2014-04-22

Applying Data Mining and Social Network Analysis Techniques in a Land Tenure Information System

Asiedu, Kwame

UNIVERSITY OF CALGARY

Applying Data Mining and Social Network Analysis Techniques in a Land Tenure Information System

by

Kwame Asiedu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

APRIL, 2014

Abstract

Network of social relationships are fundamental attribute of land tenure systems. The primary objective of this research is to develop an evolutionary TTM-based LTIS to cater for requirements dictated by the rapidly evolving social, political and economic conditions in peri-urban communities and to apply data mining and social network analysis techniques to find lineage of interest in customary systems. The method adapts agile software development approach to developing a TTM-based LTIS. Simulated data populated into the system is then extracted for mining and analysis. The results have proven that mining and analysis can facilitate visualizing complex relationships between objects, reveal hidden relationships that emerge as social change take shape as well as drive the development of an evolutionary land administration model and the LTIS developed based on such models. However the success of these methods depends on acquiring relevant data which may sometimes be manipulated to suite one's agenda.

Acknowledgements

The Almighty God Jehovah has supported me in all my endeavors including the completion of this thesis. All thanks to Him.

I would like to express my gratitude to my supervisor, Dr. Michael Barry, for giving me the opportunity to join his research team, for sharing his precious time, knowledge, wisdom, and expertise with me, for his consistent guidance, encouragement and patience for my lack of experience, and for all that he taught me. He is like a father who is concerned about my academics, health and career plans. I could not ask for a better supervisor. I would also like to thank Dr. Andrew Hunter for keeping his door open to answer my questions, and for his help and guidance throughout this research. My sincere gratitude to Dr. Reda Alhajj for his help and support especially in relation to the application of data mining and network analysis in this research.

I am deeply indebted to all the great people who always offered their help and support in proofreading the thesis. Their help and critical comments greatly improved the thesis. Thank you to Khaleel Khan, Dr. Nathan Amanquah, Dr. Reginald Daniel Duah, Amanda Lee Martin, Alan Springer, Saeed Adam, Kwame Oduro Boateng, Nsor Atindana John, and Nelson Ayamdoo.

I am most thankful and grateful for John Holmlund for his financial support. Without his exceptional generosity, this research would have been impossible.

Lastly, and most importantly, my love goes to my family. My father, Francis Adu, my mother Comfort Boahemaah Adu, my brother Eugene Kwame Adu Nyarko, and my sister Keren Happuch Adu Boahemaah, for they are always there for me, believe in me, and give me endless support.

Dedication

To whom I owe everything the Almighty God Jehovah, to my beloved parents Francis and Comfort Adu who raised me, taught me and made me the man I am today. To my brother and Sister Eugene and Keren Adu for their love and support. No words can describe my gratefulness to you and my only hope is to always make you happy and proud.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
ACID	Atomicity, Consistency, Isolation, Durability
CCDM	Core Cadastral Domain Model
CMS	Custom Management System
ER	Entity Relational Diagram
FIG	International Federation of Surveyors
FLEA	Flexible and Evolutionary Agile Approach
FTP	File Transfer Protocol
GIS	Geographical Information System
GNSS	Global Navigation Satellite System
GPRS	General Packet Radio Service
HTML	Hyper Text Markup Language
HTTP	Hypertext Transfer Protocol
ISO	International Organization for Standardization
IT	Information Technology
JUNG	Java Universal Network/Graph Framework

LADM	Land Administration Domain Model
LTIS	Land Tenure Information System
LIS	Land Information System
MDA	Model Driven Architecture
NCA	National Communication Authority
ORDMS	Object Relational Database Management System
OGC	Open Geospatial Consortium
PHP	Hypertext Preprocessor
QAP	Quadratic Assignment Procedure
RRR	Rights, Restriction, Responsibility
SNA	Social Network Analysis
SQL	Structured Query Language
STDM	Social Tenure Domain Model
TTM	Talking Titler Model
UML	Unified Modelling Language
XP	Extreme Programming

Chapter One: Introduction

1.1 Introduction

One school of social anthropology holds that sub-Saharan African tenure systems are fluid, dynamic, vague, flexible and subject to multiple interpretations (Berry 1993). Rules are contested, and may be negotiated and changed on a daily basis by individuals or institutions in charge. Membership and participation in formal social networks are critical components in a given agent's strategy for acquiring power and status, and social networks are also a major factor in multiple claims to the same piece of land (Berry 1993). People in powerful positions may assign interests in the same piece of land to a number of people who show allegiance to them, and this is the primary publicly sanctioned means of creating claims in land. Along with this, the system may be vulnerable to corrupt individuals who may impersonate land owners, selling the same piece of land to multiple unsuspecting buyers. Statutory registration and cadastral survey systems developed and used in these peri-urban communities tend to be based on rigid and static rules. Sometimes customary tenure systems are forced to merge with statutory rules. Thus interactions between individuals, families, clans, governmental agencies, chieftaincy institutions and land in these areas create complex networks that may be difficult to decipher (Barry *et al*. 2012).

In Ghana, land is managed and administered under both statutory and customary rules (Adjei 2011, Barry *et al*. 2012). Statutes are created and enforced by the government while customary rules are not written down but recognised as cultural norms dating back to pre-colonial times (Adjei 2011). These norms are continually changing (Berry 1993).

Assuming a number of enabling conditions are present, continual social network analysis and evolutionary land tenure information system development may be a design philosophy that may improve tenure security and reduce the level of conflict in these systems. The World Bank, (2003: p. 102) submits that the enabling conditions may include a simplified land procurement process, an efficient land titling system, and reliable land related policies supporting land registration and administration. In this study, the most important enabling condition will be collating reasonable land tenure data. In order to collect reasonable data, community members will have to agree to provide the data. Measures will also have to be implemented to prevent communal members from manipulating the data to suit their own agenda. If reasonable data - that is data from which entities involved in land transactions can be traced - can be acquired, social network analysis may identify and model distinct or identifiable social networks that symbolize the different forms of interests in land. Thus, social network analysis may provide a means of tracing succession, family lineage as well as recording individual interests and the changing relationships that may occur should an individual leave the network.

This thesis modified the design and redeveloped an existing evolutionary land tenure information system (LTIS), the Talking Titler system (Barry et al 2013). The evolution process was driven by applying data mining and social network analysis (SNA) techniques to LTIS data. The study applied data mining and social network analysis techniques to land tenure information stored in an ubiquitous TTM based LTIS. The scope is a test data set to calibrate and test the methodology.

Ubiquitous (web and mobile) LTIS software was written based on the Talking Titler Model (TTM) using the PostgreSql open source database. Netdriller software (Koochakzadeh et al 2011) was used to perform data mining and social network analysis to provide models of social networks based on data stored in the TTM - based LTIS. From this analysis, the Talking Titler data model may be modified to create specialized classes to reflect the social networks. The methodology challenges much of the current thinking in that, while not detracting from the importance of land objects, evolution is driven by changing personal relationships.

In peri-urban communities where rules of land entitlement are fluid - which may be as a result of merging formal and customary land tenure regimes - complex social relationships may develop as land changes hands. Conventional land registration systems are weak at modelling these complex social relationships and not designed to model situations such as customary systems in peri-urban areas which are changing due to urbanisation pressures nor are they suited to handling the complexities of informal settlements.

Data mining techniques are applied to extract social relationships from a set of land records, and in turn construct social networks from these relationships. Assuming the appropriate data are stored and can be extracted from a land tenure information system (LTIS), social network analysis may enable an analyst to visualize and analyze the relationships that underlie a local tenure system. If data are collected continually, which, as argued elsewhere, is advisable in rapidly changing and/or uncertain situations (Barry and Rüther 2005), then changes in social networks can be revealed as social change in local politics creates tensions and competition in the tenure system (Barry 1999). This in turn may reveal how the tenure system is evolving, and can be reflected in a LTIS if it is designed to evolve as the tenure system changes (Barry *et al* 2013). The data mining and SNA techniques applied here-in focuses on the existing social networks and interactions between the networks during land transactions.

However, effective analysis depends on reliable land tenure data. Therefore, to collate and store data, the flexible and evolutionary agile (FLEA) software development methodology was applied to developing an LTIS, while using an evolutionary land administration model (TTM) as a framework in terms of required functionalities and accountabilities. This LTIS on a small scale can collect, store, and relate objects, thereby aiding in handling complexities regarding data collection in uncertain situations (see section 1.2 for a description of uncertain situations). By using the agile software development approach, the system could be tailored to evolve with the changing requirements that characterize peri-urban communities.

The situational context for the study is a city expanding into a customary tenure area in Ghana, and the rationale is to improve justice and fairness as this change occurs. One likely outcome of a city expanding into customary lands as in many cases is that most vulnerable end up landless and have their agricultural livelihood expectations extinguished as powerful elites sell off land to people outside the customary lineage (Barry *et al* 2013).

The study should interest land tenure administrators and policy makers working in uncertain situations, and people who design and evaluate land tenure information systems. The study is exploratory, but this work indicates that the methodology may fill a need for modelling complex, changing social relationships in a land tenure information system that current systems fail to model adequately. The Talking Titler model design is conducive to developing a LTIS that evolves as tenure rules and practices evolve (Barry *et al* 2013). Social networks that change in response to broader change (e.g. urbanisation, housing programmes, changes in land use law and policy) may also evolve and change.

For this study, data sets were simulated based on a simplified model of a peri-urban, patrilineal customary society in Ghana experiencing urbanisation pressure. The society has a chief, a council of elders and a number of family heads. If people are members of a family, they are entitled to land as part of the customary system. Long standing tradition holds that land belongs to the living, the ancestors and future generations; it should never be sold. However times have changed and parcels for leasing and sale to strangers for residential purposes or sharecropping are surveyed by a land surveyor. The land surveyor reports to officials at the lands commission, where the survey plans are filed. In the study, the goal is to identify social networks such as family lineages in the customary system, lineage of chieftaincy institutions who set the rules for customary tenure, the key actors in the family, and the relationships between different families in the customary lineage - which is the overarching social network, and the key actors in the overarching social network (e.g. chief, elders).

1.2 Problem Context

There are numerous situations where conventional land registration models may be inadequate in securing tenure. Typical cases include informal settlements, squatting and peri-urban communities such as Bortianor, Accra.

Complexity and uncertainty in these settlements are inevitable. Barry et al. (2012) identified urbanization, market forces, power dynamics, and poverty as some issues which contribute to complexities in informal settlements and peri-urban Accra. The rich and powerful illegally procure land for residential and other purposes, depriving vulnerable indigenes of their right to inherited land. Community schisms result from competition among traditional leaders for land, coupled with opposition from the youth and communal members. In the absence of an effective and accountable land administration regime, these dynamics create chaos and confusion over land allocation practices leading to conflicts or even war in a community. Furthermore, some settlers may claim to be the rightful owner of a piece of land they have occupied illegally for a long period of time (Barry 1999).

The role of the state and its administration of land tenure can compound the difficulties. Sometimes customary rules may be ignored in a land transaction. Policies stipulated by statutory bodies, such as governmental agencies may be used to the benefit of rich and influential individuals. The policies may not be fully understood by indigenous community members, so state representatives may collaborate with corrupt individuals in grabbing land from vulnerable indigenes. The land policies and information systems may be used by a few and land expropriation may be condoned inadvertently. At other times, local communities may be forced to merge their traditional land tenure with the statutory land administration practices by governmental agencies with the goal of developing a unified land administration process. In such situations, navigating both formal (state) and customary tenure regimes may create avenues for corrupt educated individual to grab local lands. According to Berry (1993), claims to land right are mostly based on social identity, status and proof of purchase. However, there may be situations where no formal documents or evidence exist to support land transactions or aid in tracing family lineage or individual identity and status. This sometimes happens because the transaction was based on oral agreement. At other times, this may occur because land administration data used to support the transaction were collated and managed manually. In such cases, once the documents serving as evidence are misplaced or destroyed in a disaster, the poor and vulnerable are susceptible to losing inherited land.

The aforementioned scenarios in these settlements evolve rapidly and suggest the problems brought about by intense social instability. By way of contrast, in a socially, politically and economically secure and stable situation, there is no impediment to use of the land information system by the majority of the community; the system operates on a bureaucratic plane, achieves its administratively defined objectives and is insulated from the effects of corruption, power dynamics and change in social structures (Barry 1998, Fourie 1993).

1.3 Research Objective

The primary objective of this thesis is:

The development of an evolutionary TTM based LTIS driven by data mining and social network analysis.

The specific objectives of this work are listed below:

a. To develop an evolutionary customary land tenure system that can cater for requirements dictated by the changing conditions in a peri-urban community where there is conflict over leadership positions, access to land and the governance of land sales and registration.
b. To apply data mining and social network analysis techniques in finding a lineage of interest in customary systems.

The specific objectives imply a methodology for applying an evolutionary software development approach to the proposed LTIS developed based on an evolutionary land administration model (TTM). Through data mining and social network analysis techniques, the data collected in the LTIS is analyzed. The goal of this analysis is to visualize and understand the underlying social structures in these peri-urban areas, while modifying the data model based on the results. This modification in the data model, in turn, may effect changes needed in the developed system, the process expressed as a cycle, as depicted in Figure 1.1

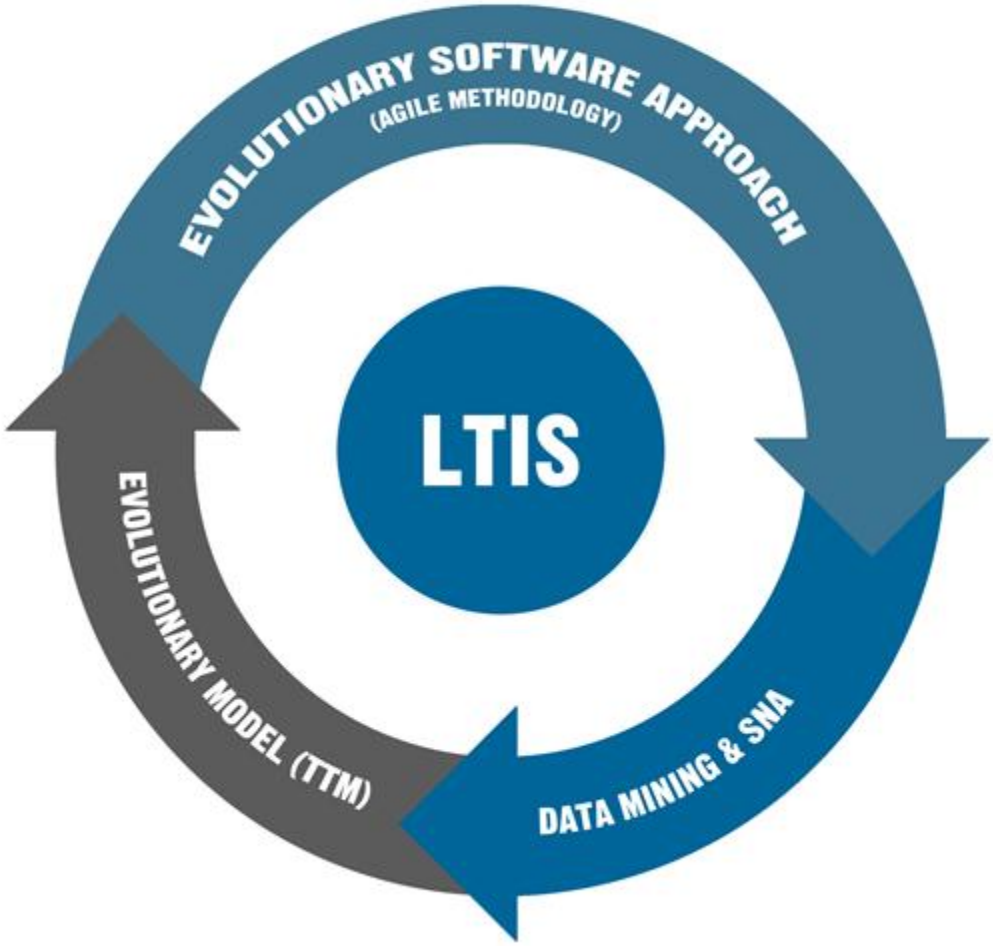


Figure 1.1: Pictorial description of the research objectives in this study

1.4 Research Questions

The following are specific questions that are addressed:

1. Why is it beneficial to develop an evolutionary LTIS using an evolutionary data model as a basis?
2. Can data mining and social network analysis techniques be used to examine changing customary systems?

3. How are the interactions, relationships, and hidden patterns between people, land and their supporting evidence in LTIS data revealed?

4. Are there relationships which conventional LTIS's are not modelling?

5. How could new, previous and unknown relationships be easily identified?

1.5 Research Method

This section describes the various activities conducted in this study to achieve the research objectives and to answer the questions posed in section 1.4. A summary of the activities is presented in Figure 1.2.

Addressing specific objective 1 required a clear understanding of the causes of complexities and uncertainties in these peri-urban communities. The approach taken here is in keeping with research conducted in peri-urban communities in Accra (Danso 2012, Barry 1998, Berry 1993).

This was followed by a detailed study of the TTM which allowed identification of the classes and attributes allowed in the model, making it suitable for use as a basis for developing a LTIS for use in peri-urban communities. Again, research was consulted to find out available land administration models and how suitable they may be for use in developing a land tenure information system (see section 1.5).

Next, the agile evolutionary software development approach, specifically Extreme Programming (XP), was adapted to developing the Ubiquitous (web and mobile) LTIS. The stages of the development and the justification for the chosen techniques and tools based on situations in peri-urban communities are explained into detail in Chapter 4. The approach taken to developing the LTIS on a robust platform is here-in referred to as an evolutionary LTIS development, in the following respects. The developed LTIS is based on an evolutionary land administration model and software development approach. The developed system will access and manipulate data in the object-relational PostgreSql database. Use of PostgreSql as the database allows the flexibility for the LTIS to evolve to incorporate and use geographical information in future; to this end, the use of its PostGIS extension provides a mapping capability.

To address specific objective 2, the TTM-based LTIS was populated with two different sets of simulated data. The first dataset was based on the scenarios described in a tutorial in the Talking Titler manual and mimics conditions in peri-urban communities (Barry 2011). The dataset was extracted, prepared and loaded into NetDriller for mining and analysis. Given that the topological relationships in this data were known beforehand, specific outcomes were expected. The aim was to test the ability of NetDriller in conducting analysis and providing meaningful results. NetDriller was used to construct a model of actual relationships found in the data. The model construction aids in visualizing the social networks and identifying family heads, lineage groups and ties into different forms of evidence. A detailed discussion of NetDriller and the data mining and social network analysis techniques employed in the software which are relevant to this study are provided in Chapter 3.

Next, a second scenario was created based on the study conducted by Barry et al. (2012) in Accra. Based on this scenario the LTIS was populated again with semi-structured data. Some of this data is real data from the Bortianor study. Using random points from this real case study data, additional data sets were simulated for analysis. The simulation was necessary because the records from the real case study were very small and cannot be used for any form of analysis. This dataset contained relationships created between people, land and media items (any form of evidence that can legally support land transactions). The scenarios contained in the dataset are described in detail in Chapter 4. The second data set was then loaded into NetDriller for mining and analysis.

Analysis using NetDriller was structured to identify pertinent information (about family heads, lineage groups, ties into different forms of evidence, stool succession, multiple land sales, and status of individuals in a family) and relevant evidentiary information from semi-structured (simulated) data in uncertain and complex situations. Based on the analysis, new main or sub-classes may be introduced into the original TTM. Following this, the developed LTIS may be modified to reflect the changes implied by the model.

Figure 1.2 briefly describes the methodology applied in this research.

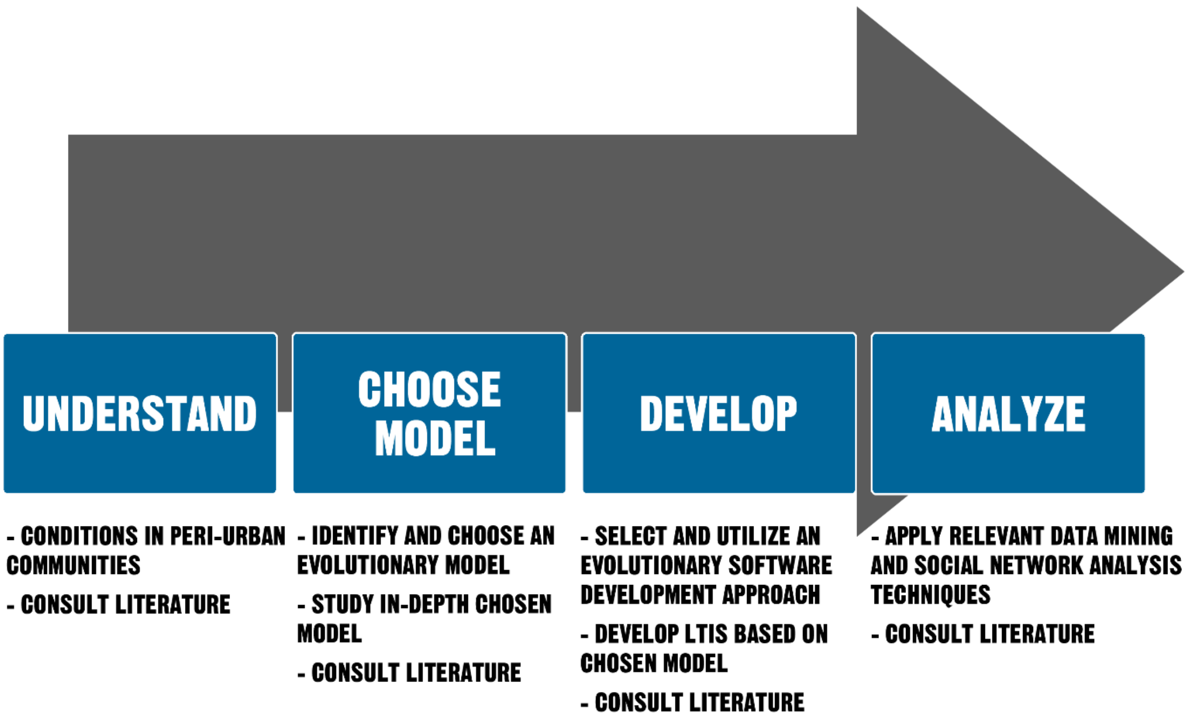


Figure 1.2: Summary of the research method followed in this study

1.5 Current State of Knowledge

The social and political conditions as well as power relations in peri-urban communities are vague, complex and dynamic (Barry 1999). Communal members are sometimes unable to understand the complexities in their own communities. For land administrators to effectively ensure tenure security in such conditions, there is the need to understand the social structures, political structures as well as the customary land tenure and its rules. Since these structures are made up of interactions between entities (e.g. people and people, people and land), creating network models in order to visualize and analyse the complex networks could aid in understanding conditions in these communities (Xu et al. 2010). Data mining and social network analysis provide techniques that can aid in constructing, modelling, and analysing the social networks that characterize peri-urban communities. The technique employs mathematical theorems with graph theory lying in the heart of studying complex networks (Xu et al. 2010, Wasserman and Faust 1994). Although the techniques have not been often applied to land tenure information systems and data, they have been successfully applied to studying changing patterns and various phenomena. These include the study of "the impact of urbanization on individual well-being", "coalition formation", and "Belief systems" (Fischer 1982, Kapferer 1969, Thurman 1980, Zachary 1977, Wellman 1979, Erickson 1988, Breiger 1981, Wasserman and Faust 1994). The technique works whenever there is a relationship that results from an interaction between entities such as individuals, family, companies, and land in a network. Land tenure information data consist of relationships created between people, land and the evidence used in support of transactions. A credible analysis in this context may enhance the modification of the components (systems and land administration models) involved in land administration to appropriately ensure justice in favour of vulnerable indigenes in peri-urban communities.

One such component is the type of land administration model that serves as the basis for developing suitable LTIS for data (structured, semi-structured or unstructured) collection and storage. To [1] harmonize and [2] produce a unified land administration system which could [3] enhance cadastral data sharing among different communities governed by different tenure rules, several standardised land administration models have been proposed (Oosterom *et al*. 2002). Examples of these models include the Land Administration Domain Model (LADM) and the Social Tenure Domain Model (STDM) (Oosterom et al. 2006). However, the systems designed based on these standardised land administration models are suited for a secure and stable situation where the information system always reflects and parallels the tenure system on the ground (Muhsen and Barry 2008). As a result, the models may not work when the conditions in the communities they are applied are vague and system requirements are almost impossible to define. With changing circumstances in peri-urban communities, it is always possible to end up with new requirements of an intended land tenure system that are a totally different from what was originally started (Barry et al. 2012). In consideration of the mercurial and variegated nature of actual property relations, an optimal strategy, therefore may be to adopt a data model where the design is grounded in the data (such as the Talking Titler Model (Barry et al. 2103)) as opposed to using a top-down approach adopted by standardized models (Barry et al. 2012).

In addition to using such an evolutionary data model, it is also important to utilise an evolutionary software development methodology in developing a LTIS. The method is important because under uncertain conditions, the initial requirements as well as the final product of a system may be unknown. When system requirements are unknown or difficult to define, an evolutionary software development approach (prototyping or trial and error) allows the system to develop astride efforts to model the ill-defined elements (Muhsen and Barry 2008, Barry *et al*. 2011, Barry *et al.* 2012). The method may not only simplify the development process but also aid in implementing system changes to reflect the dynamic conditions in peri-urban communities.

The basic idea is that, the data model and the approach to developing the LTIS should be evolutionary so as to cater properly for the changing conditions. Continual analysis of the collected LTIS data may not only help in visualizing and understanding the conditions on the ground but contribute to improving the data model by informing new or sub-classes which may in turn effect appropriate changes to the developed system.

1.6 Significance of Research

The work in this thesis is a new endeavour. Although evolutionary design and implementation of a land administration model and LTIS is often talked about (Muhsen and Barry 2008), the products of such theorizing are few in published and anecdotal sources. The application of data mining and social network analysis in this context has not been done before. The technique has the potential to foster discovery of relevant information and hidden patterns in land transactions in uncertain situations. The analysis may provide land administrators relevant information regarding communication patterns, power players, and community dynamics. This could inform intervention procedures by land administrators or anyone with authority to ensure tenure security in peri-urban areas to the benefit of marginal groups. The analyses may also reveal conflicts in how people present the network in support of their claims. It may also bring stakeholders such as policy makers, land administrators, indigenes at par with changing conditions on the ground and may ultimately contribute to improving tenure security thereby reducing the tendency of vulnerable indigenes losing their inherited land.

Additionally, the design of a land administration data model that is able to evolve with changing conditions may contribute to the local situational design of a LTIS in uncertain situations. The TTM may become a functional land administration model that could handle complexities efficiently. The improved TTM-based LTIS may allow a great deal of flexibility in such a way that data relating to people, land and evidentiary media (titles, deeds, survey plans, descriptive documents, audio records of oral testimonies, videos, photographs, valuation records etc.) can be stored and related (Muhsen and Barry 2008).

1.7 Thesis Organization

Chapter 2, which is the first part of the literature review, discusses the components needed for evolutionary land administration model and LTIS design. The chapter answers research question 1 (see section 1.4). Chapter 3, which answers research question 2 to 5 (see section 1.4) discusses the vocabulary and concepts in social network analysis. The concepts discussed are the ones utilized in this research. The chapter also explains the data mining techniques employed in the analysis. Brief descriptions of the tools available for mining and analysis are given, and the reason for choosing NetDriller is explained. Chapter 4 describes in detail the activities conducted in achieving the specific objectives and the primary objective in general. The chapter also describes the scenarios governing the data simulated and utilized in the study. The results achieved based on the exploratory experiments conducted are discussed in chapter five, as well as a brief evaluation of the developed system. The thesis concludes in chapter six with a discussion of the limitations and recommendations for further research.

Chapter Two: Evolutionary Land Tenure Information System

2.1 Introduction

The literature review is divided into two parts: [1] the development of an evolutionary LTIS using both an evolutionary land administration model and software development approach and [2] mining and analysis of LTIS data. This chapter presents the development of an evolutionary land tenure information system.

Primarily, the chapter contributes to the fulfilment of objective 1.3a by addressing research question 1 in section 1.4. This in turn addresses the first part of the primary objective, which is – the development of an evolutionary TTM based LTIS. It begins by defining what is referred to in this study as evolutionary land tenure information system design. It continues to discuss the importance of developing a LTIS based on an evolutionary land administration model using an evolutionary software development approach. It also explains why this approach is important when developing an information system to collect and store land tenure records in a peri-urban community.

Because a land tenure information system is developed using the governance structure and rights regime embedded in a given land administration model, the chapter analyses key aspects of three land administration models (LADM, STDM, TTM) for LTIS development.

2.2 Evolutionary Land Tenure Information System Design

Evolutionary land tenure information system design in the context of this thesis refers to a flexible system that can adapt and evolve with the uncertainties that characterize tenure systems in peri-urban communities. Conventional land registration systems have not yielded the desired results in peri-urban communities, because they utilized a top-down approach - both for the system design strategy and the standardised land administration model used - which works well when the social, political, and economical situations are stable and the "problem contexts are simple and well understood" (Barry *et al*. 2012). Since conditions in peri-urban communities rapidly evolve, a top-down approach may not be ideal for developing a LTIS for such situations.

Systems theories are discussed but seldom utilized in creating conventional registration systems. Participatory planning is talked about, but community members are not involved in creating these systems. Creators of conventional registration systems act as consultants and dictate what approach indigenes should adopt in dealing with uncertainties in their communities. Although these tactics may not be wrong in themselves, the creators of these systems do not do enough micro level work to model changes in uncertain conditions. It is therefore difficult to tell if indigenous communities will utilize the systems and standards proposed. Thus, in addition to providing professional guidance, community members should be involved in deciding what approach, standard, or system may work in their community (Barry *et al.*2012, Barry 2009, Barry 1999, Barry 1998).

Social conditions in peri-urban communities evolve rapidly. Chieftaincy issues, post conflict situations and urbanization may create uncertainties in land tenure management in such communities (Barry *et al*. 2012). As a result, when developing a LTIS, it may be difficult to determine the requirements of the land tenure system at the initial stage (Barry *et al*. 2012). These requirements include choosing a suitable land administration model and knowing when and which functionalities implied by the model to incorporate in the initial system. Even if some requirements are known, constant situational changes such as political instability or even war may render a previously developed system obsolete. According to Patel (2009), situational use of information system reveals purposive functional shortcomings; i.e., a system may function effectively today by satisfying all its needs but fail tomorrow as requirements change. It is therefore imperative to develop a system that may evolve with changing requirements imposed by changing conditions.

This calls for a system development approach that allows a developer the flexibility to incorporate and update a system and its requirements as conditions evolve. Evolutionary system development which accepts and embraces the uncertainty of the future is an ideal approach to developing a LTIS under uncertain conditions (Castelo and da Silver 2009). It is a ‘‘design-a-little’’, “build-a-little”, “deploy-a-little”, “learn-a-little” approach which accommodates changing requirements and is thus suited for the evolving scenarios that characterize peri-urban communities (RTO/NATO 2003). Through this approach, the system could be tailored to “handle future scenarios while addressing yesterday’s needs” (Castelo and da Silver 2009).

A conceptual evolutionary framework designed by Barry *et al*. (2012) stresses the need to create an initial prototype that will evolve with time. Since system users may be uncertain about the initial system requirements, this designed prototype could be used to verify requirements and to help users identify and articulate their needs in the system. (Barry *et al*. 2012; Boehm 1988). The system is designed and developed in an iterative fashion. As the conditions evolve and requirements change, the prototype is altered to suit evolving requirements. When the conditions on the ground eventually become stable, the prototype is hardened and the final product implemented (ISACA 2009). This kind of approach requires flexibility, meaning that changes can be introduced in the application over a given time period with little difficulty (Castelo *et al*. 2009). Flexibility reduces the difficulties in making adjustments to the system when requirements change.

Flexibility, prototyping and iterative design are necessary qualities for developing a system for any conditions. These qualities are employed in the evolutionary software development approaches and as such are utilized in developing a LTIS that can support the land administration function in complex and uncertain conditions.

2.3 Evolutionary Software Development Approach

Evolutionary software development approaches offers several methodologies that can be employed in developing a system for uncertain scenarios. The methods embrace incremental and iterative development. In incremental development, the system is developed and released in batches as opposed to releasing the entire system in one implementation. Each released batch contains an additional feature which was not present in the previous release. Iterative development on the other hand builds the system in iterations while applying adjustments based on feedback received through system testing. The system is designed, built and tested repeatedly until the desired functionality is achieved. It is regarded as the most appropriate method for dealing with complexities and risks related to software development projects (Elaine and Zimmer 1996, ISACA 2010). In the face of uncertainties - where system requirements continually evolve - incremental and iterative methods are invaluable. They offer the flexibility of going back and forth in planning, designing, implementing and testing the evolving system. This repetitive cycle allows new functionality and requirements to be incorporated at any stage in the life cycle of the system. Eventually, user requirements are met. Resources including time and money are also utilized appropriately (Elaine and Zimmer 1996). One such evolutionary methodology is the agile software development approach.

2.3.1 Agile Approach

Agility literally means the ability to move quickly or easily. Software development processes have evolved from being restricted to following a planned criterion to a flexible iterative procedure (Beck et al. 2001). Flexible iterative procedures allow a developed system to be altered any time the requirements change. Thus, the approach is suitable for developing a LTIS for uncertain conditions.

The manifesto for agile software development approach focuses on "uncovering better ways of developing software by doing it and helping others do it" (Beck et al 2001). The core values as presented by Beck et al (2001) are:

- Individual and interactions over processes and tools
- Working software over comprehensive documentation
- Customer collaboration over contract negotiation
- Responding to changes over following a plan

The Agile Manifesto incorporates twelve principles that guide the agile process (Shore and Warden 2008). This approach "recognizes people as the primary drivers of project success" (Highsmith and Cockburn 2001) and project consultants who have the ability to adapt, maneuver and respond to a continually changing environment. Critical to project success is the significance accorded to interactions between skilled professionals and stakeholders involved in the software development project over the simple application of processes and tools (Fowler and Highsmith 2001). Developing a working software that will effectively address needs and requirements requires an active, focused, high-value collaboration between end-user and developer (Fowler and Highsmith 2001; Shore and Warden 2008). The proponents of agile development methodology argue that although planning is good, "it becomes dangerous if it blinds the developer to change" in favour of system design too tightly focused on principles *per se* (Fowler and Highsmith 2001). Fowler and Highsmith's (2001) analysis of several projects shows that a software development project succeeds when the development team is agile enough to respond continually to external changes; this underscores the importance of "welcoming changing requirements even late in development" process (Shore and Warden 2008). Also, agile methodology "recognizes people as the primary drivers of project success and focuses on effectiveness and maneuverability" (Highsmith and Cockburn 2001). The principal focus of the agile process is to deliver working software (Fowler and Highsmith 2001). The aforementioned philosophy of agile makes it a suitable approach for developing a LTIS for use in peri-urban communities where user needs and system requirements continually evolve due to uncertain conditions such as conflict over leadership positions and access to land.

The approach is suitable and applicable under uncertain and evolving scenarios. Developing a customary land tenure information system to be used in a peri-urban community requires a lot of interaction between system developers and stakeholders such as land surveyors, land administrators, community leaders and family and individual title-holders. Although these stakeholders may not necessarily be end users of the developed system, the success of the system depends on understanding the conditions on the ground and the requirements, which can be achieved through interactions with the aforementioned stakeholders. Additionally, uncertain and varying conditions may change requirements that may have to be rapidly incorporated into system design. Because conditions in peri-urban communities continually evolve, it is possible to end up with new system requirements that are totally different from what was drafted at the beginning of the project. Should such drastic changes in requirements occur, it should be possible to tailor the developed system to suit these new requirements. The emphasis here is that, in order to develop a land tenure information system for uncertain situations, the software development approach must be evolutionary so that the system can be tailored to the changing requirements in these conditions. The choice of evolutionary approach however, depends on the developer and stakeholders. The flexibility inherent in the agile software development approach makes it possible to tune a developed system in unstable social contexts.

The agile approach consists of several methods including scrum, crystal family of methodologies, feature driven development, rational unified processes, dynamic system development, adaptive software development, open source software development and extreme programming (Abrahamsson et al 2002). These methods are guided by the agile manifesto and principles and could thus be applied to developing a system that will evolve with changing requirements. XP is adapted in developing the TTM-based LTIS in this study.

Extreme programming (XP) is a lightweight, efficient, predictable, low-risk, flexible, scientific and fun way of software development Beck (1999). It was chosen for the development of the TTM-based LTIS in this study because it is simple and flexible, and because it has a track record of being successful in numerous projects. Although there are no sources that show the application of XP to developing a LTIS, the method is deemed suitable for this study because it allows the creation of an "ever-evolving design that are easy to modify when plans change" (Shore and Warden 2008).

Time management is up to the system developer. Different features could be worked on at the same time. XP begins with the development of a prototype which is refined as requirements change. This makes it easy to develop a system even when little or vague requirements are defined. The system can be improved over time as requirements become clearer. It also makes it easy to add to the system when new features are defined. XP is inexpensive, requires a relatively smaller team (2 – 12 (Wells 1999)), and is suitable for relatively smaller products (Boehm 2002).

The application of XP in developing the ubiquitous TTM-based LTIS is described in chapter 4. A description of the method and its phases, according to Abrahamsson et al. (2002), Beck (1999) and Shore and Warden (2008) is provided in the next section.

2.3.2 Extreme Programming (XP)

XP facilitates successful software development in the face of vague and rapidly evolving conditions and requirements (Beck 1999b). The characteristics of XP allow room for short iteration, end-user participation and collaboration with development team, constant integration and testing, collective ownership of code and limited documentation (Abrahamsson et al 2002, Shore and Warden 2008). XP consist mainly of five stages, namely; exploration, planning, iterations to release, product ionizing and maintenance, and death.

2.3.2.1 Exploration Phase:

This is the stage where developers interact with system owners and end-users to understand the requirements and functions expected of the developed application. The system owner is responsible for funding the design and development of the system. The end-user is the actual user of the developed system. The end users, system owner, and software developers are here-in referred to as the team for the project. The system owner and end-user writes out stories describing features expected in the system and what it should do. The stories are written as and when the user identifies them. A user will continue to write these stories until there are no more. The user may sometimes rewrite stories to clarify their ideas to developers. Software developers, having read the described features, familiarize themselves with prospective tools and technology that may be appropriate for the project at hand. To test the suitability of explored and chosen tools and technology, a prototype is developed based on the described features. By using the prototype, the end-user could see some of the requested features at work and may add ideas to fine tune them.

2.3.2.2 Planning Phase:

At the planning stage, the stories are arranged in order of importance. Some features may have to be implemented earlier than others. The team collaborates and agree on which features to implement in the first released application. Here, software developers evaluate the efforts required in developing every program feature recorded on the story cards. Based on the evaluation, a schedule is drafted for the development task.

2.3.2.3 Iteration Phase:

At this stage, the system goes through several iterations. The end-user decides which stories are chosen to guide project objectives for various iterations. The initial iteration is centered on creating a system with the architecture for the whole system. The developed system is then tested for errors. In an organizational setting, the system is put into production and regarded as complete at the end of the last iteration after functional tests have been run.

2.3.2.4 Product ionizing and Maintenance Phase:

Before the developed application is released to the system owner and end-user, several tests are conducted to check the performance of the system. It is possible for stakeholders to discover at this stage that additional changes or even new features may be required. When such new changes are discovered, stakeholders decide whether to add them to the current released system or postpone and document them for future implementation. This decision help to ensure that the project does not drift, and that imprecise expectations do not cause the development efforts to fizzle. When the system is released into production, it is maintained while the development team create new iterations based on collaboration with all stakeholders involved in the project. Maintenance involves providing support for the end-user. The support may include guiding users on how to use the system and answering any additional question the users may have concerning system functionality. At this level, the rate of development is slowed down as system performance is monitored and assessed against objectives established in the founding charter.

2.3.2.5 Death Phase:

This is when end-users or system owners no longer have new stories to tell. This means there are no more additional features or changes to be made to the system design, architecture or code. The end user is satisfied with the system's performance. The system is reliable and meets the end user's needs. The final documentation of the system is written. The documentation include system architecture and configuration. According to Abrahamson et al (2002), death of the system may also occur should the application fail to live up to the standard or outcomes for which it was developed, or if the cost of future development proves prohibitive.

These phases are flexible enough to be used in developing an evolutionary LTIS. However, the development of such a system is based on a land administration model. The next section analyzes fundamental characteristics of three land administration models for LTIS development.

2.4 Land Administration Models

Various land administration models are available to inform the theoretical basis for land tenure information system development in this study. There are several land administration models that one can choose from depending on the conditions on the ground. It may however be necessary to tailor a chosen land administration model to suit the requirements dictated by the conditions on the ground. The Talking Titler Model (TTM), which is being developed by the Land Tenure and Cadastral Systems Research Group at the University of Calgary, was used in developing the LTIS for managing the customary land tenure system for peri-urban conditions.

This section provides insight into some of the land administration models being developed.

2.4.1 Land Administration Domain Model (LADM) (formerly the Core Cadastral Domain Model (CCDM))

This model was first proposed, as an attempt to developing a standard model for cadastral systems, at an FIG Congress in 2002 in Washington (Oosterom et al 2002). Various organizations have been involved in its development, including the Open Geospatial Consortium (OGC), International Organization for Standardization (ISO), International Federation of Surveyors (FIG), UN-Habitat, and Infrastructure for Spatial Information in Europe (INSPIRE) (Muhsen 2008). The contributions of the aforementioned organizations have led to different versions of the CCDM. In addressing semantic issues identified by ISO (ISO /TC211 N2385 2008), the latest version (v. 1.1) refines the model and changes its name from Core Cadastral Domain Model to Land Administration Domain Model (Hespanha et al 2008). According to ISO, the term "Land Administration" covers the scope of the model better than the term cadastral, in that land administration clearly includes legal and spatial data modelling, while the term "cadastral" implies that it is limited to spatial data only (Hespanha et al 2008). However, it is difficult to talk about the land administration domain model without mentioning the core cadastral domain model, since the LADM inherits most of its features from the CCDM.

The CCDM is comprised of three main data object classes. These classes are the RegisterObject, RRR (right, restriction, responsibility) and Person (Oosterom et al 2006). The classes and the relationship between them are represented in the Unified Modelling Language (UML). UML diagrams make it possible to show specific classes, their attributes and the restrictions on how the classes relate to each other. According to Oosterom et al (2006), the UML class diagram for the cadastral domain model contains both legal/administrative object classes like persons, rights and geographic description of real estate objects. Therefore, data could be maintained by different entities such as Municipality, Planning Authority, Private Surveyor, Cadastre, Conveyancer and/or Land Registry (Oosterom et al 2006).

The RegisterObject class represents any object that is subjected to registration by law (ISO /TC211 N2385 2008). The "RegisterObject" does not relate directly to the Person class but in a relationship mediated through the RRR (right, restriction, responsibility) class. Although a Person can have multiple Rights, Restrictions or Responsibilities, a specific defined right could be held by exactly one person. The RegisterObject class can be classified as movable or immovable. Immovable objects are also classified as land or 3D space (Oosterom et al 2006). Land objects can be grouped differently as "RegisterParcel, SpaghettiParcel, PointParcel, TextParcel, ParcelComplex or PartOfParecel" (Oosterom et al 2006). Other immovable objects are "Building Unit, NonGeoRealEstate and OtherRegisterableObject" (Oosterom et al 2006). The different specializations of immovable objects associate with one or more Persons through the RRR class (Oosterom et al 2006).

The model includes other parts called the ServingParcel. The ServingParcel relates directly to more than one RegisterParcel, serves different RegisterParcels, and is "held in joint ownership by the owners of those RegisterParcels" (Oosterom et al 2006).

The person class could be classified as a Natural Person or Non-Natural Person, the latter of which may include an organisation or group of people such as a community. A Natural Person cannot be a Non-Natural Person and vice versa.

The RRR class represents the legal and administrative aspect of the model with specializations, Rights, Restrictions and Responsibilities. The strongest right that could be held by a Person is termed ownership or freehold. Subordinate rights also exist, which allow a person to use the land in a restricted manner; security and servitude rights are also capable of representation in the CCDM. In functional terms, the CCDM includes "legal/administrative, person, immovable object specialization, surveying and geometric/topological aspects" (Oosterom et al 2006). All of the aforesaid features of the CCDM are inherited by the LADM. Figure 2.1 shows the core classes of the LADM.

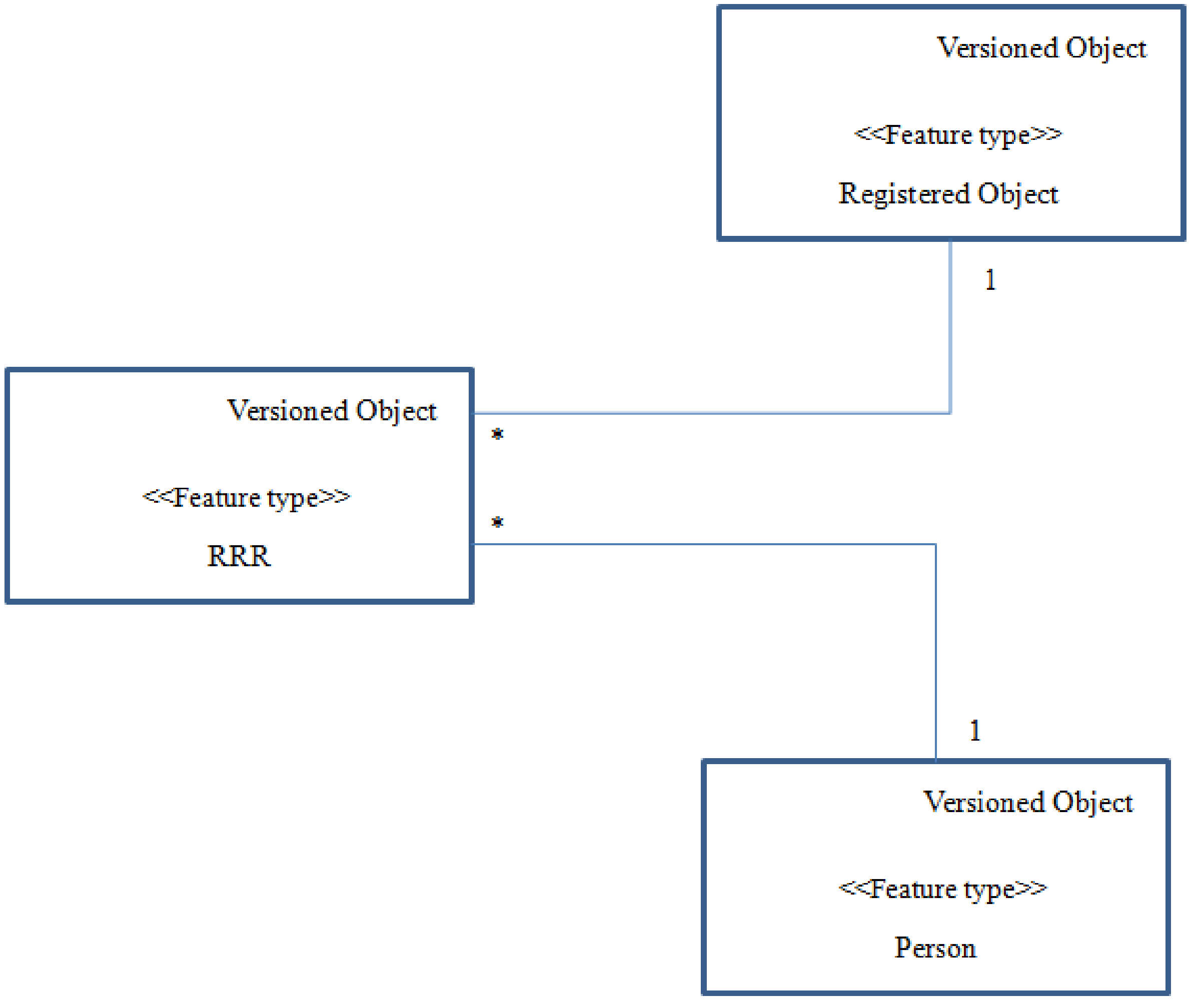


Figure 2.1: Land Administration Domain Model (LADM)

At the heart of the LADM (refer to Figure 2.1), just like the CCDM, are the RegisterObject, RRR (rights, restrictions, responsibility) and Person Classes. However, a new class called LegalNetwork which is a specialization of the RegisterObject class has been introduced into the

LADM. The class Person has an additional attribute called “Type”, which represents Natural or Non-Natural Person (Lemmen et al. 2006). Also, the model fully accommodates ISO spatial schema (Lemmen et al. 2006). The aim of the LADM in general, is to provide a basis on which efficient cadastral systems could be developed based on the model-driven architecture and to enhance internal and trans-border communication “based on the shared ontology implied by the model” (Lemmen et al. 2006).

The model is said in a broad sense to serve at least two important goals: (1) to avoid reinventing and re-implementing the same functionality over and over again as theory advances, while providing an extensible basis for efficient and effective cadastral system development based on a model-driven architecture (MDA); and (2) to enable involved parties, both within one country and between different countries, to communicate based on the shared ontology implied by the model (Oosterom et al. 2006).

However, to achieve these goals, there is the need to further improve the model and adapt it to fit specific situations. Although modelling the RRR class as standalone highlights the prominence of rights and allows the assignment of attributes to rights in the model, rights cannot exist independently (Muhsen 2008). This is because definition of rights involves other information related to immovable objects and the right holder. Therefore, the model may be deficient in modelling real rights “where a RegisterObject, regardless of its owner, holds rights in another adjacent RegisterObject” (Muhsen 2008). Also, because the LADM is a European Union model suited for secure and stable conditions, it may not be apt for uncertain situations such as communities experiencing social unrest due to post-colonial and post-conflict issues (Muhsen 2008). LTIS applications developed based on the LADM may be applicable to settings where the information system mirrors the tenure system on the ground. Such an application may fail to function efficiently in emerging economies or rapidly evolving conditions marked by ineffective governance structures, ethnic or sectarian strife, wars and conflicts over land. Accordingly, it may be necessary to consider ways to incorporate flexibility and event-based dynamics in the LADM and develop a common, but extendible application framework for land administration and survey (Hall et al. 2008).

Notably, the Land Administration Domain Model (LADM) developed to serve as the basis for LIS implementation requires specialization to fit a specific country profile (Hall et al 2005). This may result in parallel information systems development efforts rather than shared efforts between land administration experts (Hall et al 2005).

Consequently, to support the land administration function in uncertain situations, the authors of the LADM developed the Social Tenure Domain Model (STDM). The next section describes the STDM, which is based on a set of requirements that mainly mandate developing legal and technical land tools that are pro-poor; *i.e.* tools that facilitate delivery of tenure security for the poor (Lemmen et al 2007).

2.4.2 Social Tenure Domain Model (STDM)

Developed and published in Lemmen et al. (2007), the first draft version of the STDM, is based on a set of requirements presented at the 5th FIG Regional Conference in Accra, Ghana, in 2006 (Augustinus et al. 2006). The STDM aims at modelling the relationships among land and people in informal settings. Informal settings include post-conflict situations and customary areas where social rules are rooted in indigenous societies that existed prior to colonisation and which still persist. In these areas, land use and registration processes may not be the rule (Lemmen et al. 2007). According to Lemmen et al (2007), the focus is mainly on developing countries.

The STDM is a specialization of the Land Administration Domain Model. The core classes of the STDM are SpatialUnit, SocialTenureRelation and Person. STDM potentially allows the recording of “claims” and “social tenure relationships” such as tenancy into a system (Lemmen et al. 2007). The model takes care of the uncertainties that arise in right descriptions by local people through its spatial unit component. STDM records the “Social Tenure Relation” between ‘Person’ and ‘Spatial Unit’. The person and spatial unit objects do not relate directly but through the SocialTenureRelation object. The Person has specialization classes called Natural or Non-Natural Person. The class Non-Natural person includes governmental department, company or organization. Two or more persons can form a GroupPerson which represents social structures such as communities, clans and tribes (Lemmen et al. 2007). The SocialTenureRelation stores any form of interests such as ownership that a Person may have in any SpatialUnit (Lemmen et al 2007). To improve the options available in representing parcels, the SpatialUnits are grouped into categories including Parcel Family, DescriptiveSpatialUnit, and IncompleteSpatialUnit’ (Lemmen et al 2007).

Specialization of the STDM may be necessary, when applied to a specific country, to adapt and incorporate all cases as well as requirements stipulated by organisations and institutions directing a given land tenure regularisation initiative. The administrative objects of the model are the Person and SocialTenureRelations, while the geographic objects are represented by the SpatialUnit class (Lemmen et al 2007).

Like the LADM, there is no direct relation between Person and Spatial Unit but through Social Tenure Relation. In general, the STDM model assumes stability in the situations it addresses, and that the property relations in these situations are fully capable of representation. Just as in the LADM, the Person objects can be either a right holder or a surveyor. Each person can hold a share in a SocialTenureRelation which replaces the RRR class in the LADM (Lemmen at al 2007). This is necessary because the class RRR and its specializations suggest the existence of a legal basis which is not always the case in social tenure relations (Augustinus et al 2006). The term social tenure relations describes interests and claims which are not necessarily legally enforced, but which are, nevertheless, recognized within a social system in an indigenous community (Lemmen at al 2007).

The SpatialUnit class in the STDM re-uses the functionality of RegisterObject class in the LADM. It allows flexibility in defining land objects (*i.e.* areas on which social tenure relations, *e.g.* rights, are exercised) and allows the definition of land objects using methods other than accurate surveys and geometrical measurements (Muhsen 2008). The model attempts to include all possible specializations of a spatial unit that may be encountered in a given social setting. However, it is not necessarily feasible for all possible types of SpatialUnits to be known in advance. Therefore, any instance of a new type of SpatialUnit must fall under one of the specializations of the SpatialUnit class (Muhsen 2008). The STDM should therefore allow flexibility in the model to deal with the uncertainty that characterizes informal and uncertain scenarios.

Although the model is straightforward, its simplicity does not capture the complexities existing in uncertain situations. The model adopts the top-down approach which may omit some important relationships on the ground. It is not efficient to adopt a model (STDM) that does not work under some conditions without first assessing the requirements of those situations (Muhsen 2008). Therefore applications developed based on the STDM may be unable to ensure tenure security in uncertain and rapidly evolving conditions. These weakness in the STDM are the strengths of the TTM. The TTM is discussed in the next section (section 2.4.3).

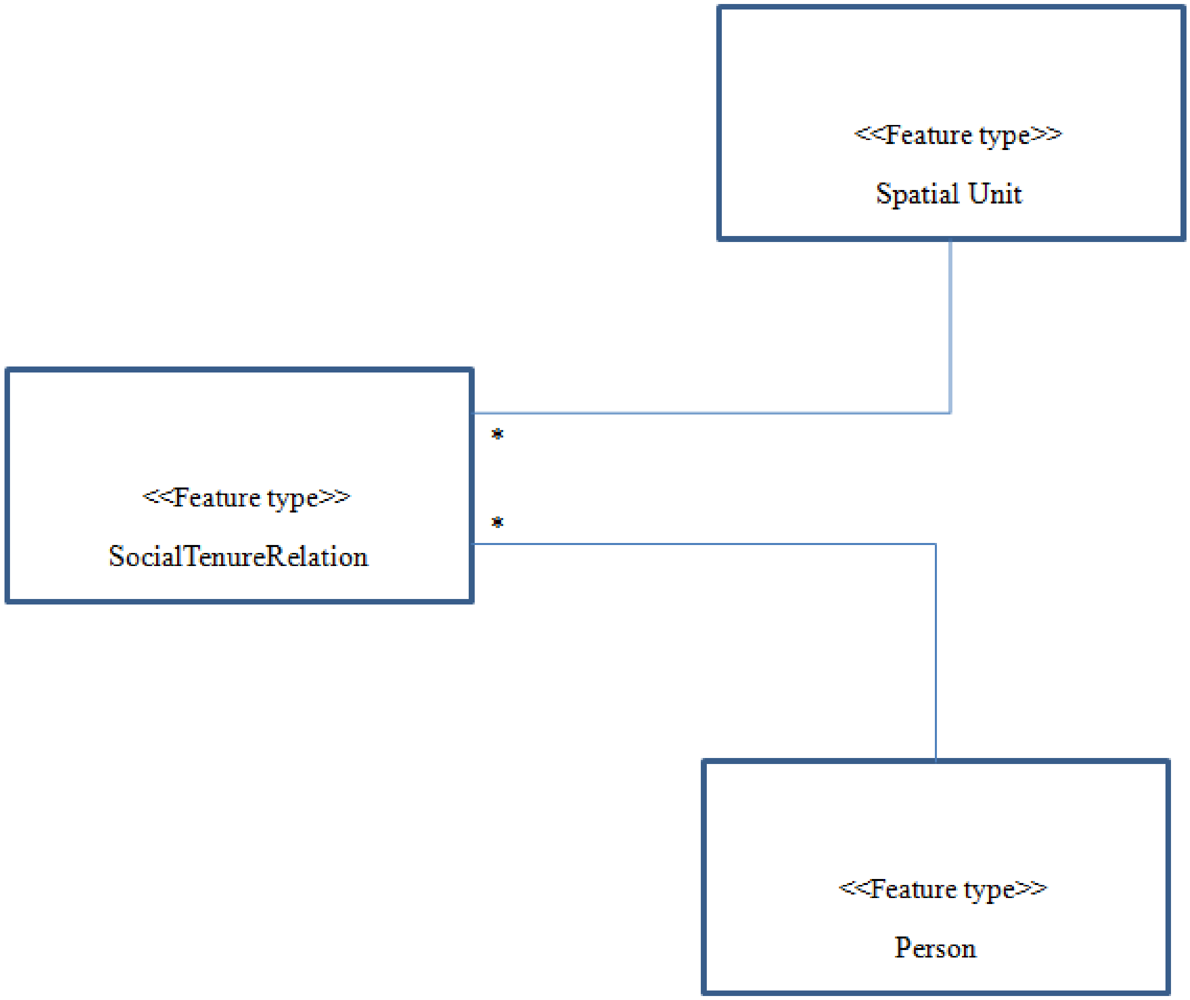


Figure 2.2: Social Tenure Domain Model (STDM)

2.4.3 Talking Titler Model (TTM)

The Talking Titler Model (TTM) is a simple and flexible land administration model designed to function in uncertain conditions such as peri-urban communities in emerging economies, informal settlements, squatting or post-conflict situation (Barry *et al*. 2012). Informal settlement is a "settlement where land is occupied according to a set of rules and processes that are not entirely legal" (Barry 1999). Informal settlers usually constitute indigenous, traditional or customary early settlers whose area may not be recognized by law (Barry 1999). They could also be a group of people who travel or change their geographical location as a result of war, natural disaster or other unforeseen occurrences. Settlers build without statutory approval for building plans although they may be owners or have permission to occupy a parcel of land (Barry 1999). Basically, “informal settlement exists on a continuum of legality” (Roux *et al*. 2009). This ranges from illegal use such as land invasion to official allocation of managed land “after the social modification of community rules” (Roux et al 2009). Squatting is when a building, any sort of structure or space is inhabited without obtained consent to use, own or rent from a rightful owner (Barry 1999). Peri-urban is when customary communities adjoin with urban societies (Barry 1999). There is a strong influence for customary to merge with statutory tenure system in peri-urban communities such as in Accra, Ghana.

The TTM has four main interrelated classes. These classes are the Person, Land or Property Object, Reference Instrument and Media (Barry et al. 2012; Muhsen and Barry 2008). The Person object or class represents an individual, company, social structure or an association. The person object is the right holder of the land or property in question. The Land or Property object or class includes dwellings, trees, parcels or trap lines (Muhsen and Barry 2008). In land administration, rights represent the legal component. These rights include ownership and lease, and simply serve as a record of any interest or relationship a person may have with the land. There are mostly legal, administrative, or evidentiary files or documents that articulate and give credence to the said rights, whether in a legal or customary system (Muhsen and Barry 2008). The Talking Titler model records every interest with its supporting evidentiary files in a Media Item class (Muhsen and Barry 2008). Through the Media class, the model accommodates structured, semi-structured and unstructured data. This allows flexibility in storing and utilizing a wide range of media files. These include video clips, audio files, maps, images, unstructured scripts and photographs (Muhsen and Barry 2008).

In peri-urban communities and uncertain conditions, it is typically difficult to understand and define user needs and requirements (Barry et al. 2012; Muhsen 2008). Under such a circumstance, the model allows the use of the Person, Land and Media item classes for data collection and relation (Muhsen 2008). However, when user requirements are clearly defined and understood, the flexibility of the model is then reduced (Muhsen 2008). The increase and reduction of flexibility in the model is done and ensured by the Reference Instrument class. Reference Instrument purposefully contains special functions applied to stored records to improve the integrity of the system, *e.g.* security functions and closing a record to editing (Muhsen and Barry 2008).

The TTM model records interests and all obligations which include special rights assigned to land or property. The model absorbs conflict and complex situations such as chieftaincy successions by allowing people-to-people relationships. This is achieved by relating a two or more Person Objects who are related in any form of relationship. However, unlike the STDM, the flexibility allowed by the TTM can result in a not so meaningful relationship created between objects. Therefore, care should be taken when relating objects through the TTM. Conversely, one can start off with any of the objects in the TTM – say media object – and then evolve with the conditions on the ground.

The design philosophy of the TTM for local level land information systems directs that system design be grounded in the data; all preconceived notions about data classes, such as person, land object and reference instrument, and the relationships between them should be subject to rigorous, continual, critical examination to handle social change and improve the accuracy of the original designed model (Barry 2012). Figure 2.3 depicts the major classes in the TTM and how they relate to each other.

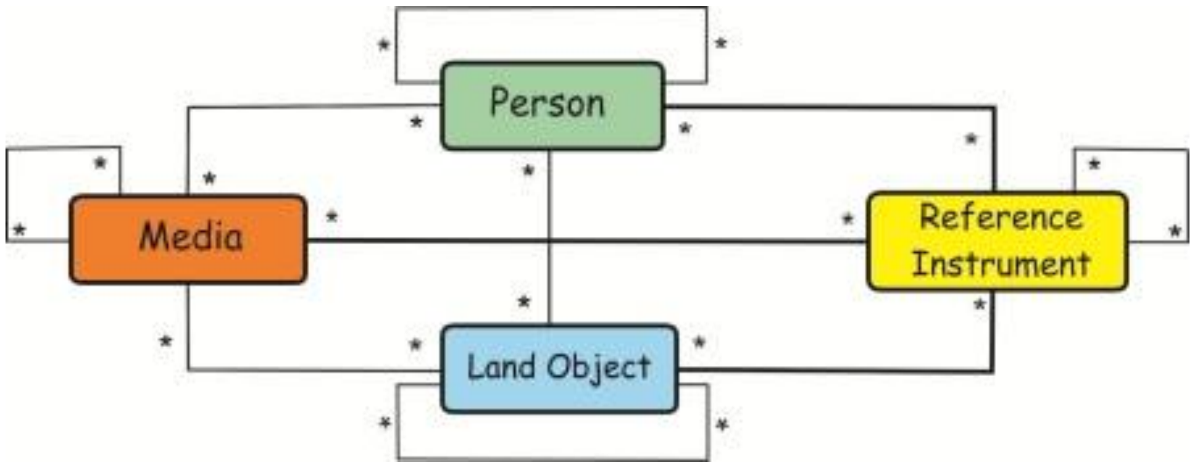


Figure 2.3: Talking Titler Model (TTM) (Barry et al. 2012 - used with permission).

2.5 Chapter Summary

This chapter explained the importance and developing a LTIS that can evolve with changing scenarios in uncertain conditions. It explained the ingredients needed for evolutionary land information system design based on an evolutionary land administration model. It also explained why an evolutionary software development approach is imperative to developing a LTIS for uncertain situations using an evolutionary land administration model. The chapter explained the agile software development approach and its phases and why the methodology is suitable for developing a LTIS for uncertain situations. The chapter concluded by describing some of the land administrations models that could serve as the basis for LTIS development. The described models were the Core Cadastral Domain Model currently known as the Land Administration Domain Model, Social Tenure Domain Model and the Talking Titler Model (TTM). TTM was chosen for this study because it is flexible and designed to suit peri-urban communities. Also it is being developed by the author's research group, the Land Tenure and Cadastral Systems Research Group at the University of Calgary. Primarily, the chapter addressed specific objective 1 and research question 1 in section 1.4.

Chapter Three: Analysis and Mining of Social Networks

3.1 Introduction

Land tenure information system data is characterized by complex network interactions between entities. The entities are mainly people (e.g. individual, organization, community), land object and media items (any form of evidence in support of land transaction). As the entities interact, different forms of relationships are created. The relationships (which could also be referred to as relations (Wasserman and Faust 1994)) between entities occur during everyday land transactions such as when a land parcel is sold, leased, bought, registered or surveyed. Examples of such relationships between entities include, but are not limited to, person-to-person (e.g. land owner sells land to buyer or land owner borrows money and uses land as collateral), person-to-land (land owner leases land out to a stranger), and person-to-evidence (land owner's title/deed provides evidence of a transaction).

Ensuring tenure security in peri-urban communities requires an accurate representation of the networks in the LTIS data model. As the number of interactions increases, a small and simple network may evolve into a larger network, because of the rapidly evolving conditions that exist in peri-urban communities. In a larger network of tenure relationships, it may be difficult to trace, or even understand and model, the inter-relationships. This could result in missing the not so obvious relationships and patterns in the LTIS data and, thus, incorporating an inaccurate model of the social network.

It is therefore essential to employ a technique that could identify important factors such as power players, and bottleneck actors (gatekeepers), and provide a pictorial representation of the representation of the network, no matter how complex, to analyze, understand and utilize LTIS data. In this regard, social network analysis and data mining techniques are utilized in this research for TTM-based LTIS data analysis to aid in answering the important questions raised in chapter one (See chapter one, section 1.4, 2-4).

This chapter discusses the two important ingredients for extraction and functional analysis of LTIS data; namely, social network analysis and data mining. It begins by defining and explaining the fundamentals of social network analysis. It then explains the perspectives of social network analysis and why the technique is useful in analyzing LTIS data for peri-urban communities. The chapter continues to explain the concepts and techniques of SNA adopted in this research. A general description of the functionalities of selected SNA tools including NetDriller are then profiled. The reason for the tools description is to provide information on some existing SNA tools and their functionalities. The chapter then concludes with a brief discussion of data mining techniques deemed necessary to augment social network analysis of LTIS data, namely, association rule mining and clustering. It explains why the techniques are useful in this regard and how they contribute to the analysis in this research. It is critical because it will aid in creating the appropriate networks embedded in the raw LTIS data. A further objective is that important relationships are not missed when networks are created during the analysis. There are however no inherent biases in choosing the SNA and data mining techniques selected for this study.

3.2 Social Network Analysis

Social network analysis is the study of social networks to gain an understanding of their structure and behavior (Srivastava 2008). A social network is a "social structure of individuals or entities related directly or indirectly based on common relation of interests such as trust and friendship" (Srivastava 2008). The characteristics of a network may include any form of relation (actual biological relation, affection, dependency, power, support, and kinship) between any type of social actors (e.g. people, organizations) through mediating entities (e.g. joint participation in formally or informally constituted groups, groups where members hold shared beliefs) (Brandes 2008). The relations in social networks are between a set of actors such as individuals and defined land parcels (Brandes 2008). Basically, a network involves a set of individual actors with ties to each other. Each tie represents a social relation. The form of social relations is multi-faceted, in that a given relation could provide connection to an entirely different network. According to Borgatti (2004), the elements that comprise a system are not as important as how these elements are put together. This means that individual actors themselves may be of little importance but that, once they interact with each other, a complex network is created that needs to be analyzed for interesting patterns. Interesting patterns include the correlation between entities in the data. In other words, the most important aspect of a network is the multitude of ways in which the actors interact and relate with each other.

Through SNA, a visual and mathematical analysis of relationships between entities within a network can be conducted (Wasserman and Faust 1994). Social network analysis maps out relationship patterns and measures the flow between actors (e.g. individual, group, animal, and organization) and other knowledge processing entities (Wasserman and Faust 1994). Actors in a network are represented as nodes and the link shows the ties, relationship or flow between the nodes (Wasserman and Faust 1994).

Wasserman and Faust (1994) present the following four perspectives of SNA which in the author's opinion make it an ideal technique for examining LTIS data:

- Actors and their actions are viewed as interdependent rather than independent, autonomous units

- Relational ties (linkages) between actors are channels for transfer or "flow" of resources (either material or nonmaterial)
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action
- Network models conceptualize structure (social, economic, political, and so forth) as lasting patterns of relations among actors

These perspectives complement Borgatti's (2004) opinion that the focus of network analysis is not on individual units but on the group of actors and the means, norms and protocols by which they relate with each other. The perspectives are seen in the way network methods are employed in SNA. Wasserman and Faust (1994) explain that network methods are centered on dyads, triads or larger systems (see section 3.2.1 for further explanation of dyads, triads and larger systems). This means that actors in a network relate or ties to the others (all or some actors) in patterns that form a structure. The structure of the network can therefore be represented by modeling the relationships between the actors. By so doing, one could understand the effect the structure has on individuals in the group and on the group as a whole. It is noteworthy that relations within a network evolve with time, and these changes can be studied through techniques and concepts of SNA.

3.2.1 Concepts of SNA

The fundamental concepts of SNA include actor, relational tie, dyad, triad, subgroup, group, relation, and network (Wasserman et al 1994). An actor is an entity (e.g. an individual person) in a network. When the collection of actors in a network is of the same type, such as the relationship between individual land owners in a community, the network is said to be a one-mode network. A two-mode network describes the situation where the collection of actors is of two different types such as the relationship between individual land owners and their respective lands. A two-mode network can be converted into a one-mode network and the process is referred to as folding. When two pairs of actors are linked to each other by some form of relationship, it is said that there is a relational tie between them. The work entailed in defining relational ties could be extensive and depends on the intricacy adduced by the network data and the purpose of conducting the analysis. A few examples are formal relations through status in the group (e.g. hereditary authority such as a chief), biological relations (e.g. kinship involving members of the nuclear, extended and adopted family group), status or locational change (e.g. migration) and transfer of material resources (e.g. lending or borrowing) (Wasserman and Faust 1994).

A dyad consists of a pair of actors and the possible relational ties existing between them. A triad is the relational ties between a subset of three actors. A subgroup of actors refers to any subset of actors and their relational ties. A group is defined as a collection of all actors in a finite set on which relational ties are determined and analyzed. Defining a group calls for identifying the restriction(s) on who does or does not belong in it. A network can contain two or more groups of actors. However, there can only be one group in a one-mode network. Different kinds of relational ties may exist among actors. The collection of a particular kind of ties existing among actors in a group is referred to as relation. Therefore, a social network is made up of a set or sets of actors and the defined social relation or relations among them (Wasserman and Faust 1994). The aforementioned concepts in SNA are greatly influenced by graph theory. The next section provides a brief insight into graph theory.

3.2.2 Graph Theory

Social network analysis is based on graph theory. To understand the results and output from any SNA would require a brief knowledge on graph theory. As explained by Wasserman and Faust (1994) and O'Madadhain et al (2005), graphs consist of a set V of n nodes or vertices with a set E of m edges between them. Vertices (also known as nodes) are usually denoted as $\{V_1... V_n\}$ whiles edges (also known as arcs, links, ties or connections) are denoted as $\{e_{i,j}\}$. The number of vertices is represented as |V| whiles that of edges is |E|. In this text, vertices (vertex) and nodes are used interchangeably. When a network is constructed, vertices or nodes represent entities or actors (e.g. a land parcel – an entity or a person – an actor) whiles edges represent relationships between nodes. Two nodes connected by an edge are adjacent to each other. The degree of a node is the number of nodes adjacent to it and is denoted as $d(n_i)$. A graph with values or weights assigned to edges is called a weighted graph. The values on the edges are called weights; these are categorical values (e.g. friendship, parent-child) and do not signify the importance of a relationship. A graph in which every edge has a head (starting point) and tail (ending point) is called a directed graph. An undirected graph has neither head nor tail. Two edges that connect to the same vertex or two adjacent vertices with the same direction are said to be parallel. The individual nodes and edges connect each other to form a graph which in turn represents a network. The properties of a graph and a network in general are explained in the next section.

3.2.3 Properties of Graphs and Networks

As explained by Wasserman and Faust (1994), the set of nodes and the existence or non-existence of edges between pairs of nodes in a graph is called a sociogram. A sociogram represents a network of nodes and their corresponding ties. Not all nodes in a graph may have ties to every other node. Sometimes there may be fewer edges than there could actually be in a graph. Thus, the density of a graph is the ratio of the number of relations currently present in the graph to the number of possible relations that could exist in the graph. The density of a graph can be used to examine the degree of cohesiveness of subgroups in a network. Also, the distance between nodes – called geodesic distance - can be used as an inferential rule to model the existence of cohesive subgroups in a network. The longest geodesic distance between a node and any other nodes in a network is referred to as the eccentricity of that node. The diameter of a connected graph is equal to the largest nodal eccentricity. A graph with a path between every pair of nodes is said to be connected. A graph whose nodes have been divided into two disjoint sets such that every edge in one set connect to a node in the other set is called a bipartite graph. However, whether a graph is connected, directed or non-directed, there are some measures that are employed in SNA to help identify the roles of individual nodes in a network.

3.3 Central Measures in SNA

Actors serve different purpose in a network, and some actors are more prominent than others, in the sense that they are the most influential in the network (Wasserman and Faust 1994). The importance of an actor is closely related to its location in the network. An actor’s location could either be to its advantage or disadvantage (Hannerman et al. 2005). According to Wasserman and Faust (1994), an actor’s prominence is mostly determined by the number of ties transmitted (out-degree) or received (in-degree) by that actor. However the best way to determine an actor’s prominence is to examine indirect ties in addition to “choices” made or received by the actor. This means that an actor’s prominence is determined by his involvement with others in the network. An actor which is extensively involved with other actors in a network may have many ties; it may lie in a central position and, therefore, may be prominent in the network in terms of controlling the way information flows and resources are shared within the network.

An actor which receives a lot of ties from others is said to be prestigious. Prestige is measured in a directional relation. Most actors in a network direct their ties to such a prestigious actor. Prestige could be either positive or negative, depending on the type of relation being measured. For instance, if the relation in question has a negative connotation, such as hate or enmity, then the prestigious actor is despised by most actors in the network. On the other hand, if the relation is positive, such as a lovely or good friend, then the inference drawn is that the prestigious actor is considered as lovely or a good friend by most actors in the network. An actor's prestige can be measured by examining the ties directed to the actor by other actors in in-degree ties (Wasserman and Faust 1994).

These qualities - importance, prestige and prominence - are directly related to the authority, influence or power an actor wields in a network. The question is how, in the context of a given network, the importance, authority, influence, or power of an actor are measured? When actors with similar features in a network form groups, how are such group structures to be understood? SNA provides methods in determining these metrics, which are referred to as centrality measures. Centrality measures include degree centrality, closeness centrality, and betweenness centrality (Wasserman and Faust 1994).

3.3.1 Degree Centrality

Degree centrality measures the importance or prominence of an actor in a network, and is related to the degree to which an actor is involved in the interactions in the network (Wasserman and Faust 1994). An actor with high degree centrality is active in the relational process and is a "crucial cog in the network" (Wasserman and Faust 1994). Such an actor has more ties and has greater advantage and opportunities than others in the network (Hannerman *et al*. 2005).

According to Wasserman and Faust (1994) and Hannerman *et al*. (2005), actors in a strategic position with many ties in a network are less dependent on other actors in satisfying their needs. By virtue of their centrality in the social network, they provide, and are perceived as, the means through which other actors transmit information or share resources which may give them some control over access to resources in the network. People usually want to associate with important people. Thus, in a directed network, an actor who receives many ties (in-degree) is said to be prominent or prestigious. Also, actors with a high number of out-degree ties are said to be influential, in that, they are able to make other actors aware of their opinions. Degree centrality is not measured only for individual actors but could also be measured for a group or a set of actors in a network, who are associated by commonality in terms of relationship with each other (e.g. individuals belonging to the same family). This also helps to determine which groups are most influential among the different sub-groups in the network. An actor's degree centrality according to Wasserman and Faust (1994) is calculated as below:

$$C_D\ (n_i) = \frac{d(n_i)}{g-1}$$

where $C_D\ (n_i)$ represents the degree centrality of node i; $d(n_i)$ represents the degree of node i; and $g-1$ represents the proportion of nodes adjacent to node n_i in the network.

3.3.2 Closeness Centrality

Closeness centrality focuses on the indirect ties an actor has with the other actors in the network (Hannerman et al 2005), and is based on how physically close an actor is to all other actors in the network (Wasserman and Faust 1994). An actor's distance from other actors is inversely proportional to the closeness centrality measure; as the distance between the actor and other actors increases, the actor's closeness centrality measure in the network decreases, and vice versa. An actor who can reach or be reached by other actors in a shorter distance may be the center of attention or reference point in sharing views or resources. Such an actor lies in a strategic position in the network and may rely very little on others for resources. Also the actor is the "closest" to all the others in the network, and can thus be reached quickly and easily by the other actors. Thus the actor may be the best medium for information transfer and resource sharing. The closeness centrality measure of an actor to others is an advantage and could be translated into power, authority or superiority (Hannerman et al. 2005). As adopted from Wasserman and Faust (1994), it is computed as follows:

$$C_c(n_i) = \frac{g-1}{\left[\sum_{j=1}^{g} d\left(n_i, n_j\right)\right]} \qquad i \neq j$$

where $C_c(n_i)$ represent the closeness centrality measure of a node *i* in the network; $g-1$ represent the maximum number of actors adjacent to the node; and i, and $\sum_{j=1}^{g} d\left(n_i, n_j\right)$ represents the sum of the distances from actor *i* to all others in the network.

3.3.3 Betweenness Centrality

Wasserman and Faust (1994) explain that the vital notion about betweenness centrality is that an actor is central if it lies between the others via their geodesics (shortest path or distance). This potentially gives the actor an increased degree of control over interactions as compared to other non-adjacent actors. According to Hannerman et al (2005), when an actor lies between two or more actors, the actor becomes a mediator through which information or resources are shared. This gives the actor an advantage over other actors in terms of authority and influence. The actor could isolate and prevent some actors from contacting and sharing information with other actors. Being a mediator provides the actor with alternative channels, making the actor less dependent and therefore more powerful than the other actors in the network. The betweenness centrality measure according to Wasserman and Faust (1994) is computed as follows:

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk} \qquad (1)$$

$$C'_B(n_i) = C_B(n_i)/[(g-1)(g-2)/2] \qquad (2)$$

In equation (1), $C_B(n_i)$ represents the betweenness index of a node i and is the "sum of estimated probabilities over all pairs of actors not including the *i*th actor"; $1/g_{jk}$ is the probability that any of the geodesics that connect two nodes j and k will be used during communication between nodes *j* and *k*; $g_{jk}(n_i)$ is the number of geodesics that connects nodes *j* and *k* by passing through a distinct node *i*.

In equation (2), $(g-1)(g-2)/2$ represents the number of pairs of actors in the network excluding the *i*th node. $C'_B(n_i)$ is the standardized betweenness centrality measure for a given node, i.

The centrality measures – degree, closeness, and betweenness - explained above and many other important measures in SNA are implemented in various social network analysis tools. The next section discusses four of the many SNA tools available. The tools discussed here are UCINET, JUNG, NetMiner and NetDriller.

3.4 SNA Tools

There are several social network analysis tools. Some are open source, while others are proprietary. In a study conducted to compare the scalability and functions of six social network analysis tools (UCINET, Pajek, Networkx, iGraph, JUNG and Statnet), Xu et al. (2010) found the tools to be relatively "complementary" and not "competitive" in that the shortcomings of one tool may be compensated for or improved by another. The responsibility therefore, lies on an analyst to research various tools to find which one may be suitable for the task at hand. The criteria for tool selection may, thus, be influenced by capabilities and cost of the tool, along with the researcher's familiarity with available tools. On the basis of cost, functionalities, and familiarity, this study selected and used NetDriller to analyze LTIS data in the Talking Titler System. Below is a description of the four SNA tools listed above.

3.4.1 UCINET

UCINET software is used for analyzing social network data (Borgatti *et al*. 2002). It can compute various measures for analysis including centrality, cohesion, brokerage and even hypothesis testing (Borgatti *et al*. 2008). It includes important basic graph algorithms needed for complex network analysis such as the Dijkstra algorithm which is used for calculating the shortest path between two nodes (Xu *et al*. 2010). A social network visualization tool called NetDraw is integrated into UCINET. Through this tool, the software is able to provide a graphic representation of networks including relations and attributes (Borgatti *et al*. 2008). The software allows flexible importation of data from Microsoft Excel, text, and VNA files for analysis (Borgatti *et al*. 2008). It can conduct analysis on both small and large data sets with a maximum network size of about 2 million nodes (Borgatti *et al*. 2002). It performs various functions needed for comprehensive analysis; data transformation (e.g. transpose, dichotomize, symmetrize, normalize, and reverse), centrality measures (multiple centrality measures at node level), graph theoretic (e.g. bipartite, incidence, line graph, multi graph, semi group) and hypotheses testing (e.g. node level regression, QAP (quadratic assignment procedure) correlation, and MRQAP) (Borgatti *et al*. 2002). Although UCINET is a capable and comprehensive complex network analysis tool, it is not adopted for the research because it is a proprietary tool and it does not utilize any data mining techniques in network construction.

3.4.2 JUNG

JUNG is an acronym that stands for Java Universal Network/Graph Framework. It is a free, open-source software library on which "applications for network data manipulation, analysis, and visualization" can be built (O'Madadhain et al. 2005). The JUNG library, similar to many SNA tools, makes use of graph theory and supports innumerable representation of nodes and their relations. JUNG can represent directed, undirected, multi-modal hyper-graphs and multi-graphs. The JUNG library generates graphs, entities, and relations annotated with metadata, thus making it easy to create analytic tools for complex data sets. JUNG also includes algorithms for data mining and social network analysis. This also means that JUNG can perform numerous functions including network clustering, decomposition, optimization, random graph generation, statistical analysis, distance calculation, flows and various vital measures such as centrality. JUNG also provides a mechanism to build tools that can interactively visualize and explore network data and filter generated graphs. (O'Madadhain et al 2005) Although JUNG is a comprehensive SNA tool, it was not chosen for the study because the JUNG libraries are utilized in NetDriller, which is the tool chosen for this research.

3.4.3 NetMiner

NetMiner is a social network analysis tool that can interactively explore, visualize and discover patterns and structure of network data (Cyram 2012). According to Cyram (2012), the tool has a comprehensive data model that can express different kinds of nodes and links with their attributes. It can create and represent network data from several social, natural and physical phenomena. The output from conducted analysis are produced both in numerical and graphical form, and allows output to be explored interactively. NetMiner makes it easier for an analyst to understand features of a conducted analysis by categorizing and presenting results of conducted analysis as a table, map, or chart report. NetMiner is capable of creating and saving sessions for users conducting various analyses, so that the user can re-use stored processes in the future, and is able to reuse output data from conducted analysis as input data for a subsequent analysis. It provides a platform where analysts can write their own scripts or algorithms for analysis.

NetMiner can manage, transform network data as well as visualize and calculate various network measures. NetMiner has several ways of transforming data into various forms including transpose, reverse, reorder, link reduction, vectorize, merge, tree construction and QAP permutation. The tool can conduct several comprehensive and detailed analyses including calculating various centrality measures (degree, betweenness and closeness) and finding connection between nodes by calculating shortest path. NetMiner has visualization algorithms and can visualize data in both 2D and 3D. The software can be used to conduct various types of statistical analysis including Anova, correlation, regression, and crosstabs on network data. NetMiner allows a connection to a database (SQL or Oracle) so that data could be extracted and analyzed. Various tools including query composer, matrix calculator and graph editor allow the analyst to extract nodes and links, perform matrix calculations and modify generated graphs.

The NetMiner viewer is a free tool, which allows users to distribute results of analysis from NetMiner and, thus, furthering the sharing and analysis of data in a wider community of practice.

Although NetMiner is a comprehensive SNA tool, it was not chosen for this study because as at the time of this study, it did not employ data mining techniques such as clustering and association rule mining (see sections 3.5.2 and 3.5.2 for definitions) in network construction. As well, NetMiner is proprietary software and, thus, not available for customization. NetMiner has a specific file format, known as .nmf, which may reduce scalability with other SNA tools. Students are permitted to download and use it freely for coursework or research for a period of six months.

3.4.4 NetDriller

NetDriller is freely available software developed to fill some of the gaps lacking in existing social network analysis tools (Koochakzadeh *et al.* 2011). Unique among this class of software tools, NetDriller is able to analyze raw data in order to construct a network by applying data mining techniques (clustering and association rule mining) to make meaningful knowledge discovery. These techniques help to identify and group nodes with similar attributes and features. Association rule mining enhances analyses of network data that contain many-to-many relationships. By applying fuzzy search on network metrics, NetDriller can identify, isolate and apply selected network metrics on specific nodes. Fuzzy search is a technique used to identify and “deal with vague or inexact facts” (Han *et al.* 2012). NetDriller is attractive as it incorporates the source code of both JUNG (described above) and Weka, while Weka is a free data mining software with tools for data classification, clustering, visualization and many more (Hall *et al.* 2009). NetDriller utilizes the java libraries and data mining techniques in these tools to enhance its capabilities. As outlined by Koochakzadeh *et al* (2011), NetDriller can analyze a pre-defined network or infer and create a network from a raw dataset. Numerous types of analyses, including filtering links, finding shortest path, bridges and cliques, clustering nodes of a network, searching and measuring network metrics at node level can be conducted on network data imported into NetDriller. This software can also track and show changes in network over time, a feature which is vital to studying evolving networks associated with LTIS data.

NetDriller was chosen for this study because the tool applies data mining (association rule mining and clustering) and social network analysis techniques in network construction and metric computations. This feature is important for this study because the TTM-based LTIS data consists of many-to-many relationships between people, land and media files. These entities (people, land, and media) relate to each other by means of some shared features. Data mining techniques such as association rule mining and clustering will aid in grouping particular entities with common features together.

The researcher's familiarity with the tool is another reason why NetDriller was chosen. Studying under Prof. Reda Alhajj, who supervised the team that developed NetDriller, allowed access to the source code and experience in writing a functionality (saving graphs and calculated metrics to a file) that is specifically applicable to this study.

3.5 Data Mining

Data mining refers to the discovery of interesting patterns and useful information from large data sets (Elmasri and Navathe 2011, Han *et al*. 2012); through various techniques, it can help to discover hidden patterns that are not immediately obvious or which cannot be revealed by simply querying the data (Elmasri and Navathe 2011). It employs statistics and machine learning techniques to discover useful information from data sets in areas including manufacturing, medical, pharmaceutical, and customer relationship management. Data mining techniques such as classification, characterization and discrimination have been applied successfully to identify shopping patterns, customer groups for marketing purposes, and to discover unusual behavior in support of fraud detection (North 2012, Han *et al*. 2012).

Choosing the appropriate technique depends on several factors including the purpose of the analysis and expected outcomes. Because land tenure data is composed of interrelated entities such as an individual land owner's relationship with the land and deed, this study was interested in applying and using a tool that incorporated association rule mining and clustering for analysis. The next section specifically describes how these two techniques work and how useful they may be in analyzing TTM-based LTIS data in peri-urban areas.

3.5.1 Association Rule Mining

Association rules are data mining techniques that identify frequent patterns in a data set (Han *et al*. 2012). This technique has been employed successfully in what is referred to as the market basket analysis where customers' purchasing habits are studied to determine which items are often purchased together (Han *et al*. 2012). With insights into purchasing patterns, recommendation could be made to customers once they purchase certain items. In other words, the technique makes it possible to predict what a customer may purchase in addition to buying particular items. The technique is explained further below according to Han *et al*. (2012) using practical examples related to this research.

In order to determine family lineage, for instance, suppose we have data about several individuals and their families stored in a relational database. If analysis on the stored data shows that, most of the time, a person related to person A is also related to person B, then there is a probability that a third party C related to A may also be related to B (referred to as the commutative principle in this study). In such a case, the following association rule may apply:

$$relation(person, \text{"A"}) \Rightarrow relation(person, \text{"B"})\ [support = 1\%, confidence = 50\%]$$

A confidence of 50% means that if a person is related to A, there is a 50% probability that the same person may also be related to B. The 1% support means that, 1% of the data being analyzed show that persons A and B are always related. The rule above will be considered interesting if the support and confidence values satisfy the minimum thresholds defined by the analyst. In general, association rule mining involves – 1.finding all frequent itemsets, that is - different items that often occur together. 2. Extracting strong association rules from the frequent itemsets.

Thus, in a dataset T containing sets of items A and B, the $support\left(A \Rightarrow B\right) = P(A \cup B)$ represents the percentage of records in T that contains both A and B. The $confidence\left(A \Rightarrow B\right) = P(B/A)$ represents the percentage of records in T containing A that also contains B. The occurrence frequency of an itemset in T (also known as the support count) is the numbers of transactions that contain the itemset. Hence, the associations rule A⇒B is expressed in terms of a conditional probability as follows:

$$confidence\left(A \Rightarrow B\right) = P(B \setminus A) = \frac{support(A \cup B)}{Support(A)} = \frac{support\_count(A \cup B)}{Support\_count(A)}$$

Association rule mining could be useful in analyzing LTIS data in that the technique will be able to identify frequent patterns as entities interact. This will make it easier to connect entities who share common interests. For instance, it could aid in identifying individuals, agencies or institutions that have similar interests in the same piece of land. If two individuals or institutions have purchased or leased some properties having some specific features, it may be easier to determine the next property these entities may want to purchase. Especially if such institutions or individuals have been involved in fraudulent purchases, then knowing the next properties they may be interested to buy, it may be possible to secure a third property. The commutative principle may not only be useful in determining family lineage but also succession in a family, clan or community at large.

3.5.2 Clustering Analysis

Clustering is the "process of grouping a set of physical or abstract objects into classes of similar objects", thus making a cluster a set of data objects that are similar to each other and, at the same time, dissimilar to data objects in other clusters (Han *et al*. 2012). Clustering may be used to discover areas of similar land use during earth observation analysis, as well as identifying groups of houses in a geographical location based on certain parameters including a house's value and type (Han *et al*. 2012). Generally, clustering involves - 1.dividing data into groups using a similarity measure. 2. Applying distinct labels to the sub-groups. Several methods are used in finding similarity or dissimilarity among objects in a dataset including determination of Euclidean distance (Han *et al*. 2012, Elmasri and Navathe 2011). The Euclidean distance measure is represented by the equation below:

$$d(i,j) = \sqrt{(xi_1 - xj_1)^2 + (xi_2 - xj_2)^2 + \cdots + (xi_n - xj_n)^2}$$

$where\ i = (xi_1, xi_2, \ldots., xi_n)$ and $j = (xj_1, xj_2, \ldots, xj_n)$ are two n-dimensional data objects (Han et al. 2012). Clustering may be useful in community detection in a network, where community refers to different kinds of groups in the LTIS datasets. Where community refers to a family, it implies that individuals from the same family could be placed in the same group or cluster based on the family association. Once clusters of entities are created, the lines of interactions both within and across the groups may become clearly visible (Han *et al*. 2012, Elmasri and Navathe 2011).

Chapter Summary

This chapter has discussed mining and analysis and their importance in analyzing LTIS data. The chapter defined and explained the concepts of social network analysis. It also described the graph theory, the properties of a graph and three central measures in SNA. It also examined some of the SNA tools, highlighting their capabilities and the reason for selecting NetDriller for the study. It explained the main data mining techniques and why they are necessary for this study. The next chapter will describe how the evolutionary software development approach (discussed in chapter 2) was utilized in developing the ubiquitous TTM-based LTIS. The chapter will also describe the methodology for utilizing the techniques described in this chapter (chapter 3) in conducting social network analysis.

Chapter Four: Methodology

4.1 Introduction

This chapter describes the activities conducted to addresses the specific research objectives (see chapter one, section 1.3). Specific objective 1 is dealt with in the first section (i) by describing the steps involved in adopting the phases of extreme programming to developing the ubiquitous TTM-based LTIS; and (ii) by explaining the rationale behind choosing the tools (PostgreSql, Apache) and the programming languages (PHP, Java, HTML) used in system development. These activities address research question 1 in chapter one. The section concludes with a description of the architecture for the design and development of the LTIS. The system architecture is the conceptual framework that outlines the structure of the developed TTM-based LTIS (Hill *et al.* 2000). It shows how the system functions which eases development, modularity and troubleshooting.

The second section of the chapter describes the activities involved in mining and analyzing the TTM-based LTIS data. These activities address questions 2, 3, 4 and 5 in chapter one section 1.4. The section describes two scenarios that governed the data stored in the TTM-based system for analysis. It then discusses the steps involving how data was populated, extracted and prepared for analysis using NetDriller. Ultimately, the section describes the activities conducted to achieve specific object 2; to determine the lineage of interest in a customary setting (see chapter one, section 1.3a).

4.2 Developing TTM-based Ubiquitous LTIS Using Agile Methodology (XP)

4.2.1 Applying XP to developing TTM-based LTIS

There are five phases involved in adopting extreme programming in developing the LTIS: exploration, planning, iteration, product ionizing and death. These activities are not done in isolation from each other but, rather, are done concurrently allowing the overall development process to benefit from feedback and synergistic effects. The exploration phase involves understanding and defining requirements for the proposed system. Planning is concerned with choosing and prioritizing requirements. The iteration phase describes the architecture, initial development and testing of the prototype for limitations. The product ionizing phase discusses system reliability, performance and maintenance. The death phase describes instances where additional features are not required in the developed system.

4.2.2 Exploration Phase

To develop a LTIS to be used in a peri-urban community, it is imperative to understand the social and economic structures and dynamics in such communities. This understanding was gained in consultation with research conducted in peri-urban societies in emerging economies with focus on Accra (Danso and Barry 2012, Barry et al 2010, Barry et al 2007, Berry 1993). The Talking Titler Manual (Barry 2011), which also contains descriptions of scenarios that mimic the conditions in peri-urban communities, was consulted to anchor the research in the context of an actual software development initiative. From this basis, having gained a general familiarity with what goes on socially, economically and politically in peri-urban communities, the initial requirements and features for the LTIS were identified.

The first requirement was to choose a data model. Since the conditions in peri-urban communities are complex and dynamic (see chapter one section 1.5 and chapter two section 2.4), a data model that could evolve with the changing conditions was necessary. Hence, the Talking Titler Model (see chapter two section 2.4.3) was chosen as the data model for the study. To effectively utilize the TTM as a framework for LTIS development, the initial classes in the TTM were studied in-depth to understand their functions as well as the attributes allowed in the model.

After choosing the data model, the initial features of the system were then documented. These features included adding and relating any of the four objects (person, land, media, and reference) allowed in the TTM in a many-to-many fashion. Based on the TTM, the media object should at least include audio, video, images in various formats, and documents in Microsoft Word, Excel and PowerPoint formats. This will make it easier to collate any form of evidentiary information (structure, semi-structured and unstructured) to support land transactions. The other features included editing and deleting records when necessary, and storing every data from the system in a relational database. Following this, a decision was made to develop a ubiquitous (web- and mobile-based) LTIS. The choice to develop a ubiquitous LTIS is due to the existence of an enabling environment for such a technology in most peri-urban areas of Ghana.

A survey conducted by the National Communication Authority (NCA) in Ghana – a government communication and regulatory authority - showed that more than 90 percent of the national population had access to internet either through a mobile phone or a computer. Also over 100 percent of the population are mobile phone users with the ability to access the internet (NCA 2012). “There are more mobile phones than people in Ghana” – said Bill Gates on his visit to the country (Gates 2012), illustrating the prevailing understanding among users, developers and regulators that the existing infrastructure readily enables a web and mobile application accessible in the country.

Developing a ubiquitous LTIS, as opposed to a stand-alone application, implies that users will not have to install any application before use. Since the application is accessible through a web browser, all a user needs is internet access. Centralizing data storage and system maintenance functions on the server hosting the developed LTIS streamlines system operation, troubleshooting and updating. The amount of expertise needed by users in accessing and using the application is also minimal. It is also cost effective to develop the application because, apart from the expertise and time needed for the design and development, there are many free but efficient tools, such as PostgreSql database and Apache web server that are readily suited to the task.

The PostgreSql object-relational database management system (ORDMS) was chosen as the database for the LTIS because it is a powerful, non-proprietary open source object-relational database system noted for data reliability, integrity and correctness (PostgreSql 2011). PostgreSql runs on various operating systems including Windows. This feature was essential to the development of the LTIS since Windows is one of the most common operating systems in most emerging economies including Ghana (Onbile 2012). PostgreSql is fully ACID (Atomicity, Consistency, Isolation and Durability) compliant and has full support for foreign keys, joins, views, triggers, and stored procedures in multiple programming languages (PostgreSql 2011). Also, the development of a LTIS that stores media items allows database size to grow from a few megabytes to several gigabytes or even a terabyte. PostgreSql is capable of handling in excess of four terabytes of data (PostgreSql 2011). Furthermore, the PostGIS extension to PostgreSql adds support for geographic objects in the object-relational database, and allows it to be used as a spatial database for geographic information systems (GIS) (PostgreSql 2011). These spatial capabilities facilitate the development of a LTIS which may in future evolve to incorporate spatial information including GNSS-coded (Global Navigational Satellite System) points in mapping out geographical locations for land administrators.

Next, it was necessary to define the requirements of a prototype that will have the functionality mentioned above and to produce its basic elements. This included creating the database schema, designing and creating the LTIS database, creating a simple user-interface with the ability to add, relate, edit, and delete records of objects where necessary. This called for a decision on which classes or sub-classes in the TTM to include or remove from the initial prototype. The person, land, media and reference classes were used to design a high level conceptual data model. Figure 4.1 shows the entity relationship (ER) diagram depicting how data is stored in the PostgreSql database. The four main objects (person, media, property or land, reference) were adopted as the major tables or relations. The relationship between objects was stored in another intermediate relation. Although the reference instrument class may not be of use at this early stage, adding it at this point had no negative connotation on the project and may save a lot of work in future should it become necessary to utilize that class. The initial prototype of the TTM-based LTIS was then developed.

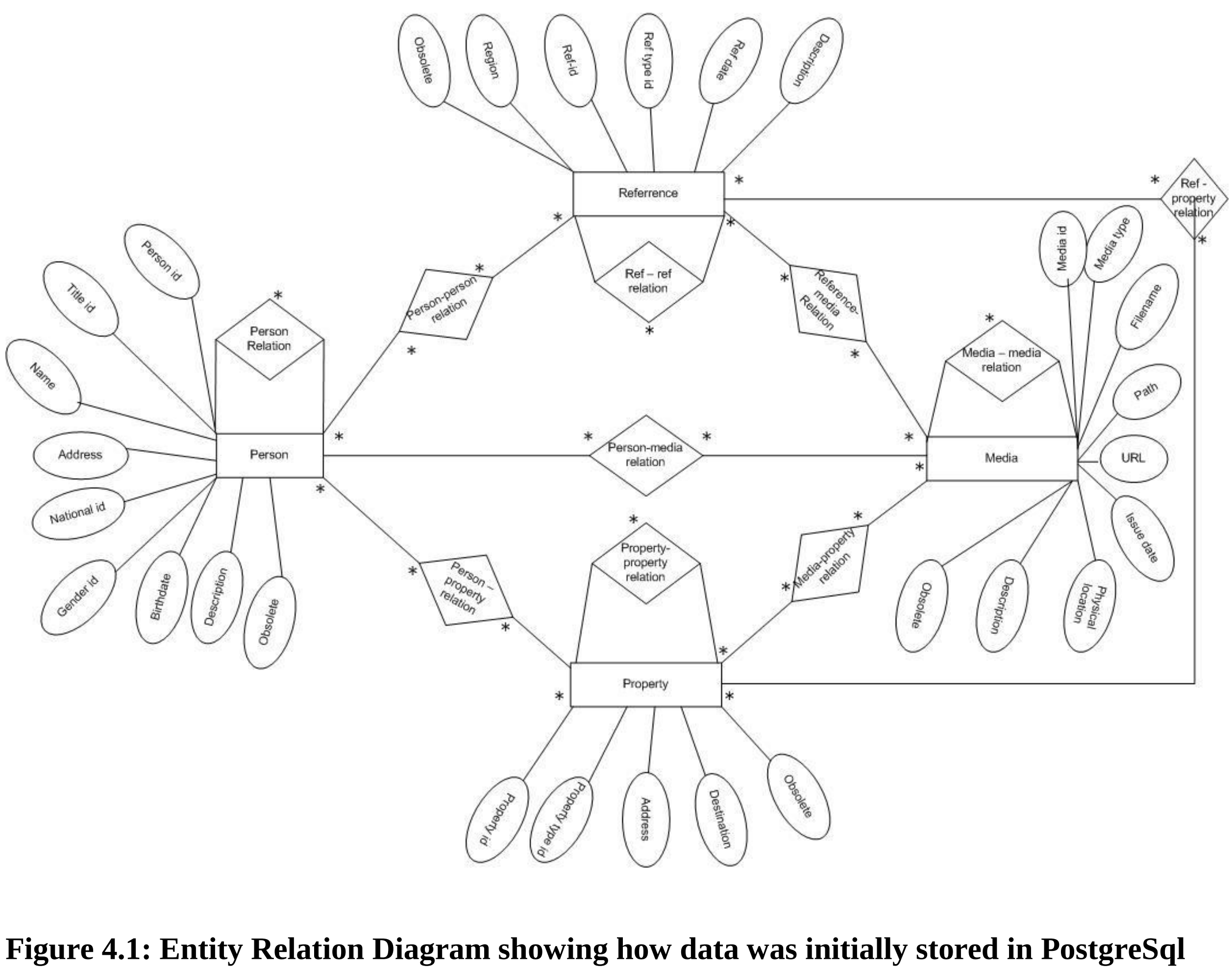


Figure 4.1: Entity Relation Diagram showing how data was initially stored in PostgreSql

4.2.3 Planning Phase

Based on the features listed in the exploration phase, the decision was made in the planning phase to execute the following in order to develop the LTIS with the desired features:

1. Create the database schema
2. Create the LTIS database
3. Create the user-interface with the ability to add objects
4. Create the user-interface with the ability to relate objects
5. Create the user-interface with the ability to edit and delete objects.

These activities were conducted during the design and development of the initial prototype.

4.2.4 Iteration Phase

The developed LTIS was taken through numerous iterations before the first version was ready for use. Meaning that, the initial features which were penned down in the exploration and planning phases were developed. The initial iteration was focused on designing the basic system architecture for the LTIS. As shown in Figure 4.2, a user or client makes a request through a computer or mobile device with internet or intranet access to the LTIS application. The user's request is sent via HTTP (Hypertext Transfer Protocol) to the Apache web service.

Apache was chosen because it is a free open source web server that supports authentication schemes and server side programming languages such as PHP. Also Apache contains popular compression methods such as an external extension module and mod-gzip (Apache 2011). This latter feature is important because of the volatility of the mobile phones and internet infrastructure in emerging economies such as Ghana; although accessible except in the most remote regions, the available speed and bandwidth of these networks may be compromised by factors such as noise on transmission lines and network failures, thus causing latency and attenuation leading to intermittent internet connections. In addition, GPRS (general packet radio service) is typically the standard wireless communication protocol used in emerging economies and, although it offers an efficient use of limited bandwidth particularly suited for sending and receiving small and large bursts of data, GPRS connections in many emerging economies may not be enhanced to provide faster third generation (3G) and EDGE throughput speed. Hence data communication via the web or mobile devices may be inhibited in terms of speed. Therefore, for the LTIS to be used efficiently in peri-urban communities amid numerous communication challenges as has been outlined above, page compression methods employed in Apache, which reduce the size of web pages served over HTTP are necessary. Other features of Apache which prompted its use for the development included virtual hosting, content negotiation, and the support of graphical user interfaces and FTP (File Transfer Protocol) (Apache 2011). As far as performance is concerned, Apache is designed to reduce latency and increase throughput, allowing it to handle a greater volume of requests, thus ensuring consistent and reliable processing of requests within reasonable time-frames (PHP 2011).

Furthermore, as shown in Figure 4.2, the storage and retrieval of data in the LTIS is enabled through the PHP server side programming language. This means that written scripts are executed on the server and sent to the client. The client receives the result of the running scripts without knowing the underlying code and the logic in the script. Hiding actual code from users, both experts and non-experts, is very important because it contributes to the security and integrity of the developed application. Also PHP is free and can be deployed on most available web servers including Apache. PHP can also be set up on a range of operating systems and platforms including Windows. It can be used with many relational database management systems including PostgreSql. PHP is fast and easily compiled (PHP 2011). Applications developed using PHP are executed rapidly (in seconds) and the results provided to the user on the fly.

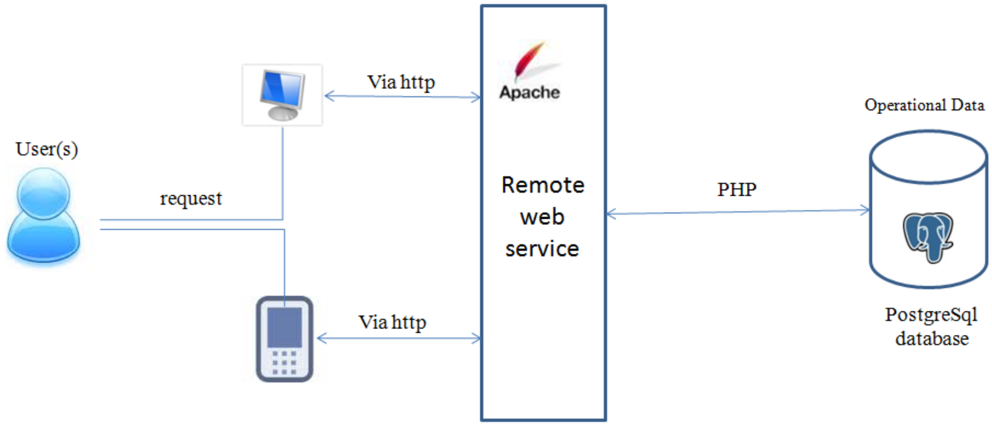


Figure 4.2: Conceptual Framework for Developing the Ubiquitous TTM-based LTIS

Further in this stage, a simulation was conducted with research colleagues and supervisor, who are familiar with the land use policies in peri-urban communities (especially in the Accra, Bortianor area). The research group were assumed to play the roles of an end user, surveyor, land administrator, and land holder. Their participation as subject matter experts was vital and was aimed at determining, from the various user perspectives, which additional features may be needed in such an LTIS. The land surveyor was interested in how survey maps are stored and related to the land in question. The land holder was interested in how the system can properly relate him and his entire family to their respective lands as well as other land transactions conducted in the family. The end user was mostly concerned with usability and ease of use of the system. The land administrator was concerned with how the system can secure tenure by creating meaningful relationships between objects, and how data stored in the system could be protected from unauthorized access and alteration.

The research group were asked to provide feedback after each iteration by using the actual LTIS and after that talking about found and imagined usefulness around a boardroom table. The familiarity of these stakeholders to the economic, political and social issues in peri-urban communities, especially Accra area revealed purposive shortcomings needed to be built into the application. The stakeholders' contributions suggested that some additional functionalities were needed. These included building some security measures into the system so that, although all users would be able to add and relate objects, not all of them should be able to edit objects that relate to other objects. No user should be able to delete an object that has already been involved in any form of object relation. Also, making changes to the database structure reduced the size of the database. Based on feedback, which was received through a research group meeting, several adjustments (explained in the section 4.2.5) were made to the developed LTIS until stakeholders were satisfied with the current system.

4.2.5 Product Ionizing Phase

Reliability and performance was essential for the LTIS to meet its current and evolving requirements. To ascertain reliability and performance – which was a priori vague and imagined - the developed LTIS was tested with fictitious data. This involved creating scenarios that mimic land transactions in peri-urban communities and then populating the system with data based on the relationships contained in the scenarios with the aim of testing the functionalities in the system.

Testing revealed loopholes, features requiring further development, and additional features needed to be added to the system. The loopholes identified were also generally in accord with the adjustments suggested by members of the research group in the iteration phase. The changes suggested and completed included validation of forms during data entry into the system. Form validation is important to ensure that users enter consistent information into the system and that all required fields are completed. Removing lookup tables and replacing their function with the Enum data type in PostgreSql streamlined database structure.

In an attempt to ensure data integrity and consistency in the developed LTIS, a simple but secure Content Management System (CMS) was built into the LTIS. Record deletion and editing were moved to the CMS. With the inclusion of the CMS, although all users are permitted to add and relate objects, only authenticated users with securely assigned login credentials are allowed to edit or delete records stored in the database. However, additional constraints were added to prevent authenticated users from deleting objects involved in any relation to other objects. Deleting a related object is not allowed due to the legal issues surrounding land administration, and controlling access to deletion capability supports the integrity of data stored in the database. Deleting an object involved in any relation means deleting an item of evidence. This may lead to the loss of a chain of evidence or title, opening the door for fraud or loss of land through negligence (Barry 2011).

The CMS is accessed through a web browser so that editing can be done anywhere and anytime by authorized users. The CMS is simple and does not require high expertise or knowledge in IT to operate. The system was deployed and refined based on additional feedback received from users (stakeholders mentioned in the description of the iteration phase, above). The goal was to assess the level of, and issues relevant to, user acceptance based on the system's usefulness and reported ease of use (Davis 1989).

Some new features and recommendations such as the drawing and storage of maps of an area in question in the database were postponed to be implemented in the future because it is out of the scope of this study and is being tackled by another researcher. Suspended ideas and requirements were noted and may be added to future release of the developed LTIS. The iterative process was continuous and consistent with evolving scenarios and requirements. As the LTIS was being developed and tested incrementally, measures were taken to maintain the system.

Maintenance is vital to ensure system reliability, availability and performance. The research group members mentioned above in the iteration phase also served as end users. It was noted that the system developer may have to continually provide support to the LTIS users. Some research group members who were assumed to be system users were provided consistent training and support so that they were conversant with the features implemented in the LTIS so as to provide feedback when necessary. The support also involved demonstrating how to use the system in adding, relating, and editing objects. It also showed how to render an object obsolete and also to set up a user with proper login credentials.

Overall, the research group assess the system's efficacy against the achievement of collecting, storing, and relating objects meaningfully and securely.

4.2.6 Death Phase:

The LTIS has not yet reached the death stage yet (explained in 2 section 2.3.2.5). Although the system has not been implemented, it is believed that as far the uncertainties in a peri-urban community continue to exist, the death stage may never be reached. The system may reach the death stage when conditions in a peri-urban community are stable; that is, when there is little or no barrier to the use of the system and is not affected by the corruption and power dynamics in the community. This will also mean that the LTIS has evolved to a stage of optimal usefulness where it is capable of supplying the required land administration functions and, therefore, no new requirements and features need be developed. In complex social and economic situations, achieving stability indeed, may never be reached. However, should this stability be achieved, then that periodic assessments and stakeholder surveys should have supplied accurate information on system efficacy, and the expected performance of the developed LTIS by stakeholders should be fully satisfactory. It is at this stage that the final documentation may be written. This documentation includes history of the major thrusts that led to the initiatives in the development process. It also discusses the difficulties and shortcoming in the system and how they were overcome. Also the system architecture and source code are included in the documentation.

4.3 Mining and Analysis of LTIS Data

4.3.1 Introduction

Once the system has been developed and is able to collect, store, and relate objects securely, the next step is to analyze the data stored. This section describes the activities conducted in analyzing the TTM-based LTIS data.

The developed TTM-based LTIS was populated with data based on two scenarios. The first data set was based on narratives and scenarios presented in the Talking Titler Manual (Barry 2011), while the second data set was inferred from a real case study conducted in Bortianor by Danso (2012). This section begins with a summary of the first Scenario In the manual, using fictitious names and relationships presented as if they were real life situations - all of which were inferred from land tenure relationship models contained in the Talking Titler Manual. The second Scenario is then described with reference to how the data was stored in the TTM-based LTIS and later utilized for analysis. The description of every Scenario is followed by how the relationships were captured and populated into the TTM-based LTIS for mining and analysis.

4.3.2 Scenario I

Adu Nyarko is a 65 year old Haminaro traditional leader at Almondvil whose status as a chief and legal right to property is recognized by state legislation. He is also the head of an extended family with siblings and offspring numbering approximately 30, the largest clan in the community. He has two daughters who expect to inherit any property he owns upon his demise. Both daughters are married with children. As Haminaro tradition demands, Nyarko as the head of the family is the custodian of the family land. When Nyarko's father, the previous chief died, Nyarko as the eldest son in the family alive succeeded his father as chief and, pursuant to both Haminaro custom and as recognized in the traditional groups land regulation, inherited this land in trust for the clan.

Due to economic hardship, Adu Nyarko borrowed money from one of his kinsmen, Yaw Duah who happens to be a money lender in the community. Nyarko used a portion of the family land as collateral, as allowed under the traditional Haminaro regulation regarding money lending. The agreement provided that, if after 20 years Nyarko is unable to pay the borrowed money with the required interest, the collateral automatically becomes the property of Duah and his family. The transaction was sealed through an oral agreement and an exchange of a ten centimeter bamboo stick as custom demands.

Nyarko also sold a portion of his land to a member of his community, Kwame Oduro. Following Haminaro custom, their transaction was sealed through an oral agreement, as witnessed by Afetor Pascal, the land surveyor. Due to old age and health issues, Nyarko became incapable of farming activities. Consequently, he gave a portion of his land to a non-Haminaro family for farming, on the condition that both families would share the produce. The stranger and his family also live on a portion of the land Nyarko allocated to them for sharecropping. Although this is a practical means of providing sustenance for both families, the experience of many villagers is that sharecroppers or their family members usually live on a particular parcel of land from generation to generation and may develop close bonds with the land owner's family such that they may be mistakenly viewed as possessing ownership rights to the land. However, Haminaro custom does not recognize such types of possessory right. Alternatively, relatives of sharecroppers may actually assert claims to the land after living and farming on it for a long period of time (say about 50 years), although no particular length of occupancy is prescribed as a minimum basis for a claim of ownership.

The only evidence, in written form, that is offered to support most transactions in Almondvil is the survey map created by a licensed land surveyor. The land surveyor is the local representative of the government in the vicinity and reports directly to two officials at the Lands Commission in the city. The land surveyor lives in the community and is also the family head of a single nuclear family. Following Haminaro tradition, he often serves as a witness during land transactions in the community. Experience of Almondvil community members shows that a corrupt land surveyor may take advantage of the vulnerable and steal their land. In other instances, his involvement in land transactions may be misused by members of his family such as sons, who may wish to claim ownership in future, asserting that their father owns the land in question.

4.3.3 Populating the LTIS with Data Based on Scenario I

In order to help Nyarko and other community members secure their land in the future, the developed TTM-based LTIS was implemented and used to record details of the transactions described above. Through interviews (simulated based on a Scenario in the Talking Titler manual), community members narrated their role in the transactions described above. The narratives were recorded (both audio and video). Sometimes some narratives were also written down in Microsoft Word documents. Photographs of the land, boundaries and landmarks as well as copies of survey maps from the land surveyor were archived in digital form as associated data items in the LTIS. The relationships between Adu Nyarko, Yaw Duah, Kwame Oduro, the Stranger, Afetor Pascal the land surveyor, and their respective family members were created in the developed system.

4.3.4 Data Preparation for Mining and Analysis Based on Scenario I

The synthetic data populated into the developed TTM-based Ubiquitous LTIS was exported through PostgreSql and prepared for mining and analysis. Data preparation which involved data selection, cleansing, and enrichment and transformation is explained below.

4.3.4.1 Data Selection

Data containing relationships created between people, land, and media (evidence) were selected. The reference instrument object was not included because no land regularisation procedures were dealt with in the scenario. The focus was mostly on the Person – Person, Person – Media, Person – Land, and Land – Media object relations. After the needed data was selected, the data was then cleansed.

4.3.4.2 Data Cleansing

Data cleansing involved removing and correcting inaccuracies in the data, including correcting typographical errors. Although the forms in the TTM-based LTIS are validated to ensure appropriated data entry, some names were spelt incorrectly and descriptions of some relationships were not so clear and so were rewritten for clarity. Mistakes in entering some addresses used were also corrected to mimic an appropriate residential address. This was fairly easy to do because the amount of data entered into the system was equally small. However, with large data sets, data cleansing could be one of the most difficult tasks in the data preparation stage.

4.3.4.3 Data Enrichment and transformation

Names representing various objects such as people, media items, and land were assigned numbers (coded). To help identify the types of relationships that exist between various data, specific values were assigned to previously defined relationship types in the data set. These values were used as the weight assigned to edges between nodes. For instance Table 4.1 shows each weight value and the type of relationship it represented for all relationships involving two people (Person to Person relationships) for Scenario I. Assigning the correct weight factors was especially critical, so that specific entities could be traced back to the raw data later on to identify what they represents in the analysis. The relationships were then mapped out to create matrices for the analysis in NetDriller.

Weight Value	Relationship
1	Parent/Son
2	Parent/Daughter
3	Husband/Wife
4	Sibling/Sibling such as brother/sister
5	Grandparent/Grandchild
6	Other (In-law, cousin, aunt, uncle, nephew)
0.1	Land surveyor
0.7	Lender/Borrower
0.8	Seller/Buyer
0.9	Land owner/Share cropper

Table 4.1: Person - Person relationship types and their corresponding weight for Scenario I

4.3.5 Scenario II inferred from case study in Bortianor

This Scenario Is simulated based on conditions in Bortianor, a peri-urban community in Accra, Ghana (Barry *et al*. 2012). Although the description is based on Danso's (2012) study, the cases described here do not refer to specific cases and people in the Bortianor community. They are fictitious and inferred so as to explore issues related to chieftaincy and land sales at both the community and family levels in a generalised peri-urban setting.

4.3.6 Scenario II

4.3.6.1 Chieftaincy

Since the death of the last chief, Bortianor has not installed a new chief for the past 18 years. This is because chieftaincy, lineage and succession are complex and contested matters and have created disputes over who assumes these high positions. There are different factions with supporters and stories to back their respective claims to the throne or stool.

Two families (referred to as families A and B) have emerged as the most powerful among the lot and are the leading contestants for the throne. These two families have been recognized almost by all communal members as the royal families linked directly to the throne, although there are other self-acclaimed chiefs belonging to neither families A or B.

According to one faction in the community, the chieftaincy is to alternate between the two families. Some (mostly supporters of family B) claim that the last chief was from family A and that, therefore, the new chief should be from family B. Another faction (mostly supporters of family A) claims that the last chief was from family B and so the new chief should be from family A. Different people tell different stories that link the late chief to either family A or B as the apical ancestor. The stories are generally biased in favour of whom the storytellers support. The heads of both leading families have pronounced themselves chiefs in the community, each of them enjoying support from sizeable factions.

Although all citizens of Bortianor trace their lineage to one ancestor, there is no documentation of lineage or succession. Older people conversant with the family lineage have been manipulated, in some cases by educated or rich individuals, to tell lineage in favour of one family or the other.

According to Bortianor tradition, in the absence of the chief, the King Maker acts as the chief. Due to these conflicts, the King Maker has been acting as the chief for the past 18 years. Consequently, some family members of the King Maker, especially his immediate offspring, believe they are of royal lineage and may be chiefs of Bortianor someday.

Currently, the Queen Mother of Bortianor is in conflict with the King Maker who is also the acting chief. The Queen Mother's educated and wealthy son is interested in becoming the next chief and has married a daughter of family A. In league with the traditional priest of Bortianor, the Queen Mother fabricated a story that simply means that the marriage alliance between her son and the daughter of family A qualifies her son to be a chief. These stories according to communal members are not in line with tradition. However, due to the money and influence of the Queen Mother and her son, they have been able to command the support of some community members. The priest keeps proclaiming the gods of the land have pronounced the Queen Mother's son as the next chief. These stories have created conflict and confusion among the royal families and community members and, as a result, they are unable to install a chief in the community.

This confusion could have been resolved if there existed a document or record that could aid in tracing lineage and succession to the stool. However, with some form of objective data, it may be possible to model and visualize the relationships between people and the stool in Bortianor.

This may be a way to document relationships in Bortianor which may also serve as a lineage population register for the community in general. Such documentation may serve as a living testimony in support of any story at any point in time in the history of Bortianor indigenes.

4.3.6.2 An Incident In family 'A' Regarding Land Sales

Due to urbanisation, land sales are a lucrative business venture in Bortianor. Wealthy people from the cities buy local lands for residential and commercial purposes. The lands in Bortianor are classified under family or stool lands. Stool lands are the lands managed and controlled by the chief. Family lands belong to specific families and is managed and controlled by the family. Sometimes family lands are improperly converted to stool land and sold to non-Bortianor indigenes – who are referred to as strangers herein. When this happens, local and vulnerable family members are dispossessed of inherited land.

Each family in Bortianor includes leaders with specific functional roles and titles with corresponding levels of status in the social order of the community. The leaders are called the Council of Elders in a family. These Elders include the family head, spokesman, secretary, and family priest if the family has a god. The family head serves as the head of the council of elders in the family and has control over the family land. Each family member is apportioned a section of the family land for farming. Should a family member die, the right to his or her portion of the family land is presumed to be transferred to his or her children. A family member so inheriting may choose to sell his or her portion of the family land to a stranger. Such a family member is obliged only to notify the family of the land sale, rather than seeking family approval. Although the family head controls the family land, he has no right to sell an apportioned land of a family member without the permission from that family member. The family head may sell a portion of the family land to a stranger after consulting and obtaining the approval of the rest of the council and family members. When such a parcel is sold, the proceeds must be shared among all legitimate family members. A legitimate family member who can enjoy proceeds of a land sale is someone who is directly related by blood to the progenitor of the family. This means that anyone related to the family through marriage alliance does not qualify to enjoy such proceeds.

Family A possesses acres of land inherited from their great ancestor. Upon request the family head apportioned a section of the land to his sister - her own share of the family land - for farming. Recently, due to urbanisation, many strangers and businesses have become interested in buying this apportioned land owned by the family head's sister because of its location. According to tradition, the sister owns the land, as it is regarded as her share of the family inheritance, and only she has the right to sell the land. She wanted to sell the land and so informed her brother, the family head, of her intention. The family head did not give his consent because of reasons best known to him. According to tradition, the sister is not required to seek her brother's approval before selling the land, she is only obliged to inform him as the family head about the land sale. Since the sister owned the land; and in line with traditional regulations, she sold the land to an entrepreneur in the community. She hired a surveyor and with the help of the family secretary, provided a survey plan and an indenture to support the sale.

At the same time the family head also sold the same piece of land to a company wanting to start business in the community. He also singlehandedly provided a survey map and an indenture to support the sale. At the same time, an impersonator in the community also sold the same piece of land to a stranger and forged an indenture to support the sale. The entrepreneur who purchased the land from the sister, that is the original owner of the land, also sold it to another private company for commercial activities. This has created conflicts and confusion over the land in question. These buyers now try to secure the land through court. However, there is not enough information to aid in resolving the issue at hand in court. The rich and influential people involved in the conflict have resorted to using land guards in protecting the land. Land guards are independent individuals hired to protect lands in Ghana. Although not permitted under law to embark on their activities, they use ruthless tactics to prevent people from developing (e.g. building upon) land parcels. A clash between land guards from competing factions can result in brutal injury and sometimes even death. The one who has more violent and strongest land guards may win the fight. It is not uncommon for one party to build today only to find their structures destroyed by the other group the next day.

Again, modelling the relationship between various family members and their land may aid in tracing family lineage. This lineage documentation may help identify who enjoys proceeds from land sales and other inheritance or properties in the family. It may help to secure land for family members should the family head or any member of the family become corrupt or greedy. Also, modelling land transactions and linking them to how people relate to the land (either through sales or through inheritance) could be a documentation that may aid in tracing people-to-land relationships. Such information may be used by any judicial authority (formal or customary) in resolving conflicts related to land sales in Bortianor.

4.3.7 Populating the LTIS with Data Based on Scenario II

Details of the transactions described above were recorded in the developed TTM-based LTIS. The scenario above mostly had to do with people-to-people relationships. For the Chieftaincy issue, it was necessary to record and model family lineage and succession to the stool. To populate the system with such data, the various individuals representing family members of the families described in the scenario were created with fictitious names. Fictitious family members were entered and related to each other in the system. In order to help model individual relationships to the stool, the stool was modelled as a single entity so that people having direct relationship to it could be identified. Each family related to the stool through the next in line to rule or the one most closely linked to the stool. This means that, as far as succession is concerned, every family has the one closely linked to the stool on top of the hierarchy with all other family members relating to him based on the family relationships. Through this mapping of inter-relations, it will be possible to model how every family as well as its members relates to the stool.

For land sales, the individuals involved in the land sale described above were related to the land base. All family members who relate to the land described in the scenario were related to the land in question. This may help identify owners and even impersonators.

4.3.8 Data Preparation for Mining and Analysis Based on Scenario II

4.3.8.1 Selection and Cleansing

The data selection and cleansing in this stage was the same as data selection done in Scenario I (sections 4.3.4.1 and 4.3.4.2) above.

4.3.8.2 Enrichment and transformation

Data enrichment and transformation was very similar to that of Scenario I (section 4.3.4.3). The only difference here was the type of relationships created to help describe the type of relationships existing between the different entities involved. The weight values assigned to existing defined relationships were also different. Below are tables showing specific weight values and the type of relationships they represent for all relationships involving two entities for Scenario II.

4.3.8.2.1 Relationships between people

Weight Value	**Relationship**

1	Parent/Son
2	Parent/Daughter
3	Husband/Wife
4	Sibling/Sibling
5	Grandparent/Grandchild
6	In-law/In-law
7	Cousin/Cousin
8	Uncle/Nephew/Niece
8.1	Step Parent/Step Child
8.3	Other(e.g. in-law's wife, father's brother's wife)
8.4	Step grandparent/Step grandchildren
0.5	Royal/Legitimate Chief
0.2	Royal/King Maker/Acting Chief
0.3	Traditional Priest/stool
0.6	Land owner/Land owner
0.7	Queen Mother/stool
0.8	Seller/Buyer

Table 4.2: Person - Person relationship types and their corresponding weight for Scenario II

4.3.8.2.2 Relationships between lands

Weight Value	**Relationship**

1	Emanated From/Related to

Table 4.3: Land - Land relationship types and the corresponding weight for Scenario II

4.3.8.2.3 Relationships between people and land

Weight Value	**Relationship**
1	Ownership
2	Lease
3	Partial rights
4	Collateral
5	Survey(or)
6	Claimed
7	Disputed
9	May inherit (all/portion)

Table 4.4: Person - Land relationship types and corresponding weight for Scenario II

4.3.8.2.4 Person to Land to Sales Relationships

The coding of this relationship is best explained by the following example. An aggregate relationship factor of 1.8 means that parcels 1 was either sold or bought by the entity (person) involved in such a relationship. That is, 1 represents the parcel number or the parcel in question and 0.8 represent the entity's (person) relationship with that parcel. Since 0.8 in Table 4.1 represents a given seller or buyer, 1.8 means that, for the two nodes or individuals involved in the relationships, one of the nodes represents the person who sold parcel 1 and the other node represent the person who bought parcel 1.

2.1	Parcel 2 surveyed
2.8	Parcel 2 sold/bought
2.6	Parcel 2 owned

Table 4.5: Person – Land - Sales relationship types and the corresponding weight for Scenario II

4.3.9 Mining and Analysis of Data for Scenarios I and II Using NetDriller

The resulting data containing the relationships between entities were loaded into the NetDriller (Chapter 3 section 3.4.4 for information on NetDriller). NetDriller constructs a network or graph from the data sets based on the relationships detected between objects. Through the graph, the various groups and subgroups (referred to as communities) in the network were identified. Also, the relationships between detected communities were noted from the generated graphs. The networks were treated as a single or multi-mode network depending on the objects involved. The bi-partite graphs linking each object were then found.

In the case of Scenario I, the resulting networks were sometimes folded to find the relationship between instances of the objects involved (see chapter 5 for the folded networks). Also, the betweeness centrality measures were calculated to detect the central actors in the communities. For Scenario II, closeness and degree centrality measures were computed in addition to betweeness. The values assigned to previously defined relationship types in the data set were used as the weight on edges between nodes. The generated relationship types depicted in the graph aided in determining the kinds of ties existing between individuals, key actors and within communities in general.

4.4 Chapter Summary

This chapter described the methodology for first developing the TTM-based ubiquitous LTIS. The chapter described how XP was adopted to developing the LTIS based on the TTM. Throughout the phases of XP, the reason for developing a web and mobile LTIS to be used in a peri-urban community was explained. The system architecture was described as well as various changes that needed to be made in the system during development.

The chapter also described the activities involved in analysing the LTIS data. The second section of the chapter described the scenarios that governed the data entered into the developed LTIS. Two scenarios were described in the chapter, focussing on family relationships, chieftaincy issues, succession, and land sales. The first scenario was based on the social and political conditions in peri-urban communities described in the Talking Titler system manual. The second does not depict the situation in an actual community, and was inferred from conditions in Bortianor, Accra, Ghana.

The chapter described how data based on the described scenarios were populated into the system and how the data was later extracted, prepared and loaded into NetDriller for mining and analysis. The next chapter will discuss the results of the system development and simulation activities conducted in this chapter.

Chapter Five: Experimental Results

5.1 Introduction

This chapter discusses the results of the activities conducted in Chapter 4, in two sections. Section one shows the results of applying the agile software development approach (extreme programming) to developing the TTM-based LTIS. The section evaluates the developed system and explains how the current functionalities in the system operate. Section two discusses the results of applying data mining and social network analysis techniques – through use of NetDriller - to analyzing the LTIS data that was populated into the system based on the two scenarios described in Chapter 4 (section 4.3.2 and 4.3.6).

5.2 Evaluation of the Developed Ubiquitous LTIS

The ubiquitous LTIS is designed as a dynamic and flexible tool for collecting and storing land tenure records. It allows the management of data relating to land tenure such as land surveys and communal land records. It is developed to help reduce some of the land-related challenges encountered in peri-urban communities such as multiple sales and land appropriation which leads to tenure insecurity. The system may contribute to reducing these challenges by collecting, storing and relating land transactional data with appropriate parties (e.g. land owners and buyers) involved. It may also aid in building a family record system for indigenous communities.

The developed TTM-based LTIS is a web application with some functionalities adapted on a mobile device. The ubiquity of the application may be useful, in that it is installed at a central location and can be accessed everywhere in the world via the web through a computer or a mobile device with internet connectivity. Data is collected, stored and manipulated at this central location, rather than being distributed or scattered in different locations. This ensures data security and makes back up and data restoration uncomplicated in case of disaster. Updates to the developed application can also be made easily and quickly.

A user needs no complicated configuration on a computer to access and use the application. This is very important to reduce the amount of expertise required in using the system. The system allows users to add information about all the four objects (Person, Media, Property and Reference) incorporated in the Talking Titler Model. Data related to each object can be collected and related with every other object in a flexible but efficient manner. To improve efficiency in extracting data relations, numerous scenarios implied in the TTM involving Objects are considered and incorporated into the system. For instance, the person object can be a land holder, the neighbour of a parcel being surveyed, a partial right holder in a parcel (servitude or lease), or a company, trust or business entity. At the same time, the object can be related to itself or any of the other three types of entities allowed in the system. A person object may have an identity document or their image, and records of them giving oral evidence may appear in various media items such as photographs or videos used in support of a land transaction. These are stored as media objects and topologically related to the person in question. A wide range of document types can be stored in this way, including survey plans, geodetic control coordinate files, satellite images, topographic plans, aerial photographs, videos and audio or sound files, geophysical survey data, written documents, scanned documents and maps. Where media objects relate to a person and to a property, the person and media are related to the property. Information about properties is stored in a property object.

The property object may include a land parcel and any other physical property present on it. Valid parcel type include those that are being surveyed or affected by a survey, and which have attached rights over another parcel, such as harvesting rights during specific times of the year. To accommodate land regularization and formalization, the system through the reference object is able to store title and/ or deed, property file, rent card or occupation permit and also relate them to any of the other objects.

Relating added entities through the system may help land administrators in answering some important questions such as - which people have registered or recorded rights over a parcel or object? Which people and physical objects feature in certain multimedia files and how are they related to other objects? What title or deeds are affected by a survey and what interpersonal relationships exist in a particular population that may give rise to land rights or *prima facie* expectations of land rights? These and many other questions may be answered by identifying the objects in question and how they have been related in the system and, thus, may help land administrators to manage land and related resources efficiently in peri-urban communities.

5.2.1 A Walk through the Developed LTIS Using XP

To begin using the system, a user clicks on one of the menus that allow the addition of an Object such as a Person, Property, Reference or Media. The user is then presented the form to add an object. When the form for adding (Figure 5.1) an object is filled and submitted, a user is shown all added objects and is given the options to relate (Figure 5.2) the added object to any other object. Assuming that a media object is added, then before relating the added media object to another object, the user may view any uploaded media file already stored in the application including the one just uploaded by the user. The user can view the uploaded media to confirm that the intended media has been stored. By viewing the media item before relating it to any other object, the user is also able to ensure that the intended media is related to the appropriate object. The uploaded media file is stored in a folder and all data related to it are stored in the database. The application allows users to relate only stored media files with other related objects in the database. Figure 5.1

illustrates a form that collects information on a media object while Figure 5.2 shows the options available to a user in viewing and relating the added object with other objects.

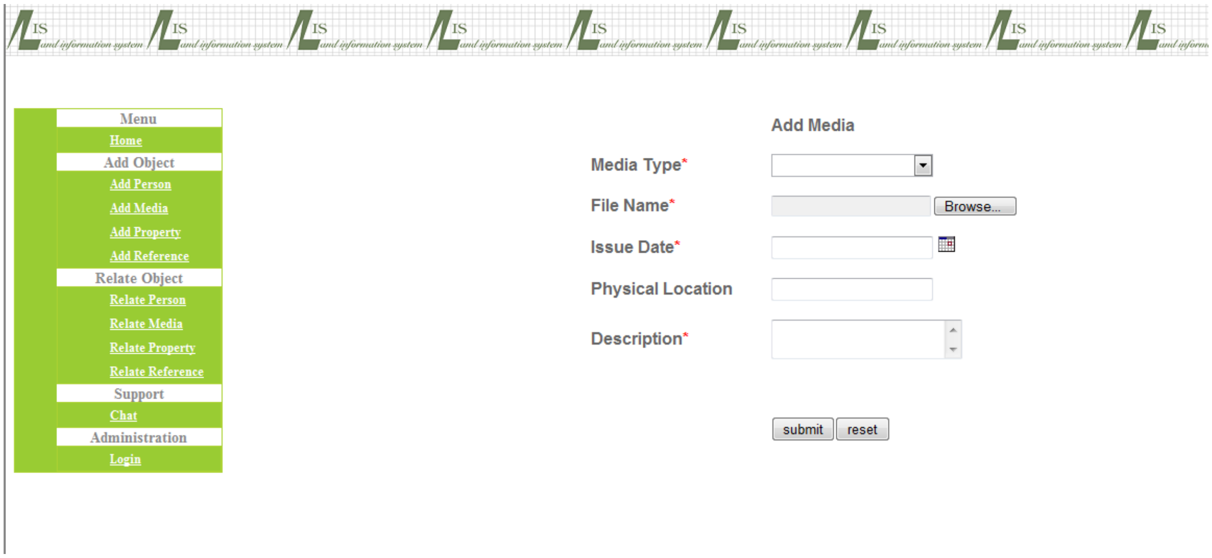
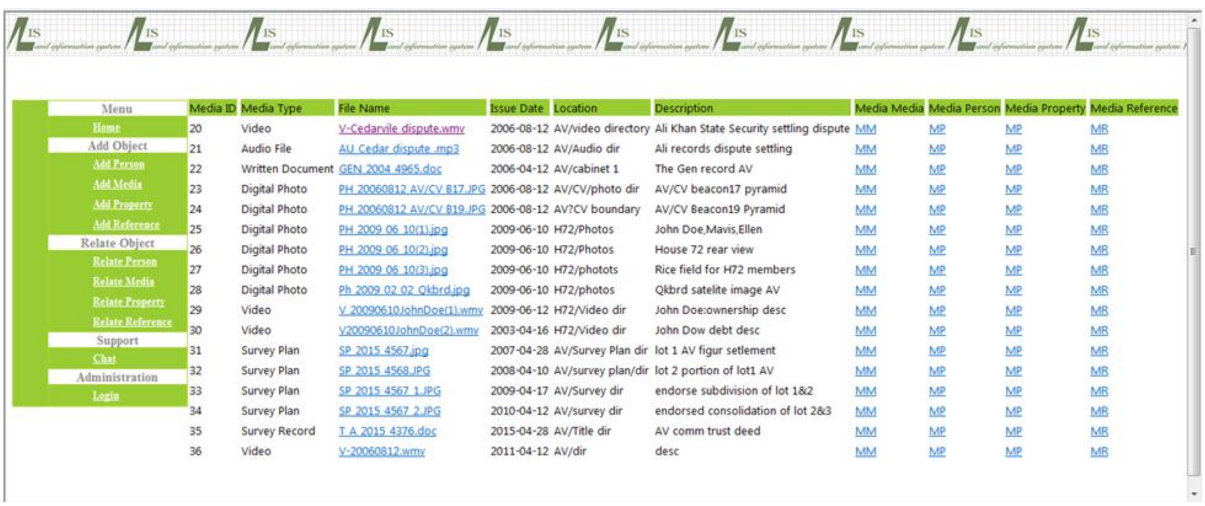


Figure 5.1: A form to add a media object



Media ID	Media Type	File Name	Issue Date	Location	Description	Media Media	Media Person	Media Property	Media Reference
20	Video	V-Cedarvile dispute.wmv	2006-08-12	AV/video directory	Ali Khan State Security settling dispute	MM	MP	MP	MR
21	Audio File	AU Cedar dispute .mp3	2006-08-12	AV/Audio dir	Ali records dispute settling	MM	MP	MP	MR
22	Written Document	GEN 2004 4965.doc	2006-04-12	AV/cabinet 1	The Gen record AV	MM	MP	MP	MR
23	Digital Photo	PH 20060812 AV/CV B17.JPG	2006-08-12	AV/CV/photo dir	AV/CV beacon17 pyramid	MM	MP	MP	MR
24	Digital Photo	PH 20060812 AV/CV B19.JPG	2006-08-12	AV?CV boundary	AV/CV Beacon19 Pyramid	MM	MP	MP	MR
25	Digital Photo	PH 2009 06 10(1).jpg	2009-06-10	H72/Photos	John Doe,Mavis,Ellen	MM	MP	MP	MR
26	Digital Photo	PH 2009 06 10(2).jpg	2009-06-10	H72/Photos	House 72 rear view	MM	MP	MP	MR
27	Digital Photo	PH 2009 06 10(3).jpg	2009-06-10	H72/photots	Rice field for H72 members	MM	MP	MP	MR
28	Digital Photo	Ph 2009 02 02 Qkbrd.jpg	2009-06-10	H72/photos	Qkbrd satelite image AV	MM	MP	MP	MR
29	Video	V 20090610JohnDoe(1).wmv	2009-06-12	H72/Video dir	John Doe:ownership desc	MM	MP	MP	MR
30	Video	V20090610JohnDoe(2).wmv	2003-04-16	H72/Video dir	John Dow debt desc	MM	MP	MP	MR
31	Survey Plan	SP 2015 4567.jpg	2007-04-28	AV/Survey Plan dir	lot 1 AV figur setlement	MM	MP	MP	MR
32	Survey Plan	SP 2015 4568.JPG	2008-04-10	AV/survey plan/dir	lot 2 portion of lot1 AV	MM	MP	MP	MR
33	Survey Plan	SP 2015 4567 1.JPG	2009-04-17	AV/Survey dir	endorse subdivision of lot 1&2	MM	MP	MP	MR
34	Survey Plan	SP 2015 4567 2.JPG	2010-04-12	AV/survey dir	endorsed consolidation of lot 2&3	MM	MP	MP	MR
35	Survey Record	T A 2015 4376.doc	2015-04-28	AV/Title dir	AV comm trust deed	MM	MP	MP	MR
36	Video	V-20060812.wmv	2011-04-12	AV/dir	desc	MM	MP	MP	MR

Figure 5.2: View and relate added object (in this case a media object) to any other object

Next, the user is presented a form (see Figure 5.3) where the selected object can be related to the intended object. To help users in selecting the right media files for a relation, a description is added and tagged to the file name in the drop-down menu and presented to the user for selection. The user can also add a clear description of the relationship created between objects. For instance in Figure 5.3, a survey plan (media file) has been related to the right holder (person object) of the surveyed parcel. To ensure and confirm that objects have been related correctly, the data about the created relation is fetched and shown to the user. The user can then view the records of the objects that have been related to each other. If the created relationship is between an object and a media object, such as in Figure 5.3, the user does not only have access to seeing and reading the records and description about the created relationship but can as well view the media involved in the relationship. The ability to view objects before and after relating them allows the user to confirm that the correct relation(s) have been represented. This also reduces the tendency to incorrectly relate objects. It may also aid in identifying mistakes when relating objects so that the necessary corrections are made.

Correctly adding and relating objects gradually builds a land tenure record system which may support the management of land related resources, aid in conflict resolution in uncertain and complex situations where land rights and ownership are not regularised.

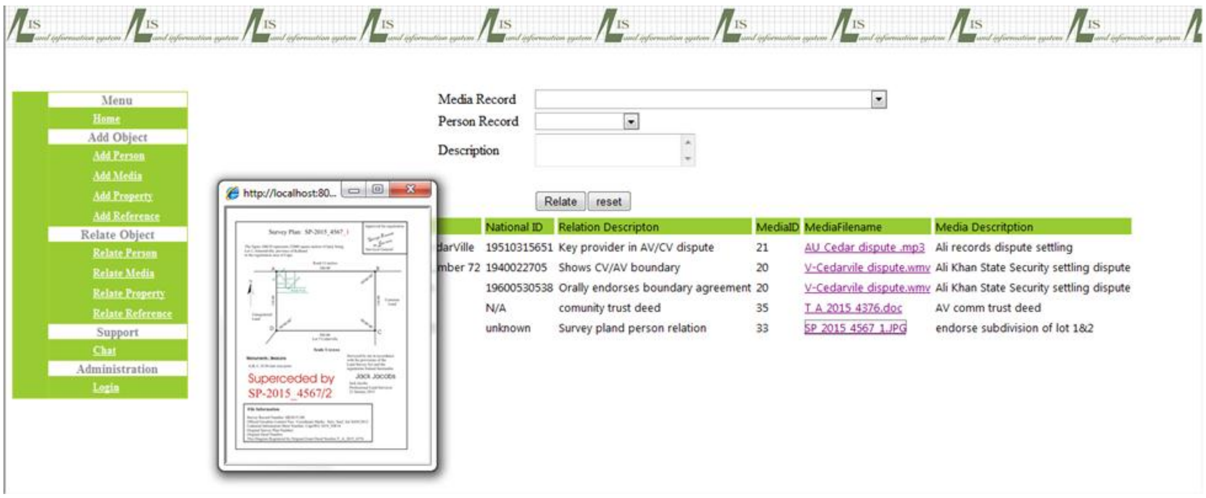


Figure 5.3: A survey plan (media object) that has been related to the right holder (person object) of the surveyed parcel.

In order for users to assist and support each other while using the system, an open source instant chat utility (PHP free chat) has been incorporated into the LTIS (Figure 5.4). Through this feature, users can ask and answer questions while adding and relating objects. Also, communications occurring between users are stored in the system, which may serve as a future ancillary reference if required. Experts can easily answer questions posed by non-experts and provide information from any location to other users in the field. For instance, if a user in the field needs to know the physical location of a survey plan or even the file number of a survey plan that has not yet been entered in the system, the user can ask another user at the office through the chat utility to retrieve the needed information. In this way, inexperienced users can receive training from experienced ones to solve actual difficulties as they occur. This function can also be used by a system administrator in providing technical support on the fly to users. Users can work together as a team irrespective of their location.

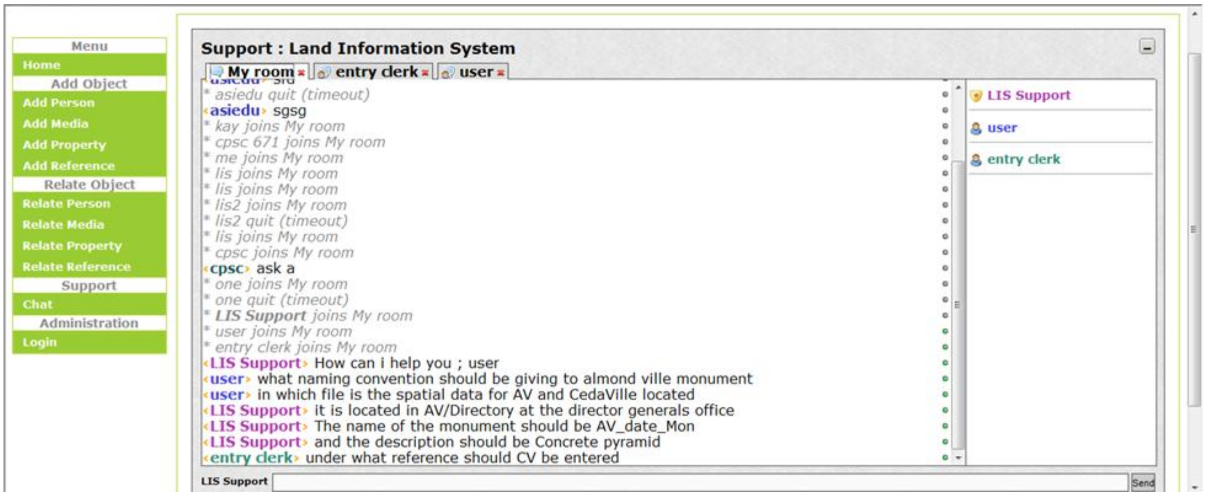


Figure 5.4: Instant communication among users

Additionally, to ensure security and integrity of data stored in the database, a simple but secure content management system (CMS) is built into the TTM-based LTIS. Although users can add and relate objects, only authenticated users with securely assigned login credentials are allowed to edit or delete records in the system. However, if a particular object has been related to other objects, even authenticated users are not allowed to delete the record. Preventing deletion of records makes them available for future reference, thus contributing to the integrity of data stored in the database. This may also facilitate the recording and analysis of metadata, or heuristics, to augment the relations that are made in the system.

As implied by the adoption of the TTM methodology, the system itself is an evolutionary process; thus, it is a challenge to initially define and know user requirements. When requirements are not clear, flexibility in the system allows it to act as a data collection tool to inform system development (Muhsen 2009). However, when user requirements are finally understood, the flexibility of the system is reduced. This is done through the reference instrument class. The reference instrument relates evidentiary files and legal documents to other objects to help formalize and regularize the tenure system in a community. The reference instrument also contains functions that reduce the flexibility but which maintain the security of data stored in the application. Hence, authenticated users are allowed to render a record obsolete or to close a record if it is no longer needed. For instance, if a survey plan is superseded by a new one (see figure 5.3 above), the previous survey plan could be rendered obsolete. Once an object is rendered obsolete, users are no longer able to relate that object to any other objects. However, all previous relationships are retained in the system. Figures 5.5, 5.6 and 5.7 show the login functions of the LTIS and the options available to an authorized user.

The incorporation of a CMS provides many advantages. The CMS is based on a web browser so that editing can be done anywhere and anytime by authorized users. The CMS is simple and therefore does not need high level expertise in IT to manage.

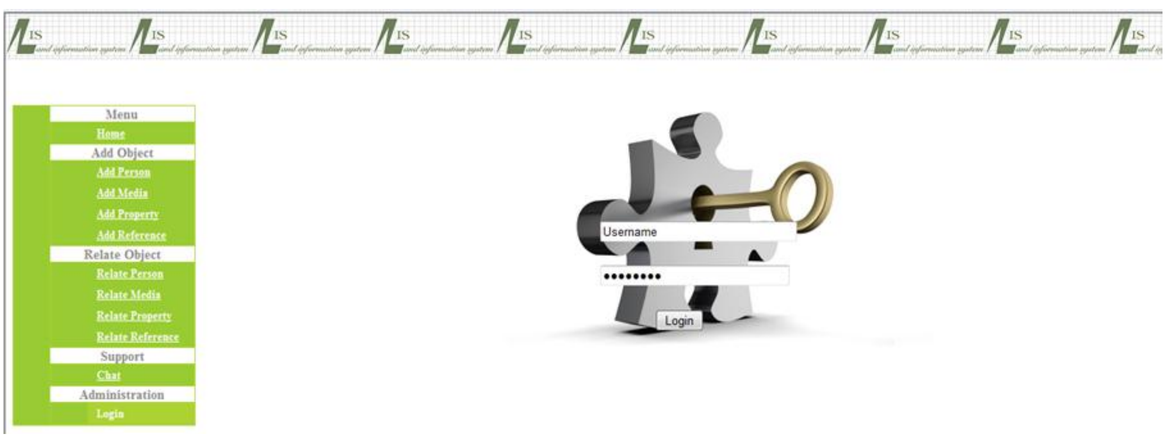


Figure 5.5 Login to the secure content management system (CMS)

Figure 5.6 Options available to authorised users to view, edit or delete objects that are not involved in any relationship

Figure 5.7: A form to edit and/or close a record

Some functionalities of the LTIS have been adopted on a mobile device. Users are allowed to use the mobile application to add objects. In areas where internet access such as broadband is not available, mobile users are still able to use mobile applications or browse the internet. With accessibility to a LTIS on a mobile device, a land administrator is not restricted to a particular location but, instead, carries the application along while on the move. Currently, the mobile application allows users to add person, property and reference objects from a mobile device. Further developments may allow users take photographs or record videos and upload on the spot to the database from a mobile device. Figures 5.8 shows an emulator that is run to show how the developed LTIS works on the mobile application. The object to be added (in this case, the person object as shown in the emulator) is successfully integrated into the central database and a successful record added confirmation notice is previewed to the user.

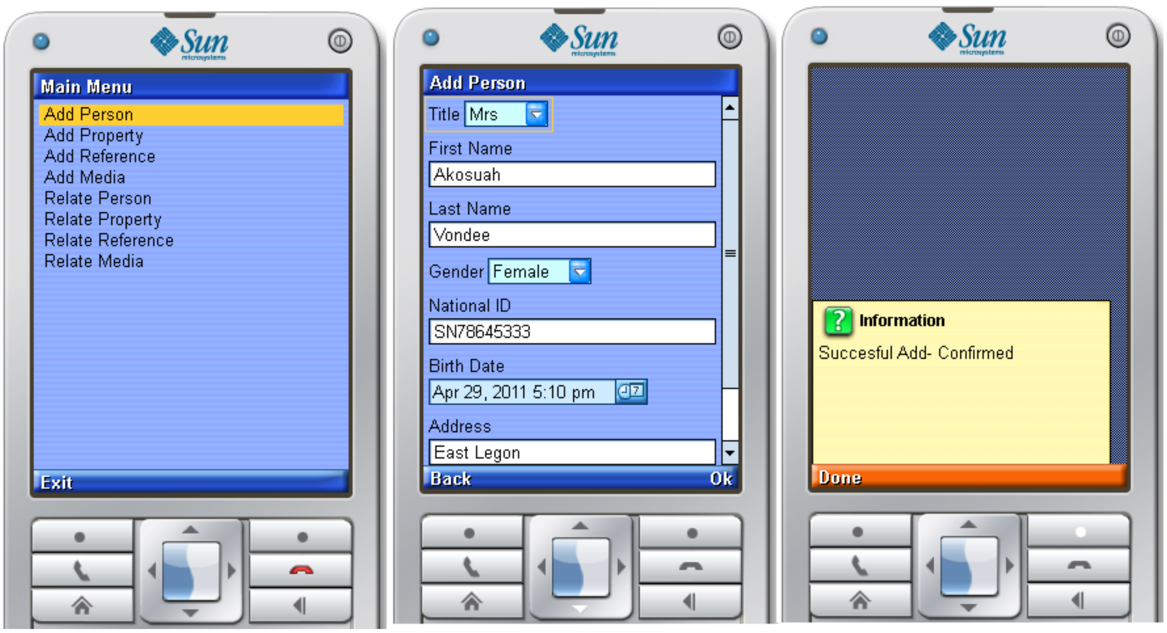


Figure 5.8: Adding a person record through the mobile application

The developed LTIS is a means to collect and store LTIS data in peri-urban communities. The emphasis is not on the technology (web and mobile), but on the ability to efficiently store information about objects involved in land transactions. Once the system has collected, related and stored information about objects, the next stage is to analyze the LTIS data. The next section discusses the results of applying data mining and social network analysis to the LTIS data.

5.3 Results of Mining and Analyzing LTIS Data

This section presents and explains the results of data mining and network analysis of the LTIS data. The section describes the results of both scenarios I and II (see sections 4.3.2 and 4.3.6). In both cases, a description of the relationships and the resulting graphs are first provided. This is followed by an interpretation of the relationship graph. As stated earlier in Chapter 4, the outputs of the analysis are all from NetDriller.

5.3.1 Scenario I

5.3.2 Person- Person Relationship Graphs

The graph of the Person – Person relationships generated from the data is portrayed in Figure 5.9. The various clusters represent discrete communities; a community in the case of Figure 5.9 being an extended family. Each node represents an individual person. Each edge (line in Figure 5.9) represents a relationship between two persons. The weight of the edge (the number on a line in Figure 5.9) is a code for the type of relationship. The weight on every edge is preceded by the cluster number to which the edge belongs. This means that an edge with a weight of 5 (see Table 1 in Chapter 4) belonging to cluster 1 will be labelled as 1.5. In the graph the network analysis suggests that nodes 1, 20, 24, 9, and 30 appear to be the central and important nodes in the network. Calculations of the betweenness centrality measure confirm this (Faust and Wasserman 1994) (see section 3.3.3 for a discussion on the betweenness centrality measure).

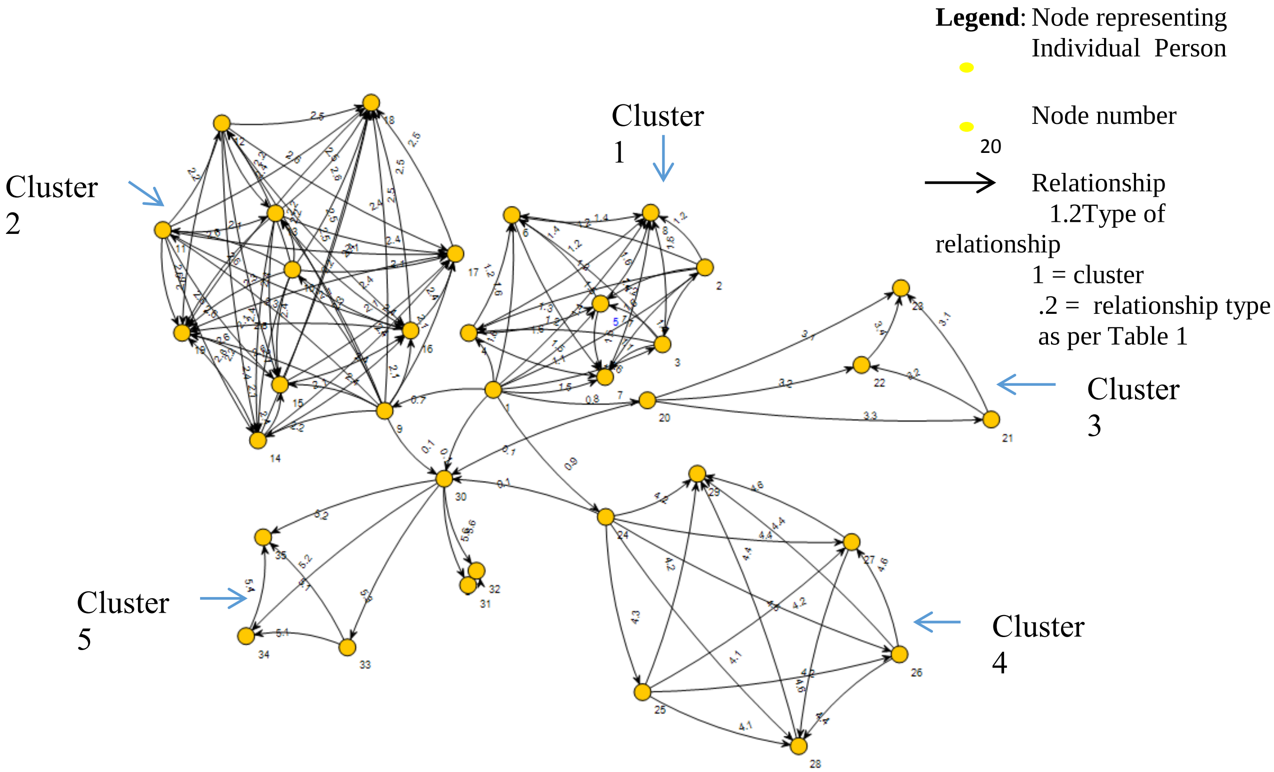


Figure 5.9. Social Network Graph of Person-Person relations

5.3.3 Parcel – Parcel Relationships

Network analysis of object-object relationships can show how land objects are created and how they change over time, and can thus be used to show a chain of subdivisions and consolidations that can be traced back to the parent parcels. The dataset contains a single parcel that has been sub-divided into several lots by a surveyor. In Figure 5.10, each node represents a parcel while the edges signify relationships between parcels. In this instance the weights on the edges have no meaning, but the arrows show the direction of the relationship and the progression of the chain of parcel configurations. All the lots are subdivisions of a single parcel, represented by node 1 in Figure 5.10. Nodes 2 and 3 represent subdivisions of Lot 1, which are represented by the edges 1-2 and 1-3. Parcels 2 and 3 were consolidated to create parcel 4, as represented by edges 2-4 and 3-4. The edge 1-4 shows that there is a relationship between parcel 1 and 4. An examination of the data will reveal that parcel 4 originates from parcel 1.

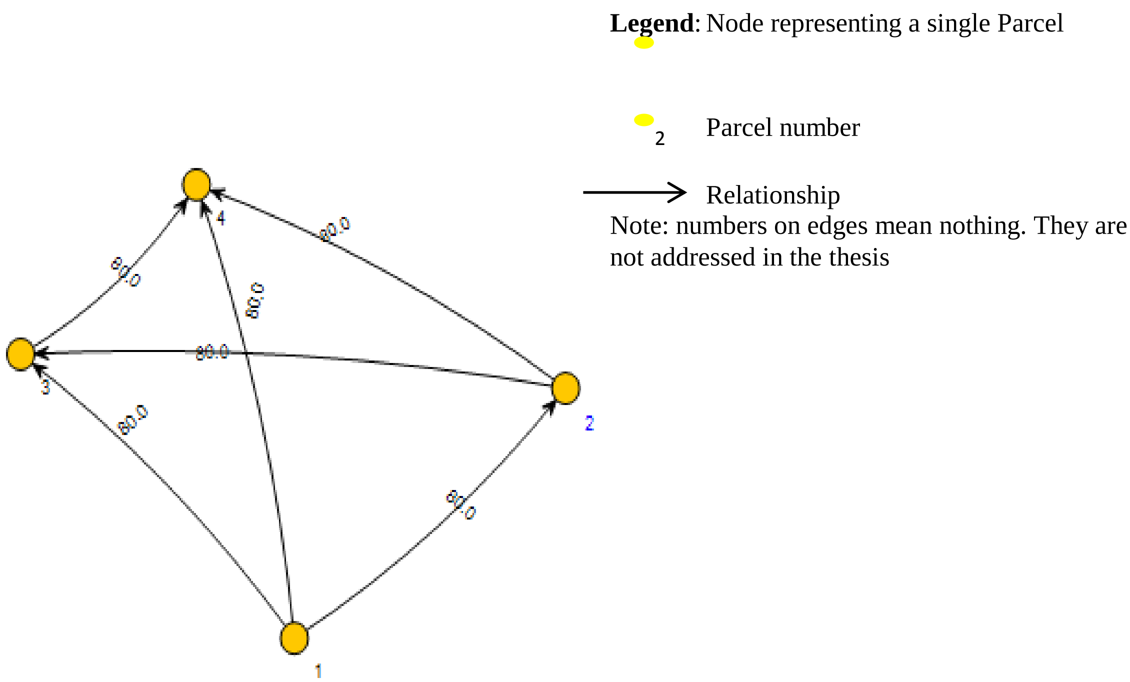


Figure 5.10: Graph of Lots (Land Object) Parcel-Parcel Network

5.3.4 Person – Media Object Relationships

Evidence supporting transactions and relationships between a person and land objects are stored as media files and are handled by the media object. Media items include a variety of data types such as video, audio, pictures, survey plans, and documents. Media files may be related to each other, perhaps because together they support the same transaction; e.g. a paper document may represent a title and a video showing the ceremonial handing over of ownership is evidence that the transaction occurred; otherwise, they may be part of the chain of title. Figure 5.11 shows the results of analyzing the Media-Media relationships in the data set. The nodes represent specific media files while edges show a relationship existing between two files.

The graph and the centrality measure shows that Media Items 19 (media file describing transactions by the chief – the original traditional leader) and 2 (survey plan of the original parcel showing deductions for subdivisions) are important and part of the evidence needed to support transactions in the dataset. They are analogous to documents showing the root of title in a deeds registration system which is a library of documents that indexes both parcel and personal information. The edges linking them depict the transactions that have occurred in the chain of title or chain of transactions in a customary system.

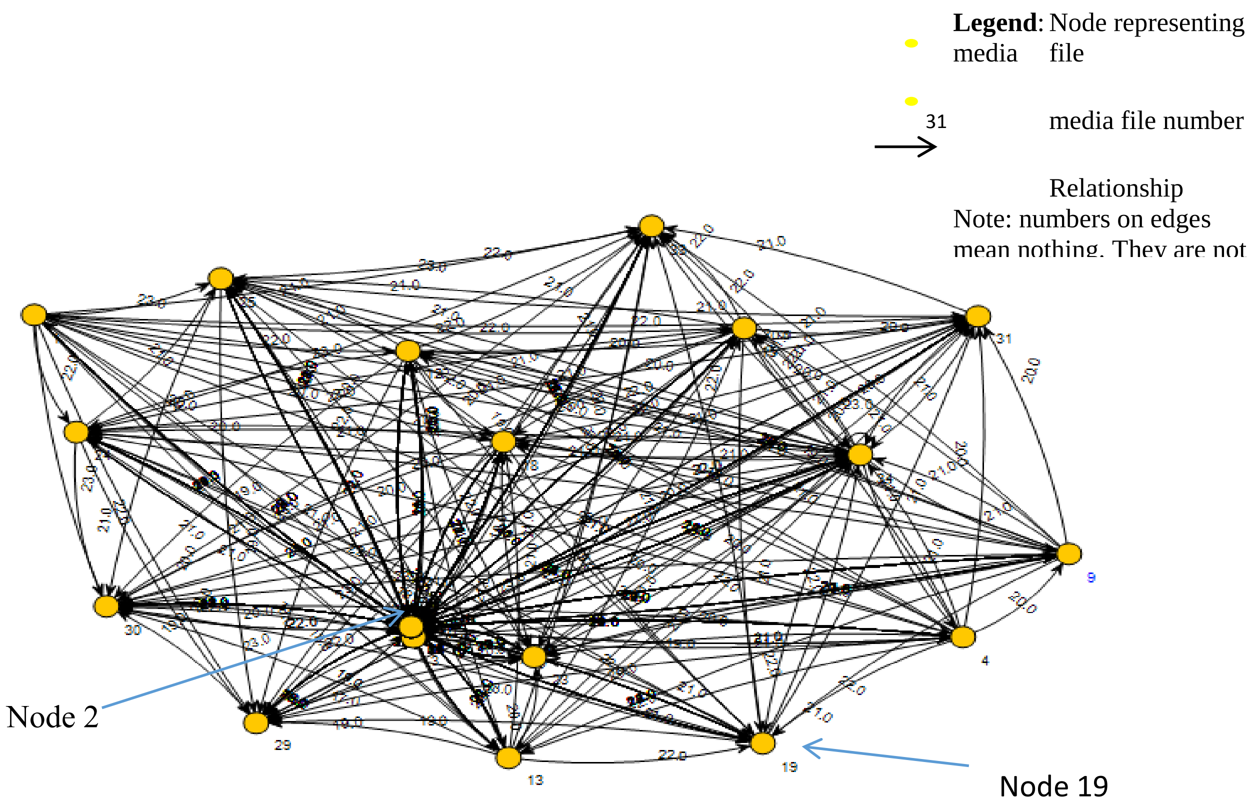


Figure 5.11. Media – Media relations derived by folding the Person – Media relations.

5.4 Interpreting the Graphs – Scenario I

Analysis using social network analysis and data mining techniques identified family heads, lineage groups, interests in land, and ties into various forms of evidence in the data. These are based on the Person – Person, Person – Media, Person – Land, and Land – Media object relationships stored in the database. A description of the relationships that were revealed in the analysis is provided below.

5.4.1 Family Heads

The graph in Figure 5.9 shows how individuals are related to each other, and the clusters represent families in the lineage and other groups. In the sample data, there are five groups, representing four families within the lineage and one group of outsiders. The five communities are joined to each other through central nodes. The centrality measures and graphic image indicate that nodes 1, 9, 20, 24, and 30 are the most important in the network because they link sub-groups in the network together to all the other nodes in the cluster. If these bottleneck nodes are removed, the links between clusters will not be shown and the communities in the network appear to be fragmented.

The links between clusters show which people in a family have a relationship with other families. In the sample data, nodes 1, 9, 20, 24, and 30 represent the family heads. The graph indicates that the family heads play a gatekeeper role (according to the data stored in the TTM-based LTIS), as there are no other links between clusters. Node 30 represents a family head who is also a land surveyor, as shown by the weight (0.1) on the links to node 30. Nodes 33 – 35 represent his family members. They are not part of the customary lineage, and so this is a cluster of outsiders. The two isolated nodes 31 and 32 that are linked to node 30 are officials to whom he reports as a land surveyor.

In Figure 5.9, Nodes 1 and 9 are linked to one another. An examination of the data shows nodes 1 and 9 as the heads of families 1 and 2. Node-1 borrowed money from Node-9 in an oral agreement using a parcel of land as collateral. There was no document to reflect this. It was merely stored as a relationship between two persons in the TTM-based LTIS. It is reflected in edge 1-9 with a weight of 0.7 (borrower/lender). Node-1 later sold a portion of his family land to family 3 through Node-20 (represented by edge 1-20 in Figure 5.9 with a weight of 0.8 (Seller-Buyer) in Table 1). The edge 1-24 represents a similar transaction, where node 1 gave a portion of his land to node 24 for share cropping. This relationship is also reflected in edge 1-24 with weight 0.9 (land owner/share cropper).

In Figure 5.9, all family heads are linked to Node-30, the land surveyor who also acts as a local authority representative for the Lands Commission in the vicinity. This is because the land surveyor is always present to survey the land and also acts as a witness during every land transaction in the community.

Assuming that: (i) Scenario I occurred within a 10 year period and that all the data collected are correct, trustworthy and are supported by community members; and that (ii) the records are up to date and every change in the data over time is reflected in the network represented by the graph; then the graphs of the analysis could serve as a way to record and visualize a population register in a community. Should that happen, as described in Scenario I, assuming that, after a long period of farming on the land Node 1 gave to Node 24 for sharecropping, a family member of node 24, say node 26, would like to claim ownership of the land in question, this register will help identify which parties have the rights to that land. The record (e.g. an audio recording of the transaction process) may be used as evidence. It will be easier to identify node 26 as a fraudster and support such a finding with evidence through the visualization in the graph. The same will apply to any family member of node 30 or any of the officials. In the same way, if it has not been 20 years yet, then the oral agreement between node 1 and 9 still holds, which means that the network is still valid. Thus no family member in family 2 can lay claim to the land used as collateral by node 1. Should any change occur before or after the 20 years, the change will be effected in the data which will then be reflected in the created network.

5.4.2 Lineage Groups

As shown in Figure 5.9, the type of relationship existing between individuals in a family makes it easier to trace the lineage group or a person's family tree. For instance, Node-6 is a daughter of Node-4 (weight 2) who in turn is a sibling to Node-1 (weight 4). This makes Node-6 a niece of Node-1 (weight 6). Just as described in Scenario I and depicted in Figure 5.9, Node 1 is the chief and head of his family. This is why Node 1 is seen as the parent node relative to several child nodes. All the child nodes are family members of Node-1. As the family size increases, the size of family 1 will keep increasing. The changes in the family will be reflected in the network. Assuming this change continues for a long period of time, say 50 years, family members will always be able to trace their ancestry. No matter how complex the change is, family members will be able to trace their relationship to other members of their family through to their ancestors.

5.4.3 Interests in Land

Folding the two-mode Person – Land relation into a one-mode Person-Person relation reveals interesting relationships. In Figure 5.12 some nodes becomes outliers in the resulting network. This means that, although these individuals relate to node 1's land in some fashion, the case study data suggests that they currently have no claim to any of the land in question no matter their family or lineage group. They are family members of strangers such as a share cropper or the land surveyor. As shown in Figure 5.12, all the outlier nodes belong to families other than the landholding lineage family. This means that although related to the land, they are not right holders and cannot entertain the hope of inheriting family one's land.

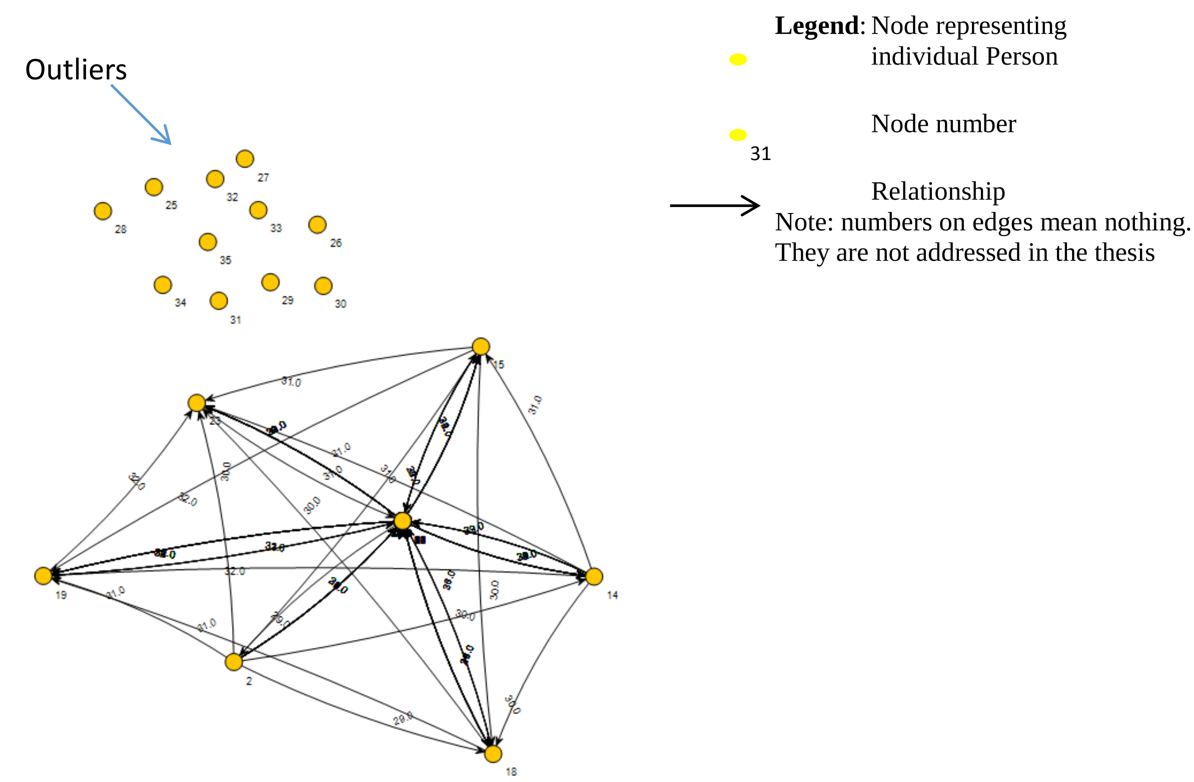


Figure 5.12: Person-Person relation based on the Person to Land relation.

5.4.4 Ties into different forms of evidence

Examining ties between different items of evidence and how they relate to land and people may address numerous questions such as - who was involved in which transaction and what evidence is there to support that they were involved (or not involved) in a particular transaction?

Figure 5.11 represents the two-mode Person – Media relations folded into the one mode Media – Media relation network. This may reveal which media files contain the most evidence in the dataset relevant to land tenure administration. As noted above media files 19 and 2 are very important.

In contrast, folding the two-mode Person – Media relation into the one-mode Person – Person network reveals individuals who are related to some items of evidence but are not exactly heirs or owners of the land in question. These individuals may be officials or family members of officials involved in land transactions but have no right of ownership. Such individuals are portrayed as the isolated nodes in Figure 5.13. It can be seen that none of those nodes belong to family one. They are either members of the sharecropper or surveyor's family.

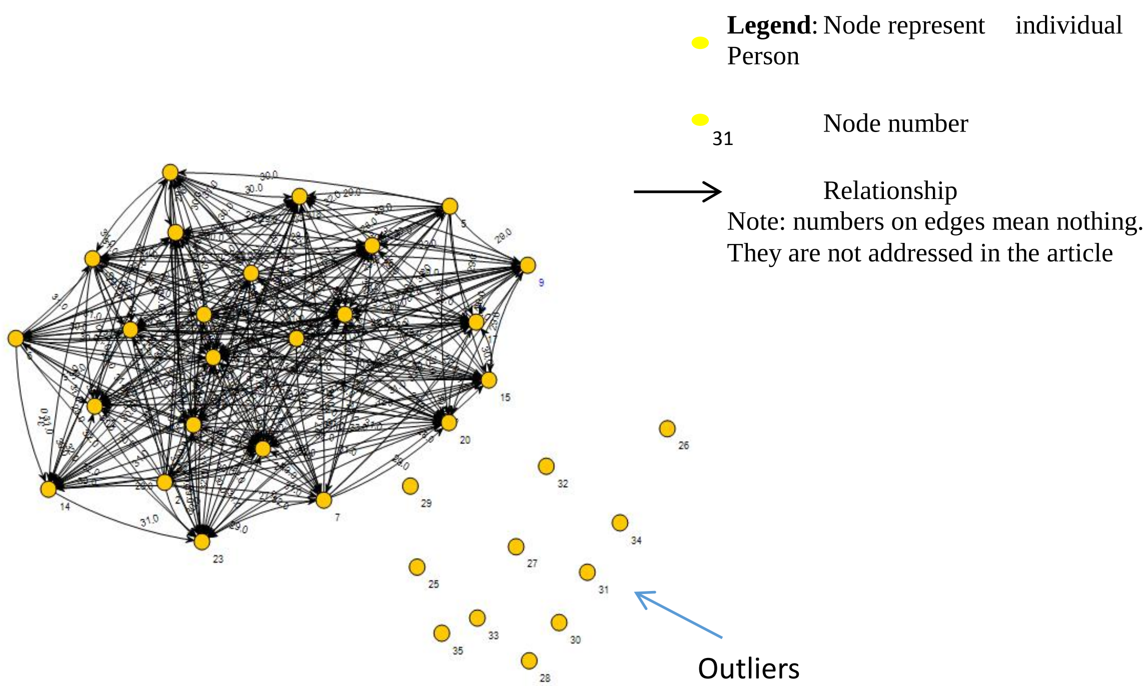


Figure 5.13: Person – Person relation based on the Person – Media relation

5.5 Scenario II

The analysis of the LTIS data based on Scenario II is described below; it is focused on the person - person and person - land relationships.

5.5.1 Person-Person Relationship Graph Representing Family Lineage and Stool Succession.

The Person – Person relationships from the data are shown in Figure 5.14 in the form of a graph. The graph represents family lineage and succession to the stool in Bortianor. The various clusters represent different families related to the stool. All the nodes except one (Node 87) represent individuals. Node 87 represents the stool in Bortianor. The network could have been a two mode network but, just as described in Chapter 4 (section4.3.8.2.1), in order not to deviate from using any other object apart from those allowed in the TTM, the stool was also modelled as a person and related to all the other individuals in the families who are in direct contact with the stool. Since it is about succession, the next in line to rule in a family was modelled as the parent node to whom all other family members relate. Each edge represents a relationship between people, including the stool. The weights on the edges represent the relationship type (see Table 4.2 in Chapter 4). The outliers represent individuals in the community who are not related to the families who have direct contact to the stool and thus are not related to the stool in any way.

Although not so clear in the graph, the calculated centrality measures reveal the important nodes in the graph to be nodes 87, 1, 33, 52, 79, 70, 85, 86, 18, 66, 77, 12, 5, 81, and 78. A reference to data will explain why these nodes are important in the network. This will be explained in the graph interpretation section. The calculated centrality measures are betweenness, closeness, and degree.

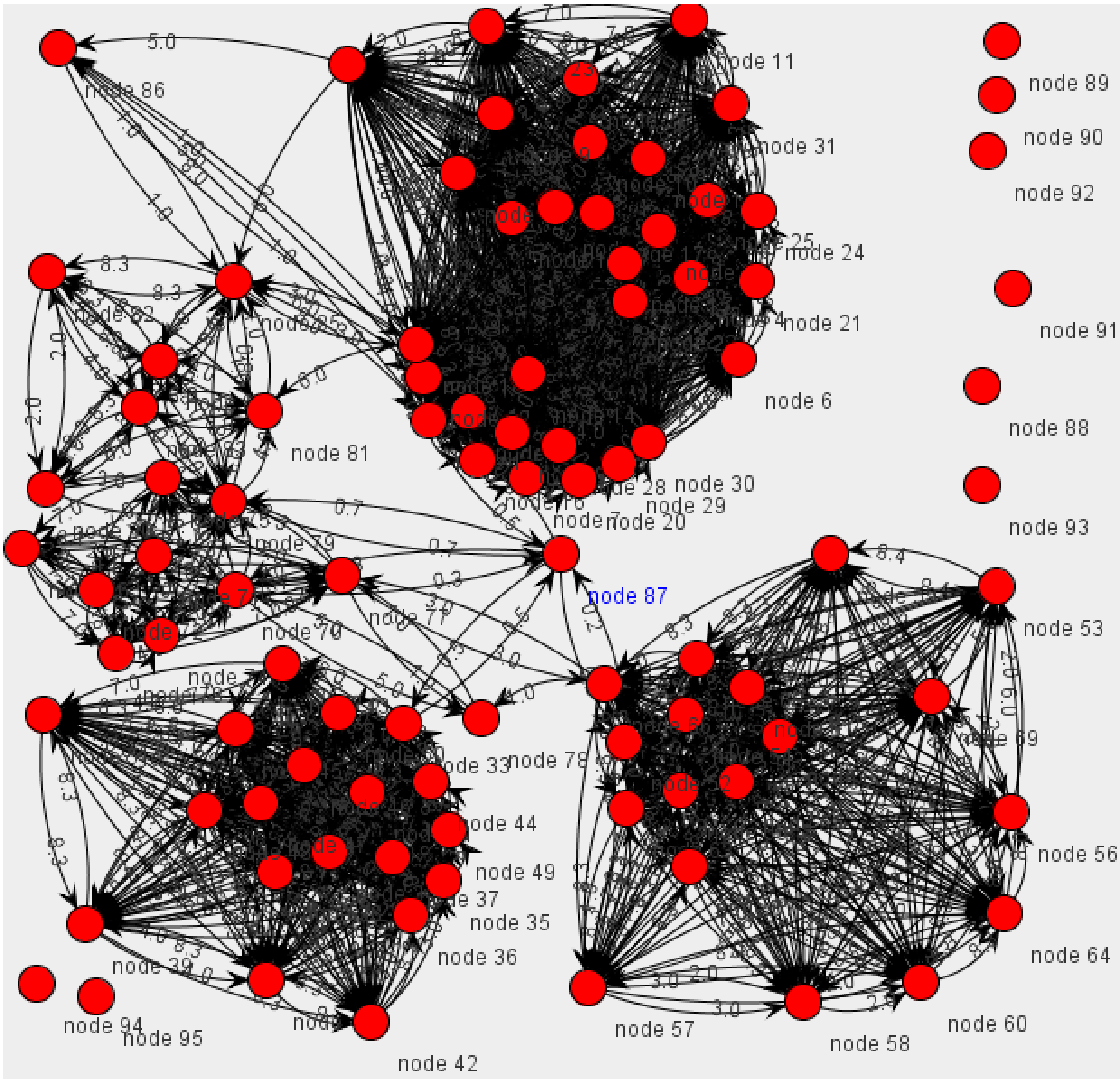


Figure 5.14. Social Network Graph of Person-Person relations for succession

5.5.3 Person – Parcel Object Relationships (Family A)

Person and land objects relations can record and show which individuals in a family own which parcels of family land. Analysis of this type of relations can also show who in the family are directly related to the owners of the said lands and could hope to inherit the land in future or even enjoy proceeds from the land sale at the family level. A proper record at the family level could later be extended to the whole communal membership and their relationship to individual lands in the community. This could also be useful in recording how land changes hands over time and can be used to show the chain of transactions that can be traced back to the original owner. In Figure 5.15, the lands belonging to family one are modelled to show who in the family is directly related to which of the family lands. Nodes 1, 2, 3 and 4 represent the lands in family one. All the other nodes represent individuals in the family and to which portion of the family land they are related.

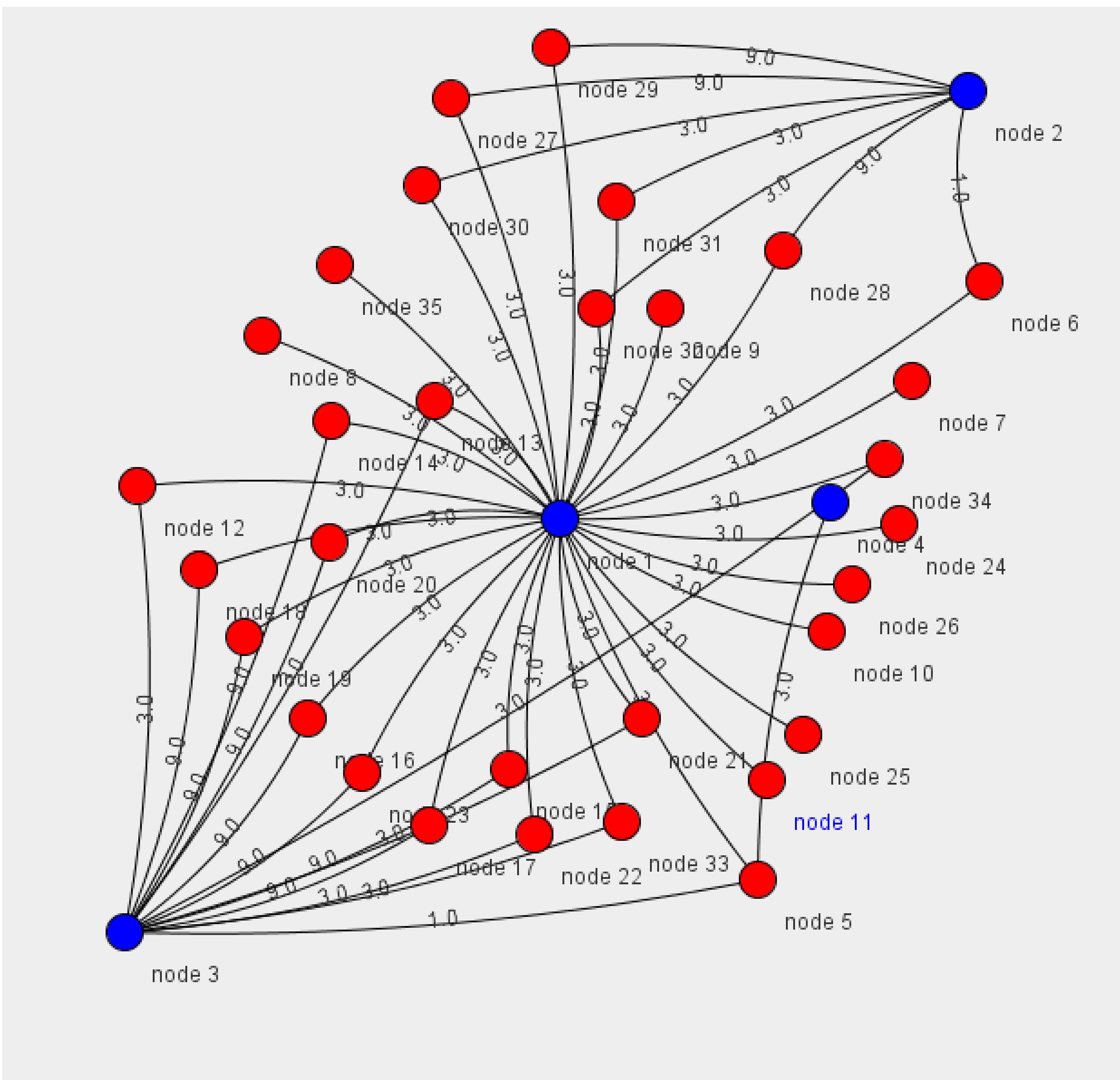


Figure 5.15: Parcels in Family A and how individual family members relate to particular parcels

5.5.2 Person- Person Relationship Graph Representing Land Sales in Family A.

This graph represents the person-to-person relationship modelled to reflect land sales in Family A. Every node in the figure represents an individual. The edges show which individuals are related to whom, as far as the recorded land sale is concerned. That is, who conducted land transactions with whom, and which parcels are involved in the transaction? The weight explains the type of relationship existing between the individuals. This relationship type is a special one as explained in Chapter 4 (section 4.3.8.2.4). Unlike graphs in Scenario I, weight on the edges is not preceded by the cluster number to which the individuals involved in the relationship belong. Rather, the weight is preceded by the number representing the parcel in question. The number that follows represents the person – person relationship type. For instance 2.8 means that the parcel in question is parcel 2 and the relationship between the individuals (Nodes) involved in the transaction is seller/buyer (0.8) relationship.

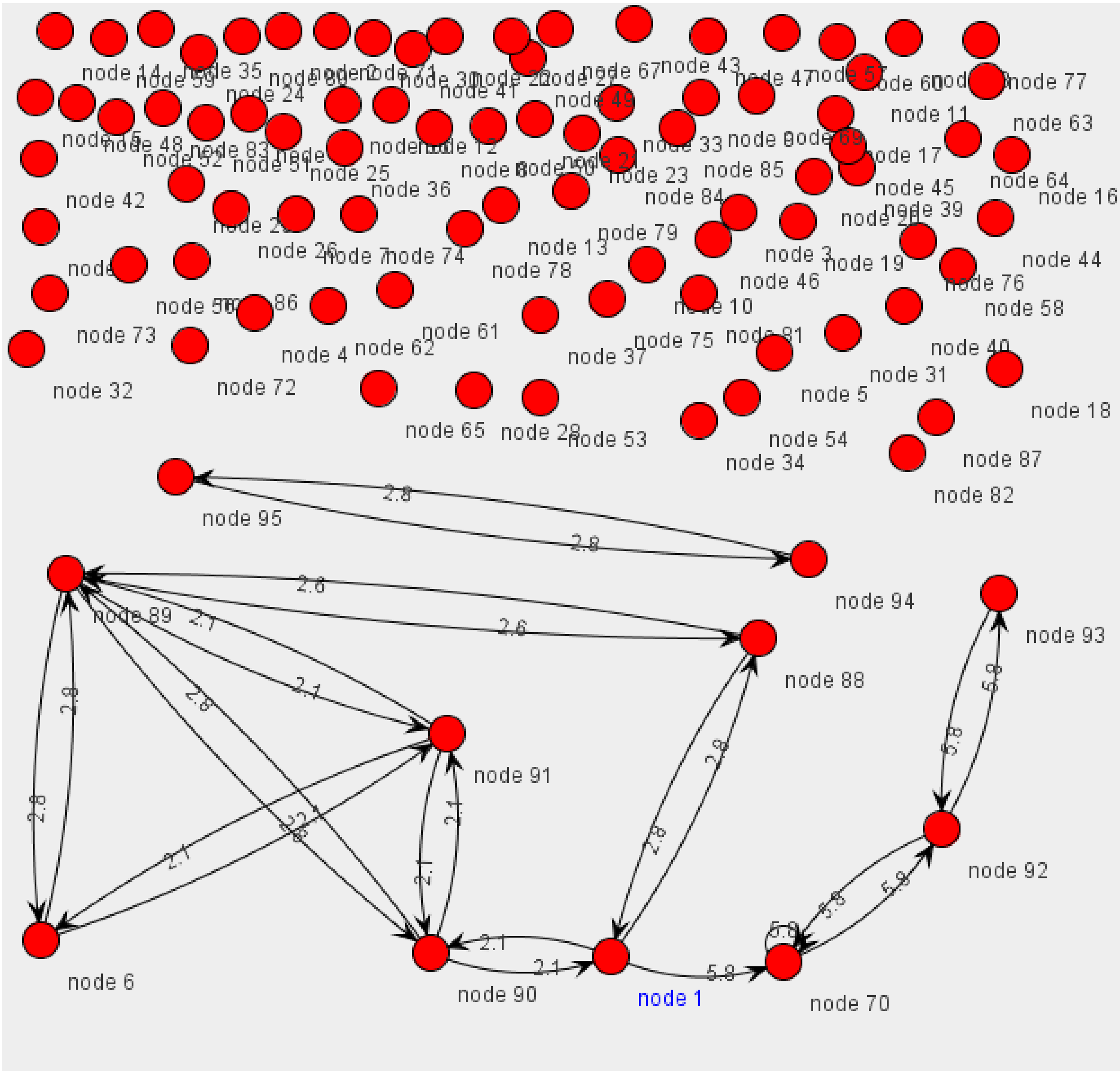


Figure 5.16: Land sales in a family in relation to all other community members

5.5.4 Parcel – Parcel Relationships (The lot of Family A)

As explained in section 5.3.2, above, network analysis can show a chain of consolidations of lands over time and can help trace subdivisions back to the parent parcels. Family one has a single land parcel, represented by Node 1 in Figure 5.17. Portions of this land have been allotted and are owned by individual family members. Each node in the figure represents a parcel while the edges represent relationships between the parcels. An examination of the data reveals that Nodes 2, 3, and 4 are all parcels that originating from the single larger parcel - Node 1. The calculated centrality measures support this fact. The weight (1.0) on the edges represent the relationships (emanated from/related to) between parcels. That is, Nodes 2, 3 and 4 emanated from Node 1.

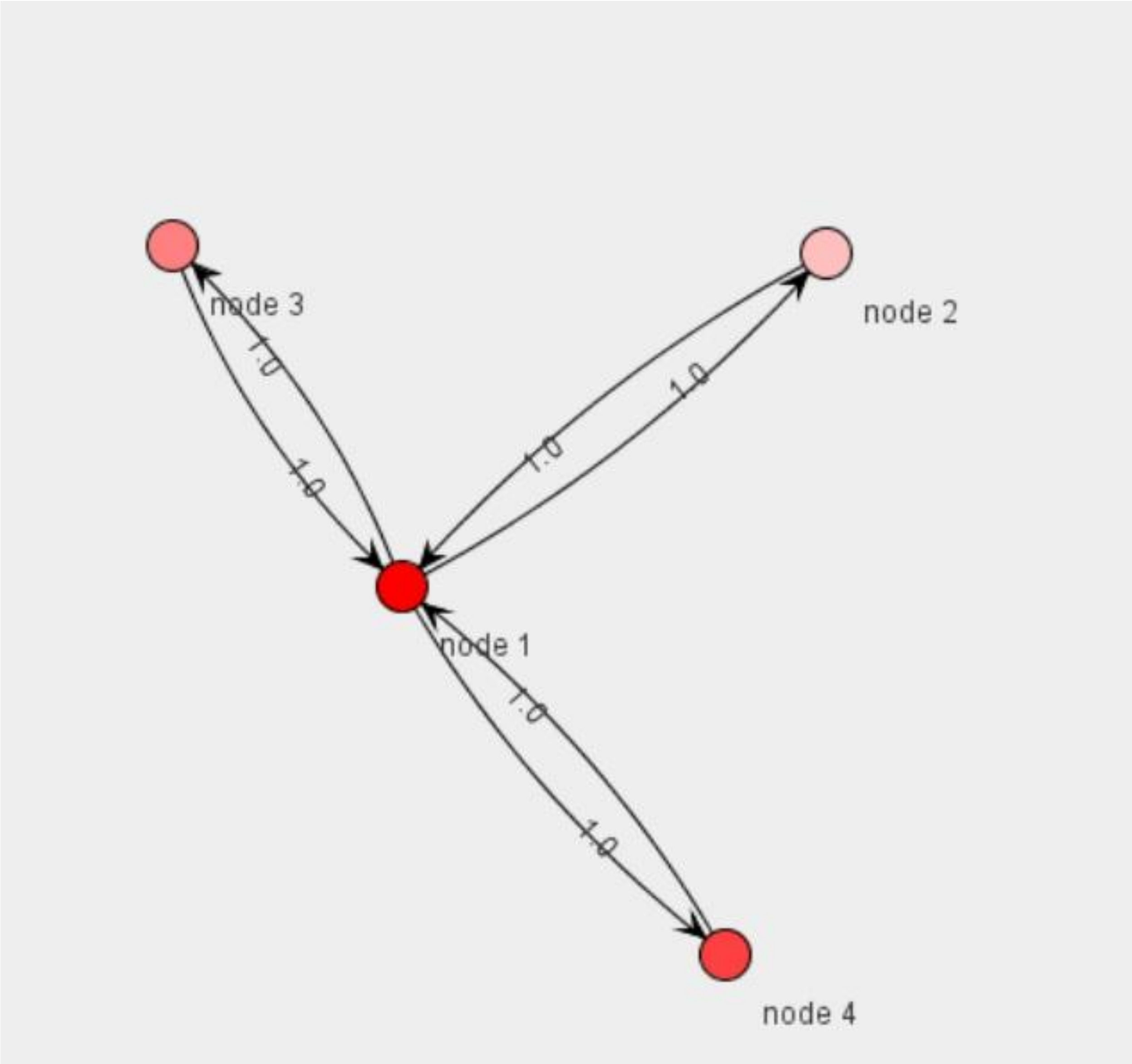


Figure 5.17: Graph of Family A's Lots (Land Object) Parcel-Parcel Network

5.6 Interpreting the Graph – Scenario II

The analysis of the social networks based on Scenario II was oriented on finding relationships between the individual's lineage groups, succession, status in family, inter-marriage between families, and multiple sales. These are based on modelling the Person – Person relationships stored in the database.

5.6.1 Lineage Groups

As shown in Figure 5.14, although the next in line to rule in the kinship group from the family was modelled as the parent node, it still shows who is related to whom and how these individuals are related. Although the parent node may not be the ancestor of all other family members, he can still trace his lineage to all his relatives and other family members can also trace their lineage to him (parent node) and their affiliation to other relatives. As in section 5.3.1 the weight between edges which describes the type of relationship aids in tracing lineage. The weights that are relevant to this section are tabulated in table 5.2. For instance, Node 1, who is the next in line to rule in Family A, is a brother (weight 4) to Node 5 who is also potentially the next in line (after Node 1) to rule as chief. Node 5 is married (weight 3) to Node 12 and they both have children, one of whom is node 18. This means that Node 18 is the nephew of Node 1. This form of lineage model could not only be used to trace family lineage but could also aid in determining who is next in line to rule, depending on the rules set in the family for choosing a successor.

5.6.2 Stool Succession

In Figure 5.14, Node 87 represents the stool of the community. All other nodes represent individuals in royal families or other families who are currently related to the stool based on Scenario II. The nodes relate to the stool through either the next in line to rule or the family member who is currently directly related to the stool. As shown in the graph, all nodes relate to the stool through either Node 1, 33, 53, 70, or 79. These nodes are the bottlenecks in the network. The weights on edges that connect each of the bottleneck nodes to the stool aid in determining the type of relationship that a given individual and his or her family may have, in general, to the stool.

At the top of the family hierarchy, Nodes 1 and 33 are the legitimate heirs (weight 0.5) to the stool. Although the scenario did not describe exactly how the two royal families came up with these two individuals as the next in line to rule, it is easy to pin-point them in the network. This also means that, individuals related to Nodes 1 and 33 belong to the royal family. Since only males – according to the tradition in the kingship group – can become chiefs in the customary society that was used as a model for this exercise, males from these royal families may entertain the hope of becoming chiefs someday.

Assuming that the rule in choosing a successor in Family A is that, once Node 1 becomes the chief, the next in line to rule should be Node 1's brother. If the brother is unable to rule due to death or any other circumstance, then he could appoint his eldest son to rule in his place. If Node 1 has no brother but, rather, has a sister, according to tradition, the eldest male child of node 1's eldest sister could rule. If such an assumption holds, then this register (graph) could aid in tracing the next chief. For instance as shown in the graph (Figure 5.14), Node 1 has two siblings: one sister (Node 6) and a brother (Node 5). He has three wives and has procreated with each of them. Since only males can become chiefs in the community and Node 1 has only one brother, it holds that, after Node 1, the next person in the family directly in line to the stool from Family A is Node 5. If, for some reason, Node 5 is unable to rule, then he could appoint the elder of his sons (Node 14 or 15). Similarly, if it happened that Node 1 had no brother (that is, Node 5 did not exist in the network), then Node 29, the only son of Node 6 (the only sister of Node 1) would then be the next in line to rule in the family. This type of analysis could provide an objective way of tracing succession in Family A. The other royal family could also decide their next chief using a similar process, depending on the rules governing chief selection in the family and the community in general.

Node 70 is the traditional priest of the whole kinship group. The traditional priest is viewed as the mediator between the chief and the greater community, in general, and the gods. His family members, especially the male children, have the prospects of becoming traditional priests. From the data, he has two wives (Nodes 72 and 73; weight 3). He has 5 children (Nodes 76, 77, 73, 74, 75) with both wives, all of whom are females. However, as shown in the Figure 5.14, based on the relationship their family has to the stool, they can only hope to be a priestess to the stool and not chiefs themselves.

Node 79 is the Queen Mother in the community. As the Queen Mother she is directly related to the stool. Females in her family can hope to be Queen Mothers since such a position can be held only by a woman. The Queen Mother's family is not a royal family from which a chief is selected, so males in this family cannot hope to be chiefs of the community. In Scenario II, the Queen Mother, with the help of the traditional priest, is trying to pronounce her only son (Node 83) as the chief. Again, assuming that the graph in Figure 5.14 faithfully and unambiguously represents the succession register, then it can easily be shown that Node 83 does not belong to a royal family from whom chiefs are selected and so cannot be a chief or even an acting chief. This could be used among the traditional council in settling any form of confusion in the community.

Node 52 is the King Maker and the current acting chief (weight 0.2) in the community. As tradition demands, in the absence of a chief, the King Maker acts as the chief until a legitimate chief has been installed. Since the acting chief may serve as the head of the community at a point in time, the position of the King Maker cannot be held by a female. Males in the King Maker's family may hope to be King Makers someday, but not chiefs of the community. They could only be acting chiefs if and only if there is no chief installed in the community at that particular point in time. In actual fact (as described in Scenario II), the current King Maker has been acting as a chief for 18 years. Assuming that he dies and, by the time of his death, there is still no chief installed in the community, a new King Maker may be selected based on the rules set out in the family for selecting King Makers. If, for instance, the rule says that, in the absence of the King Maker, the eldest male in the family is pronounced the next King Maker, it would mean that Node 64 will be the next King Maker and acting chief in the community. If King Makers act as chiefs for a long period of time, it is possible for future generations from the King Maker's family to think or at least to try to assert that they could be chiefs someday. This is another case where the succession register may become useful in explaining any misconception these future generations may have. In such a case, it would be possible to trace through the network and the changes recorded in the network to let them know their roots and the role they and their family play in the community. This could be a way of educating future generations as well as resolving any confusion related to succession in the community in general.

5.6.3 Intermarriage

From the succession graph (Figure 5.14) it can be seen that there have been inter-marriages between families. There is a marriage alliance between the King Maker's family and the traditional priest's family. Node 77, who is the daughter of the traditional priest (Node 70) is married to Node 66 (a son of the King Maker's brother (Node 64)) from the King Maker's family. The couple (Node 77 and Node 66) have a son (Node 78). The centrality measures of Nodes 66, 77 and 78 are high and this intermarriage explains why. They lie at a central location and are important in the network because they join together sub-groups (two different families) within the network. The question now is, to which family does this male child (Node 78) belong? Since he is a male and is part of the King Maker's family, could he be the King Maker in future? And could he also be the traditional priest in future? Determining what role such a child could play in the community, and whether he could be a priest or a King Maker will depend on the rules set in the community.

Although the community practices a patrilineal form of inheritance, such rules are negotiable and may have to be set explicitly and documented to help avoid conflicts in the future. If such rules exist, they could be incorporated and modelled to show the relationships as the network evolves. The most important aspect here is that, the analysis is able to show how these families merge, the individuals involved and the ties between them. As the network grows, the merging of these two families can always be traced and used in addressing any form of confusion.

Another form of marriage alliance is between the Queen Mother's family and Family A (one of the royal families). A male nephew of the Queen Mother (Node 85) got married to a female nephew (Node 18) of Node 1. The couple also have a son (Node 86). This also explains why nodes 18, 81, 85 and 86 are central in the network. They also link together two sub-groups (two families) in the network. In this case, another important question arises: Does this male child (Node 86) qualify to be a legitimate chief in the community someday? Again, the answer to this will also depend on the family and community rules relating to such situations. However, no matter the rules established in the community, the network can show how the families merge and aid in determining the status (royal or not) of the child in the genealogy of the community. The rules could be incorporated into the network as the network grows, based the types of relationships defined. This could also aid in resolving any confusion over governance in the future.

Assuming that the graph is a register that specifically records information on succession in Bortianor, then it could aid in determining the genealogical credibility of potential chiefs in the community. Assuming that these graphs were created prior to the conditions described in Scenario II, the graph can automatically be used to tell who belongs to a royal family, the role of certain individuals in the community, and whether they could be chiefs or not.

5.6.4 Proceeds Beneficiaries and Land Owners

In a family, although the family land ostensibly belongs to all family members in common, parcels of family land are apportioned among family members. Modeling of the network of property relations between family members and their respective apportioned lands in the context of, and in relation to, the entire family land base will help determine who owns which part of the family land and who enjoy proceeds from a land sale in the family. This could be extended to the community in general to identify which communal members and even strangers in the community are not related to specific family lands and so could neither enjoy proceeds from land sale, nor even lay claim to particular land parcels in the community. As shown in Figure 5.17 the family land of Family A has been sub-divided into at least four sections. Node 1 represents the entire family land and the other three sections (Nodes 2, 3, and 4) are the sub-divisions of the entire family land. From Figure 5.15 it is easier to identify who in the family owns or has exclusive use of particular parcels in the family. Since the entire family land belongs to all family members, all individuals belonging to Family A have partial rights in common (weight 3) in the family land (Node 1). It is also clear to see which portions of the land have been assigned to particular individuals in the family. Lot 2 (Node 2) was apportioned to the sister of the family head, Node 6. This means that she (Node 6) owns (weight 1) that portion of the family land. Her nuclear family members (Nodes

27, 28, 29, 30, 31, and 32) have partial interest (weight 3) in Lot 2. Her three children (Nodes 27, 28, and 29) stand to inherit (weight 9) that portion of land in the future.

Assuming a portion of the family land (Node 1) is converted into stool land and the family is compensated with a sum of money, the graph in Figure 5.15 helps to identify who in the family can enjoy proceeds from this transaction. As shown by this graph all nodes in the family (representing family members) can have a share of the amount of money given as compensation to the family. Consider the sale of one of the lots, say lot 2. Since that parcel has been assigned by the family and is owned by Node 6, the proceeds of the sale will be shared among the nuclear family members of Node 6, who are Nodes 27, 28, 29, 30, 31, and 32. In this case, if anyone is not included in this graph, that individual is deemed not to be a family member and cannot enjoy proceeds or any form of benefit from the family land.

5.6.5 Multiple Sales

Modelling relationships between individuals involved in land transactions (in this case sales) can help trace ownership, fraudulent activities and contested status in the lineage group. Figure 5.16 represents the sale of a particular land parcel in Family A. From the figure, it can be seen that Lot 2 (which is owned by Node 6) has been sold (weight 0.8) to an entrepreneur (Node 89) spoken of in the Scenario II. During the land transaction, a land surveyor (Node 91) was hired to survey the land and also to serve as a witness. The transaction was sealed and so lot 2 is now “owned” (exclusive use under lease) by the entrepreneur (Node 89). The same piece of land was sold to another stranger (Node 88) by the family head (Node 1), who also managed to get a land surveyor (Node 90) to survey the land and to act as a witness. As a result, Nodes 88 and 89 both claim to be legitimate owners of Lot 2. The entrepreneur (Node 89) then sold (weight 0.8) the land (Lot 2) he had purchased from Node 6 to a company (Node 90). The same piece of land has also been sold by an impersonator (a fraudster) (Node 94) to a stranger (Node 95). Figure 5.16 also shows that, the family head of Family A has sold a portion of the family's land (parcel number 5) to Node 70, who in-turn sold it to a stranger (Node 92) who also sold it to a company (Node 93).

Comparing Figures 5.16 and 5.15 shows that Node 6 is the actual owner of Lot 2. This means that, Node 89 purchased the land from the rightful owner (Node 6) while node 88 purchased the land from Node 1, who although is the family head do not own the land (Node 2 in Figure 5.15) and so have no right to sell it. From Figure 5.15, it clear that node 94 is not a family member of Family A and so cannot sell any of that family's land. Node 94 is in no way associated with Nodes 6 or 1. This has been clearly shown in Figure 5.16 as the network between Nodes 94 and 95 are separated from the other group in the network. All other nodes (individuals) in the community who have not been involved in any way in land sales as far as the land of family A or lot 2 is concerned, are shown as outliers in Figure 5.16.

This could aid in identifying who the actual owners of particular lands are and who has the right to sell. It could also show who may have been defrauded in land transactions. Properly structured for the purpose, the developed system could be a useful analytical tool during conflict resolution in a judicial committee.

5.7 Chapter Summary

The chapter has addressed both the primary and specific research objectives and questions. It portrayed the unique results of developing software and a land tenure information system that were, in themselves, evolutionary systems, both based on an evolutionary land administration model (TTM). It explored the functionalities embedded in the developed ubiquitous TTM-based LTIS. It showed how the system can be used to collect, store and create meaningful relationships between objects allowed in the TTM while ensuring data security and integrity.

The chapter also described the results of applying data mining and social network analysis techniques in analysing LTIS data stored in the TTM-based LTIS. The data used is similar to a deeds registration system, where a library of documents are stored that represent land transactions and relationships between people and the land, and their relationships involving land transactions. Thus, the system provides notice to the world of the rights dealt with by the records and, thus gives some level of credibility to the claims to rights. It also includes in the database digital records of oral agreements that are not otherwise formally accepted in conventional land registration systems. It is a simple trial model of a relational database application to document property relations. Applying network analysis and data mining techniques to the data revealed the topological footprints of family heads and lineage groups; identified who has the right to claim certain interests in land, who owns specific parcels of family lands, who benefits from proceeds of a land sale, and the existence of multiple sales of a single land parcel; incorporated and established connections to various forms of evidence; highlighted intermarriages between families; and produced a stool succession register. Network analysis portrayed the different social and physical relationships in the tenure model.

An unexpected outcome of the exercise is the potential usefulness of the methodology in analysing a deeds registry system to isolate errors and inconsistencies in indexing and in the records themselves. For example, if a relationship has not been entered correctly into a database or a document, network analysis may reveal different relevant relationships and so the error may be identified.

This work indicates that the methodology has potential in analysing complex tenure relationships, providing of course that the data can be obtained.

Overall, this work shows that social network analysis may be useful in improving the efficiency of land record systems and in communicating change in complex, transforming situations.

Chapter Six: Conclusion and Future Work

6.1 Introduction

This chapter summarizes the research presented in this thesis by giving account of the main points. It also discusses how the main points contributed to addressing the research objectives (see chapter 1 section 1.3) and questions (see chapter 1 section 1.4). The chapter continue to describe the limitations of the study and concludes with recommendations for future work.

The research explored the application of data mining and social network analysis techniques to land tenure data stored in a ubiquitous LTIS developed based on the TTM. This was accomplished by first applying the agile software development approach to developing a LTIS that used the regimes and functionalities embedded in the TTM. Mining and analysis was then conducted on simulated data stored in the LTIS. The data populated in the system were simulated to mimic the economic, political, and social conditions that exist in peri-urban communities. These conditions were established from literature and from a case study in a peri-urban community in Accra conducted by fellow researchers in the Land Tenure Systems Group at the University of Calgary (Berry 1993, Barry 2009, Barry 2011, Danso 2012). A summary of the chapters are provided below.

Chapter 1 provided a background to this study, described the problem and the context, outlined the objectives, posed the questions this study seeks to answer and the research methodology. Chapter 2, which is the first part of the literature review, discussed evolutionary land tenure system development. The chapter explained the advantage of matching an evolutionary land administration model with an evolutionary software development approach to developing a LTIS that could support the land administration function in peri-urban communities. The second part of the literature review was presented in chapter 3 where the concepts of data mining and social network analysis were described. This included an explanation of the data mining and social network analysis techniques that could aid in analyzing land tenure data in order to reveal pertinent information as well as hidden patterns in the data. The chapter also provided information on some of the SNA tools that could aid in conducting such analysis. Chapter 4 described the activities carried out to achieve the objectives in this research. This included a description of how the agile software development approach was adopted in developing a LTIS, how data was simulated for the study as well as how the analysis was conducted. Chapter 5 then discussed the results of the experimental work conducted in the research explaining the findings and output of the analysis. This chapter presents the key conclusions and limitations of this study and makes recommendations for future work.

6.2 Conclusion

This section outlines the results in relation to the primary and specific objectives of this study.

> **Specific objective (a):** To develop an evolutionary customary land tenure system that can cater for requirements dictated by the changing conditions in a peri-urban community where there is conflict over leadership positions, access to land and the governance of land sales and registration.

This objective was achieved by using an evolutionary land administration model (TTM) as the basis for developing an evolutionary LTIS using an evolutionary software development approach (extreme programming of the agile software development). To achieve this, research was consulted to understand the social, political and economic conditions in peri-urban communities. The purpose of this consultation was to understand how land transactions are conducted and administered as well as the effects of merging customary and statutory land administration regimes on indigenes in peri-urban communities. It was realized that the conditions can be volatile and subject to rapid change, possibly leaving indigenes in peri-urban communities landless.

This meant that the dynamic conditions in peri-urban communities required a land administration model that is flexible enough to evolve with these conditions. Hence the TTM was studied in-depth to understand the regimes and functionalities implied by the model. Also the LADM and STDM was studied to know and understand the functionalities embedded in these models and how they may be leveraged as the basis for developing a LTIS. Although evolutionary LTIS development is mostly spoken of, very few models achieve this in practice. Most land administration models are not created to help cater for such conditions in peri-urban communities. They utilize the top down approach that does not take into account that, one can end up with a system totally new and unrelated to what was originally intended (Barry 2012). The conclusion was drawn that, to be able to cater for the rapidly evolving conditions in peri-urban conditions, it is imperative to utilize a flexible land administration model such as the TTM - that could evolve with the conditions - as the blue print in developing a LTIS.

This was followed by applying the agile evolutionary software development approach, specifically Extreme Programming (XP) to developing a LTIS. This approach was chosen because the conditions in peri-urban communities were found to be vague and, as such, defining system requirements when developing a land tenure system may be difficult and sometimes impossible. Thus, an evolutionary software development approach may help tailor the system to suit the evolving conditions. Following this, a ubiquitous LTIS that can be used to create meaningful relationships between entities was developed to collect and store land tenure data in peri-urban communities.

It is therefore beneficial to develop an evolutionary LTIS using an evolutionary data model as the basis to prevent the system from becoming obsolete, in that, it could be tailored to suit the rapidly evolving conditions that characterize peri-urban communities. Thus a combination of an evolutionary land administration model (such as TTM), evolutionary software development approach could aid in developing and evolutionary land tenure system for peri-urban communities.

Specific objective (b): To apply data mining and social network analysis techniques in finding a lineage of interest in customary systems.

This objective was achieved by applying data mining and social network analysis techniques to data stored in the developed LTIS to discover relationships between entities. To accomplish this, two scenarios that mimic the social, economic and political conditions, as well as the family structure and the way land transactions are carried out in peri-urban communities were created.

The first scenario was inferred from the Talking Titler manual (Barry 2011) while the second was based on a study conducted in Bortianor, Accra by Danso (2012). The first scenario was focussed on the relationships between people, land and the respective evidence used to support land transactions. The second scenario focussed on the people-to-people relationships.

Different social network analysis tools were analyzed to identify their capabilities and functionalities in order for them to be utilized for analysis. The main functionality required was a tool that utilizes data mining techniques in network construction and that contains most of the algorithms for social network analysis; most importantly the creation of multi and single mode networks, community detection, creation of bipartite graphs, and computing centrality measures. These features were necessary because the TTM-based LTIS data consists of many-to-many relationships between entities (people, land and media files) which relate to each other by means of shared features. Data mining techniques such as association rule mining and clustering is necessary to identify and group entities with common features together. Computing network metrics on the graphs (network) created in the process aid in identifying power players and bottleneck actors in the network.

The tools analyzed were UCINET, JUNG, NetMiner, and NetDriller. It was found that, although all of the tools have the appropriate functionalities in computing network metrics and conducting functional network analysis, few of the tools utilized data mining techniques in network construction. Thus, NetDriller was chosen as the tool for mining and analysis.

Based on the first scenario, the LTIS was populated with fictitious data and the relationships between the entities were created. The data was then extracted and loaded into NetDriller for mining and analysis. This first analysis was to test the ability of NetDriller in identifying pertinent relationships in land tenure data. The betweenness centrality measure was calculated to identify important nodes in the network. Analysis of data from the first Scenario Identified family heads, lineage groups, interests in land, and ties into different forms of evidence in the data.

The LTIS was again populated with data generated from the second scenario which was created based on the Bortianor case study. Analysis of this data revealed lineage groups, stool succession, intermarriages, beneficiaries of proceeds from land sale and multiple land sales.

It was proven and concluded that data mining and social network analysis techniques could be used to examine changing customary systems, reveal patterns between people – perhaps hidden patterns, identify supporting evidence during land transactions as well as new, previous and unknown relationships between entities. The relationships revealed in this study are seldom modelled by conventional LTIS. One of such of relationships is the relationships between beneficiaries of a property.

In all, mining and analysis identified family heads, lineage groups, interests in land, ties into different forms of evidence, stool succession, intermarriages, proceeds beneficiaries and parties involved in multiple land sales.

Primary objective: The development of an evolutionary TTM-based LTIS driven by data mining and social network analysis.

The fulfillment of the two specific objectives above led to achieving the primary objective. Developing a TTM-based LTIS and applying data mining and SNA techniques to analyze the data stored in the system revealed ways in which the TTM could be improved. Based on the analysis it was realized that the TTM could be improved with the inclusion of two specialized classes. The first is a class for listing beneficiaries while the second is for defining land access processes. A specialized class for listing beneficiaries may keep account of individuals in a family or community who may inherit or benefit in any way from a property. A class for land access processes will define regimes that will oversee the use of any particular parcel based on the defined rights an entity (Person) may have to the land in question. These new specialized classes incorporated into the TTM could then be fused into the developed LTIS and tested for efficiency.

Ultimately, the study revealed that the TTM could help ensure equity in peri-urban communities. Also mining and analysis techniques can aid in visualizing and uncovering hidden patterns and could aid in improving an evolutionary land administration model.

6.3 Limitations

The analysis depends on relevant data. Acquiring relevant data from the real world could be a challenge. The data collected based on stories told by community members could be biased and subjective to one's personal affiliation. Stories may evolve and may be manipulated to suite one's agenda. It is therefore imperative to find a way to verify the authenticity of the stories told in peri-urban communities in order not to bias the results of the analysis in favor of other people's agenda. At other times due to privacy and protection of personal information, people may not freely give up their personal (genealogical, property) information, thus the data may not be available.

Larger datasets yield larger networks and may thus be difficult to visualize all nodes at a go. Since nodes may be clustered together, one may have to move them around in order to view clearly the relationships that exist between nodes. Also some skill is required in analysing the graphs.

6.4 Future Work

This section makes recommendation for further development of the current LTIS as well as analysis of the data.

Although the TTM-based LTIS has been tested with simulated data, it will be best to install and use it in a peri-urban community with data from a real case study. As outlined in section 6.3, there are competing factions with their respective stories told based on their preferences (which may be concocted). Thus in a real world, measures should be put in place to ensure that, for instance, a kingship graph generated based on the data in the system is objective, scientific and otherwise credible, especially in light of the subterfuge, in-fighting and trickery that abounds in matters in peri-urban communities. These measures could be in the form of procedures for conducting interviews and entering the relational data. Such procedures could serve as safeguards (such as asking the same question three ways) put in place to make certain that the system contains the absolute last word of the hereditary relationships described by community members. It will also be vital to utilize authorities (such as community leader, civil courts, constituted traditional courts, council of elders) as arbitrators, so that the system may not be regarded as usurping the current means and authority behind arbitrating land related issues where there is ambiguity as to hereditary rights in the community.

Through this the system may be tested against a real world situation in a peri-urban community. Since conditions rapidly evolve, the method of applying an evolutionary software development approach to developing an evolutionary LTIS based on an evolutionary data model could also be tested out for applicability and efficiency.

Also the developed TTM-based LTIS could be improved by incorporating a mapping component into the application. In this case, when relationships are created between other entities and a land parcel, users could be allowed to draw the actual demarcation on a virtual map of the area in question, which will then be stored in addition to the evidence stored and related in the system. It is imperative that, the map be drawn to scale so that it could be a valid representation of the actual parcel in the real world. It was for this reason that PostgreSql was used in developing the current LTIS. Also the mobile application could be fully developed and extended so that it could be used on smart phones as well other operating systems in addition to windows.

Furthermore, mining and analysis of the land tenure data in this study was mostly focussed on the people-to-people and briefly on the property and media relationships. Although the reference instrument class allowed in the TTM was included in the developed system, its relationships with other entities were not included in the analysis for this study. There are numerous relationships that exist in peri-communities. New relationships are also created day in and day out as lands are transacted. The family structure as well as the rapidly economic, social and political changes in these communities warrants further analysis. Conducting further analysis to include all entities could aid in creating a repository of meaningful relationships that exists between entities and that could be leveraged in creating a register for families and communities in general. It could also create a genealogy-based register for a common property regime which allows leases for the exclusive use of specific parcels. Future experimental work could focus on types of relationships that might flow from a social anthropological study of a lineage group combined with a local set of land records.

The current analysis conducted in this research was done in a tool separate from the TTM-based LTIS. What if the algorithms for mining and analysis were embedded in the Talking Titler system and an automated way of data preparation was included in the application? The LTIS could then become a "one stop shop application", in that, it will collect, store and create meaningful relationships between entities, provide a pictorial representation of the created networks that exist in the data so that the not so clear relationships could be easily identified as well as a means to analyze the data by computing various metrics on the network. With such a system, changes in the network which may result from the changes in the created relationships in the land tenure data could be visualize instantly or near real time.

Future research could extend the application of data mining to include face detection. Through this, the faces of individuals in media files (pictures, videos) could be detected, authenticated and related to other entities. As land related data become more readily available and a mixture of complex data increases in volume, platforms such as Hadoop could be used in developing LTIS and analyzing the land related data stored in such systems.

Bibliography

Abrahamson, P, Salo, O, and Ronkainen, J 2002, Agile Software Development Methods Review and Analsis, Espo 2002. VTT Publications 478.107 p.

Arko, A-A, 2011, Adapting land administration to the institutional framework of customary tenure, the case of peri-urban Ghana, PhD thesis, Delft University Press.

Apache 2011, viewed 22nd August 2011, <http://httpd.apache.org/ABOUT_APACHE.html>

Augustinus C, Lemmen C and van Oosterom P 2006, Social Tenure Domain Model Requirements from the Perspective of Pro-Poor Land Management. Promoting Land Administration and Good Governance 5th FIG Regional Conference, Accra, Ghana, March 8-11, 2006.

Barry M, 1998. *Secure Land Tenure for Informal Settlement Communities: The Effectiveness of the Cadastral System in Cape Town*. In: Proceedings of the International Conference on Land Tenure in the Developing World with a focus on Southern Africa. Cape Town, South Africa, 27-29 January 1998.

Barry, M B 1999. Evaluating Cadastral systems in Periods of Uncertainty: A Study of Cape Town's Xhosa-speaking Communities. PhD thesis, University of Natal, Durban.

Barry M, 2009. *Standard for Oral Tradition Evidence: Guidelines from First Nations Land Claims in Canada.* In: Federation Internationale des Geometres (FIG), 7th FIG Regional Conference. Hanoi, Vietnam, 19-22 October 2009.

Barry M, 2009. *Dysfunctional Communal Property Association: The Case of Elandskloof South Africa.* In: Federation Internationale des Geometres (FIG), FIG Working Week 2009: Surveyors Key Role in Accelerated Development. Eilat, Israel, 03-08 May 2009.

Barry, M 2011. Talking Titler Manual, viewed 15th March 2011,

<http://ucalgary.ca/mikebarry/talkingtitler/downloads>

Barry, M., Roux, L., & Danso, E., 2012. Land Registration Usage Theory: A Case Study in Ghana. Viewed 10th October 2013,

<http://www.landandpoverty.com/agenda/pdfs/paper/barry_full_paper.pdf>

Barry, M, Molero, R & Muhsen, A-R 2013, Evolutionary land tenure information system development: the talking titler methodology, viewed 3rd August 2013, <http://www.fig.net/pub/monthly_articles/february_2013/feb_2013.html>

Barry, M, Muhsen A-R and Molero R 2011. Alternative Land Tenure Information Systems: The Talking Titler System - an Evolutionary Design based Methodology. Proceedings AfricaGeo, Cape Town 31 May – 2 June 2011.

Barry, M, and Rüther H., 2005. Data Collection and Management for Informal Settlement Upgrades. *URISA Journal*, 17(1), 43-52.

Barry M, and Khan K, 2005. *Law and Policy Implications of Multimedia Land Records: The Talking Titler Project*. In: Federation Internationale des Geometres (FIG), FIG Working Week and 8th International Conference of the Global Spatial Data Infrastructure. Cairo, Egypt, 16-21 April 2005.

Barnes, J.A. and Harary, F, 1983, Graph Theory in Network Analysis, Social Networks 5, Elsevier Science Publisher B.V, pp. 235-224.

Beck, K, 1999, Extreme Programming Explained: Embrace Change, ISBN: 0201616416, 224 p.

Beck, K, Beedle, M., Bennekum, V.A. and Cockburn, A, 2001, Manifesto for Agile Software Development, Software Engineering Practitioners Approach, Roger Pressman.

Berry, S. 1993, *No condition is permanent: the social dynamics of agrarian change in sub-Saharan Africa*, University of Wisconsin Press, Madison.

Boehm, B, 1998, A Spiral Model of Software Development and Enhancement, IEEE Computer 21, 5, 61-67.

Boehm, B, 2002, *Get Ready for Agile Methods, With Care*, IEEE Computer, January 2002, pp. 64-69

Borgatti, S, 2002, NetDraw: Graph Visualization Software. Harvard, Analytic Technologies.

Borgatti, S.P, Everett, M.G. and Freeman, L.C. 2002. UCINET for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Borgatti, S.P, 2006, Identifying Sets of Key Players in a Social Network, Comput Math Organiz Theor, Business Media, Springer, LLC 2006.

Borgatti, S.P, 2002, Stopping Terrorist Networks: Can Social Network Analysis really contribute? Sunbelt International Social Networks Conference 13-17. New Orleans.

Borgatti, S, Halgin, D., and Dejordy., 2008, An Introduction to UCINET and NETDraw, 2008 NIPS UCINET and NETDraw Workshop, Organizational Studies Depatment, Carroll School of Management, Boston College.

Borgatti, S, 2004, Social Network Analysis, Carroll School of Management, Boston College, http:faculty.ucr.edu/~hanneman/nettext/.

Brandes, U, 2008, Social Network Analysis and Visualization, IEEE Signal Processing Magazine, Springer, pp.147 – 151.

Breiger, R.L, 1981c. The social class structure of occupational mobility. *American Journal of Sociology*. 87, 578-611.

Castelo, R and da Silver, M M, The Value of Flexibility, Contemporary Theoretical Models in Information Systems, pp. 141-163.

Cederbaum, I, 1984, Some Application of Graph Theory to Network Analysis and Synthesis, IEEE Transactions on Circuits and Systems, Vol. Cas-31, No.1, pp. 64-68.

Cockburn, A, 2002, Agile Software Development, Addison Wesley, 2002.

Cyram Co. NetMiner4, Viewed 29th October 2012,<http://www.netminer.com/sub01/sub01.php>

Danso, NE 2013, Peri-urban Land Tenure in Ghana (Accra): Case Study of Bortianor', MSc thesis, University of Calgary, Calgary.

Davis, F D. 1998, Perceived usefulness, Perceived ease of use and user Acceptance of Information Technology, MIS Quarterly, Vol. 13 (No. 3), pp. 319-340, Sep., 1989

Dale, P, 2000. "The Important of Land Administration in Development of Land Markets-A Global Perspective in Land Market and Land consolidation". Central Europe IV, TU – Delft, pp.31-41.

Elmasri, R. and Navathe, S.B, 2001, *Fundamentals of Database Systems*, Addison-Wesley Publishing Company 2010.

Elliott, G, 2004, *Global Business Information Technology: An Integrated Systems Approach*, Pearson Education.

Erickson, B, 1988. The relational basis of attitudes. In Wellman, B., and Berkowitz, S.D. (eds), *Social Structures: A Network Approach,* pages 99-121. Cambridge, England: Cambridge University Press.

Faust, K, 1997. Centrality in Affiliation Networks. Soc. Networks vol 19, 1997, pp 157-191.

Fourie, C. 1998, The Role of Local Land Administrator: an African perspective, Land Use Policy, 15(1).

Faust, K, and Wasserman, S, 1992, Centrality and Prestige: A Review and Synthesis. Journal of Quantitative Anthropoloy. 4, 23-78.

Fischer, C.S, 1982, *To Dwell Among Friends: Personal Networks in Town and City.* Chicago: University of Chicago Press.

Fourie C.D, 1993. A New Approach to the Zulu land Tenure System: An Historical Anthropological Explanation of the Development of an Informal Settlement. Unpublished PhD thesis, Rhodes University, Grahamstown, South Africa.

Fowler, M and Highsmith, J, 2001, The Agile Manifesto, Software Development, August 201.

Fowler, M 2005, The New Methodology, viewed 18th January 2014,

<http://www.martinfowler.com/articles/newMethodology.html>

Freeman, L, 2005, Graphic Techniques for Exploring Social Network Data, Models and Methods in Social Network Analysis, New York: Cambridge University Press, pp. 248-269.

Gates, B, 2012, viewed 20th September 2013,

<http://www.thegatesnotes.com/Topics/Health/Three-Reasons-Im-Excited-About-the-Future-of-Ghana >

Han, J, Kamber, M. and Pei, J, 2012, *Data Mining Concepts and Techniques,* Morgan Kaufmann Publications 2011.

Hanneman, R, Robert, A., and Riddle, M., 2005, Introduction to Social Network Methods. Riverside, CA: University of California, Riverside. Viewed October 20th 2013, < http://faculty.ucr.edu/~hanneman/nettext/>.

Hall, M, Frank, E, Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H., 2009, The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.

Hall, R, 2003, Rural restitution: evaluating land and agrarian reform in South Africa. Paper 2. Programme for Land and Agrarian Studies, University of the Western Cape, Cape Town, South Africa.

Hay, G, and Hall, G. B. 2009. Architecture for an open source land administration systems. FIG working week, Eilat, Israel, 3-8 2009.

Hespanha J,van Oosterom P, Zevenbergen J, and Dias G 2006. A modular standard for the cadastral domain: Application to the Portuguese Cadastre. *Computers, Environment and Urban Systems*, 30 (2006) 562–584.

Hespanha, J, J. van Bennekom-Minnema, P. van Oosterom, and C. Lemmen. 2008. The model driven architecture approach applied to the land administration domain model version 1.1 – with focus on constraints specified in the object constraint language. Paper presented at FIG Working Week 2008, Integrating Generations.

Hill, J, Szewczyk, R, Woo, A, Hollar, S, Culler, D. and Pister, K 2000, System Architecture Directions for Networked Sensors, Cambridge, MA, USA.

Highsmith, J, 2002, *Agile Software Development Ecosystems.* Adisson-Wesley, Boston, MA.

Highsmith, J. and Cockburn, A, 2001, Agile Software Development: The Business of Innovation, IEEE Computer Society.

ISACA 2009, Certified Information Systems Auditor, CISA Review Manual 2009.

ISACA 2010, Certified Information Systems Auditor, CISA Review Manual 2010.

ISO /TC211 N2385. 2008. New work item proposal, geographic information - land administration domain model (LADM), international organization for standards, Lysaker, Norway. filed 2008.

Kapferer, B, 1969, Norms and the manipulation of relationships in a work context. In Mitchell, J.C.(ed), *Social Networks in Urban Settings*, pages 83-110. Paris: Mouton.

Koochakzadeh, N, Sarraf, A, Kianmehr, K, Rokne, J & Alhajj, R., 2011, NetDriller: A Powerful Social Network Analysis Tool, *11th IEEE International Conference on Data Mining Workshops*, Vancouver, pp. 1235-1238.

Lemmen, C., and van Oosterom, P., 2006. Version 1.0 of the core cadastral domain model. Paper presented at XXIII FIG Congress – shaping the change, Munich, Germany.

Lemmen, C., Augustinus C., van Oosterom P., and van der Molen P. 2007. The social tenure domain model – design of a first draft model. Paper presented at Strategic Integration of Surveying Services, FIG Working Week 2007.

Lindstrom, L. and Jeffries, R., 2003, Extreme Programming and Agile Software Development Methodologies, CRC Pres.

Martino, F and Spoto, A, 2006, Social Network Analysis: A Brief Theoretical Review and Further Perspective in the Study of Information Technology, Psychology Journal, Volume 4, Number 1, pp. 53-86.

May, E L, Zimmer, A B 1996, The Evolutionary Development Model for Software, Hewlett-Packard Journal, viewed 17 August 2013,

< http://www.hpl.hp.com/hpjournal/96aug/aug96a4.pdf >

Mirkin, B, 2011, Core Concepts in Data Analysis: Summarization, Correlation and Visualization, ed. 1, Springer.

Muhsen, A. 2008. Developing Multimedia Land Record Systems. MSc. University of Calgary.

Muhsen A-R., and Barry M. B., 2008. *Technical Challenges in Developing Flexible Land Records Software*. In: Federation Internationale des Geometres (FIG), FIG Commission 3 Workshop on Spatial Information Management Towards Environmental Management of Mega Cities. Valencia, Spain. (Revised version in Surveys and Land Information Science 68(3), 171-181).

Mohsen, J and Abolhassani, H 2006, Different Aspects of Social Network Analysis, 2006 IEEE/WIC/ACM International Conference on Web Intelligence, Computer Society, IEEE.

North, M., 2012, *Data Mining for the Masses*, Creative Common Attibution 3.0 License.

National Communication Authority, 2012, viewed 24th June 2013, <http://www.nca.org.gh/

O'Madadhain, J, Fisher, D., Smyth, P., White, S., and Boey, Y-B., 2005, Analysis and Visualization of Network Data usisng JUNG.

<http://jung.sourceforge.net/doc/JUNG_journal.pdf>

Onbile 2012, Most Popular OS in Africa, viewed 10th August 2013,

< http://www.onbile.com/info/most-popular-os-in-africa/>

Patel, N V 2009, The Theory of Deferred Action: Informing the Design of Information Systems for Complexity, Contemporary Theoretical Models in Information Systems, pp. 164-191.

PostgreSql 2011, viewed 22nd August 2011, <http://www.postgresql.org/about/ >

PHP 2011, viewed 22nd August 2011,<http://php.net/manual/en/intro-whatis.php>

Roux, L and Barry, M 2001. Using Video Imagery in Land Tenure Information Systems: A Study of the Algeria Communal Property Association. CONSAS 2001 Conference, Cape Town.

Roux L, Barry M, and Whittal, J, 2009.*Analytical Methodologies to Guide Interventions Informal Settlements*. In: Federation Internationale des Geometres (FIG), FIG Working Week 2009: Surveyors Key Role in Accelerated Development. Eilat, Israel, 03-08 May 2009.

Schindler, S, 2008, Agile Software Development Methods and Practices in Austrian IT-Industry: Results of an Empirical Study, Computer Society, IEEE.

Smith, J, 2001, A Comparison of RUP and XP. Rational Software White Paper. http://www.rational.com/media/whitepapers/TP167.pdf.

Srivastava, J, 2008, Data Mining for Social Network Analysis, IEEE ISI Invited Talk (III), Intelligence and Security Informatics, Taipei, pp. XXXIII-XXXIV.

Shore, J. and Warden, S., 2008, *The Art of Agile Development*, Oreilly, USA.

Thurman, B, 1980. In the Office: Networks and coalitions. *Social Networks*. 2, 47-63.

Van Oosterom P, Lemmen C, Ingvarsson T, van der Molen P, Ploeger P, Quak W, Stoter J. and Zevenbergen J 2006, The Core Cadastral Domain Model, Computers, Environment and Urban Systems, 30(2006), 627-660.

Van Oosterom, P, and Lemmen, C., 2002. Towards a standard for the cadastral domain: Proposal to establish a core cadastral data model. Paper presented at COST Workshop 'Towards a Cadastral Core Domain model', Delft, The Netherlands.

Wasserman, S. and Faust, K, 1994, Social Network Analysis: Methods and Applications, Cambridge University Press.

Well, D. 1999, When Should Extreme Programming be Used? Viewed 1st December 2013, < http://www.extremeprogramming.org/when.html>

Wellman, B, 1979, The community question: The intimate networks of East Yorkers, *American Journal of Sociology*. 84, 1201-1231.

World Bank 2003, Ghana Land Administration Project, Project Appraisal Document. Washington, D.C. (World Bank): 103.

Xu, K, Tang, C, Tang, R, Ali, G, Zhu, J, 2010, A Comparative Study of Six Software Package for Complex Network Research, *Second International Conference on Communication Software and Networks*, Singapore, pp. 350-354.

Zachary, W.W., 1984. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research.* 13, 452-473.