A FRAMEWORK FOR ENHANCING PRIVACY IN LOCATION BASED SERVICES

JANE WAIRIMU MUGI

MASTER OF SCIENCE

(Computer Systems)

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY.

2016

A framework for Enhancing Privacy in Location Based Services

Jane Wairimu Mugi

A thesis submitted in partial fulfillment for the degree of Master of Science in Computer Systems in the Jomo Kenyatta University of Agriculture and Technology.

2016

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University.

Signature…….. Date ……..

Jane Wairimu Mugi.

This research has been submitted for examination with our approval as University Supervisors.

Signature ……. Date …….

Dr. Michael Kimwele.

JKUAT, Kenya.

Signature ……. Date …….

Dr George Okeyo.

JKUAT, Kenya.

DEDICATION

To God, my refuge and my strength. I might not know where the life's road will take me, but walking with you, God, through this journey has given me strength.

My late parents, my parents were a great inspiration, and their great commitment to education will always act as a motivation for me to put a lot of effort in my studies.

My children, Bobby and Jacintah, you are everything for me, without your love and understanding I would not be able to make it.

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to my supervisors Dr George Okeyo and Dr. Michael Kimwele for providing valuable guidance throughout the course of my project. I am immensely grateful to them for their advice and support without which I would not have come up with this work. I express my special appreciation and acknowledgement to my children for their love, cooperation, understanding and sacrifice throughout the work. I thank God for giving me knowledge and good health throughout the work. Last but not least, I thank all those who directly or indirectly contributed to the success of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

GPS:	Global Positioning System.
GSM:	Global System for Mobile Communication
RFID:	Radio Frequency Identification
Wi-Fi:	Wireless Fidelity
LBS:	Location Based Services
SSL	Secure Socket Layer
AZ:	Anonymizer.
ASR:	Anonymized Spatial Region.
CS	Candidate Set.
FCC:	Federal Communications Commission.
ATM:	Automated Teller Machine.
KNN:	K-Nearest Neighbour.
PIR:	Private Information Retrieval.
SC:	Secure Coprocessor.
POI:	Point of Interest.
ICT:	Information Communication Technology.

CAK:	Communication Authority of Kenya.
LSP:	Location Service Provider.
AT&T:	American Telephone & Telegraph
T-Mobile:	Telkom Mobile
JVM:	Java Virtual Machine
RDBMS:	Relational Database Management System

ABSTRACT

Re**c**ently anonymity in location based services has attracted a great deal of attention. This is due to the emerging location-detection devices together with ubiquitous connectivity that have enabled a large variety of location based services. Users have to reveal their location information to access location based services. However, this may threaten the user's privacy. Different solution approaches have been proposed to tackle this problem. K-anonymity approach has been studied extensively in various forms, however it is only effective when the user location is fixed. When a user moves and continuously sends the rectangular regions containing her location to the Location Service Provider (LSP), the LSP can still approximate the user's trajectory if it takes into account the overlap of consecutive rectangles, which poses a threat to the trajectory privacy of the user. This study reviewed recent advancement for offering K-anonymity and the most prevalent methodology. The outcome of the review is a framework that enhances privacy in Location Based Services. This framework ensures that user privacy is enhanced for both snapshot and continuous queries. The efficiency and effectiveness of the proposed framework was evaluated and the results indicates that the proposed framework has high success rate and good run time performance.

CHAPTER ONE

INTRODUCTION

1.1 Background

Due to the rapid advances in positioning technologies such as Global Positioning System, Global System for Mobile Communication, Radio Frequency Identification and Wireless Fidelity (802.11b/g/n), mobile devices are often equipped with geo-located and wireless communication capacities. These recent development of ubiquitous devices have led to the development of a new class of services known as Location Based Services (LBS), that are tailored to the current location of the individual querying the service. LBS can be defined as a service that takes as input the current location of a user (generally acquired through a mobile device carried by this user) and tailors its output depending on the acquired location data (Gambs et al., 2013).

LBS can access, combine, and transform contextual information and more specifically location information, in order to personalize the service provided to the user. For example, LBS can be used for resource discovery, path finding, real time social applications or location-based gaming. When people use LBS to support them in their daily tasks, their position is usually acquired automatically through mobile equipments they carry with them, thus these systems continuously monitor and reveal information about the location of their users as the position of these mobile systems is essentially the same as the users of such systems (Gambs et al., 2011).

Users with location-aware mobile devices can issue location-based snapshot or continuous queries to a database server at anytime and anywhere. Examples of snapshot queries include “Where my nearest petrol station is” and “what are the

restaurants within one mile of my location", while examples of continuous queries include "continuously report my nearest police car" and "continuously report the taxis within one Mile of my car"(Mokbel et., al 2006).

Although location-based services promise safety and convenience, they threaten the security and privacy of their customers. With untrustworthy servers, an adversary may access sensitive information about an individual's based on their issued location-based queries. E.g. an adversary may check a user's habit and interest by knowing the places she seeks (Chow & Mokbel 2006).

Due to the nature of spatial queries, LBS needs the user position in order to process her requests. LBS makes spatial data available to the users through one or more location servers (LS) that index and answer user queries on them. Examples of spatial queries could be "where is the closest hospital to my current location?" or which pharmacies are open within a 1km radius? In order for the LS to be able to answer such questions, it needs to know the position of the querying user. There exist many algorithms for efficient spatial query processing, but the main challenge in the LBS industry is of a different nature. In particular, users are reluctant to use LBSs, since revealing their position may link to their identity. Even though a user may create a fake identity to access the service, her location alone may disclose her actual identity. Linking a position to an individual is possible by various means such as publicly available information such as telephone directories. User privacy may be threatened because of the sensitive nature of accessed data e.g. inquiring for pharmacies that offer medicines for diseases associated with a social stigma, or asking for nearby addiction recovery groups. Another source of threats comes from less sensitive data e.g. gas station, shops, restaurants, that may reveal the user's interest and shopping needs, resulting in a flood of unsolicited advertisements through e-coupons and personal messages (Mouratidis & Yiu 2010).

To solve this problem the K-Anonymity concept is adopted, when a user *u* wishes to pose a query, she sends her location to a trusted server the anonymizer (AZ) through a secure connection. The later obfuscates her location, replacing it with an Anonymzing Spatial Region (ASR) that encloses a user. The ASR is then forwarded to the LS. Ignoring where exactly the user is, the LS retrieves (and reports to the AZ) a Candidate Set (CS) that is guaranteed to contain the query results for any possible user location inside the ASR. The AZ receives the CS and reports to user the subset of candidates that corresponds to her original query. In order for the AZ to produce valid ASRs the user send location updates whenever they move through their secure connection. (Mouratidis & Yiu 2010).

The ASR construction at the AZ (i.e., the anonymization process) abides by the user's privacy requirements. Particularly specified in an anonymity degree K by user, the ASR satisfies two properties that ASR must contain user and at least other k-1 users, and even if the LS knew the exact locations of all users in the system, it would not be able to infer with a probability higher than 1/k who among those included in the ASR is the querying one .

In the ASR LS must produce an inclusive and minimal CS. Inclusiveness demands that CS is a superset of users query results; this property ensures that user receives accurate and complete answers. Minimality, on the other hand, requires that the CS contains the minimum number of data objects, without violating inclusiveness. Minimality ensures that CS transmission (from the LS to the AZ), and its filtering at the AZ do not incur unnecessary communication and processing overheads (Mouratidis & Yiu 2010).

The advantage of this approach is that it incurs low communication cost between the client and the anonymizer. This study developed an effective and efficient framework to protect privacy of users in location based services. K-anonymity approaches are appropriate for preserving the privacy of the users who request LBS services. The identity of the requester can be easily revealed based on the participants of his/her anonymity set (Mokbel et al., 2006).

1.2 Statement of the problem

The emerging location-detection devices have enabled a large variety of LBS which has greatly enriched our mobility experience. Unfortunately, LBS may threaten the users privacy. Malicious attackers may collude with LBS provider to steal user location information and cause serious consequences. Thus, beyond the benefits they provide, users have started to be worried about the privacy of LBS (Zhong et al., 2005). Most of the existing approaches are only effective when the user location is fixed while querying for POI. When a user moves and continuously sends the rectangular regions containing her location to LSP, the LSP can still approximate the user's trajectory if it takes into account the overlap of consecutive rectangles, which poses a threat to the trajectory privacy of the user.

1.3 Justification

Offering location based services has led to privacy violations caused by sharing sensitive location information, leading to an urgent need for research on privacy. Although a few solutions have been proposed to address the privacy concerns in various aspects, there has not been any comprehensive study of the problem. There is urgent need for real time query processing framework that must efficiently process large numbers of continuous and snapshot queries (Khoshgozaran et al., 2011).

1.4 Broad objective

The overall objective of this research was to develop a framework that enhances user privacy in Location Based Services (LBS).

1.5 Specific Objectives

1. Identify the challenges in protecting privacy in LBS.
2. Investigate various models and techniques that have been used in protecting privacy in location based services.
3. Formulate a framework that protects the users from privacy attack using K-anonymity model.
4. Evaluate the framework for efficiency and effectiveness.

1.6 Research questions

1. What are the challenges in protecting privacy in LBS?
2. What are the various models and techniques that have been used in protecting privacy in location based services?
3. How will the conceptual framework be formulated?
4. How will the framework be evaluated?

1.7 Scope of Study

The study examined privacy in location based services and various techniques and models used in protecting privacy. The essentials examined included K-anonymity concept, in the design and implementation of a framework that enhances privacy in Location Based Services.

1.8 Structure of Thesis

Chapter 1: presents the background of the problem under study and objectives of the study. The chapter presents Location based services and its privacy challenges.

Chapter 2: It gives a review of related literature that formed the basis of the research. The chapter gives a review of related literature on models and techniques of Location Based Services.

Chapter 3: Describes the methodology used in the research.

Chapter4: Describes results and discussion of thesis.

Chapter 5: Covers conclusion, recommendations and further research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview of Location-based services

Location-based services emerged from a mandate called the enhanced 911 (E911) Mandate (Bellavista et al., 2008).The Federal Communications Commission (FCC) established this mandate to require mobile network operators to provide accurate location information of wireless 911 callers to emergency services personnel, enabling them to locate and respond quickly to these wireless 911 callers much more quickly (Federal Communication Commission, 2001). Because of this mandatory policy, mobile network operators (e.g., AT&T and T-Mobile) and mobile device manufacturers made large investments to incorporate advanced positioning capabilities into their mobile and mobile devices (Bellavista et al., 2008).

Businesses began developing commercial location-based services to exploit the potential of these positioning capabilities. Location-based service is an information or entertainment service, which is accessible with mobile devices through the mobile network and uses geographical position information provided by the mobile devices (Quercia et al., 2010). LBS provide information that is specific to a given location (Rainer & Cegielski, 2012). Hence, location-based service is a form of mobile services, which is a service provided through a wireless Internet-enabled device.

There are five necessary elements for location-based services: mobile devices, communication network (the mobile network which sends user data and service request from the mobile terminal to the service provider), the positioning component, service and application provider, and data and content provider (Jiang, & Yao, 2006)

The interaction of these five key elements of location-based services enable a mobile user to search for the nearest business or service in their location proximity e.g., receive alerts, find a friend, locate taxis, service personnel, doctors, and rental equipment; schedule fleets; track objects such as packages and train boxcars; find information such as navigation, weather, traffic, and room schedules; and automate airport check-ins (Rainer & Cegielski, 2012).

Schiller & Voisard (2004), identified six categories of location-based services applications infotainment services, tracking services, selective information dissemination services, location-based games, emergency support services, and location-sensitive billing.

Table 2. 1 : Categories of LBS Applications

SERVICE CATEGORY	EXAMPLE APPLICATION
Infotainment services	Finder applications (e.g., route, location, friend, store, restaurant, gas station, and parking) Information requests (e.g., tourist, travel, news)
Tracking services	Goods, vehicle, and fleet People (e.g., child care, elderly, sick, and offenders) Security of entities (e.g., cars) Maintenance and assistance Workforce dispatching Supply-chain and inventory
Selective information Dissemination	Targeted content dissemination (e.g., advertisements)
Location-based Games	Treasure hunts Scavenger hunts
Emergency support Services	Emergency 911 ambulance, fire, police dispatching Roadside assistance
Location-sensitive Billing	Call billing Toll payment Purchase of goods and services

2.2 LBS Information Delivery Mechanism

According to (Xu et al., 2010), there are two types of information delivery mechanisms commonly used for LBS – pull or push. The pull-based LBS is a type of LBS which is user initiated. Users request for specific information or service by voluntarily providing their location information. For example, users provide their location information to receive real time navigational requests to the nearest auto-teller machine. In contrast, the push based LBS is a type of LBS which is service provider initiated.

Through positioning technologies in mobile devices, a service provider is able to sense the location of users and sends relevant information or service to the user based

on the user's location. (Schiller & Voisard, 2004)users receive location-specific information, either with prior consent (subscription-based) or without prior consent (non-subscription-based), without having to actively request for it. For example, location information is used to target and send related advertisements to the user when the user is near a store.

2.3 LBS Communication Model

The LBS communication model consists of three layers, a positioning layer, a middleware layer, and an application layer (Schiller & Voisard, 2004). The positioning layer is responsible for calculating the position of a mobile device with the help of position determination equipment and the geospatial data in a geographic information system. The calculated position is then passed directly to an application. Recently mobile network operators have introduced a middleware layer between the positioning layer and the application layer to reduce the complexity of service integration, saving operators and third-party application providers' time and cost for application integration. This middleware layer manages the interoperability between networks for location data. The application layer comprises of all those services that request location data.

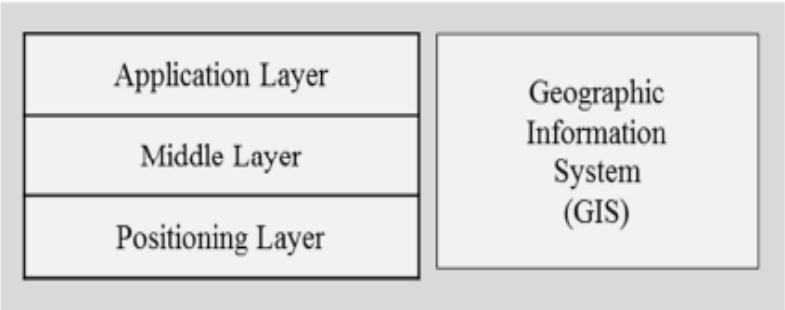

Figure 2. 1 : Communication model

2.4 Challenges in Location Based Services

Since its introduction in 2001, after the E911 mandate as a non-consumer application, Location-based services are transitioning into a mature technology (Duck ham et al., 2007). However, location-based services have not been well received by consumers. There are several challenges that have affected the development and growth of location-based services. These challenges have slowed down the wide-scale adoption of location-based services, which is critical for its success. Such challenges include (Dhar &Varshney, 2011).

Pricing for LBS poses a big challenge to all the players in the mobile ecosystem including customers, advertisers, and marketing vendors as well as advertising service agencies. The pricing for voice and data services including applications differ considerably. The network operators generally have per-minute and flat rate charging models for voice services and open ended pricing models for data services.
Currently, the consumers have to pay on the basis of airtime and amount of data, and this has not been widely accepted by consumers.

Privacy for LBS poses a big challenge .Mass adoption of location based services will depend on how well network operators protect privacy of consumers. Unsolicited mobile advertising is generally considered intrusive and some consumers dislike marketing efforts that encroach their privacy (Dhar &Varshney, 2011).

Adoption of LBS will largely depend on how the network operators charge for value-added mobile data services. The success of mobile advertising will depend on the volume of traffic using the wireless network and associated location- based services. With this nascent and emerging market comes some uncertainty, which may cause low advertising revenues initially. Although there are quite a few location-based services currently available, the use of wireless data services is far below the expectation particularly in the U.S. Consumers view these services as 'nice to have' rather than 'must have' and are not willing to pay the additional fee for many of these niche services Other technological challenges include a need for development of a middleware to support multiple diverse technologies with different frequencies, protocols, and location accuracy. Other challenges include dynamic location changes, cost of communications, battery and computing power, continued wireless connectivity, download speed, and smaller screen size and lack of standards (Dhar &Varshney, 2011).

2.5 Location Privacy Approaches

There are a number of approaches in the literature to solve the problem of privacy protection. Privacy protection aims at preventing the server from knowing the user's original query information (e.g., identity, location). The existing privacy protection approaches are categorized into into: spatial anonymization, Obfuscation and private retrieval methods. Spatial

anonymization requires an anonymizer in the architecture, whereas the other two approaches adopt the (conventional) client-server architecture. The three approaches includes:

2.5.1 Spatial Anonymization

A trusted third-party server, called an anonymizer, is placed in-between the client and the server. The anonymizer enlarges an exact user location q into a (superset) cloaked region Q_0 so that it contains q and also the locations of K-1 other users. This way, the server cannot distinguish q from other user locations in Q_0. Upon receiving Q_0, the server processes a range NN query. (Hu et al., 2005) in order to return a candidate set that contains the nearest data point for any location in Q_0. The anonymizer then refines the candidate set and reports the actual result to the client. The advantage of this approach is that it incurs low communication cost between the client and the anonymizer.

Mokbel et al., (2006) propose efficient heuristics for computing a superset of the candidate set, (Hu & Lee, 2006) develop the range kNN algorithm for computing the minimal possible candidate set for a rectangular cloaked query, and (Kalnis et al., 2007) study the computation of the candidate set for a circular cloaked query. After receiving the candidate set from the server, the anonymizer derives the actual result for q and sends the result to the user. Like other previous studies (Gedik & Liu,2008) the location cloaking algorithms account for snapshot user locations only (Mokbel et al., 2006) Neither of them considers continuous queries and trace analysis attacks. Anonymity-based defenses aim at preserving the anonymity of the issuers so that an adversary is not able to associate private information present in the requests with a specific individual. The defenses transform the so-called quasi-identifier information in requests so that the issuer becomes indistinguishable in a sufficiently large group of users (anonymity set).

(Gruteser & Grunwald (2003) and (Chow *et al.*, 2006) have based their approaches on K-anonymity. The main idea behind this concept is to send a box of locations instead of only the true one, whereby the probability to guess the user's location is always less than 1/K. Most of techniques relying on K-anonymity use a middleware (the anonymizer). This anonymizer is a third party responsible for creating a Cloaking Region (CR), which contains the true user's location, as well as K-1 other neighbors. With such a technique, a typical scenario can be a user trying to localize the nearest bank. The user sends his/her requests (including his/her credentials) to the anonymizer through a wireless net-work. Thereafter, the anonymizer, which keeps the locations of all current users, authenticates the requester first and chooses a set of K-1 neighbors to create a CR(Croaked Region) that can be sent instead of the user's position. This way, the risk of violating the user's privacy is reduced by making it difficult to locate the position that has triggered the process (since the server is answering the whole CR).

This approach suffers from several drawbacks. Firstly, the users' data is still revealed to a third party (the anonymizer) and thus the problem of preserving the user's privacy has not been solved. That is, we still have no guarantees that the anonymizer cannot be misused if a malicious hacker gains access to it. Secondly, the anonymizer needs to update the current location of all the subscribed users repeatedly, which will require a permanent communication and remote monitoring of the users, which is a clear violation of the users' privacy. Finally, the robustness of these approaches depends totally on having a relatively big number of neighbors at the time of receiving the requests. Therefore, depending on a middleware is far from being a perfect solution to secure location-dependent queries and hence any secure solution need to communicate directly with the Location Based Server without any intermediate parties (Gahi et al., 2012). (Kido et al.,

2005) proposed a new technique to hide users' location and trajectory by sending several queries instead of only one. The technique depends on creating several fake queries with fake identities in addition to the real query, thereby; the LBS server will not be able to identify. Apparently, the perfection of this mechanism depends on the number of fake re-quests generated; the more fake queries generated the more robust and secure the system becomes.

The problem with this technique is that as the number of requests sent out by a user grows, the LBS may suspect that it is under an attack and thus the requests may be ignored. Moreover, receiving a big number of requests can slow down the server's response time significantly. The advantage of spatial anonymization is that the communication cost between the client and the anonymizer is optimal. The anonymizer and the K-anonymity model exhibit several disadvantages. First, the anonymizer becomes a performance bottleneck because it needs to serve all its subscribed users, as well as maintaining accurate records of their locations. Second, the anonymizer is vulnerable to malicious attacks. An example is the collusion attack, where the adversary subscribes to the anonymizer multiple times with fake user locations. Then, the adversary can exclude those fake locations from the cloaked region Q0 generated by the anonymizer, this way identifying a user with higher probability. Third, it heavily depends on the distribution and density of other mobile users. When the user is located in a sparse region or few users are subscribed to the anonymizer, it needs to construct a very large cloaked region Q0 such that it contains K user locations.

This incurs high processing time at the server and the anonymizer. Alternatively, a K-anonymous region can be derived through peer-to-peer communication (Chow et al., 2006). Users close

together form a group and set their cloaked region as a rectangle containing them. The drawback is that group formation and maintenance incur communication latency.

2.5.2 Obfuscation

Obfuscation adopts the traditional client-server architecture. It avoids the disadvantages associated with the anonymizer (Ardagna et al., 2007). The client is responsible for enlarging the user's location q into a hidden set Q_0. Specifically, Q_0 can be represented as a simple region or a discrete set. (Cheng et al., 2006). The server returns the candidate set of Q_0 to the client, which then computes the actual result from the candidate set (Xu et al., 2010). However, this approach also has drawbacks. In case Q_0 is a simple region, it is hard to control the candidate set size and the client communication cost especially when Q_0 falls into a dense area. On the other hand, if Q_0 is a discrete set, then it cannot survive location guesses by the adversary (Kido et al., 2005).

A connected obfuscated query Q0 is a simple region (e.g., a rectangle or circle) that contains the user's location q. It can be processed by server-side techniques employed in K-anonymity solutions. For instance, the processing of a rectangular query and a circular query can be handled by the proposal of (Kalnis et al., 2007).The only difference is that no anonymizer is used now. Upon receiving the candidate results from the server, the client is responsible for refining them into the actual result. The disadvantage of using a connected obfuscated region Q0 is that it is hard to control the communication cost, especially when Q0 overlaps with a dense region (of data points).

Various techniques have been proposed for the client to construct Q0 from q. The study of (Ardagna et al., 2007) takes location positioning inaccuracy into account, models the user

location as a circular region, and develops several geometric operators for deriving obfuscated regions. (Xu et al., 2010) study the obfuscation for a continuously-moving query user.

Cheng et al (2006), process range queries on a private dataset that contains the (obfuscated) locations of all users. The client enlarges the user's exact location q into a circular region Q0 based on the user's requirement on the region area and the coverage of sensitive facilities (e.g., a hospital). The server manages the circular regions of all users. Upon receiving a range query, the server computes a candidate result set and derives the probability/confidence of each candidate being an actual result. The quality of the result set is summarized by a quality score that combines the confidence value of each candidate. Although the method of (Cheng et al.,2006) can be generalized to the case part of the cloaked range query so a candidate is still associated with a probability. Here all other users' locations are precise points; each candidate only qualifies as result for part of the cloaked range query so a candidate is still associated with a probability.

2.5.3 Private Information Retrieval (PIR)

Private Information Retrieval also operates on the client-server architecture (Ghinita et al., 2007). It was first proposed by (Chor et al., 1998) in information theoretic setting, which also proves that any theoretical PIR scheme has a lower communication bound equal to the database size. However, relaxing the problem to computationally bounded adversaries, a PIR framework is proposed by (Kushilevitz & Ostrovsky1997) based on the quadratic residuosity assumption. The client and server follow a secure two-party computation which allows the client to privately retrieve the i^{th} bit from a bit string of size n owned by the server. (Ghinita et al., 2007) build a framework on top of this PIR protocol to enable location privacy. The key strength of the

technique proposed by (Ghinita et al., 2007) is the perfect privacy guarantee against the most powerful adversaries and correlation attacks. This stringent privacy guarantee comes at the cost of executing expensive protocols that result in significant communication and computation overhead.

Mouratidis& Yiu (2012), propose the first PIR-based method for shortest path queries. Such queries are used, for example, in location-based services that provide driving directions from the user's location to a desired destination (through the road network of a city or country.

The client encodes its original query q into an 'incomprehensible' query q'. Next, the server computes the encoded result of q} blindly, and then the client derives the actual result from the encoded result. This approach offers the strongest privacy guarantee when compared to spatial anonymization and obfuscation approaches (Khoshgozaran et al., 2007) However, it requires specialized algorithms at the client and the server that are hard to implement and do not utilize existing spatial indexes. They also need a trusted party to perform pre-processing on the dataset beforehand (Indyk &Woodruff 2006).

Unlike obfuscation, PIR offers cryptographic privacy guarantees, based on reductions to problems that are either computationally infeasible or theoretically impossible to solve.PIR is generally resource-intensive. However, recent PIR protocols achieve practical response times (in the order of seconds over Giga- byte databases (William & Sion 2008), and have been successfully applied to private spatial queries in Euclidean space (Khoshgozaran et al., 2011) As yet, there has been no PIR-based solution for shortest path computation.

2.6 Techniques and Models for Offering Privacy in Location Based Services.

2.6.1 Mix Zones

Mix zones approach was proposed by (Beresford et al., 2004) to define areas called mix zones, where all user positions must be hidden such that the user position is not known within these zones. This is achieved by not sending any position updates within a zone. If a user enters a mix zone, the user identity is mixed with all other users in the zone by changing pseudonyms to protect user identities. Therefore, an attacker cannot relate different pseudonyms of the users even by tracing the entry and exit points of a Mix zone. The MobiMix approach proposed by (Palanisamy et al., 2011) applies the mix zone concept to road networks. They take into account various context information that can be used by an attacker to derive detailed trajectories such as geometrical and temporal constraints. It has the advantage of location and sampling accuracy but operation lack in multiple responder.

2.6.2 Dummy-Q

The dummy based location privacy proposed by (Jensen &Yiu2010).A user centric technique for query privacy protection which operates solely on the user side and doesn't require any trusted third party, the key idea is to confuse the adversary by issuing counterfeit queries with varying service attributes but the same (real) location, henceforth referred to as dummy queries, along with each real query issued by the user. While the notion of dummy queries appears simplistic, the challenges are plenty and intricate especially in the scenario of continuous LBS queries (Liu et al., 2008). In the proposed techniques, a user sends the true user location mixed with fake locations, *i.e.* dummies, to the LBS provider. The user extracts the necessary information from

the reply message while the service provider can only learn vague details of the user location. However, the problem of the existing dummy based approach is that they only consider the snapshot scenario. Adversaries can comprehend the user movements when tracking data during long-term observations. The sequence of location data can be used to discover a rough trajectory by applying data mining techniques.

Challenges of Dummy-Q

A critical requirement for dummy generation is that the dummy service attribute values must be generated in a judicious manner so as to remain consistent with the query context i.e. the location where query is issued. E.g. while users on a coastal location may often query for beaches, the same service attribute value may be quite rare around a desert area. If such adherence to the trend of queries is shunned then the adversary may be able to exclude certain service attributes according to common sense and thereby identify the real query.

One must insert the same(dummy) service attribute values over different snapshots of a continuous LBS query, in order to prevent the adversary from inferring the most frequent value as the real one(note that the real value has to be included at all snapshots) This combined with the first challenge requires the dummy insertion process to take into consideration the user's possible future locations as otherwise the dummy values inserted in the beginning may later be excluded by the adversary according to the context of future queries.

One must minimize the number of inserted dummy queries because each consumes additional overhead for issuing the query and waiting for the answer. The real and dummy queries have to be issued sequentially, because many LBS servers (e.g. yahoo!) deny queries issued with time

interval shorter than a threshold (Claudio et al., 2005). The real query cannot always be issued before the dummies as otherwise the adversary may identify the real query based on timing (Brinkhoff et al., 2002). Limited storage and computational capacity of mobile devices from which many LBS queries are issued and therefore privacy protection must be enforced. The query context has to be stored locally on a mobile device because of the design choice of user centric Dummy-Q, the major challenge is to store and retrieve the query context information in an efficient manner i.e. with minimal storage and computational overhead (Pingley et al., 2011).

2.6.3 Location K-Anonymity Model

Anonymity was first discussed in relational databases, where published data (e.g. census medical) should not be linked to specific persons (Wortman et al., 1989). K-anonymity has been used widely recently. A relation satisfies K-anonymity if every tuple is indistinguishable from at least K-1 other tuples with respect to a set of quasi identifier attributes e.g. date of birth, gender, zip code that can be linked to publicly available data to identify individuals (di Vimercati, 2011).K-anonymity concept has been adopted in most previous work done on location-based services.. In K-anonymity framework a user sends his location and query to the anonymizer through a secure connection the anonymizer removes the id of the user and transforms his location through a technique called cloaking. Cloaking hides the actual location by a K-anonymity spatial region (K-ASR or ASR), which is an area that encloses the client that issued the query as well as at least k-1 other users. The anonymizer then sends the ASR to the LBS which return to the anonymizer a set of candidate results that satisfy the query condition for any possible point in the ASR. The LBS may be compromised. I.e. an adversary may have complete knowledge of all queries received by the LBS (Kalnis et al., 2007). The advantage of this approach is that it incurs low communication cost between the client and the anonymizer.

Location cloaking in general seeks to prevent an attacker from being able to match queries to particular users and to thus compromise their privacy. The attacker may be in a position to observe traffic flowing through the network or even be situated at the LBS provider endpoint. One popular cloaking technique is based on the principle of k-anonymity where a user is hidden among k-1 other users. Queries from multiple users are aggregated at an anonymity server which forms an intermediary between the user and the LBS provider. The central anonymity server can provide spatial and temporal cloaking functions, so that an attacker will encounter difficulty matching multiple queries that are observed with users at particular locations and at particular points in time. Many cloaking solutions for location privacy suggests either a central anonymity server as described (Gruteser et al., 2003), or other means such as decentralized trusted peers (Chow et al., 2006), or distributed k-anonymity (Zhong et al., 2009).

The chief problem is that the anonymizer server must normally be part of the trusted computing environment and present a single point of vulnerability. If it is successfully attacked or the collusion with the LBS occurs then the location of all users may be divulged. Although a cloaking technique by itself is advantageous in that it does not result in increased computational cost on the server it can carry with it a high communication cost from the LBS provider to the client, this can mean a large and unacceptable penalty for mobile phone users. If a reduced sample population results from the number of users in a particular geographic area it may not suffice to satisfy the desired degree of anonymity. If the anonymity server delays execution of a request until the k-anonymity condition is is satisfied, then this delay may prove to be unacceptable to the user from a feature interaction point of view (Olumofin et al., 2010).

Location k-anonymity is a traditional technique to provide both location and query privacy simultaneously. With this technique, a LBS query is issued to the LBS server via a trusted third party. The third party arguments a user's location to a cloaking region, which geographically covers not only the user who issues the query but also

k-1 other users and then transmit the query to the LBS server. Since all the k users report the same cloaking region in their queries, the adversaries cannot distinguish the location or service attribute of any user from the received queries (Bettini et al., 2005).

2.6.4 K Nearest Neighbor(kNN) Query

Existing approaches to *k* nearest neighbor (*k*NN) computation in spatial networks can be divided into two types: approaches that compute *k*NN queries by incrementally scanning the network until *k* neighbors are found, and approaches that apply some form of pre-computation and "compute" *k*NN queries by looking up data collected in pre-computed data structure. Both types of approaches assume that the spatial network is represented by graph-like data structures.LBS rely on k nearest neighbor query that retrieves the k data points closest to a user location q. (Hjaltason et al., 1999) A server-side database stores a set of points of interest. To retrieve the nearest data point, the user's mobile device sends its location q to the server. The server then computes the (nearest neighbor) result and returns it to the mobile client. The moving k nearest neighbor query, which computes one's k nearest neighbor set and maintains it while at move, is gaining importance due to the prevalent use of smart mobile devices such as smart phones. Safe region is a popular technique in processing the moving k nearest neighbor query. It is a region where the movement of the query object does not cause the current k nearest neighbor set to change. Processing a moving k nearest neighbor query is a continuing process of checking the validity of the safe region and recomputing it if invalidated. The size of the safe region largely decides the frequency of safe region recomputation and hence query processing efficiency. Existing moving k nearest neighbor algorithms lack efficiency due to either computing small safe regions and have to recompute frequently or computing large safe regions with a high cost. As a complication to this scenario users may wish to avoid disclosing their exact locations to the server (Yiu et al., 2011).

2.6.5 Space Twist Framework

The Space twist framework aims to offer location privacy for K nearest neighbor (kNN) queries at low communication cost without requiring a trusted anonymizer.The client specifies a fake user location called an anchor, which utilizes incremental NN query processing at the server. The server returns data points to the user incrementally in ascending order of their distances from the anchor (Hjaltason et al., 1999). Space twist rectifies the shortcomings of k nearest neighbor queries. This approach is flexible, needs no trusted middleware and requires only well-known incremental NN query processing on the server. Space twist ensures that no duplicates are retrieved. It offers a server side ring ranking technique that reduces the communication cost of exact queries. A delayed termination technique that reduces the communication cost of exact queries. However Space Twist may fail since it cannot guarantee K-anonymity

The server-side functionality can be implemented by a SQL query. The computational effort of the server can be significantly reduced by including the clause4 into the SQL query as shown in the query below where m is (an upper bound on) the number of data points to be retrieved by the client, q the actual user location, q'is the anchor location, P is the the set of points of interest/data points, we propose to store the point set P as a relational table with schema (id; x; y),where id is an identifer, x and y are location coordinates. Each row corresponds to a point p in P. Then, we denote the coordinates of the anchor location q0 by qx0 andqy0 (where an underscore indicates a given value). The distance dist(q0; p) can then be expressed as in equation (1):

$$\sqrt{(x - -qx') \_ (x— -qx') + (y - -qy') \_ (y - -qy')} \quad \text{..... (1)}$$

Thus, the INN operator can be implemented by the following SQL query.

```
SELECT id, x, y

FROM P

ORDER BY (x-_qx')_(x-_qx') + (y-_qy')_(y-_qy') ASC
```

Figure 2. 2 : Incremental Nearest Neighbor SQL query

The ORDER BY clause uses the squared distance because it preserves the ordering of distances and is faster to compute than dist(q0; p). The query is a ranking query, and it can be executed efficiently

2.6.6 The New Casper

This is a framework in which mobile and stationary users can entertain location-based services without revealing their location information. Casper consists of two main components the location anonymizer and the privacy-aware query processor. The location anonymizer hides the users exact location information into cloaked spatial regions based on user-specified privacy requirements. The privacy-aware query processor is integrated inside the location-based database server in order to deal with the cloaked spatial areas rather than the exact location information. Casper maintains an anonymizer and a privacy aware query processor. For a successful anonymization and cloaking, a pyramid structure is maintained. The cells in the region contain the number of mobile users active in the cell. If the current region/cell of user cannot satisfy the value of K or Amin then a neighboring cells are considered.

Casper has a superb run time performance, but it does not return the smallest cloaking region and also it is expensive (updates and cloaking) to maintain the

pyramid structure. Casper employs a grid-based complete pyramid data structure that hierarchically decomposes the spatial space into H levels where a level of height h has 4h grid cells. The root of the pyramid is of height zero and has only one grid cell that covers the whole space. Each pyramid cell is represented as (cid, N) where cid is the cell identifier while N is the number of mobile users within the cell boundaries. The pyramid structure is dynamically maintained to keep track of the current number of mobile users within each cell. Thus updating and cloaking is very expensive due to maintaining the pyramid structure (Stenneth & Yu 2012).

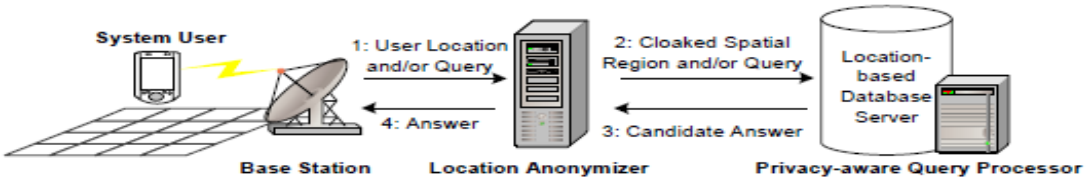

Figure 2. 3 : Casper System Architecture

2.7 Summary

There are various models and techniques for offering privacy in location based services; Mix zones, Dummy-q, K-Nearest Neighbor, Space twist and the New casper. Mix zone has the advantage of location and sampling accuracy but operation lack in multiple responder. While the notion of dummy queries appears simplistic, the challenges are plenty and intricate especially in the scenario of continuous LBS queries. K Nearest Neighbor Query, a server-side database stores a set of points of interest. As a complication to this scenario users may wish to avoid disclosing their exact locations to the server. Space twist framework aims to offer location privacy for K nearest neighbor (kNN) queries at low communication cost without requiring a trusted anonymizer. The client specifies a fake user location called an anchor, which utilizes incremental NN query processing at the server. However Space Twist may fail since it cannot guarantee K-anonymity. The New Casper consists of two main components the location anonymizer and the privacy-aware query processor. The

location anonymizer hides the users exact location information into cloaked spatial regions based on user-specified privacy requirements. The pyramid structure is dynamically maintained to keep track of the current number of mobile users within each cell. Thus updating and cloaking is very expensive due to maintaining the pyramid structure.

Useful obfuscation can be provided at hardware level and underlying frameworks but it is impossible due to presence of multiple LBS Providers accepting location data. Anonymization can only work if service providers can come to a mutual agreement on location data handling and disclosure. It is possible to track user movement even if user is not publishing their location publically this happens when an attacker collude with the LBS. The process to standardize and mandate a common standard in use of location Based Services is still in infancy. With rising concern about privacy and re-identifications, usefulness of location based services will be severely challenged (Singh, 2013). Most of the PIR-based approaches for location privacy rely on hardware-based techniques, which typically utilize a secure coprocessor (SC) at the LBS server host (Hengartner et al., 2007). This hardware creates a computing space that is protected from the LBS, to realize query privacy. A major drawback of SC-based PIR is that it requires the acquisition of specialized tamperproof hardware and it usually requires periodic reshuffling of the POIs (Point Of Interest) in the database, which is a computationally expensive operation (Ilievet al., 2005).

While creative solutions have been proposed to solve the location privacy problem, there are still many challenges to be addressed. Devising a framework that while ensuring perfect privacy, can very efficiently respond to various spatial queries dealing with both static and dynamic objects is still an open problem and far from what the existing approaches offer (Khoshgozaran et al., 2010

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

This chapter provides the methodology that was used in the research. It covers target population, research method, design, testing techniques, tools for analysis, technology for development and proposed architecture.

3.2 Research Design

In this study the researcher adopted a descriptive survey design. A descriptive survey research study was preferred since it has the dimension of investigating possible relationships between two or more variables. The descriptive survey design is ideal since it is concerned with making accurate assessment of the inference, distribution and relationship of the phenomenon (Edwards & Gillies 2006). According to (Gay,1981) descriptive research is a process of collecting data in order to answer questions concerning the current status of the subject in study.

Some test was conducted for the purpose of testing the proposed framework.

3.3 Target Population and Sample

The population of interest in this study was mobile phone users in Kenya who use location based services applications to search for Point of Interest (POI). According to survey titled Mobile Life Study, conducted by TNS Global and whose results were released in 2013, about 74 per cent of Kenyans who participated in the survey as respondents indicated their willingness to share their location status online. The researcher adopted simple random sampling technique Random sampling is a sampling design in which k distinct items are selected from the n items in the population in such a way that every possible combination of k items is equally likely to be the sample selected (Thompson, 2012).

3.4 Instrumentation

The type of data used in the research study was primary data collected through a questionnaire. According to (Kothari 1984) primary data is original information collected for the first time. To ensure reliability of data collected a pre-test of the questionnaire was done to determine whether the respondents understood the questions correctly and where the questions did not seem clear enough, the necessary adjustments were made. The questionnaire distributed to mobile phone users contained both open ended questions as well as close ended questions. Questionnaire was chosen because of its simplicity of administration and high reliability as advocated by (Babbie 1993). The items on the questionnaire were developed on the basis of the objectives of the study. The questionnaire contained two sections; Section A addressed background information and Section B addressed privacy concern in Location Based Services.

3.5 Reliability and Validity of the instrument

Reliability refers to a measure of the degree to which a research instrument yields consistent results or data after repeated trials. This type of reliability is referred to as Test-Retest. Test and retest simply put, is that you should get the same result on test 1 as you do test 2 when the two tests are administered after a time lapse. Retest involves two administration of the measurement instrument (Yin, 2003). The instrument was pre-tested for their reliability where a number of 3 mobile phone users were chosen for pre-test as a sample for pre-test should be small. Alpha coefficient was used to test reliability of the instrument whereby a coefficient of 0.70 or more was acceptable as advocated by (Fraenkel & Wallen2009). A high Cronbach alpha coefficient (0.7 and above) implies that the items correlate highly among themselves, that is, there is consistency among the items in measuring the concept of interest. The sample for pre-test was also used to test data validity. The validation of the instrument was aimed at ensuring the instrument was measuring what they were intended to measure (Kathuri & Pals, 2007).

3.6 Data Collection Procedures

The instrument was administered by the researcher where the mobile phone users were required to respond to questions asked by the researcher hence any clarification regarding the instrument was easily addressed.

3.7 Development of Proposed Framework

An anonymizer system was developed using scala (JVM-based), the developed anonymizer included a continuous and a snapshot query anonymizer that would help in cloaking both snapshot and continuous queries. An Android mobile application system was created with a user interface that would help mobile users to search for POI. The mobile application system was developed using Android developer.

3.8 Experiment and Test

To evaluate the efficiency and effectiveness of our framework the following evaluation criteria was used.

3.8.1 Success Rate

According to (Stenneth.L& Philip S. 2012), success rate is one of the most important evaluation criteria. The main goal of any anonymization server is to maximize the number of messages that can be successfully anonymized with the personalize quality of service and privacy requirement desired. The success rate is measured as the ratio of the number of successful anonymized request, by the total number of incoming mobile request. A success rate of 100% implied that all the requests that were sent by the mobile clients were safely anonymized. Let N be the total number of mobile requests sent by mobile clients to the anonymization server. The number of mobile request that can be anonymized successfully by the anonymization server is Mt as shown by (2).

$$N/M_t * 100 = 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

This implies that all the queries sent to the system will be anonymized successfully and none will be dropped. In our evaluations we refer to this property as the *success rate* of the proposed framework.

3.8.2 Performance Measure

According to (Stenneth.L& Philip S. 2012), an algorithm with a lower cloaking time does better, because the cloaking time is a measure of the temporal complexity. Efficient cloaking implies that the algorithm spends less time processing the incoming mobile requests from the mobile clients. We define a function startTime (cloakingAlgorithm as shown below), that returns the time the cloaking algorithm start the anonymization process. Also, endTime (cloaking Algorithm as shown below), which returns the time the cloaking algorithm completes the anonymization process shown by (3).

Cloakingtime=endTime(CloakM(ms))-
startTime(CloakM(ms)…….(3)

3.9 Summary

The chapter includes the methodology used in the research. The researcher adopted a descriptive survey design. A descriptive survey research study was preferred since it has the dimension of investigating possible relationships between two or more variables. The descriptive survey design is ideal since it is concerned with making accurate assessment of the inference, distribution and relationship of the phenomenon. The framework was developed that included an anonymizer system which was developed using scala, and a mobile application system which was developed using android developer. An experiment was conducted to evaluate the framework for efficiency and effectiveness. The evaluation criteria used was success rate and performance measure.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Data Analysis

The study utilised first hand data which comes from the chosen respondents who answered the survey-questionnaires administered to them .In order to answer the research questions, the data

collected needs to be thoroughly analyzed. Yin, (2003) explains that every investigation should start with a general analytic strategy, allowing the researcher to decide what to analyze and why.

The questionnaires were edited for completeness and consistency before processing. Editing helped in detecting errors and omissions and which were corrected to ensure that maximum data quality standards were achieved. Data was then coded to enable responses to be grouped into categories. Coding involved assigning numbers so that the responses could be grouped into number of classes or categories. Data analysis was then carried out using the Statistical Package for Social Sciences (SPSS). The data collected was first subjected to descriptive statistics which included frequencies, percentages, means, and standard deviations. Inferential statistics was also very important for the study, in this case Pearson's moment correlation coefficient was used to determine the magnitude of relationship between two variables. A positive relationship means that an increase in one variable leads to an increase in another variable and vice versa. The responses were tabulated and subsequently presented by use of tables, bar charts and pie charts.

4.1.1 Data Description

The target sample size was 50 respondents who were selected randomly and their age bracket was to be between 18yrs and 60 years. Though we were not able to meet our target sample size we managed 74% who are 37 respondents in our study.

4.1.2 Type of Location Based Services applications in use

The study showed that majority(100%) of the respondents were familiar with Google maps while only 5.6% used map quest but 94.4% of the same didn't know. Nobody knew about Yelp with some of them having used Google-Ingree. We chat and Face book Check in was moderately used. This implies that the mostly used LBS applications used in Kenya is Google map and face book check in.

Table 4. 1 : Type of Location Based Services applications in use

	Yes	No
	Row N %	**Row N %**
Types of location based service/applications that you use-Google map	100.0%	**.0%**
Types of location based service/applications that you use-MapQuest	5.6%	**94.4%**
Types of location based service/applications that you use-Yelp	.0%	**100.0%**
Types of location based service/applications that you use-Face book check in	52.8%	**47.2%**
Types of location based service/applications that you use-Google ingree	20.0%	**80.0%**
Types of location based service/applications that you use-we chat	27.8%	**72.2%**

4.1.3 Trustworthiness of Location Based Services

The study showed that 77.8% of the respondents trust LBS while 22.2% don't trust it. This implies that LBS are not trusted due to lack of privacy.

Table 4. 2 : Trust in Location Based Services

	Yes	No	
	Row N %	**Row N %**	
Do you trust location based service provider	77.8%		**22.2%**

4.1.4 LBS Applications Usage

The study indicated that 10 (27%) of the respondents use LBS applications 3 times a week while the least number of respondents use LBS with 5.4% and 8.1% do not use LBS at all. This implies that LBS is used on average use in weekly basis that is 3 times in a week.

Table 4. 3 : LBS Application Usage

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<1	2	5.4	5.4	**5.4**
	1	5	13.5	13.5	**18.9**
	2	3	8.1	8.1	**27.0**
	3	10	27.0	27.0	**54.1**
	4	5	13.5	13.5	**67.6**
	5	2	5.4	5.4	**73.0**
	>5	7	18.9	18.9	**91.9**
	I don't use LBS	3	8.1	8.1	**100.0**
	Total	37	100.0	100.0	

.4.1.5 Importance of Location Based Services

The study showed that 55.6 % of the respondents think that LBS applications are good; 0% thinks that they are poor and 16.7% think that they are Excellent.

Table 4. 4 : Importance of Location Based Services

	Poor	Moderate	Good	Excellent
	Row N %	Row N %	Row N %	**Row N %**
How valuable do you think location based services/apps are?	.0%	27.8%	55.6%	**16.7%**

4.1.6 Privacy risk in sharing location information

The study showed that 10; 84% of the respondents perceive a high risk to sharing location information as compared to 16% who do not perceive any level of risk. This implies that most people perceive a high risk while sharing location information.

Table 4. 5 : Privacy Risk in Sharing Location Information

	Yes	No
	Row N %	**Row N %**
Do you perceive any risk to privacy with sharing your location information	84.0%	**16.0%**

4.1.7 Risk perceived in sharing location information

The study showed that 63.6% of the respondents perceive risk associated with insecurity while 27.3% of the respondents don't trust their service providers because they think it lacks privacy and 9.1% of the respondents trust their location service providers. This implies that there is high privacy risk perceived by users while using location based services.

Table 4. 6 : Risk perceived in sharing location information

	Insecurity	**I trust my location service provider**	**Lack of privacy**
	Row N %	Row N %	**Row N %**
Explain any risk perceived	63.6%	9.1%	**27.3%**

4.1.8 Control over personal information

The study showed that majority 89.3% need control over their personal information while only a few 10.7% do not need control over their personal information. This implies that most people who use Location Based Services need control over their personal information for security purpose.

Table 4. 7 : Control over personal information

Table 4.3.0	Yes	No	
	Row N %	Row N %	
Do you need control over your personal information	89.3%		**10.7%**

	To feel secure	For privacy purpose	I trust the privacy provided	I don't know
	Row N %	Row N %	Row N %	**Row N %**
why do you need control over your personal information	42.3%	53.8%	.0%	**3.8%**

4.1.9 Adoption of Location Based Services Applications

Information was sought concerning various reasons that would make people adopt Location based Services applications. The study showed that 40.5% of the respondents agreed that they would adopt LBS applications if the application has a third party seal certifying and obeying CAK rules. This is a clear indication that most people would adopt LBS if Communication Authority of Kenya controls the application regarding privacy .45.9% agree they would adopt LBS if the industry is self regulated and has a governing body. 32% strongly agree that they would adopt LBS applications if they have control over how their information is collected and used, 38.9% agree that if the privacy control settings were robust and highly

configurable they would adopt LBS applications, This is a clear indication that if the privacy control settings are robust and highly configurable in an application people would adopt use of Location based services. 32.4% strongly agree that if the privacy protection literature is easily accessible when protected they would adopt LBS applications.

Table 4. 8 : Adoption of Location Based Services Applications

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly agree**
I would be more willing to adopt location based services applications if:					
The application has a third party seal certifying it obeys the rules set by CAK regarding LBS privacy	16.2%	.0%	18.9%	40.5%	**24.3%**
The industry is self regulated and has a governing body.	2.7%	2.7%	32.4%	45.9%	**16.2%**
I have control over how my information is collected and used.	13.5%	5.4%	24.3%	24.3%	**32.4%**
The privacy control settings were robust and highly configurable.	8.3%	2.8%	19.4%	38.9%	**30.6%**
Theprivacy protectionliteratureis easilyaccessibleand protected.	5.4%.	5.4%	27.0%	29.7%	**32.4%**

4.10 Tracing user movement

Information was sought concerning tracing user movement by LBS. The study showed that 80.6% feel that they can be traced and that collected information be used to interfere with their privacy while 19.4% feel it is safe and cannot be used against them. This implies that most users have a feeling that Location Bases Server cannot be trusted.

Table 4. 9 : Tracing user movement

	Yes	**No**
	Row N %	**Row N %**
Do you feel that location based server can keep traces of your movement and use that information to interfere with your privacy	80.6%	**19.4%**

4.11 Clear mobile phone cache memory

The study showed that it is evident that most of the people 15(40.5%) of the respondents never clear their memory cache while 9(24.3%) of the respondents clear their memory cache weekly and 12(32.4%) clear their memory cache daily. This implies that most people don't clear their phones cache memory, and that information can be used to interfere with their privacy.

Table 4. 10 : Clear mobile phone cache memory

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Daily	12	32.4	33.3	**33.3**
	Weekly	9	24.3	25.0	**58.3**
	Never	15	40.5	41.7	**100.0**
	Total	36	97.3	100.0	
Missing	System	1	2.7		
Total		37	100.0		

4.12 Privacy of future LBS

Information was sought concerning future of LBS, It is evident that 78.1% believe that privacy will be improved in the future while 6.3% are not sure about the future and 15.6% think it will more violated than before. This implies that if privacy improves in the future more people are likely to use it.

Table 4. 11 : Clear mobile phone cache memory

	it will be improved	Not sure	It will be more violated
	Row N %	Row N %	Row N %
What is your opinion regarding privacy of future LBS	78.1%	6.3%	15.6%

4.13 Sharing location information

Information was sought concerning sharing location information. The study showed that 90.6% of the respondents are comfortable sharing information with family but 9.4% of the same are not comfortable sharing with their families, 11.1% of the respondents are comfortable sharing information with anyone but 88.9% of the same

do not. 75.9% of the respondents are comfortable sharing with friends but 24.1% of the same do not. 60.7% are not comfortable sharing information with nobody but 39.3% are comfortable sharing with anybody. This implies that most people would share location information with family and friends.

Table 4. 12 : Sharing location Information

	Yes	**No**
	Row N %	**Row N %**
who would you share location information with- family	90.6%	9.4%
Friends	75.9%	**24.1%**
Anyone	11.1%	**88.9%**
Nobody	39.3%	**60.7%**

4.14 Improved privacy techniques encourages use of LBS

Information was sought concerning improving privacy through technology and the study showed that 91.7% of the respondents would be encouraged to use LBS while 8.3% would not be encouraged to use it. This implies that if privacy is guaranteed to the users they would be encouraged to use LBS.

Table 4. 13 : Improved privacy Techniques

	Yes	**No**
	Row N %	**Row N %**
if technology could guarantee the privacy of your location would this encourage you to use a Location Based Service	91.7%	**8.3%**

4.15 General awareness of the LBS prior to the survey

Information was sought on general awareness of LBS prior to the survey. The study showed that

78.4% of the respondents knew about LBS but 21.6% of the respondents did not have a general awareness of the LBS prior to the survey. This implies that most people were aware of the LBS services and the survey has also improved the general awareness of LBS.

Table 4. 14 : General Awareness of the LBS prior to the survey

	Yes	**No**
	Row N %	**Row N %**
Prior to this survey did you have a general awareness of LBS	78.4%	**21.6%**

4.16 Summary of Data Analysis

From the analysis results, it shows that most people perceive high risk while using location information. There was also a clear indication that if privacy control settings are robust and highly configurable in an application people would be more willing to use location based services. The results also indicated that if privacy techniques are improved most people would be encouraged to use LBS. There are other challenges experienced by LBS users like sharing location information, which has raised privacy concerns. Therefore, analysis results justified a need for a framework for enhancing privacy in LBS.

4.2 Framework for Location Privacy

The proposed framework consists of mobile clients, Anonymizer, LBS and a Spatio-Temporal database as shown in figure 4.1.

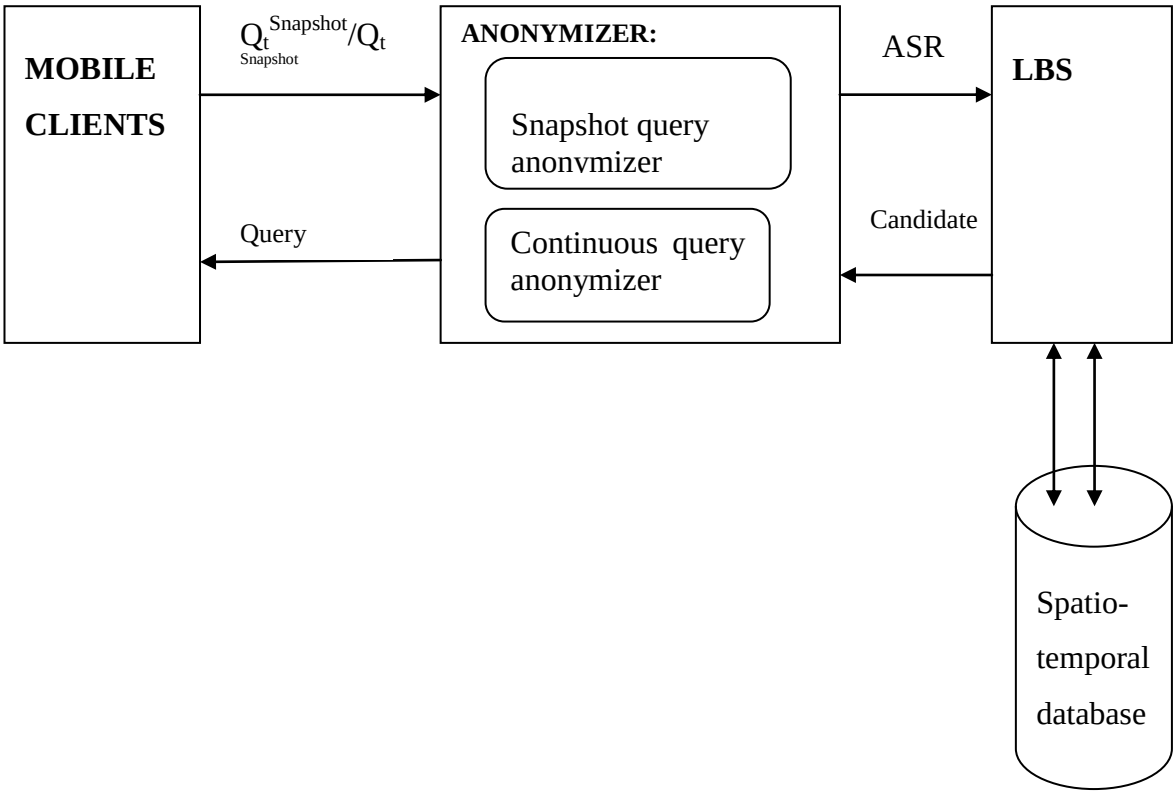

Figure 4. 1 : Framework for Location Privacy

4.2.1 Anonymizer

The anonymizer receives the location of all the mobile clients. The physical location, computed by the mobile device, is sent to the anonymizer server with the query. The anonymizer sends to the LBS an Anonymized Spatial Region (ASR) instead of the actual user location. This procedure is called cloaking. Cloaking hides the actual location by a K-anonymity spatial region (ASR), which is an area that encloses the client that issued the query as well as at least k-1 other users. The anonymizer consists of two components, the snapshot query anonymizer and the continuous query anonymizer.

Snapshot Query Anonymizer Location-based snapshot query ($Q_t^{Snapshot}$) is a query submitted at time t, for example "where is my nearest swimming pool". When a mobile user submits snapshot query incorporating positioning information such as current location (latitude and longitude) as a parameter of request to the AZ, the AZ then cloaks the client's query location point into a region containing K - 1 other mobile user request. The cloaking process for a snapshot is presented below.

First the anonymizer receives the current location (latitude and longitude) of the user together with the snapshot query. Second the cloaking algorithm takes the radius of the region (current user location in metres) and converts it into degrees. Third the cloaking algorithm creates a random generator and two random points are generated within the region. The new random points are the new latitude and the new longitude, which are then forwarded to the LBS.

Continuous Query Anonymizer Location-based continuous query ($Q_t^{continuous}$) is a sequence of continuous queries submitted at discrete time points, for example "Continuously report my nearest police station". When a mobile user submits a continuous query incorporating positioning information such as current location (latitude and longitude) at specified time interval to the AZ, the AZ then cloaks the client's query location point into a region containing K - 1 other mobile user request at time t_0, then the AZ schedules the next cloaking process at time t_1, where $t_0 < t_1$. The LBS continues receiving the cloaked regions until all the jobs in the scheduler are finished.

The cloaking algorithm that describes the cloaking process is shown in figure 4.3:

Step 1: Set regionRadius to current configuration

Define location object as;

Step 2: Get initial location latitude

Step 3: Get initial location longitude

Step 4: Convert regionRadius to degrees

// start of cloaking process

Step 5: Generate a random number

Step 6: Compute the product of (step 4 value) and the square root of (step 5 value)

Step 7: Compute diameter as pi*2*(step 5 value)

Step 8: Compute cosine of step 7 value

Step 9: Compute the product of step 8 value and step 6 value

Step 10: Compute the sine of step 7 value

Step 11: Compute the product of step 10 value and step 6 value

Step 12: Divide step 9 value by cosine of step 3 value

Step 13: Compute new latitude as Sum of step 12 value and step 2 value

Step 14: Compute new longitude as sum of step 11 value and step 3

Step 15: Return step 13 value as latitude and step 14 value as longitude of new location

// end of cloaking process

Figure 4. 2 : Cloaking Algorithm

4.2.2 Mobile Device (Clients)

Mobile device includes mobile phone, PDA, and other devices such as laptops with positioning capabilities. First, each mobile device computes its physical location from the GPS or Wi-Fi component on the device. Mobile users specify the Point of Interest they desire to search from the user interface of their mobile device in the proposed system. Mobile clients search for Point of Interest, it sends a query (search term) to the Anonymizer.

4.2.3 Location Based Services (LBS)

Provides access to location data sources for example the Google places. The LBS receives ASR, ignoring where exactly the user is, the LBS retrieves (and reports to the AZ) a Candidate Set (CS) that is guaranteed to contain the query results for any possible user location inside the ASR. The AZ receives the CS and reports to user the subset of candidates that corresponds to her original query.

4.2.4 Spatio-Temporal database

The spatial temporal database processes both the snapshot and continuous queries that it receives from the LBS. The spatial temporal database captures both spatial and temporal aspects of data. Spatio-temporal databases is very important because many real world applications like, location based services, Geographic information systems, etc need to store real world data which shows spatial as well as temporal characteristics, into database. Many data objects in real world have attributes related to both space and time, and managing them using existing Relational Database Management System (RDBMS) is complex and in-efficient, as these objects which show spatio-temporal behavior are multi-dimensional in nature. Spatio-temporal concepts, embodies both spatial and temporal concepts. Spatio-temporal Database concepts need to accommodate both spatial and temporal DB concepts while

implementing spatio-temporal Database. Spatial databases store the spatial attributes, which have space related properties usually the temporal extent is not present. Spatial databases offers, spatial data types, spatial data models and spatial query capabilities. Temporal databases represent attribute of objects which are changing with respect to time i.e. like functions with continuous range or functions which represent discrete values at different points in time. Temporal databases aims at capturing the temporal aspects of the systems in real world into DBMS. Temporal DBMS have a built-in time concept incorporated in it to handle queries related to varying time.

4.3 Framework Implementation

This section describes the design and implementation that is used to realize the framework.

4.3.1 Design

UML which is a graphical modeling language was used in the design of the system. Diagrams such as class diagrams and sequence diagrams were used to describe major elements.

Class Diagram The class diagram displays the static structure of the system, it shows how the various individuals (data, things, and people) link up to each other. The classes can be linked to each other like they can be dependent, combined, specialized or packed. Figure 4.4 shows the various classes for an anonymizer, they include the Application, the AnonUtility and LBS. There is an association relationship between all the classes, since they are communicating to each other. An Application can send one or more points of interest to the AnonUtility, and the AnonUtility can send one or more query results to the Application. The AnonUtility forward one ASR and the LBS returns one candidate set.

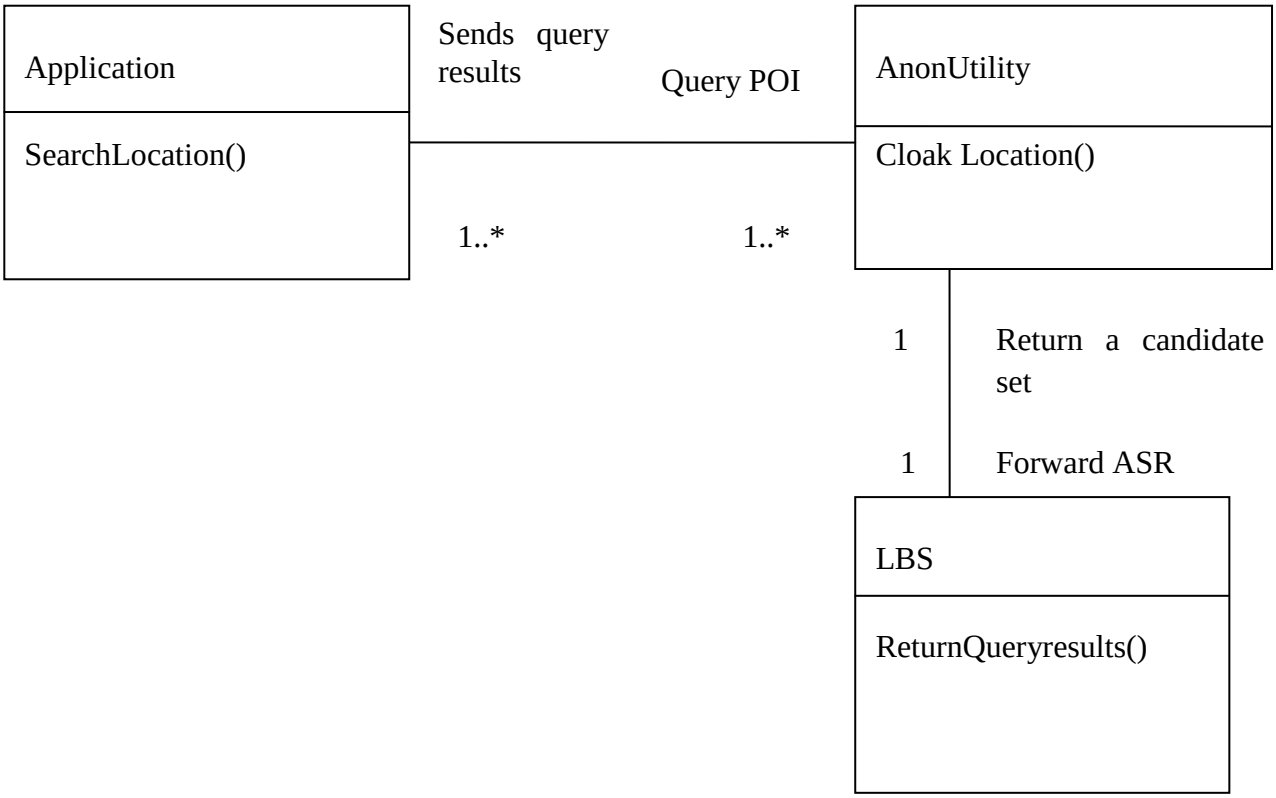

Figure 4. 3: Class Diagram

Sequence diagram shows an elaborated flow for a specific use case or even just part of a specific use case. . Tells how objects interact with each other i.e. how messages are being sent and receive between objects. A sequence diagram has two dimensions: The vertical axis shows time and the horizontal axis shows the objects.

Figure 4.4 shows the activities of a user querying for location based services. The connections between two objects are shown by an arrow along with communication messages. The vertical line shows the lifeline of the object. The user searches for a Point of Interest through an application interface, the anonymizer cloaks the location and forward the Anonymized Spatial Region to the Location Based Server. On receiving the ASR, the LBS forwards a set of candidates to the user containing the results of the query.

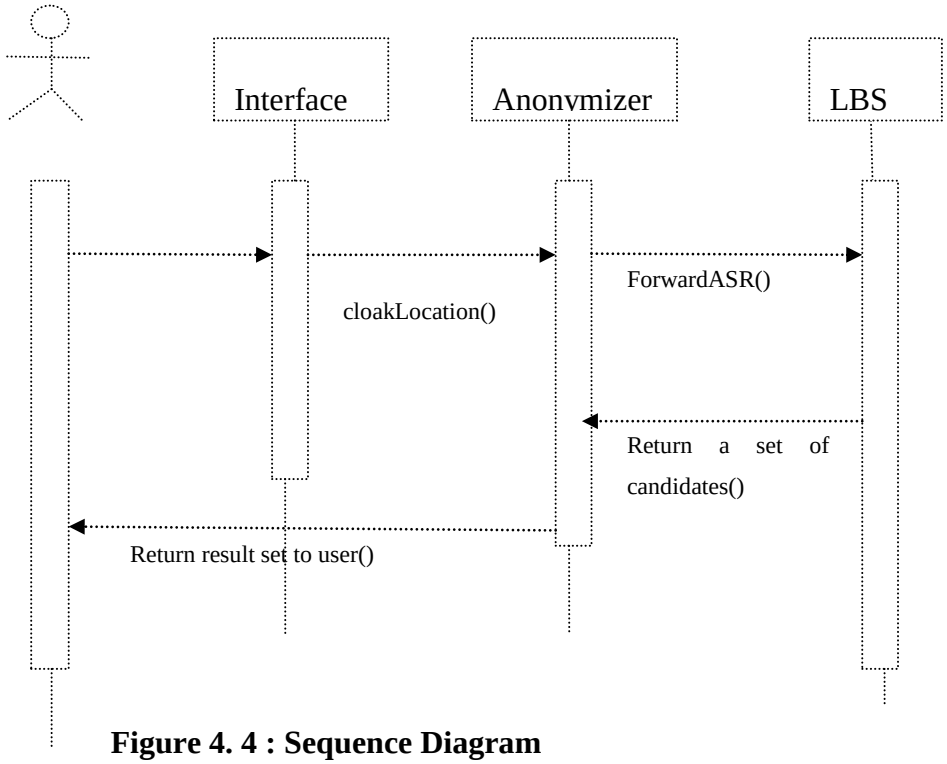

Figure 4. 4 : Sequence Diagram

4.3.2 Anonymizer Architecture

The anonymizer architecture uses the k-anonymity concept; the mobile clients communicate with third-party LBS providers through the anonymity server. The anonymity server is a secure gateway to the LBS providers for the mobile clients. Each message sent to an LBS provider contains the current location information of the mobile client. Upon receiving a message from a mobile client, the anonymity server hides the location information through spatio-temporal cloaking, and then forwards the anonymized message to the LBS provider. Spatial cloaking refers to replacing current location by a spatial range, and generating a random point.

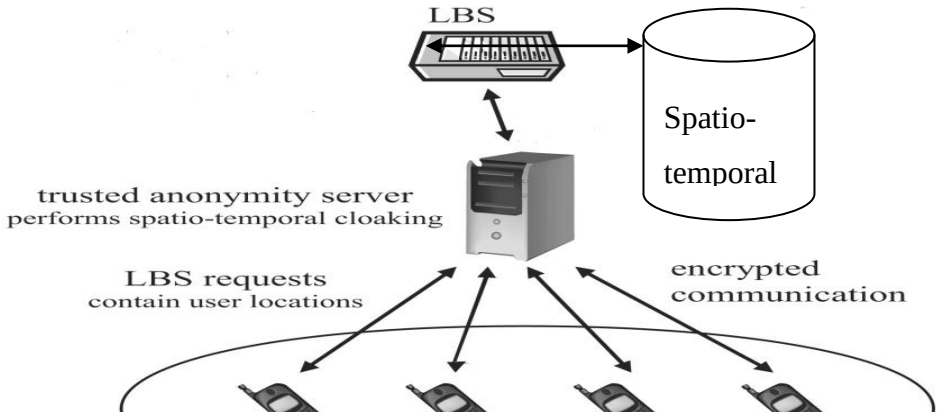

Figure 4. 5 : Anonymizer Architecture

4.3.3 Anonymizer System Implementation

The system to perform the anonymization was implemented using scala (JVM Based). The anonymizer server was Implemented using scalar (JVM Based) and was hosted in a cloud server that ensured the availability of the service. The cloud server was running on Linux and to connect to the server an encrypted communication was used.

4.3.4 Mobile Application system Implementation

In order to test the framework, this thesis developed a mobile application that mobile users uses to query for Point Of Interest (POI). The system to query for POI was implemented using Android developer, and it was installed on mobile phones and was used by mobile phone users to search for POI. Figure 4.6 shows the main search interface for the mobile android application. The user enters the search term in space provided as shown below. The search term is the name of any Point Of Interest for example hotel, restaurant swimming pool, shop, hospital etc. After typing the search term the mobile user then touches on search button to start searching.

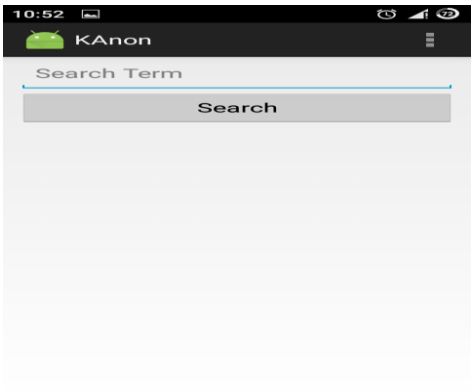

Figure 4. 6 : Mobile Application Search Interface

Search results for a Cyber- Figure 4.7 shows search result set for a cyber within Juja in Kenya. The latitude, longitude and vicinity of the location are returned. The returned results show the location (latitude and longitude) of different cyber cafe in Juja in Kenya. The result set shows that the returned results are at different location points meaning they are within a cloaked region. This is shown by different values of latitude and longitude. The mobile user gets information of the cyber he is interested in from the result set.

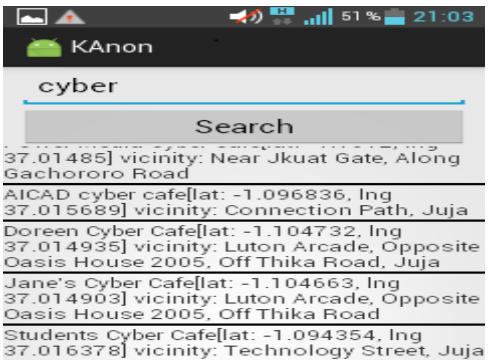

Figure 4. 7 : Search results for a cyber

Search results for a Bank- Figure 4.8 shows search result set for a Bank within Juja in Kenya. The latitude, longitude and vicinity of the location are returned. The returned results show the location (latitude and longitude) of different Banks in Juja in Kenya. The result set shows that the returned results are at different location points

meaning they are within a cloaked region. This is shown by different values of latitude and longitude. The mobile user gets information of the bank he is interested in from the result set.

Figure 4. 8 : Search results for a bank

Search results for a Saloon- Figure 4.9 shows search result set for a salon within Kenyatta market Nairobi Kenya. The latitude, longitude and vicinity of the location are returned. The returned results show the location (latitude and longitude) of different saloons. The result set shows that the returned results are at different location points meaning they are within a cloaked region. This is shown by different values of latitude and longitude. The mobile user gets information of the saloon he is interested in from the result set.

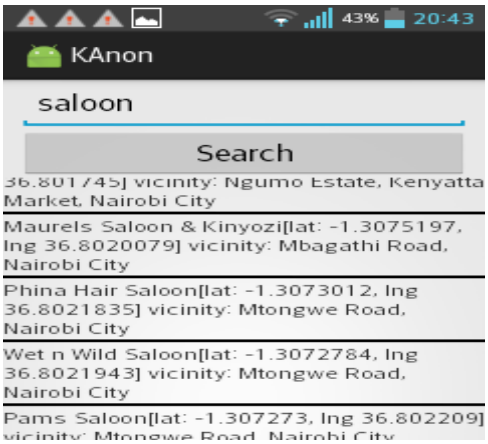

Figure 4. 9 : Search results for a saloon

Search results for a Swimming Pool- Figure 4.10 shows search result set for a swimming pool within Kenyatta market Nairobi Kenya. The latitude, longitude and vicinity of the location are returned. The returned results show the location (latitude and longitude) of different swimming pools. The result set shows that the returned results are at different location points meaning they are within a cloaked region. This is shown by different values of latitude and longitude. The mobile user gets information of the swimming pool he is interested in from the result set.

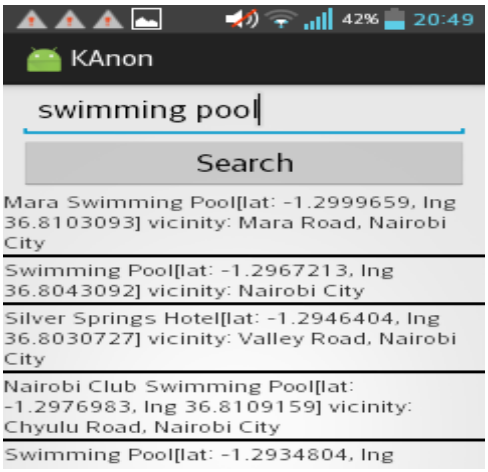

Figure 4. 10 : Search results of swimming pool

4.3.5 Activity log at anonymizer

Figure 4.11shows log activity on how the anonymizer received the query request from mobile phone, the cloaking process, forwarded the ASR to the LBS and finally returning result set to the user. Information on anonymizer was logged to display cloaking process. From the logged information it shows that the anonymizer is a trusted server because it does not forward the real user location to the location server but it forwards an anonymized spatial region.

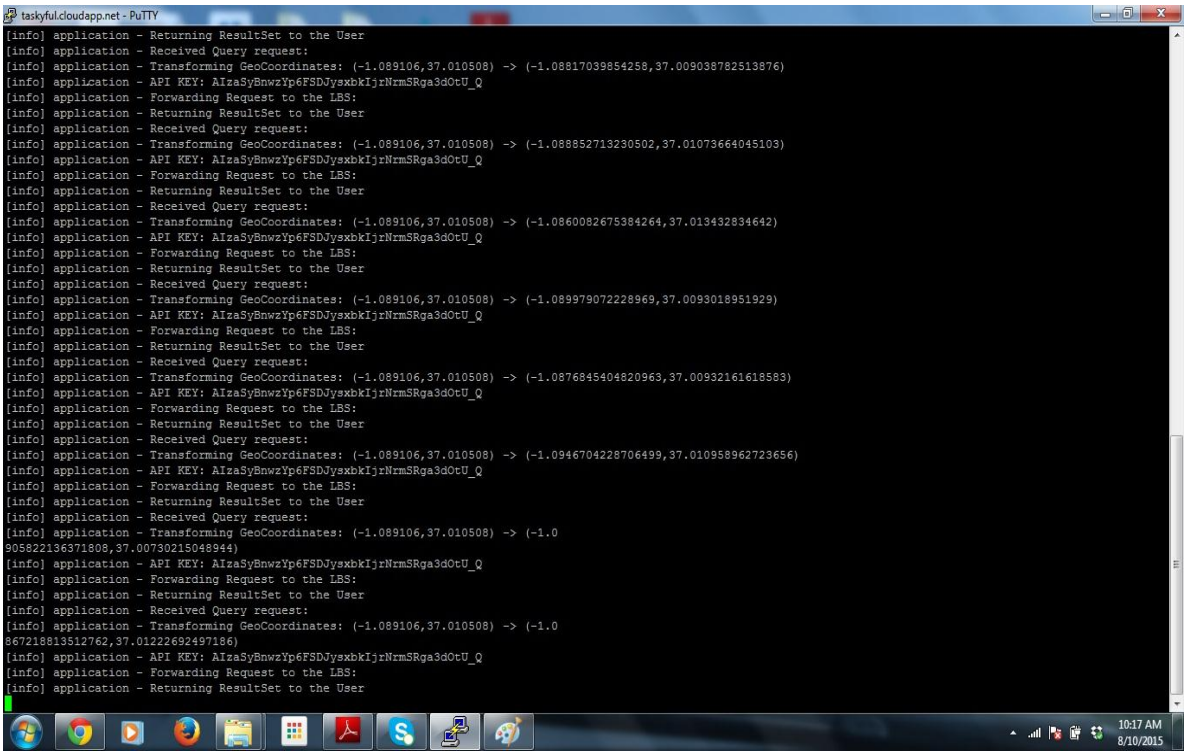

Figure 4. 11 : Activity log at anonymizer

4.4 Evaluation of Framework

To evaluate the framework an experiment was conducted on a windows machine running Pentium(R) duo-core CPU 2.10 GHz processor with 4GB of RAM. Jmeter software was installed and used for simulation by sending 100 http requests to the anonymizer server. Simulation was done starting with 10 http requests adding 10 http requests after every request and observing the performance until 100 http requests were sent. The statistics were returned that showed the performance through graphs and tables as shown below.

4.4.1Experimental results for Success Rate

The experimental results for success rate are shown through the Jmeter summary report that was generated when 100 mobile request were simulated to the server. In

most of the request that was made the error % was zero, because most of the requests ran successfully as shown in the table 4.15.

Table 4. 15 : Generated Summary report of 100 Http requests

NO OF HTTP REQUEST	AVERAGE RESPONSE TIME	MAX TIME	MIN TIME	ERROR %
10	3498MS	6749MS	956MS	0.00%
20	1246MS	3580MS	563MS	0.00%
30	6136MS	11250MS	564MS	0.00%
40	4621MS	52773MS	563 MS	0.00%
50	8877MS	113129MS	564MS	0.00%
60	12704MS	463022MS	564MS	0.00%
70	6313MS	140468MS	581MS	1.43%
80	8424MS	122276MS	579MS	0.00%
90	6684MS	119033MS	572MS	1.11%
100	6232	57962MS	568MS	2.00%

Average: Average is the average response time for that particular http request. This response time is in milliseconds.

Min: Min denotes to the minimum response time taken by the http request.

Max: Max denotes to the maximum response time taken by the http request.

Error %: This denotes the error percentage in samples during run. Most of the mobile request generated 0.00% error, this implies that most of the request ran successfully. Error % shows the percentage of failed requests per 100 requests made,

this means the said percentage of requests have not been able to complete successfully.

4.4.2 Experimental results for Performance Measure

The experimental results for success rate are shown through the response time graph that was generated after simulating request to the server. The run time performance of algorithms is measured as the cloaking time. An algorithm with a lower cloaking time does better as shown by (2).

Cloakingtime=endTime(CloakM(ms))–startTime(CloakM(ms).

The graph below shows the response time when 100 http requests were sent to the anonymizer server.

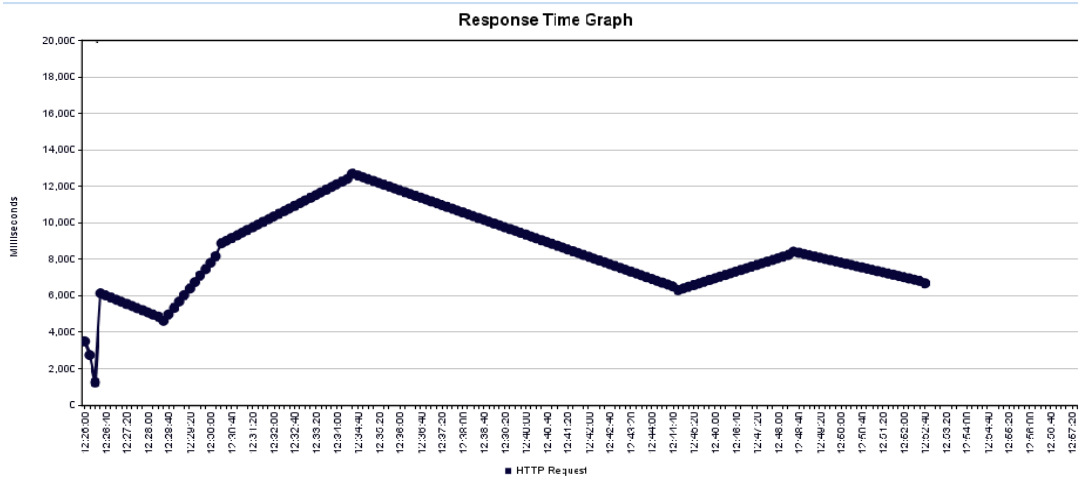

Figure 4. 12 : Response Time Graph

Cloaking time of 100 http requests was computed as: Cloaking time = endTime (CloakM (ms)) – startTime (CloakM (ms) as shown by (2).

endTime (CloakM(ms))=6500ms

startTime (CloakM(ms)=3950ms

Cloaking time=6100ms-3900ms Cloaking time=2200ms

1000 milliseconds make up one second.

2200ms/1000ms=2.2s.

It took 2 seconds to cloak 100 requests.

4.5 Discussion of Results

After evaluating the performance measure of our proposed framework by simulating a 100 http request to the server, it indicates that our proposed framework takes 2 seconds to cloak 100 http requests. This was observed in figure 4.12 in the response time graph. An algorithm with lower cloaking time does better It therefore indicates that the proposed framework has a good run time performance. After evaluating the success rate of our proposed framework by simulating a 100 http requests it indicates that the success rate of most of the request sent to the server is 100%. In most of the request that was made the error % was zero, because most of the requests ran successfully as shown in table 4.15. Therefore the experimental results show that our proposed framework is effective and efficient.

This research compared the experimental results of the proposed framework against the new Casper framework a pyramid based approach (Mokbel et al., 2006). This is a framework in which mobile and stationary users can entertain location-based services without revealing their location information. Casper consists of two main components the location anonymizer and the privacy-aware query processor. The location anonymizer blurs the users exact location information into cloaked spatial regions based on user-specified privacy requirements. The privacy-aware query processor is integrated inside the location-based database server in order to deal with the cloaked spatial areas rather than the exact location information. Casper maintains

an anonymizer and a privacy aware query processor. For a successful anonymization and cloaking, a pyramid structure is maintained. The cells in the region contain the number of mobile users active in the cell. If the current region/cell of user cannot satisfy the value of K or Amin then a neighboring cells are considered. Casper has a superb run time performance, but it does not return the smallest cloaking region and also it is expensive (updates and cloaking) to maintain the pyramid structure. Casper employs a grid-based complete pyramid data structure that hierarchically decomposes the spatial space into H levels where a level of height h has 4h grid cells. The root of the pyramid is of height zero and has only one grid cell that covers the whole space. Each pyramid cell is represented as (cid, N) where cid is the cell identifier while N is the number of mobile users within the cell boundaries. The pyramid structure is dynamically maintained to keep track of the current number of mobile users within each cell. Thus updating and cloaking is very expensive due to maintaining the pyramid structure. From the previous research conducted (Mokbel 2007) it showed that the New Casper dropped over 98% of the total mobile requests sent to the anonymization server. After evaluating the success rate of our proposed framework by simulating a 100 http requests it indicates that the success rate of most of the request sent to the server is 100%. In most of the request that was made the error % was zero, because most of the requests ran successfully as shown in table 4.15.

After experimental evaluation it has shown that our proposed framework is efficient and effective. It shows that it has a good run time performance and a high success rate. From the previous research conducted it showed that the New Casper dropped most of the total mobile requests sent to the anonymization server whereas in our proposed framework experimental results indicates that the success rate of most of the request sent to the server is 100% as shown in table 4.15.

CHAPTER FIVE

CONCLUSIONS, RE COMMENDATIONS AND FURTHER RESEARCH

5.1 Conclusions

Recently, location-based services have become very popular, mainly driven by the availability of modern mobile devices with integrated position sensors. Prominent examples are points of interest finders or geo-social networks such as Facebook Places, google places and google maps. However, providing such services with private user positions may raise serious privacy concerns if these positions are not protected adequately. It was for this reason that this thesis sort to carry out a research project to:

1. Identify the challenges in protecting privacy in LBS.
2. Investigate various models and techniques that have been used in protecting privacy in location based services.
3. Formulate a framework that protects the users from trajectory privacy attack.
4. Evaluate the framework for efficiency and effectiveness.

We concluded by highlighting the achievement of each objective based on our research findings.

The first objective was achieved by exploring the various models and techniques that have been used in protecting privacy in location based services, through literature review, where the researcher looked at strength and weaknesses of each model and technique. Chapter two gives detailed information on literature review.

By studying related work on various models and techniques their strength and weaknesses it was possible for the researcher to identify existing gaps and strengths on protecting user privacy in location based services. The gaps found in the

literature, provided guidance in development of the framework. The data analysis results justified the need to develop a framework for enhancing privacy in LBS.. Chapter 4 provides information obtained from the survey.

The framework was integrated into a mobile application system and an anonymizer system to ensure effective implementation. For the purpose of investigating if the proposed framework was effective and efficient, an experiment was conducted and the experimental results showed that the framework has a good success rate and a good run time perfomance.

5.2 Recommendations

The findings described in chapter four is evident that a framework is for enhancing privacy in Location Based Services will encourage many people to use the services Most people have not been able to use Location Based Services due to threat in privacy. The framework is a very resourceful tool and can be used by mobile phone users to search for Point Of Interest. The approach discussed in this paper performs much better than most of the techniques and models. For example the new casper that employs the pyramid structure is dynamically maintained to keep track of the current number of mobile users within each cell. Thus updating and cloaking is very expensive due to maintaining the pyramid structure. From previous studies pyramid based Casper dropped over 98% of the total mobile requests sent to the anonymization server (Mokbel et al., 2006). From the experiment conducted on the proposed framework it showed that 80% of the total mobile requests sent to the anonymization server were anonymized successfully.

The proposed framework is used by both snapshot and continuous queries thus it's able to protect the user from correlation attack. The anonymizer is a performance bottleneck and therefore it should be replicated and deployed behind a reverse proxy server to avoid a single point of failure.

5.3 Further Research

Future research can address more attack scenarios, such as attacks based on user preferences. Assume that each user is interested in certain types of queries, e.g., cyber, restaurants, bar etc. An attacker may use the additional knowledge to infer the query source. To prevent this, users can be classified into groups according to their interests. Then, spatial diversity would take into account these groups when forming ASRs; i.e., an ASR should contain users with similar interests, from the same group. The framework can also be integrated in target content dissemination application e.g. advertisement .

REFERENCES

Ardagna, A., Cremonini, M., Capitani di Vimercati, S. D., & Samarati, P. (2007). Privacy enhanced location-based access control. In M. Gertz & S. Jajodia (Ed.), *The Handbook of Database Security: Applications and Trends* (pp. 531-552). Springer-Verlag.

Babbie, E. (1995.)*The Practice of Social Research* (7th ed.). Belmont, CA: Wadsworth.

Bamford, W., Coulton, P., & Edwards, R. (2006). Location-bsed Mobile Blogging. In *Information and Communication Technologies, 2006. ICTTA'06. 2nd* (Vol. 1, pp. 111-116). IEEE.

Bellavista, P., Kupper, A., & Helal. (2008). Location-Based Services: Back to the Future. *IEEE Pervasive Computing , 7* (2), 85-89

Beresford, A. R., & Stajano, F. (2003). Location privacy in pervasive computing. *IEEE Pervasive computing*, (1), 46-55.

Bettini, C., Jajodia, S., Wang, X.S. (2000). *Time Granularities in Databases, Data Mining, and Temporal Reasoning,* Heidelburg. Berlin: Springer

Bettini, C., Wang, X. S., & Jajodia, S. (2005). Protecting privacy against location-based personal identification. In *Secure data management* (pp. 185-199). Springer Berlin Heidelberg.

Brinkhoff, T. (2002). A framework for generating network-based moving objects. *GeoInformatica*, *6*(2), 153-180.

Chapelle, O. & Li, L. (2011) An empirical evaluation of Thompson sampling. In NIPS.

Cheng, R., Zhang, Y., Bertino, E., & Prabhakar, S. (2006, June). Preserving user location privacy in mobile data management infrastructures. In *Privacy Enhancing Technologies* (pp. 393-412). Heidelburg. Berlin: Springer

Cheng, R., Zhang, Y., Bertino, E., Prabhakar, S.: Querying private data in moving object

Chor, B., Kushilevitz, B., Goldreich, O. & M. Sudan. Private information retrieval. ACM, 45(6):965–981, 1998.

Chow, C. Y., Mokbel, M. F., & Liu, X. (2006, November). A peer-to-peer spatial cloaking algorithm for anonymous location-based service. In *Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems* (pp. 171-178). ACM

Dhar, S., & Varshney, U. (201 1, May). Challenges and Business Models for mobile Location-Based Services and Advertising. *Communications of the ACM, 54*(5), pp. 121-129.

Duckham, M., Mokbel, M., & Nittel, S. (2007). Special Issue on Privacy Aware and Location-based Mobile Services. *Journal of Location Based Services, 1*(3), 161-164.

Edwards, R., & Gillies, V. (2006b) A qualitative analysis of parenting and social capital: comparing the work of Coleman and Bourdieu‘, *Qualitative Sociology Review*, II(2), published on-line: www.qualitativesociologyreview.org.

Cheng, R., Zhang, Y., Bertino, E., & Prabhakar, S. (2006, June). Preserving user location privacy in mobile data management infrastructures. In *Privacy Enhancing Technologies* (pp. 393-412). Springer Berlin Heidelberg.

Fleisch, E., Weinberger, M., & Wortmann, F. (2015). Business models and the internet of things. In *Interoperability and Open-Source Solutions for the Internet of Things* (pp. 6-10). Springer International Publishing.

Fraenkel, R. & Wallen, E. (2009). *How to Design and Evaluate Research in Education*. New York. McGraw-Hill Companies.

Gahi, Y., Guennoun, M., Guennoun., .Z & El-Khatib, K., (2012). Privacy Preserving Scheme for Location-Based Services,*Journal of Information Security*, Volume 3, 105-112.

Gambs, S., Marc-Olivier, K., & Miguel Nunez Del Prado Cortez (2013). De-anonymization attack on geolocated datasets. In 12th IEEE Inter. Conf. on Trust, Security and Privacy in Computing and Communications, pages 789–797.

Gay,L.(1981).Educational research competencies for analysis and application(2nd ed.), Columbus, Ohio: Char!esE. Merrill.

Gedik, B., & Liu, L. (2008). Protecting location privacy with personalized k-anonymity: Architecture and algorithms. *Mobile Computing, IEEE Transactions on, 7*(1), 1-18.

Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C., & Tan, K. L. (2008, June). Private queries in location based services: anonymizers are not necessary. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 121-132). ACM.

Gruteser, M., & Grunwald, D. (2003, May). Anonymous usage of location-based services through spatial and temporal cloaking. In *Proceedings of the 1st international conference on Mobile systems, applications and services* (pp. 31-42). ACM

Hjaltason ,G. R., & Samet, H. (1999)Distance browsing in spatial databases.*TODS*, 24(2):265–318.

Hjaltason, G. R., & Samet, H. (1999). Distance browsing in spatial databases. *ACM Transactions on Database Systems (TODS)*, *24*(2), 265-318.

Hu, H., Xu, J., & Lee, D. L. (2005, June). A generic framework for monitoring continuous spatial queries over moving objects. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data* (pp. 479-490). ACM.

Indyk, P., & Woodruff, D. (2006). Polylogarithmic private approximations and efficient matching. In *Theory of Cryptography* (pp. 245-264). Springer Berlin Heidelberg.

Kalnis, K., Ghinita, K., Mouratidis, K., & Papadias, D., (2006.) Preserving Anonymity in LocatioBasedServices. Technical Report TRB6/06, National University of Singapore.

Kathuri, N.J. & Pals, A.D. (1993). Introduction to Education Research. Education Media Centre, Egerton University.

Khoshgozaran, A., & Shahabi, C. (2007). Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy. In *Advances in Spatial and Temporal Databases* (pp. 239-257). Springer Berlin Heidelberg.

Khoshgozaran, A., Shahabi, C., & Shirani-Mehr, H. (2011). Location privacy: going beyond K-anonymity, cloaking and anonymizers. *Knowledge and Information Systems*, *26*(3), 435-465.

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

Kido, H., Yanagisawa, Y., & Satoh.,T. (July 2005). An anonymous communication technique using dummies for location-based services. Proceedings of the 2nd International Conference on Pervasive Services (ICPS '05); Santorini, Greece, pp. 88–97.

Kothari, C.R. (1984).*Quantitative Techniques*, 2nd ed., New Delhi: Vikas Publishing House Pvt. Ltd.

Kushilevitz, E. &Ostrovsky,R. (1997)Replication is not needed: single database, computationally private information retrieval. In FOCS ’97: Proceedings of the

Mokbel, M. F., Chow, C. Y., & Aref, W. G. (2006, September). The new Casper: query processing for location services without compromising privacy. In *Proceedings of the 32nd international conference on Very large data bases* (pp. 763-774). VLDB Endowment.

Morgan, K., Xu, H., Gupta, S., & Shi, P. (2009). Balancing User Privacy Concerns in the Adoption of Location-Based Services: An Empirical Analysis across Pull-Based and Push- Based Applications. iConference (iSociety: Research, Education, and Engagement). University of North Carolina-Chapel Hill.

Mouratidis, K., & YIU, M. (2012) Shortest Path Computation with No Information Leakage.. *Proceedings of the VLDB Endowment: 38th VLDB 2012, August 27-31, Istanbul, Turkey*. , 5 , 692. Research Collection School Of Information Systems.

Olumofin, F., Tysowski, P. K., Goldberg, I., & Hengartner, U. (2010, July). Achieving efficient query privacy for location based services. In *Privacy Enhancing Technologies* (pp. 93-110). Springer Berlin Heidelberg.

Palanisamy, B., & Liu, L. (2011, April). MobiMix: Protecting location privacy with mix-zones over road networks. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on* (pp. 494-505). IEEE.

Pingley, A., Zhang, N., Fu, X., Choi, H. A., Subramaniam, S., & Zhao, W. (2011, April). Protection of query privacy for continuous location based services. In *INFOCOM, 2011 Proceedings IEEE* (pp. 1710-1718). IEEE.

Rainer, R. K., & Cegielski, C. G. (2012). *Introduction to Information Systems* (4e ed.) New York: Wiley.

di Vimercati, S. D. C. (2011). Access Control Policies, Models, and Mechanisms. In *Encyclopedia of Cryptography and Security* (pp. 13-14). Springer US.

Schiller, J., & Voisard A., (2004.). *Location based services* (LBSs), Morgan Kaufmann, London, UK.

Singh, S. R. (2013). *Essays on Corruptible Markets, Strategic Certification and Online Peer Effects* (Doctoral dissertation, University of California, Berkeley).

Stenneth, L., & Yu, P. S. (2012). Mobile Systems Privacy:'MobiPriv'A Robust System for Snapshot or Continuous Querying Location Based Mobile Systems. *Transactions on Data Privacy*, *5*(1), 333-376.

Stenneth, L., Wolfson, O., Yu, P. S., & Xu, B. (2011, November). Transportation mode detection using mobile phones and GIS information. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 54-63). ACM.

Xu, H., Teo, H. H., Tan, B. C., & Agarwal, R. (2009). The role of push-pull technology in privacy calculus: the case of location-based services. *Journal of Management Information Systems*, *26*(3), 135-174.

Xu, J., Tang, X., Hu, H., & Du, J. (2010). Privacy-conscious location-based queries in mobile environments. *Parallel and Distributed Systems, IEEE Transactions on*, *21*(3), 313-326.

Yin, R. K. (2003). Case study research: Design and methods (3rd ed.). Thousand Oaks, CA: Sage.

Yiu, M. L., Jensen, C. S., Huang, X., & Lu, H. (2008, April). Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services. In *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on* (pp. 366-375). IEEE.

Zhong, S., Li, L., Liu, Y. G., & Yang, Y. R. (2004). Privacy-preserving location-based services for mobile users in wireless networks. *Department of Computer Science, Yale University, Technical Report ALEU/DCS/TR-1297*.

APPENDIX B: QUESTIONNAIRE

QUESTIONNAIRE INSTRUMENT

FOR:

FOR FRAMEWORK FOR ENHANCING PRIVACY IN LOCATION BASED SERVICES.

SERIAL No…...…

Date of interview DD ☐☐ **MM** ☐☐ **YY** ☐☐

This research is purely academic information collected through this research will be confidential and will solely be used for that purpose. I wish to communicate information about the survey to you. Should you be interested, please indicate your email address on the first page of the questionnaire. Please take a moment of your time to answer the survey questions. I will appreciate your frank and critical response to this questionnaire.

INSTRUCTIONS

1. This Questionnaire consists of **one Section**. Please answer all questions in this section
2. Do not indicate you **Name** on the questionnaire
3. Make sure that you tick within the box.

PRIVACY CONCERN FOR LOCATION BASED SERVICES USERS

1) What are the types of location based services/applications that you use?

a) Google map ☐

b) MapQuest ☐

c) Yelp ☐

d) Face book check in ☐

e) Google ignore ☐

f) Wechat

2) Do you trust location based service provider?

☐ Yes ☐ No

Why or Why not_____

3) How valuable do you think location-based services/apps are?

☐ Poor ☐ Moderate ☐ Good ☐ Excellent

4) On average how many times per week do you use location-based services/applications?

<1

1

2

3

4

5

>5

I do not use location based services

5) Do you perceive any level of risk to privacy with sharing your location information with a mobile application provider? Explain.

_____

_____

_____

6) Do you feel you need more choice and control over your personal information while using location based services and why?

_____

_____

_____

7) To what extent are you concerned about 3rd party companies having access to the location of your mobile phone/Smartphone without your permission? How familiar are you with the possibility that an application might collect personal information without your knowledge? Please provide your level of agreement to the following statement. Tick your level of agreement in the space provided in the table.

Statement	**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
I would be more willing to adopt location based services applications					
If the application has a third party seal certifying it obeys the rules set by CCK regarding LBS privacy.					
The industry is self regulated and has a governing body.					
I have control over how my information is collected and used.					

The privacy control settings were robust and highly configurable.					
The application made privacy protection literature easily accessible when needed.					

8) If some or all of the above (QN 15) privacy protection buffers were in place or implemented would you adopt the use of LBS?

☐ Yes ☐ No ☐ May be

9) What features of the LBS do you find could be improved?

_____

_____

10) Perceived usefulness: Location based services have been important in finding Point of Interest (POI).

☐ strongly disagree ☐ disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

11) Do you feel that location based server can keep traces of your movement and use that information to interfere with your privacy?

☐ Yes ☐ No

12) How often do you clear your mobile phone/Smartphone cache or memory?

Daily ☐

☐

Weekly ☐

Never

13) What is your opinion regarding privacy of future LBS?

_____

___

14) Is privacy an issue you would consider when using a Location Based Service?

☐ Yes ☐ No

15) Who would you share location information with?

Family	Yes	☐	No	☐
Friends	Yes	☐	No	☐
Anyone	Yes	☐	No	☐
Nobody	Yes	☐	No	☐

16) If technology could guarantee the privacy of your location, would this encourage you to use a Location Based Service?

☐ Yes ☐ No

17) Would you use LBS if there was an application that consolidates the services?

☐ Yes ☐ No

Why or Why not _____

18) Prior to this survey did you have a general awareness of LBS?

☐ Yes ☐ No

APPENDIX C: ANONYMIZER CODE IMPLEMENTATION

```
 object AnonUtility  {

valregionRadius=current.configuration.getDouble("anon.regionradius").get

def cloakLocation(location: Location): Location={

val x0=location.lat

val y0=location.lng

val radiusInDegree=regionRadius/111000f

val r= new Random()

val u=r.nextDouble

val v=r.nextDouble


val w= radiusInDegree*Math.sqrt(u)

val t= 2 * Math.PI * v

val x = w * Math.cos(t)

val y= w * Math.sin(t)


val new_x= x/Math.cos(y0)
```

```
val selectedX=new_x + x0

val selectedY=y +y0


Location(selectedX,selectedY)
```

APPENDIX D: MOBILE APPLICATION CODE IMPLEMENTATION

```
class MainActivity extends Activity with ConnectionCallbacks with
OnConnectionFailedListener{


  var dialog:Option[ProgressDialog]=None

  var adapter:Option[ArrayAdapter[Place]]=None

  var mGoogleApiClient:Option[GoogleApiClient]=None

  var location: String="-1.089106,37.010508"

  var lat:Double= -1.089106

  var lng:Double= 37.010508

  var radius=5000.toString

override def onCreate(savedInstanceState: Bundle): Unit = {

    super.onCreate(savedInstanceState)

    setContentView(R.layout.activity_result)

   val btnSearch=findViewById(R.id.btn_search).asInstanceOf[Button]

    val txtTerm=findViewById(R.id.txt_term).asInstanceOf[EditText]

    val listView=findViewById(R.id.lst_results).asInstanceOf[ListView]
```

```
  val arrAdapter=new ArrayAdapter[Place](this,android.R.layout.simple_list_item_1,new util.ArrayList[Place]()){

    override def getView(position: Int, convertView: View, parent: ViewGroup): View = {

      var row:View=convertView

      if(row==null){

        val inflater=getContext.getSystemService(Context.LAYOUT_INFLATER_SERVICE).asInstanceOf[LayoutInflater]

        row=inflater.inflate(R.layout.item_template,null)



      }



      val place=getItem(position)

      val name=place.name

      val lat=place.geometry.location.toString

      val strPlace=s"$name[$lat]"

      val lbl=row.findViewById(R.id.txt_place).asInstanceOf[TextView]

      lbl.setText(strPlace)

      row
```

```
  }
}
listView.setAdapter(arrAdapter)

val queue= Volley.newRequestQueue(this)
buildGoogleApiClient
btnSearch.setOnClickListener(new OnClickListener {
  override def onClick(v: View): Unit = {
    val
uri=Uri.parse(appConfig.baseUrl+"/search").buildUpon().appendQueryParameter("la
t",lat.toString).appendQueryParameter("lng",lng.toString).appendQueryParameter("r
adius",radius).appendQueryParameter("term",txtTerm.getText.toString).appendQuer
yParameter("types","all").toString
    val success=new Response.Listener[String]{
      override def onResponse(json: String)={
        dialog.get.dismiss()
        Log.e("Search Results",json)
        val results=new JSONObject(json).getString("results")
        val tType=new TypeToken[java.util.List[Place]](){}.getType
        val gson=new GsonBuilder().create()
```

```
            val result:java.util.List[Place]=gson.fromJson(results,tType)


            Toast.makeText(getApplicationContext,result.size+"
matches",Toast.LENGTH_SHORT).show()


           arrAdapter.clear();

           arrAdapter.addAll(result)

            arrAdapter.notifyDataSetChanged()



          }

       // Continous query implementation

ef continoussearch(lat: Double,lng: Double, radius: String,term: String, types: String,
mins: Int){

     Logger.info("Received Query request: ")


      val timeStep=mins*60*1000

      val loc=Location(lat,lng)

      val conRequest=ContinousRequest(term,l continousActoroc)
```

```
    val continousActor = system.actorOf(Props(classOf[ContinousSearchActor], this))
    system.scheduler.schedule(
      0 milliseconds,
      timeStep milliseconds,
    continousActor,
    conRequest)
  }
 class ContinousSearchActor extends Actor{

   def receive={

    case ContinousRequest(term,loc)=> {
```