ON THE SOCIAL NETWORK PERSPECTIVE OF ROAD-FREIGHT FACILITIES

J.W. JOUBERT

Industrial and Systems Engineering, University of Pretoria, South Africa, 0002
CSIR Built Environment, PO Box 395, Pretoria, South Africa, 0001

ABSTRACT

Transport is the price we pay for establishing spatially dispersed firms. Logistics and supply chain literature abound with contributions describing, often qualitatively, the chain structure between firms. Usually only a single focal company is at the centre of the network, and ties between the nodes (companies) are based on contractual agreements and formal relationships. This paper takes a slightly different network perspective: we consider the actual number of vehicle movements between facilities, and infer a weighted social network among the facilities visited by commercial vehicles. Since commercial vehicles travel continuously between facilities, and in doing so contribute to congestion, the objective is to gain an understanding of where key facilities in the vehicle chains are.

The paper provides an overview of social network analysis (SNA) and reviews the state-of-the-art in applying SNA to a supply chain context. The methodology to identify facilities from general freight movement is illustrated, and we present a spatial analysis of key network players in Gauteng, chosen for its economic significance and unique omnidirectional through traffic. Contrary to other research contributions and general expectation, the main freight facilities are not situated on the periphery of the urban areas, hence competing for the limited available road space.

1 INTRODUCTION

While supply chain researchers and practitioners are dealing with the challenge of “how can we as a firm better compete?” transport planners are trying to answer an aggregate question: “how can we provide better supporting infrastructure so that firms and individuals can participate in the economy, and compete more efficiently?” The latter decision is considered amidst the uncertainty caused by the various competing objectives of the firms and other road users. The objective in this paper is to link the domains of supply chain with that of transport planning using social network analysis. Although many freight models deal with commodity movement, we argue in this paper that it is not the commodities that contribute to congestion, but rather the commercial vehicles carrying those commodities. Hence, understanding where the vehicles move and conduct its business may yield valuable insight in order to improve the modelling of freight in our transport models.

This paper makes three contributions. The first contribution is to introduce a number of key concepts from social network analysis (SNA) and to relate these to supply chain management (SCM). Secondly, to propose a novel methodology to extract a social network from actual vehicles movements. This is achieved by inferring the locations where commercial vehicles perform activities, and then using the actual vehicle trip between two facilities as a proxy for a social interaction. The third contribution is a novel analysis of a social network generated from the Global Positioning System (GPS) data from more than 30 000 vehicles in Gauteng.

In Section 2 a number of related work are reviewed, and the context of this paper is established. The methodology used to extract the social network is presented in Section 3, while the resulting network is depicted and discussed in Section 4. The paper is concluded in Section 5 with a discussion and research agenda for the application of SNA to transport planning, specifically related to commercial and freight vehicles.

2 RELATED WORK

Hensher and Figliozzi (2007) acknowledge that freight transportation companies, and their associated activities, have developed and evolved a lot more rapidly than what we as modelling and research communities have been able to keep up in terms of the modelling of freight transportation. This is even more pronounced for commercial vehicles in general, i.e. when service vehicles are included.

The field of supply chain management (SCM) is built on the metaphor that companies form relational chains: upstream with suppliers, the suppliers' suppliers, all the way to the raw material source; and downstream with customers, customers' customers, all the way to the final consumer. These entities are often referred to as supply chain partners, and the intensity of the relationship depends on the impact that the two partners have on one another. The number of tiers up- and downstream depends on the extent of the case study and objectives concerned. In each study there is usually a single focal company around which the supply chain is configured. The objective is then to design, or redesign the supply chain so as to maximise the benefit for the focal company. Operations among supply chain partners are integrated; information sharing is considered; supplier and customer development in terms of financial support and technological innovation is contemplated; all in an effort to ensure that the flow of goods through the chain is optimised.

But, the SCM perspective is usually very biased in favour of the focal company. Should the supply chain be mapped, or designed, for another company in the chain (upstream or downstream from the focal company), the result might be quite different.

Still, companies invest millions into the development of its supply chains and Autry and Griffis (2008) introduced supply chain capital as a metric to assist companies to evaluate the benefit (or loss) of investing in its supply chain.

Although the relationships between firms might be contractual, we argue that they could be considered social since they are most often voluntary. The idea of a mere linear chain of firms linked with one another has been challenged for many years. Lazzarini et al. (2001) already integrated different supply chains into supply networks. In applying a more network-oriented approach, the door is opened to the large array of network theory tools to map and analyse the relationship between firms.

The premise in this paper is, whether we agree with the way in which SCM is applied, or its perceived benefit, that companies do connect with one another in the form of relationships: be it informal connections between individuals of two companies, the one supplying material to the other; or formal service level agreements. We argue that however formal/informal the relationship between two firms, there is more often than not a vehicle moving between two facilities associated with each of the firms. The vehicle might be performing a (regular) service, delivering goods to the customer firm, or collecting from the supplier firm.

The field of social network is well developed and many research publications are dedicated to the field. The idea that entities with social relations meet up with one another by transporting themselves to a jointly agreed upon venue is not new at all. Studies such as those completed by Hackney and Marchal (2009) and Kowald et al. (2009) are analysing the travel patterns of commuters to participate in social activities.

A number of authors have acknowledged the importance of behavioural insights gained from studying commercial firms, but it is only recently that Liedtke (2009) developed the first (to our knowledge) model that predicts the truck movements as a result of inter-organisational relationships.

Borgatti and Li (2009) comprehensively discuss the scope of social network analysis in a supply chain context, distinguishing between continuous and discrete relationships at the highest level. Figure 1 shows our interpretation of their typology of the types of ties that may exists between firms, and highlights the tie considered in this paper. Much of network theory contributions are based on continuous relationships that are often said to be pre-social: relationships that may exist without any real social interaction among the entities.

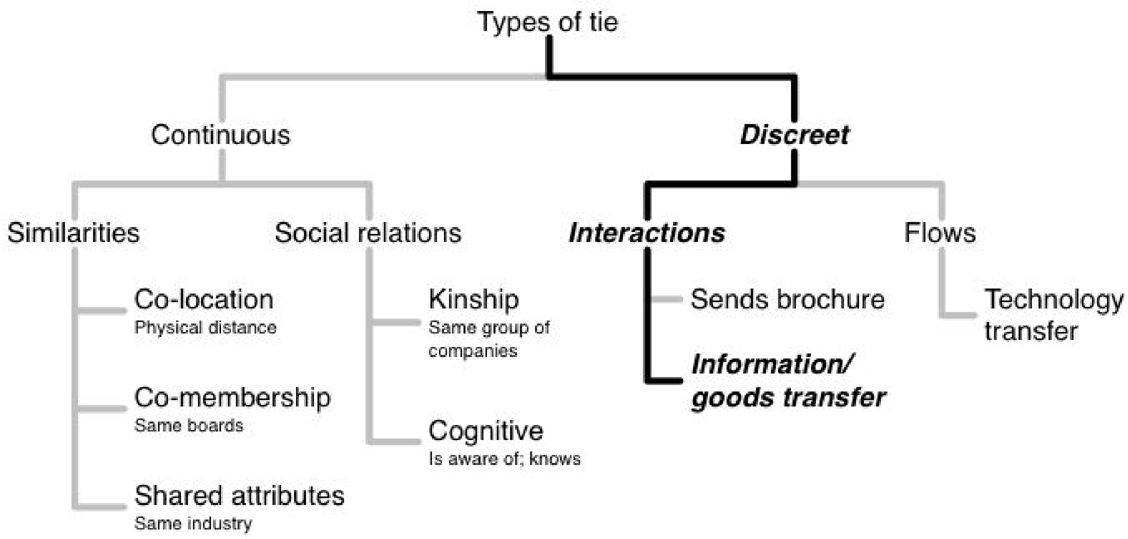


Figure 1 - Typology of the types of ties among firms (Adapted from Borgatti and Li (2009)). The highlighted path indicates the choice of tie pursued in this paper.

In this paper we agree with Borgatti and Li (2009) that when measuring discrete interactions, we assume some kind of social relation between the entities. In fact, we not only agree, in the next section we actually infer the social network from discrete interactions between firms. We also acknowledge that multiple types of relationships may exist simultaneously between firms, and between different individuals of the firms—a phenomenon referred to as multiplexity. For example, two companies may both be subsidiaries of the same parent company (continuous social relation through kinship); serve the same automotive industry (continuous similarity through shared attributes); and ship components and parts from the one to the other for assembly (discreet interaction through goods transfer). The basic unit of data in network analysis is the dyad, the direct connection between two entities. In the next section we define the measured dyad, and present the method for extracting the network resulting from the identified dyads.

3 METHODOLOGY

In this paper the interest is in systematically considering the relationships amongst all the firms in an area. More specifically, since each firm can have multiple facilities (buildings) in the study area such as plants, warehouses, and head offices, we are interested in the relationship amongst all the facilities in the study area.

Our choice for a dyad (the basic unit of analysis in a network) is the direct movement of a commercial vehicle between two facilities. The premise is that if a commercial vehicle travels from one facility to another, there ought to be a purpose for the direct trip, and such a purpose is founded on a social relationship between the two facilities. We acknowledge at this point that a vehicle may incidentally visit two facilities consecutively without a real purpose, and hence without a real social relationship existing between the two facilities. One such example is a small truck rented by a company to collect a parcel from supplier A with spare parts destined for plant C. After collecting the parcel, the driver may have received an instruction to also collect the marketing material from the printers at B for the company’s marketing division that are located on the same premises as the plant C. The vehicle’s activity chain is then A → B → C. Although there is no real relationship between the spare parts supplier A and the printers at B, a social interaction is inferred.

We argue in this paper, and this will be further substantiated later in Section 4, that although such incorrect relationships are inferred, they play a very small role in the bigger picture of identifying key players.

Also, each relationship between two facilities is strengthened with every vehicle that moves directly between the two facilities. Incidental trips may thus create relationships where they ought not, but such incidental relationships will remain weak.

The activity locations of commercial vehicles were inferred in Joubert and Axhausen (2010a) from the raw Global Positioning System (GPS) logs of more than 30 000 vehicles over a period of 6 months. An activity started when the vehicle signalled an engine-off trigger, and the activity ended when the vehicle started again, sending an engine-on signal. In their paper, the authors discussed in how activity extraction is influenced by parameter sensitivities, and the interested reader is referred to their work for a detailed discussion.

An activity chain is then made up of all consecutive minor activities, i.e. activities that have a duration of 300 minutes or less. Activities exceeding 300 minutes are referred to as major activities, and make up the end points of activity chains. The more than 30 000 vehicles considered by Joubert and Axhausen (2010a) resulted in activity chains with more than 10-million activities over the study period.

The identification of activity locations are dependent on the accuracy of the GPS instrument, and activity locations were first clustered using a density-based clustering algorithm. Joubert and Axhausen (2010b) explain the algorithm, and propose a thorough calibration and validation procedure to identify the correct facility locations. Their results suggest that if at least 15 activities occur in a 30m radius, the centroid of the activities is a good approximation of the location of the facility where the activities take place.

Once facilities are identified, we consider the activity chains of each vehicle, and more specifically the links between each pair of activities. An example is shown in Figure 2.

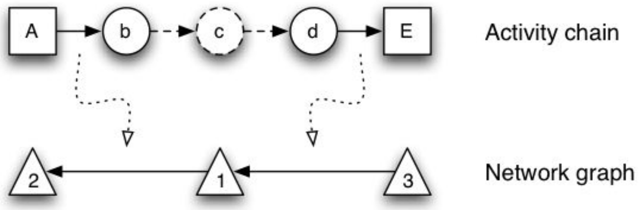


Figure 2 - An example of extracting the social network dyads from an activity chains.

Consider the activity chain $A \rightarrow b \rightarrow c \rightarrow d \rightarrow E$ where upper case letters indicate a major activity (squares in Figure 2), and lower case letters a minor activity (circles in Figure 2). During the clustering exercise some activities will be associated with facilities, and others not. Assume for the purpose of this example that both *A* and *E* are associated with facility 1; activity *b* with facility 2, and activity *d* with facility 3. Unassociated facilities merely suggest that the activity density was not sufficient to be of interest and considered a facility within the current study. Activity *c* was not associated with any facility since it was conducted in an area where very few activities occurred.

To extract the social network among the three facilities, we consider each pair of consecutive activities in the activity chain. The vehicle travelled from facility 1 to facility 2 (activity *A* followed by activity *b*). We infer a social relationship with a specific direction from facility 1 to 2.

The next activity pair, $b \rightarrow c$, is between a recognised facility and no facility, and we hence do not establish any social relationship between facilities. The same hold for the activity pair $c \rightarrow d$.

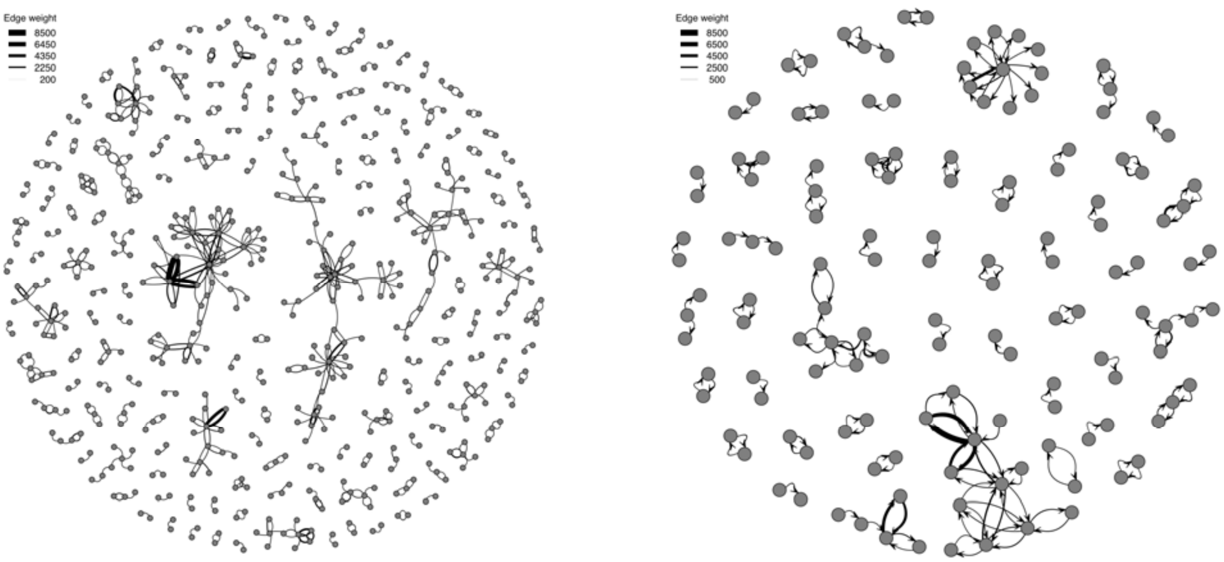
Since both activities d and E are associated with facilities, we establish a social link between facility 3 (activity *d*) and facility 1 (activity *E*) in the direction $3 \rightarrow 1$. Whenever a directed activity pair exists between two facilities where a social tie already exists, the weight of the social tie is increased with one unit.

Continuing in this manner for all vehicle chains of all vehicles thus allows for the creation of a weighted directed network on which social network analysis can be performed. In the next section we report on the social network extracted for Gauteng, South Africa.

4 SOCIAL NETWORK IN GAUTENG

The extremely large data set of commercial vehicle activities resulted in more than 43 000 identified facilities. Amongst these facilities existed more than 1.3-million social ties with weights between 1 and 8 468 over the six-month period. A tie with a weight of 8 468 relates to an average of more than 54 direct trips per day between the two facilities in a specific direction. To put the distribution of weights into perspective, the 99th, 99.5th and 99.99th weight percentiles were 29, 47 and 533, respectively.

It should be clear that visualising such extensive networks is no mean feat, and unless performed with a specific goal in mind, will yield very little insight. As two examples, we visualise two social graphs extracted from the complete network in Figure 3, courtesy of R Development Core Team (2005) and Butts (2010, 2008).



(a) Minimum weight of 200. (b) Minimum weight of 500.

Figure 3 - Social graphs extracted from the complete network of facilities.

Each dot in the graph represents a facility, while the thickness of the line connecting two facilities is an indication of the weight of the social relationship between the two facilities, and in a specific direction. Although the direction is known, we've omitted the use of arrows in Figure 3b since it results in cluttered images. The location of the nodes in the social graph has no relation to the geographic location of the facility. Although possible, it would be quite difficult to distinguish between the different components if we plotted all nodes at their correct location.

Each social graph is made up out of small groups of connected facilities. Each small group within the graph is called a component.

To extract the social graphs, we eliminated all social relationships with a weight smaller than a given threshold (200 in Figure 3a and 500 in Figure 3b). This allowed us to only consider relationships that resulted from frequent vehicle movements between facilities, for example. When low-weight ties were removed, many facilities remained unconnected to any other facilities or components. Such unconnected facilities are referred to as an isolate. For purposes of the visualisation, we then removed all isolates from the graph, and only components of two or more facilities are shown.

Both Figures 3a and 3b contains weak components. That is, whenever two facilities are connected in either direction, they are assumed to be part of a component. There need not be a relationship in both directions before two facilities are considered to be part of a component. Also, if facility *A* is connected to facility *B* in the direction $A \rightarrow B$, and *C* is also connected to facility *B* and in the direction $C \rightarrow B$, then all three facilities are part of the same component, even though no direct connection between facilities *A* and *C* in either direction exist.

4.1 Social network analysis

Once the network is established, one can perform a multitude of social network analyses (SNA), for example those suggested by Borgatti and Li (2009). However, mere analysis simply because it is possible is of little use, and most often yields very little insight. The objective of this paper is not to perform an array of analyses, but to provide a platform and methodology to show how one can, firstly, generate a social network amongst firms and, secondly, that generating such networks are indeed feasible on a large scale, even at provincial level as this paper has shown.

That said, we argue here that some examples of possible analyses are useful in setting a research agenda. A key concept in social network analysis is that of centrality. Centrality is a measure of the importance of a specific node, in our case facility, in the network. In SNA the centrality is based on the structural position of the facility, and not necessarily the geographic position.

However, to identify important facilities may mean different things to different people, and hence various centrality concepts have emerged, and are continuously emerging. All centrality concepts have its objective to assist in identifying important players. The premise is that if, for example, one wants to disseminate vital information throughout the network, the information will spread throughout the network the quickest (and widest) if the information is initially passed to the most central (important) player in the network.
Three centrality concepts are presented in this paper:

Degree centrality use the (weighted) number of social relationships that a facility has as a measure of its importance. If, for example, facility A has social relationships with five other facilities, it is considered more important and more centrally located in the graph than a facility B that only has a single relationship with one other facility. The relationships can also be weighted in the calculation of degree centrality. Consider then, in the example, each of facility A's five relationships to have a weight of one, then it would have a weighted degree centrality of 5. If facility B's single relationship, however, would have a weight of 10, then facility B will have a weighted degree centrality score of 10 and be considered more central and important in the graph than facility A. Degree centrality thus suggests that the level of connectedness is a good indication of a node's importance. A facility with a degree centrality score is well-connected in the network and may be an important hub where many, seemingly non-related facilities, are actually connected. If one looks for opportunities for load consolidation, such facilities scoring high on degree centrality ought to be a good start. Facilities with more relationships are argued to have access to large amounts of information, making them important nodes to gain information from, and also to share information with.

Eigenvector centrality is another measure of importance. It suggests that those nodes that are connected to well-connected nodes (those having high degree centrality scores) are more important than those that are connected to the same number, but lesser well-connected nodes. Again, the weighted relationships can be taken into account in the calculation. In directed graphs, Eigenvalue centrality is often replaced by an alpha centrality expression that is slightly different than that of the Eigenvalue. If facilities with high degree centrality scores are said to have access to large amounts of potentially valuable information, nodes with high eigenvector centrality, i.e. connected to many well-connected nodes, can be argued to have access to even more information via their links with the well-connected facilities.

Betweenness centrality for a node indicates on how many shortest paths of other facility pairs the node lies. A high betweenness centrality score suggests that the facility has exclusive/high control over the flow throughout the network. Structurally such facilities are important: if they slow down their operation, or are removed from the network, the throughput effect in the network will be more adversely affected than when they had lower betweenness scores.

In Figure 4 the results of the three centrality scores calculated for the Gauteng Network are presented, indicating both the scaled score, and the geographic location of the facilities.

(a) Degree centrality. (b) Eigenvector centrality. (c) Betweenness centrality.

Figure 4 - Centrality scores for facilities within the Gauteng envelope.

The scaling of the semi-transparent circles reflect the relative scoring, with the largest circles indicating the highest scoring (most prominent) facilities. Once key players are identified, decision makers can target them more effectively with policy interventions, for example. Most notable from the three scores is the location of key (important) players. One might, according to international literature such as Hesse and Rodrigue (2004), expect key freight players to be located on the periphery of the metropolitan areas. Yet, the majority of key facilities, with the exception of the high-scoring degree centrality facilities outside Gauteng, are located very centrally in Gauteng, and not on the periphery.

5 CONCLUSION

In this paper we contributed to the young, yet growing body of knowledge that attempts to improve freight modelling by understanding the social and behavioural dimensions of relationships amongst firms.

We have shown how a social network can be extracted and inferred from the actual movement of commercial vehicles between facilities, and we've used the direct trip of as vehicle from one facility to another as a good proxy for a social relationship between the two facilities. The paper is a valuable contribution in terms of the methodology proposed for establishing the social network, and is also novel in applying the methodology to a large data set for Gauteng, depicting that the methodology is both implementable and insightful.

Two main items are suggested for further research. The first is to establish and analyse the land use of the various facilities that are members of high degree components. That will allow researchers and transport planners to understand what the extent of vehicle activities are that are generated by different land use types. Analysis based on observed activities only allow for improved understanding of the status quo. If we understand the land use, and on the premise that land use drives logistics, then we can argue that as land use changes, we would be able to better predict the resulting logistics activities. With such prediction capability we would be able to model future scenarios of commercial vehicle movement more accurately and realistically.

The second agenda point would be to deepen the social network analyses itself. An ever growing body of knowledge is developing within the sociology fields, and also in private vehicle and commuter transport in relation to social networks. Understanding the developments in other areas will indicate opportunities to extend those areas to the commercial vehicle movement and supply chain analysis.

6 ACKNOWLEDGEMENTS

The author is grateful to the South African National Research Foundation (NRF) who was partly funding the research under Grant FA-2007051100019.

7 REFERENCES

Autry, CW and Griffis, SE, 2008. Supply chain capital: The impact of structural and relational linkages on firm execution and innovation. Journal of Business Logistics, 29(1), p.157–173.

Borgatti, SR and Li, X., 2009. On social network analysis in a supply chain context. Journal of Supply Chain Management, 45(2), p.5–22.

Butts, CT, 2008. Network: a package for managing relational data in R. Journal of Statistical Software, 24(2). Available online from http://www.jstatsoft.org/v24/i02/.

Butts, CT, 2010. SNA: Tools for Social Network Analysis. R package version 2.0.

Hackney, JK and Marchal, F, 2009. A model for coupling multi-agent social interactions and traffic simulation. In 88th Annual Meeting of the Transport Research Board.

Hensher, D and Figliozzi, MA, 2007. Behavioural insights into the modelling of freight transportation and distribution systems. Transportation Research Part B: Methodological, 41(9), p.921–923.

Hesse, M and Rodrigue, J, 2004. The transport geography of logistics and freight distribution. Journal of Transport Geography, 12(3), p.171–184.

Joubert, JW and Axhausen, KW, 2010a. Inferring commercial vehicle activities in Gauteng, South Africa. Forthcoming in Journal of Transport Geography.

Joubert, JW and Axhausen, KW, 2010b. On the social networking of commercial vehicles. Arbeitsberichte Verkehrs- und Raumplanung, 601, IVT, ETH Zurich, Zurich. Available online from http://www.ivt.ethz.ch/vpl/publications/reports.

Kowald, M, Frei, A, Hackney, JK, Illenberger, J, and Axhausen, KW, 2009. Collecting data on leisure travel: The link between leisure acquaintances and social interactions. In Applications of Social Network Analysis.

Lazzarini, S, Chaddad, F, and Cook, M, 2001. Integrating supply chain and network analyses: The study of netchains. Journal on Chain and Network Science, 1(1), p.7–22.

Liedtke, G, 2009. Principles of micro-behavior commodity transport modeling. Transportation Research Part E: Logistics and Transportation Review, 45(5), p.795–809.

R Development Core Team, 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.