation.[3][7]

CONCLUSION

LBS have revolutioned the digital communication and shaped a new era of loction aware services. These services and platforms are very engaging and entertaining but they pose significant privacy risks due to lack of any central autority to manage location data. In social networks and online media location sharing is mostly a user action thus requiring a high degree of user undestanding of privacy risks. While useful obfuscation can be provided at hardware level and underlying frameworks it remains largely unused due to presense of multiple LBS providers accepting location data. Anonymization concepts can only work if service providers can comply to a mutual agreement on locataion data handling and disclosure. As proved already it is possible to track user movements, even if user is not publishing their location publically. The process to standardize and mandate a common standard is still in infancy. Apart from some implementations like Cache, none of them address location issues as efficiently. With raising concern about privacy and reidentifications, usefulness of location based services will be severely challanged for public applications like taxi managements and trafiic planning. Tradeoff between anonymity and personalized services through LBS are currently too high, it will still take some time to adopt a common policy which can provide sufficient granularity for service alongwith user protection.

REFERENCES

1. A.J. Bernheim Brush, John Krumm and James Scott. Exploring End User Preferences for Location Obfuscation, Location-Based Services, and the Value of Location. *Proceedings of the 12th ACM international conference on Ubiquitous computing, 2010, ACM Press (2010), 95-104.* doi>10.1145/1864349.1864381
2. Philippe Golle and Kurt Partridge, On the Anonymity of Home/Work Location Pairs. *Proceedings of the 7th International Conference on Pervasive Computing, 2009, Springer-Verlag Berlin, Heidelberg (2009), 390-397* doi> 10.1007/978-3-642-01516-8_26
3. Shahriyar Amini , Janne Lindqvist , Jason I. Hong , Maladau Mou , Rahul Raheja , Jialiu Lin , Norman Sadeh , Eran Tochb, Caché: caching location-enhanced content to improve user privacy. *ACM SIGMOBILE Mobile Computing and Communications Review, v.14 n.3, 2010, ACM New York, NY, USA (2011)* doi>10.1145/1923641.1923649
4. Adam Sadilek, Henry Kautz, Jeffrey P. Bigham, Finding Your Friends and Following Them to Where You Are. *Proceedings of the fifth ACM international conference on Web search and data mining, 2012, ACM New York, NY, USA (2012), 723-732* doi>10.1145/2124295.2124380
5. Eran Toch, Super-Ego a framework for privacy-sensitive bounded context-awareness. *Proceedings of the 5th ACM International Workshop on Context-Awareness for Self-Managing Systems, 2011, ACM New York, NY, USA (2011), 24-32* doi>10.1145/2036146.2036151
6. Iris A. Junglas, Richard T. Watson, Location based services. *Communications of the ACM - Urban sensing: out of the woods CACM Homepage archive Volume 51 Issue 3, March 2008, ACM New York, NY, USA, 65-69* doi>10.1145/1325555.1325568
7. Gabriel Ghinita , Panos Kalnis , Ali Khoshgozaran , Cyrus Shahabi , Kian-Lee Tan, Private queries in location based services: anonymizers are not necessary. *Proceedings of the 2008 ACM SIGMOD international conference on Management of data, June 09-12, 2008, Vancouver, Canada, ACM New York, NY, USA (2008)* doi>10.1145/1376616.1376631
8. Paolo Bellavista , Axel Küpper , Sumi Helal, Location-Based Services: Back to the Future. *IEEE Pervasive Computing, v.7 n.2, p.85-89, April 2008, 85-89* doi>10.1109/MPRV.2008.34
9. Toby Xu , Ying Cai, Feeling-based location privacy protection for location-based services. *Proceedings of the 16th ACM conference on Computer and communications security, November 09-13, 2009, Chicago, Illinois, USA, ACM New York, NY, USA (2009), 348-357* doi>10.1145/1653662.1653704
10. Leighton Evans, Location-based services: transformation of the experience of space. *Journal of Location Based Services Volume 5, Issue 3-4, 2011, 242-260* doi>10.1080/17489725.2011.637968
11. Subhankar Dhar , Upkar Varshney, Challenges and business models for mobile location-based services and advertising. *Communications of the ACM, v.54 n.5, May 2011, ACM New York, NY, USA (2011), 121-128* doi>10.1145/1941487.1941515
12. Jing Yuan , Yu Zheng , Liuhang Zhang , XIng Xie , Guangzhong Sun, Where to find my next passenger. *Proceedings of the 13th international conference on Ubiquitous computing, September 17-21, 2011, Beijing, China, ACM New York, NY, USA (2011), 109-118* doi>10.1145/2030112.2030128
13. Vassilis Kostakos , Jayant Venkatanathan , Bernardo Reynolds , Norman Sadeh , Eran Toch , Siraj A. Shaikh , Simon Jones. Who's your best friend?: targeted privacy attacks In location-sharing social networks. *Proceedings of the 13th international conference on Ubiquitous computing, September 17-21, 2011, Beijing, China, ACM New York, NY, USA (2011), 177-186* doi>10.1145/2030112.2030138