Social Network Analysis and Data Mining of Land Tenure Information in the Talking Titler System

Kwame ASIEDU and Michael BARRY, Canada

Key words: social network analysis, land administration, land information system, talking titler model

SUMMARY

Networks of social relationships are a fundamental attribute of land tenure systems. In changing situations, individuals and groups may belong to complex, multi-layered, changing and perhaps conflicting social networks at any particular time. Standard land administration models serve as blueprints for developing land tenure information systems (LTIS), but many of them are derived from western administration systems and fail to accommodate some of the complex realities of many customary systems and unstable, changing states of affairs such as post conflict situations or the major changes that occurred in post-apartheid South Africa. The paper describes exploratory data mining and social network analysis (SNA) techniques of simulated social data stored in an ubiquitous (web and mobile) Land Information System (LIS) based on the Talking Titler Model (TTM). The Talking Titler Model is a flexible land administration model suited for securing tenure in complex and uncertain situations. The design philosophy for local level land tenure information system development is that the design should be grounded in the data, and evolves according to what local circumstances dictate. All preconceived notions about data classes, such as person, land objects and reference instrument (e.g. title deed), and the relationships between them should be subject to rigorous, continual, critical examination to handle social change and improve the suitability of the original design model to local circumstances. Social network analysis may facilitate visualizing complex relationships between objects (people to people, people to land, and people to particular data types or items). Social network analysis and data mining may reveal relationships that were not identified in the initial model or uncover relationships that emerge as social change takes shape. It may also identify errors and inconsistencies in a land records system.

Social Network Analysis and Data Mining of Land Tenure Information in the Talking Titler System

Kwame ASIEDU and Michael BARRY, Canada

INTRODUCTION

Conventional land registration systems are weak at modelling 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 (Barry and Rüther 2001, 2005). This paper describes an exploratory test of data mining and social network analysis techniques to land tenure information stored in a web based version of the Talking Titler system. The scope is a test data set to calibrate and test the methodology and the Netdriller software.

We used data mining and social network analysis techniques to extract social relationships from a set of land records, and in turn constructed 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 paper 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 early work indicates that the methodology may fill a need for modelling complex, changing social relationships in a land tenure information system that current systems fails 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.

The situational context for the experimental work 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 is in many cases is the 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* 2012).

There is a need for land tenure information systems that handle complex, changing situations such as rural customary lands that transform into conurbations due to urbanisation pressures. The second author's research group has been working on the Talking Titler land tenure information model as a design and testing tool for unusual land tenure situations for a number of years (Barry et al 2002, Roux and Barry 2001). One challenge is how to handle social change in the information model, as the effects of social change may not be factored into land reform and land tenure administration projects during major social and political transformation (Walker 2008).

One social change theory holds that conflict and competition over access to power, land, and other resources likely accompany this transformation and coalitions will form and dissolve as part of this change. Thus social networks will change as coalitions change (Barry 1999, 2006). Tenure systems can transform rapidly as tensions between individualised tenure and communal systems unfold, rules are created and manipulated, entrepreneurial individuals appropriate land, and perhaps dispose of it to outsiders (Fourie 1993).

A land professional in Ghana ventured that a manifestation of this is that once a customary area is urbanised, a major cause of litigation is that the same piece of land may be sold multiple times. One cause is a consequence of contested status within a family lineage and attempts to manipulate and change the traditional rules concerning status. Strangers purchase land off a family member whose status to alienate family land is contested. If many family members do this, then they may sell or attempt to sell the same piece of land, claiming that they are empowered to do so in terms of the current traditional rules as they choose to interpret them (Barry - Ghana Interview 46-09).

Assuming the situation permits the relevant data to be collected - which we acknowledge may often be unrealistic in practice - we are arguing that the rules, lineage relationships and statuses that are assigned to different segments within a lineage may be stored in a LTIS. Social network analysis may be applied to examine changes or claims to change in status and claims to changes in the rules as social change occurs. Contested status may be revealed, and so outsiders may have a better understanding of how dealings in land are supposed to occur, which in turn should lead to more fair land administration.

For this initial experimental work, a data set was created based on a simplification 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 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 experimental work to date, the objective was to identify social networks such as family lineages in the customary system, the key actors in the family, 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).

In brief the following methodology was applied. Based on the Talking Titler Model (TTM), software for an ubiquitous (web and mobile) LTIS was developed, which interfaces with a PostgreSql database. Test data were simulated based on a real case study of a rapidly changing tenure system peri-urban to Accra, Ghana (Barry and Danso *in press*), augmented by data from the tutorial in the Talking Titler manual (Barry 2005) and stored in the LTIS. NetDriller software integrates data mining and machine learning techniques in model construction and analysis (Koochakzadeh *et al* 2011), and was used to extract relevant relationships to visualise social networks. In our experimental work we applied a clustering coefficient algorithm in the software to detect communities (families and other groupings). A social cluster such as an extended family is a small network (community) that is part of the larger network. The entire social network comprises a set of clusters (communities), which may or may not be linked (figure 2). As part of the process, the software detected key actors in each community (e.g. family head) and in the main network (e.g. chief) based on calculations of the betweenness centrality measure (Faust and Wasserman 1994).

The paper is structured as follows. Section 2 is a description of the major classes of the TTM used in the model. Following this, section 3 describes the research methodology, while section 4 discusses the experimental work. Section 5 analyses and interprets the (network) graphs, and the paper concludes in section 6.

2. TALKING TITLER MODEL (TTM) AND THE DATA SET

The Talking Titler Model is a tool for designing and prototyping LTIS's for complex, insecure, uncertain, and unstable tenure situations. It has four main interrelated classes namely: Person, Land or Property Object, Reference Instrument and Media Items. As is more fully explained in Barry *et al* (2013), the system is designed to evolve from one or more of these classes into a LTIS resembling forms that may resemble aspects of some currently LTIS's in use. It may also evolve into a form of LTIS that is completely new.

Figure 1 The Main Classes of the Talking Titler Model (TTM)

Our interest in this paper is primarily the different relationships in the Person class (i.e. social networks) and the Media class and how they relate to land objects. Media items, i.e. data items, can be stored in a number of structured, semi-structured and unstructured formats, such as land titles, survey plans, minutes of meetings, descriptive documents, audio recordings, and videos. Media items can be related to one another and to items in any of the other classes in the model (Barry *et al* 2013, 2007).

The Person class represents an individual, group or entity (e.g. a juridical person) with an interest in land. A range of relationships can be created between people (e.g. family ties, lineage, clan and sub-clan relationships) that may have a bearing on their relationships with a set of land objects (e.g. trees, land parcels, hunting areas, grazing areas and hunters' trap lines).

The social network focus in this article is on extracting relationships in the Person class which may apply to the land tenure system, which might then be used to update the LTIS or merely provide information on changes in the tenure system. We will also show that the system can be used to visualise the chain of subdivisions and consolidations of a land parcel over time.

3. METHODOLOGY

The network analysis is based on graph theory (Faust and Wasserman 1994) as per figures 2 – 7. Building on the discussion above, the case study data were stored in the LTIS based on the TTM. The data were than extracted from the database, formatted and loaded into NetDriller for analysis.

The data included people`s names (we have used node numbers instead of names), the land that they lived on, how they acquired the land (purchase, part of lineage, rented, inherited), the parties and agents (land surveyors, chief who signed the deed) to the transaction, and evidence of how tenure might be defended (e.g. deeds, survey plans). Family relationships were included, but the different statuses that an individual might have in the family (e.g. power to alienate land) were not included in this preliminary experiment. This is an area for further work.

NetDriller was set up to mine and analyse Person – Person, Person – Media, Person – Land, and Land – Media object relationships in the relational database and to model the social networks. Using graph theory terminology, the data set was treated as a multi-mode network. Each relationship type was considered a 1-mode (e.g. Person – Person) or 2-mode relationship (e.g. Person – Media). The bi-partite graph linking each object was created from these modal relationships (a two mode network as per figure 2). A set of clusters (communities or families) emerged as part of a broader network representing the customary group.

Inter-community relationships may take a number of forms. For example, in figure 2, one relationship linking communities is transactional; two family networks are linked by one family head borrowing money from another.

Portrayed in Table 1, nominal weights are assigned to relationship types in the data set. In the network, these weights are assigned to the edges (lines representing the relationships in figure 2) between nodes (representing individuals). Table 1 shows each weight value and the type of relationship it represents. As weights are nominal (i.e. categorical), their numerical values do not signify a level of importance.

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

Table 1. Types of Relationships and Weights Assigned to Them

4. EXPERIMENTAL WORK

4.1 Person- Person Relationship Graphs

The graph of the Person – Person relationships generated from the data is portrayed in figure 2. The different clusters represent different communities which in turn represents different families. Each node represents an individual person. Each edge represents a relationship between two persons. The weight of the edge is a code for the type of relationship. The weight on every edge is preceded by the cluster number that that edge belongs to. This means that, an edge with a weight of 5 (see table 1) belonging to cluster 1 will be labelled as 1.5. In the graph in figure 2 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).

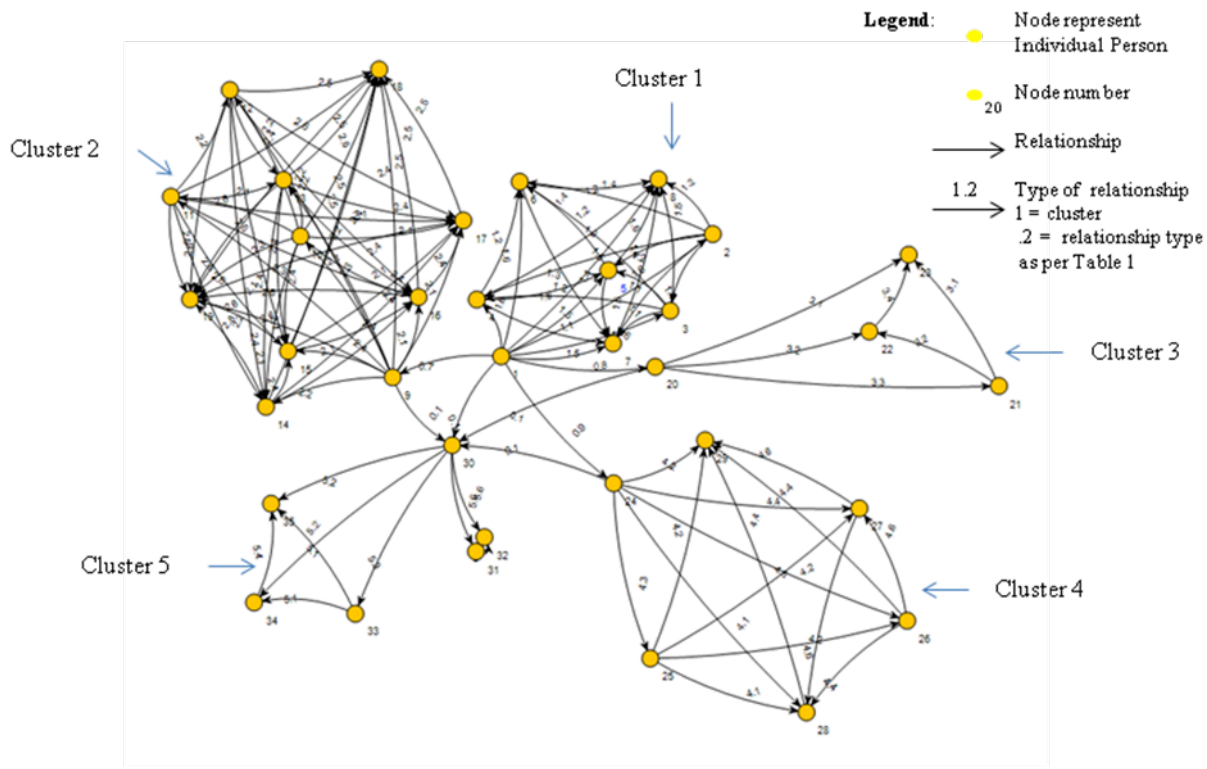

Figure 2. Social Network Graph of Person-Person relations

4.2 Parcel – Parcel Relationships

Network analysis of object-object relationships can show how land objects are created and 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 3, each node represents a parcel while the edges signify relationships between parcels. All the lots are subdivisions of a single parcel, represented by node 1 in figure 3. 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.

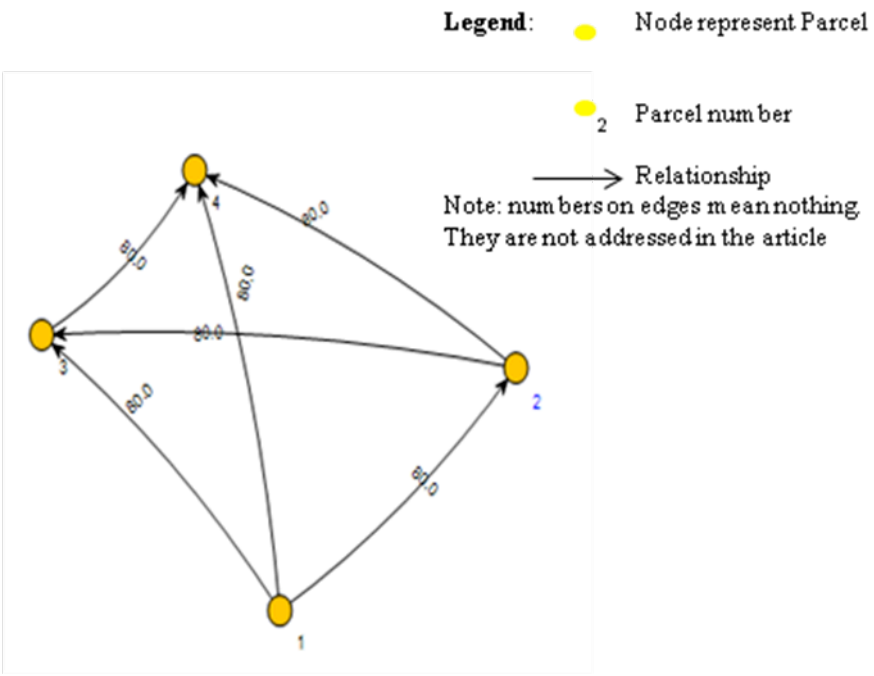

Figure 3 Graph of Lots (Land Object) Parcel-Parcel Network

4.3 Person – Land Object Relationships

Evidence supporting transactions and relationships between the 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 handing over is evidence that the transaction occurred), or they may be part of the chain of title. Figure 4 shows the results of mining the Media-Media relationship in the data set. The nodes represent specific media files while edges show a relationship exists between two files.

The graph in Figure 4 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 any transaction 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.

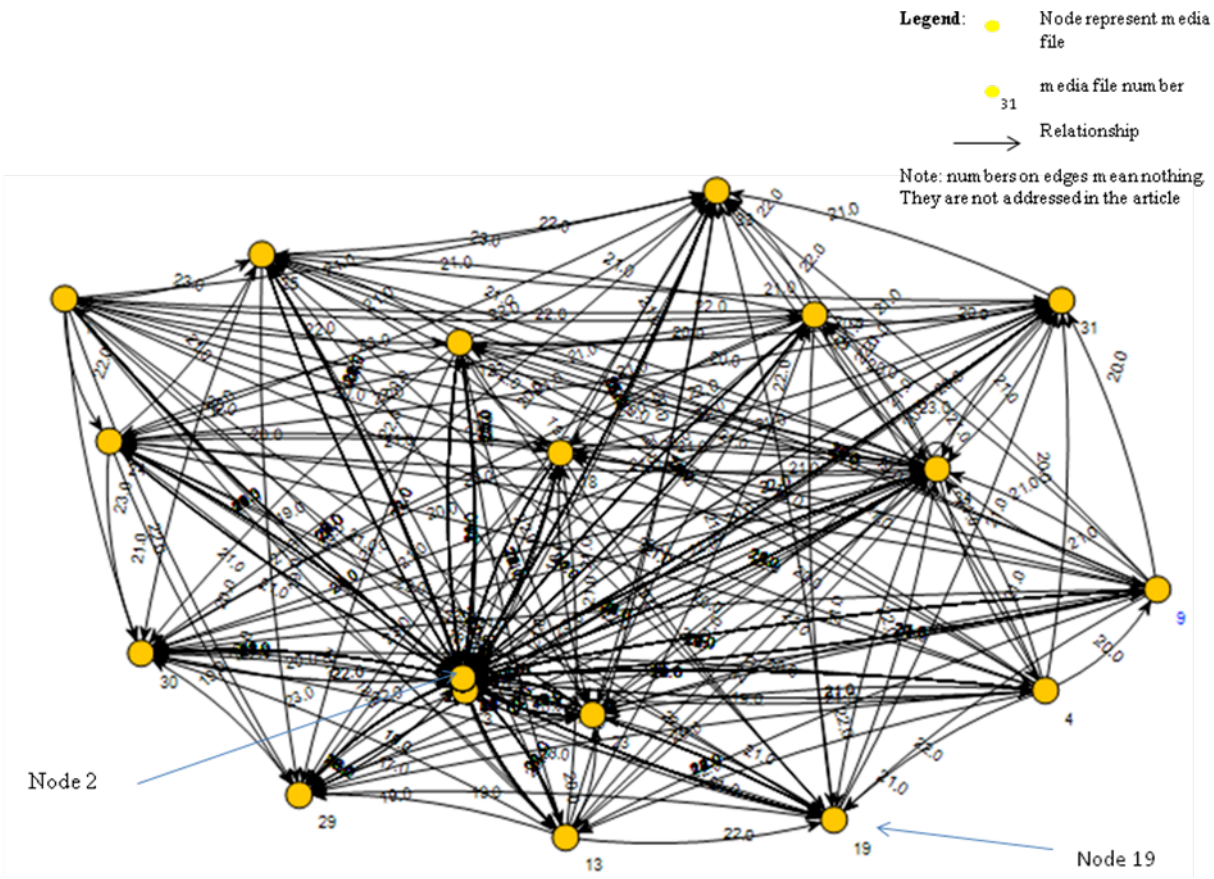

Figure 4. Media – Media relations derived by folding the Person – Media relations.

5. INTERPRETING THE GRAPHS

Data mining and network analysis identified family heads, lineage groups, and ties into different forms of evidence in the sample data. These are based on the Person – Person, Person – Media, Person – Land, and Land – Media object relationships stored in the database. We briefly describe some of the relationships that the data mining and network analysis revealed.

5.1 Family Heads

The graph in Figure 2 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 are linked to all the other nodes in the cluster. If you remove them, the links between clusters will not be shown.

The links between clusters show which people in a family have a relationship with other families. In our 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 LTIS), as there are no other links between clusters. Node 30 represents a family head who is also a land surveyor, and the links to node 30 show this. 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 2, nodes 1 and 9 are linked 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 Talking Titler database. 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 2 with a weight of 0.8 (Seller-Buyer) in Table 1). The edge 1-24 (weight 0.8) represents a similar transaction.

In the graph, 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.

5.2 Lineage Groups

As shown in Figure 2, 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).

5.3 Interests in Land

Folding the two-mode Person – Land relation into a one-mode Person-Person relation reveals interesting relationships. In Figure 6 some nodes becomes outliers in the resulting network. This means that, although these individuals relate to the land in some fashion, the data suggests that they have no claim to land rights to any of the land in question no matter their family or lineage group. They are family members of a stranger such as a share cropper or the land surveyor.

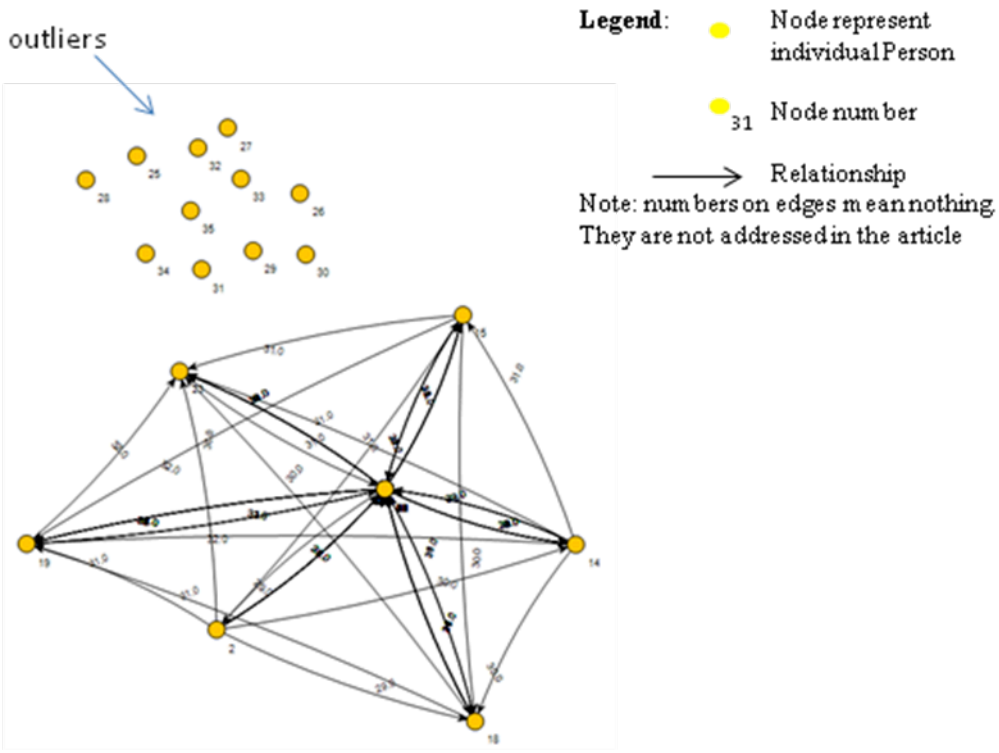

Figure 6. Person-Person relation based on the Person to Land relation.

5.4 Ties into Different Forms of Evidence

Examining ties between different items of evidence and how they relate to land and people may address the question; 'Who was involved in which transaction and what evidence is there to support that they were involved in a particular transaction?'

Figure 5 depicts the two-mode Person – Media relations folded into the one mode Media – Media relation network. This should 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 7.

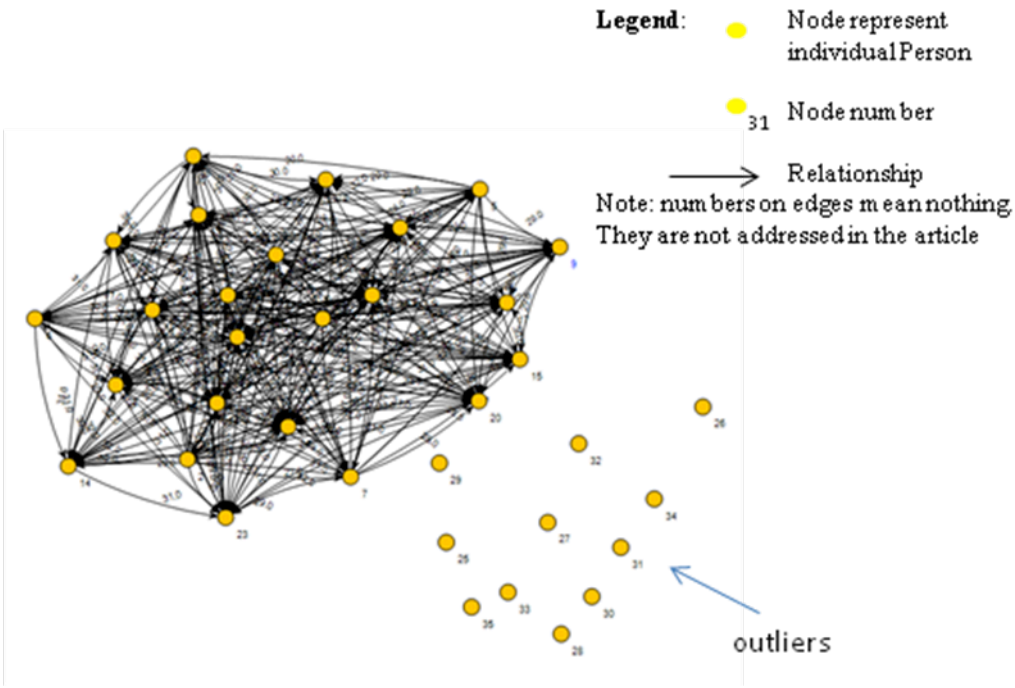

Figure 7 Person – Person relation based on the Person – Media relation

6. CONCLUSION

The Talking Titler model and the manner in which the data were organised for the experimental work is similar to a rudimentary 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. A distinction is that undocumented oral agreements and other forms of unstructured data such are also included in the database in this simple trial model. The experimental work, which relies on coding the data items and the relationships effectively using keywords and phrases, shows that data mining and network analysis may reveal family heads, lineage groups, who has right to claim said interests in land and ties into different forms of evidence. Network analysis accurately portrayed the different social and physical relationships in the tenure model.

An unexpected outcome of the exercise is the methodology may be useful as a simple tool to analyse a deeds registry system to isolate errors and inconsistences 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 early work indicates that the methodology has potential in analysing complex tenure relationships, providing of course that the data can be obtained. Future experimental work will 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. An attempt will be made to reveal changing rules, changing status and conflicts over power as a tenure system evolves.

Overall, this early 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. It may also be useful in applying the Talking Titler methodology to evolutionary land records development, as per Barry *et al* (2013) where the final form of the information system may be completely different to systems currently in use. Social network analysis could well uncover relationship patterns that are very different to the ones currently modeled in land records systems.

REFERENCES

Barry M B 1999. *Evaluating Cadastral Systems in Periods of Uncertainty:* A Study of Cape Town's Xhosa-speaking Communities.

Barry M and Rüther H 2001. Data Collection and Management for Informal Settlement Upgrades. Proceedings of the International Conference on Spatial Information for Sustainable Development, Nairobi, Nairobi, FIG, TS13.2.

Barry M, **Roux L**, **Barodien G** and Bishop I 2002. Video-evidencing and Palmtop Computer Technology to support formalising Land Rights. *Development Southern Africa* 19(2), 261-272.

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

Barry M 2005. Talking Titler Manual. http://ucalgary.ca/mikebarry/talkingtitler/downloads (Accessed 20th Feb,2013).

Barry M 2006. Formalising Informal Land Rights: the case of Marconi Beam and Joe Slovo Park. *Habitat International*, 30(2006), 628–644.

Barry M, Hunter A, **Muhsen A** 2007. *Scalable Land Tenure Record Systems*. , FIG Com3, Proceedings of Workshop. Informal Settlements - Real Estate Markets Needs related to Good Land Administration & Planning, 28-31 March 2007, Athens.

Barry M, Roux L, and Danso E 2012. Land Registration Usage Theory: A Case Study in Ghana. World Bank Conference on Land and Poverty, Washington, D.C. 23 – 26 April 2012.

Barry M, Muhsen A-R and Molero R 2013. Alternative Land Tenure Information Systems: The Talking Titler System – an Evolutionary Design based Methodology. FIG Article of the Month February 2013. http://www.fig.net/pub/monthly_articles/february_2013/feb_2013.html

Barry M and Danso E. *in press.* Tenure Security, Land Registration and Customary Tenure in Peri-urban Accra: A Case Study. *Land Use Policy*.

Danso N E 2012. *Peri-Urban Land Tenure in Ghana (Accra): Case Study of Bortianor.* MSc Thesis, University of Calgary UCGE Reports 20373. http://ucalgary.ca/mikebarry/Publications

Fourie C D 1993. *A New Apprach to the Zulu Land Tenure System:* A Historical Anthropological Explanation of the Development of an Informal Settlement.

Koochakzadeh N, Sarraf A, Kianmehr K, Rokne J, Aljaj R 2011. NetDriller: A powerful Social Network Analysis Tool, 11th IEEE International Conference on Data Mining Workshop (1235 – 1238).

Roux L and Barry M 2001. Using Video Imagery in Land Tenure Information Systems: A Study of a Communal Property Association in the village of Algeria, South Africa . *Geomatica* 55(1), 47-55.

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

Walker C 2008. Land Marked, Land Claims and Restitution in South Africa. Jacana Media, Johannesburg.

BIOGRAPHICAL NOTES

Kwame Asiedu recently completed a MSc in the Department of Geomatics Engineering in the Land Tenure and Cadastral Systems Research Group at the University of Calgary. He obtained a BSc degree in Computer Science and Psychology from the University of Ghana, Legon in 2007.He is a Microsoft Certified Professional, holds a diploma as a Computer Hardware Engineer and an MIT-AITI certificate in Java programming and entrepreneurship. Kwame has worked as a technical support specialist, information systems auditor, database administrator and a software architect with experience in web and mobile application development. His research is focussed on ensuring tenure security in uncertain and complex situations through the application of Social Network Analysis and Data Mining techniques to Ubiquitous (web and mobile) land information systems.

Professor Mike Barry holds the Chair in Land Tenure and Cadastral in the Geomatics Engineering Department at the University of Calgary, where has been since the end of 2002. Prior to this, he was at the Department of Geomatics at the University of Cape Town His research interests are in land tenure, cadastral systems, and land registration effectiveness. Mike has worked in the field as a land surveyor, consulted, or done field research, in Canada, Botswana, Ghana, Iraq, Indonesia, Lesotho, Malawi, The Netherlands, Nigeria, The Philippines, Somaliland, South Africa, Swaziland, Zambia and Zimbabwe. Mike has BSc(Survey) and MBA degrees from the University of Cape Town and a PhD from the University of KwaZulu-Natal.

Contacts:

Mr. Kwame Asiedu
Department of Geomatics Engineering
University of Calgary
2500 University Dr.N.W.
Calgary, Alberta T2N 1N4
Canada
Email: kasiedu@ucalgary.ca
Tel +1 (403) 903-8718

Dr Michael Barry
Chair in Land Tenure and Cadastral Systems
Department of Geomatics Engineering
University of Calgary
Website: http://www.ucalgary.ca/mikebarry/
Tel +1 (403) 220-5834
Fax +1 (403) 284-1980