

A Network Formation Model for Social Object Networks

Somayeh Koohborfardhaghighi and Jörn Altmann

Abstract Social networks can be differentiated according to the type of entities that are represented within them. Entities in human networks can act strategically to maximize their own payoffs during interactions with other humans. However, entities in social object network (e.g., SaaS service network) are not able to perceive the environment and act strategically upon that at any time. This paper contends that existing network formation models lack sufficient attention to social object networks. Therefore, we propose a new network formation model, through which we are able to explain how a SaaS service network emerges during the service composition procedure by service developers. The new network formulation model not only considers the usage frequency and reputation but also the similarity of the functionalities of the main SaaS services. It also explains how social objects (e.g., SaaS services) benefit from establishments of links between each other in the network.

Keywords Software-as-a-service network · Network formation model · Social object networks

1 Introduction

The main software services (i.e., Software-as-a-Service (SaaS) services), which are typically owned by big enterprises, can be used by others through application programming interfaces (APIs). An API provides a way for applications to work with each other and deliver the information or functionalities users or other services need.

In spite of the variety of SaaS services in the market, service developers are moving fast in combining the existing services (i.e., in building and delivering new

S. Koohborfardhaghighi (✉) · J. Altmann
Technology Management, Economics, and Policy Program, Dept. of Industrial Engineering,
College of Engineering, Seoul National University, Seoul 151-744, South Korea
e-mail: skhaghighi@snu.ac.kr

J. Altmann
e-mail: jorn.altmann@acm.org

ones) through APIs. This process is called service composition, which is one of the most important features of service-oriented architecture.

The existing SaaS services belong to different service type categories (e.g., advertising, finance, Internet, social networking, music, games). Due to the different service types and the diversity of different types of APIs that go into the market, the process of service composition becomes more and more challenging. In order to deliver a new business solution to the market, we can augment functionalities of existing services. For example, composing a new social networking service may require utilization of the services delivered through the APIs of Facebook, Twitter, or Flickr. Therefore, we can state that the main APIs of big enterprises are the foundations of newly generated services in the market.

Basically, the SaaS service network emerges from strategic integration of basic SaaS services by service developers. The process of web service composition creates a network, in which social activities (i.e., the use of APIs) are visualized. The emerging network can be seen as a graph that highlights similar or different types of SaaS services through its links. One thing to keep in mind is that the emerging network has a meaningful structure, through which interactions of objects lead to certain outcomes for them.

In literature, we can find strategic network formation models, in which payoffs of individuals in a network are the result of their strategic interactions with others (Jackson and Wolinsky 1996). However, existing models are mainly applicable to human networks. This paper contends that they lack sufficient attention to social networks of objects (e.g., SaaS service networks). The lack of responsiveness in objects limits entities to interact strategically with one another quickly. However, the human role is still the key for proper understanding of how a network of objects emerges (e.g., through the combination of functionalities of services by service developers). Other network formation models have also been proposed to explain the characteristics of networks (Barabasi and Albert 1999). Barabási–Albert model is the best-known example in this category. However, the method of growth in scale-free network is only based on the relative frequencies of existing nodes' degrees but not on the benefit and the cost of a link establishment.

Taking each of the discussed key elements about service composition procedure into account, this article aims to address the following research questions:

1. What does the structural analysis of the SaaS service industry bring?
2. Can current network formation models produce the topology of SaaS service networks?
3. What kind of network formation model for social object networks is needed?

With respect to our research questions, in this paper, we propose a new network formation model for social object networks that explains how a SaaS service network emerges during the service composition procedure. The new formulation not only considers the use frequency and reputation but also the similarity of the functionalities of the SaaS services. It also explains how social objects (e.g., SaaS services) benefit from establishing links among each other in the network.

The remainder of this paper is organized as follows: In Sect. 2, we discuss related works and theoretical background on strategic network formation models. In Sect. 3, we define a social object network and compare it with social networks of

humans. We also present the proposed new social object network formation model. Finally, we conclude the paper with a discussion in Sect. 4.

2 Theoretical Background

There are two branches of literature that are related to network formation models. The first branch focuses on strategic network formation models, while the second one comprises the growth models of networks.

In the first category of literature, researchers look at a network as a network structure emerging from strategic interactions of the individuals that are located in it. Strategic network formation models assume that network entities are rational and that opportunity seeking actors establish links with other actors if they find the link establishment beneficial. Therefore, each entity in the network obtains a payoff from its interaction with others. Such values are derived from a payoff function, which can be defined in different ways. For example, it can be defined in terms of the number of connections to other actors or the distance to other nodes (i.e., closeness centrality Buechel 2008). A comprehensive introduction to social and economic networks is presented by Jackson (Jackson 2008). He presents a few strategic network formation models, in which choices of entities have certain impact on the topological features of networks. Two examples are the symmetric connection model and the co-author model (Jackson and Wolinsky 1996).

The symmetric connection model is a model, in which entities communicate with their direct contacts, and the value obtained through their interactions is proportional to their distance to the target entities. As this distance increases, the obtained payoff value decreases. The co-author model focuses on the collaboration among individuals, and the payoff is due to the collaboration with direct contacts. There is no payoff for communicating with indirect contacts. The structural changes in the proposed models are the consequence of the composition of the strategies of rational entities (i.e., here humans). Therefore, they are applicable to social networks of humans (Sect. 3.4). Variations of symmetric connection models are also presented in literature. They are degree-distance-based versions and have positive and negative externalities (Möhlmeier et al. 2013; Morill 2011). Jackson and Rogers assumed geographic costs for forming links, and entities are grouped into so called islands (Jackson and Rogers 2005). The connection costs are low within an island and high across islands. Further variations of connection models with geographic costs are investigated by Johnson and Gilles and by Carayol and Roux (Johnson and Gilles 2000; Carayol and Roux 2005, 2009). Other studies that identify the effects of actors' interaction on social networks can be grouped into network games (Galeotti et al. 2010), public good provision, and bargaining in networks.

From the view point of network growth models, which fall into second category of research, proposing a model to produce predetermined structural properties is a challenging issue (Koohborfardhaghighi and Altmann 2014). Nonetheless, different network formation models have been proposed to produce desired structures. The preferential attachment growth model, introduced by Barabási–Albert (1999),

is capable of generating power law degree distributions. In this model, new nodes are more willing to link to high-degree nodes during the growth of a network. The method of preferential attachment growth is based on the relative frequencies of existing nodes' degrees, but not on benefits and costs of link establishments. Therefore, it is not suitable for object-based social networks. Besides, random network growth models, in which new nodes can randomly choose the attachment point with the same probability for all nodes (Erdős and Rényi 1959, 1960, 1961), are also not suitable for object-based social networks. As mentioned in Sect. 1, the SaaS service composition is not the result of random selections of services.

Although a lot of research has been performed on strategic network formation models, they are mostly applicable to networks of humans. The lack of flexibility and spontaneity in networks of social objects limits strategic interactions. In short, social object networks are social network, where the interconnectivities among the objects have been taken from strategic manipulations of humans. For example, if we think about a SaaS service network, the only way of depicting how it emerges is to identify specific patterns in the data related to new SaaS service compositions by developers. Therefore, for a proper understanding of the formation of social object networks, the human role is still the key. The fact to remember is that human beings do not connect the objects in a random way. They use and combine them to get value out of them. Consequently, we can state that literature cannot explain social object networks, making a new network formation model for social object networks necessary.

3 Proposed Social Object Network Formation Model

3.1 Definition of Social Object Networks

Social networks can be differentiated according to the entities (i.e., humans or objects), which are located in them. Two types can be distinguished: social networks of humans and social networks of objects. Entities in human networks can act strategically to maximize their own payoffs through interactions with others. Since time, attention, and other resources are limited in a human network, the number of connections to other entities determines the strength of interactions. The more connections exist, the lesser resources are available per connection in average. However, entities of social object networks (e.g., SaaS network) are not able to act strategically upon changes in the environment after they have joined the network. This lack of responsiveness in objects limits them to interact strategically with one another quickly. Furthermore, in contrast to human networks, actors of a social object network can have a large number of connections to other actors.

The distinction between a social object network and a human network is that the collaboration between an entity A and an entity B does not depend strongly on the degree of connections to others. Therefore, a SaaS network can have an unlimited number of connections with other entities. A new service can simply combine the

functionalities of existing services. However, the influence of humans on the formation of social object networks should not be neglected, if new objects join the social object network. That is to say, actors in social object networks receive payoffs, depending on how they have been arranged by humans when they joined the network.

3.2 Definition of the Software-as-a-Service Network

The SaaS network is defined as a set of nodes, representing software services that opened up their APIs for other services to connect to them, and a set of links between these nodes. A link indicates the existence of a composite service that uses the nodes. Therefore, the creation of a composite service yields a graph of those SaaS services that are used for the development of the composite service (Kim and Altmann 2012, 2013; Kim et al. 2011). The generated network is a weighted graph, in which the weight of the link between SaaS service i and SaaS service j indicates the degree of collaboration between the two SaaS services. The collaboration can be considered a complementarity feature of the process of human needs satisfaction.

We gathered the data from www.programmableweb.com, which lists information about SaaS services. The information collected includes the name of the service that offers open APIs, the name of the existing services used, the launch date, and service ratings of users. We collected this information from February 2006 until August 2013. In total, we observed 1153 SaaS services with open APIs that have been used for service composition by developers. The main characteristics of a composed SaaS service (i.e., mashup) is its combination and aggregation of content from more than one source (i.e., from existing SaaS services and the developer) to create a single new service. An example of the SaaS network is shown in Fig. 1. Since the resulting network is large, we filtered the whole network to only show those top 20 SaaS services that have the highest node degree. We consider this network to be the core of the SaaS network. Figure 1 depicts the links of the SaaS network with different thickness, according to their weights (i.e., the number of composite services that used them). The core of the SaaS service network shows the service developers' use patterns of the most frequently used SaaS services. It shows that the most attractive SaaS service from the developer's perspective has been Google Maps. Service developers have mainly been interested in integrating Google Maps with micro blogging services (e.g., Twitter), social networking services (e.g., Facebook), photo sharing services (e.g., Flickr), and video sharing services (e.g., YouTube).

3.3 Network Formation Model

Existing network formation models mostly depict scenarios, in which network participants are humans and have desires and needs for efficient resource allocations among themselves. A payoff functions of previous models map individuals' degree of connections to utilitarian values during the course of individuals' link

establishments, the consequence is that the more connections exist, the lesser resources are available per connection in average. However, many social objects do not have such a limitation. Besides, as previous formulations do not consider the frequency of interaction among individuals, the same payoffs will be generated for two individuals with similar connectivity patterns. However, we believe that entities with same connectivity patterns can make different contributions and, therefore, receive different payoffs. What really matters is the frequency of interactions and the relevance, quality, and novelty of information produced as the result of link establishments. Although this issue is true for both human networks and social object networks, the latter is the main focus of the work presented in this paper.

In detail, social objects differ along many dimensions and such differences often arise on the basis of preexisting functionalities or services they provide. This suggests that networks of objects form because of the advantages to the end user through the combination of different functionalities they provide. In SaaS service networks, for example, the benefits come from developing new services by combining existing services. The users' benefits are subject to the type of services combined and their contributions. For example, if the combination of two SaaS services i and j of the same service category does not lead to an added functionality that satisfies the user, there is no high benefit to such service integration.

Therefore, we need a network formation model for capturing the cost and benefit of networking among social objects with respect to relevance, quality, and novelty

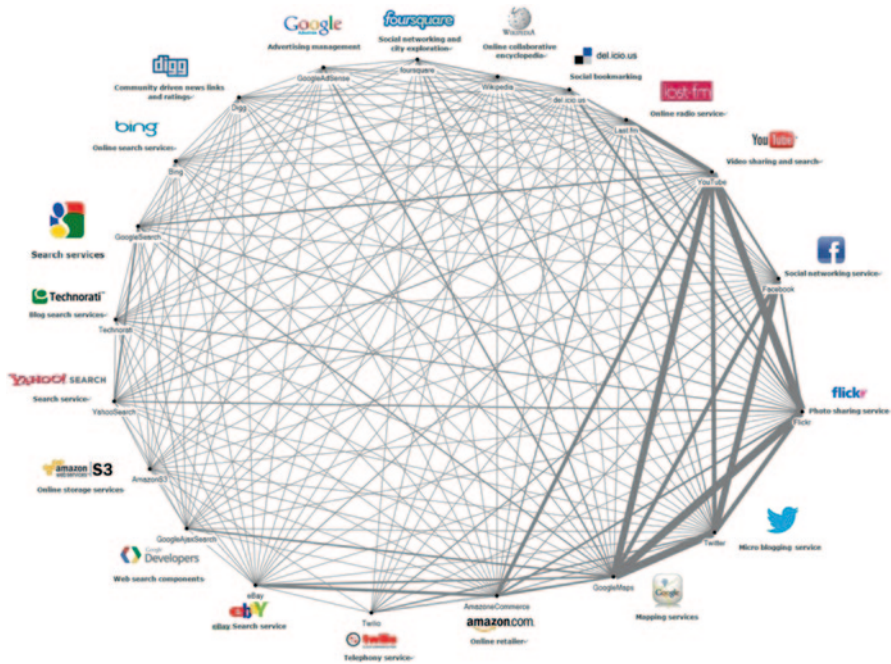


Fig. 1 The core of a SaaS service network

of information provided to and perceived by users. Specifically, if social objects are linked together and used, utility is created. If there is no link between two objects, no utility is obtained by the social objects. The detailed amount of benefit depends on the frequency of interaction, reputation of the parties, and the type of services being integrated.

Due to the above-mentioned differences between social object networks and human networks, we propose a network formation model for social object networks that distinguishes cost and payoff for establishing links when actors initially establish new relationships. We formulate the expected net utility L_i for SaaS service i from the link establishment with SaaS services j as follows:

$$L_i = \alpha_{i,j} P_{i,j} \quad (1)$$

In the same way, the expected net utility L_j for SaaS service j can be formulated (Eq. 2). The expected net utility of the new service n is given in Eq. 3.

$$L_j = \alpha_{j,i} P_{j,i} \quad (2)$$

$$L_n = V_n - C_{\{i,j,n\}} \quad (3)$$

Parameter $P_{i,j}$ is the undiscounted payoff that a SaaS service i receives, when its functionalities are integrated with SaaS service j through service composition and $C_{\{i,j,n\}}$ is the cost for integrating services i , service j , and value add of the new service n . The amount of expected payoff by combining the two SaaS services i and j depends on the probability of successful service integration. We introduce parameter α that defines the evolution dynamics of the obtained payoff. This parameter evolves gradually over time and captures the amount of changes in the net utility of SaaS services. Therefore, with probability $\alpha_{i,j}$, service developers benefit from integrating the services i and j . Furthermore, service developers would not invest any effort in a SaaS service composition, if the payoff V_n is smaller than the integration cost $C_{\{i,j,n\}}$.

Consequently, the total net utility $L_{i,j}^n$ generated through the establishment of a new SaaS service n that builds on SaaS services i and j can be formulated as shown in Eq. 4:

$$L_{i,j}^n = \alpha_{i,j} P_{i,j} + \alpha_{j,i} P_{j,i} - C_{\{i,j,n\}} + V_n \quad (4)$$

Based on our observation in Fig. 1, we propose to define α as a function of service developers' use patterns of different SaaS services. A high degree of collaboration (i.e., strong edge thickness) between SaaS service i and j shows the explicit interest of service developers on the integration of both services. In addition to this, as shown in the following formula, we normalize the degree of collaboration with the largest degree of collaboration of SaaS service j with its k direct contacts so that its value stays in the range $[0, 1]$.

$$\alpha_{i,j} = \frac{\text{deg_of_col}_{i,j}}{\max_k \text{deg_of_col}_{j,k}} \tag{5}$$

Similar to the co-author model, the focus of our model is on the collaboration among entities. There is no payoff for non-collaborating entities. Therefore, at the beginning, the utilities of objects are assumed to be 0. As soon as a link is established between two objects i and j , α can be calculated to determine the actual amount of benefit from the service composition. It should be noted that the proposed formulation produces asymmetric payoffs for two SaaS service i and j . Therefore, a newly entered SaaS service i can only expect an increase in its own payoff during the interaction with SaaS service j , if it increases the level of collaboration with SaaS service j . Otherwise, the payoff will be largely discounted through a small value of the collaboration degree $\alpha_{i,j}$.

Equation 4 can be further generalized by accounting for a combination of any number of SaaS services of the set S of SaaS services. Therefore, we formulate the expected net utility function for any new service n as follows:

$$\begin{aligned} \overline{L}_{l \subset S}^n &= \sum_{j,i \in l; j \neq i} (\alpha_{i,j} P_{i,j} + \alpha_{j,i} P_{j,i}) - C_l + V_n \\ &= \sum_{j,i \in l; j \neq i} (\alpha_{i,j} rep_j + \alpha_{j,i} rep_i) - \sum_{j,i \in l; j \neq i} sim_{i,j} + V_n \end{aligned} \tag{6}$$

In the expected net utility function sim_{ij} can be considered a representative of the service integration cost. The more similarity exists, the more effort it takes to isolate the redundant components of the services. Similarity is a numerical value indicating the degree of similarity between the two SaaS service i and SaaS service j . If we consider $x = (x_1, x_2, x_3, \dots, x_n)$ to be the profile of an object described by n characteristics, then, with the help of existing vector-similarity measures (e.g., the cosine or the Jaccard similarity measure), we can compute the similarity between the two profile vectors. If the similarity measure is 1, the cost of SaaS service integration is the highest. The comparison of similarity measures is not the focus of our current work but will be considered as an extension of the current study. We use Jaccard Similarity Measure (JSM) to capture the degree of similarity between two SaaS services. The JSM for calculating the sim_{ij} is given as:

$$sim_{ij} = \frac{SaaS\ service_{i_j1}}{(SaaS\ service_{i_j1} + SaaS\ service_{i_0j_1} + SaaS\ service_{i_j0})} \tag{7}$$

$SaaS\ service_{i_j1}$ represents the total number of attributes, where $SaaS\ service_i$ and $SaaS\ service_j$ both have a value of 1. $SaaS\ service_{i_0j_1}$ represents the total number of attributes, where $SaaS\ service_j$ has a value of 1 and $SaaS\ service_i$ has a value of 0.

$SaaS\ service_{i,j_0}$ represents the total number of attributes where $SaaS\ service_i$ has a value of 1 and $SaaS\ service_j$ has a value of 0. The JSM calculates the fraction of profile attribute matches between the profile of SaaS service i and the profile of the SaaS service j . The result of the calculation is a numerical value indicating the degree of similarity between SaaS service profiles.

We consider rep_j as the obtained payoff $P_{i,j}$ of SaaS service i during service integration with SaaS service j . Parameter rep_j is the predicted global reputation value of SaaS service j based on the issued ratings of users on newly developed services. The reputation of the integrated services reflects on the new service n . In order to calculate the payoff of SaaS service i obtained from a new web service composition with SaaS service j , we predict the global reputation of SaaS service j , rep_j . Therefore, first we extract a neighborhood $N \subseteq M$ of the total set of mashups M , in which SaaS service j has been used. Once N has been determined, we extract the issued ratings of user on all mash ups $n \in N$ and calculate the average. The resulting value represents the reputation of SaaS service j . Parameter r_n represents the collected ratings from users on newly developed mashup services.

$$rep_j = \frac{\sum_{n \in N} r_n}{|N|} \quad (8)$$

The proposed formulation is based on the idea that we can consider a new SaaS service composition to be successful, if the issued rating of users on the use of the service is positive. Therefore, we require feedback values for each of the mashups that uses the service.

Considering the SaaS network shown in Fig. 1, it seems that SaaS services that fall into different service categories (e.g., Google Maps and Flickr) are highly appreciated for service integration by service developers. Contrarily, the degree of collaboration between Yahoo search and Bing is quite low (Fig. 1). Therefore, we derive that the establishment of a link between two SaaS services in the same category is inefficient and costly.

4 Conclusion

In this paper, we discussed about existing network formation models and argued that they lack sufficient attention to social network of objects (e.g., SaaS service networks). We point to the fact that humans' desires and needs for efficient resource allocation is the reason for strategic link establishments in their social networks. Since time, attention, and other resources are limited, humans prefer to act strategically in their network to maximize their own payoffs.

However, this feature does not apply to social networks of objects. The SaaS service network, for example, can be considered as a network with potentially unlimited collaborations between its SaaS services. Each collaboration represents a new

service with new functionalities to users. Furthermore, we believe that entities in a social object networks with same connectivity patterns can make different contributions and receive different payoffs. What really matters is the frequency of interactions and the relevance, quality, and novelty of information produced as the result of link establishments. Therefore, we proposed a new social object network formation model that explains how a SaaS service network emerges during the service composition procedure.

The proposed network formation model is able to capture the cost and benefit of collaboration among objects with respect to relevance, quality, and novelty of information provided to and perceived by users. We captured the core component of a SaaS service network and, based on our observations, setup the new formulation. The new formulation not only considers the use frequency and reputation but also the similarity of the functionalities of the SaaS services. It also explains how social objects (e.g., SaaS services) can benefit from establishing links among each other in the network.

Due to the fact that the cost and benefit terms are introduced in the model, someone may argue that the proposed model is one extension of the symmetric connection model. However, as it is discussed in (Jackson 2008), the only efficient network structure of the symmetric connection model is a star network. Therefore, the symmetric connection model is unable to explain the topology of a SaaS network, which is depicted in Fig. 1. Furthermore, since the frequency of collaborations among entities is not considered in the symmetric connection model, symmetric payoffs will always be generated for two individuals with similar connectivity patterns. In short, what really matters in the symmetric connection model is the number of connections or the distance to other parties. Our model, however, considers the frequency of collaborations, reputation of an entity, and the type of services being integrated to calculate the obtained benefits during link establishment.

The co-author model with the focus on the collaborations among individuals depicts a scenario, in which network participants have desires and needs for efficient resource allocations. Each entity receives low payoff, if the degree of their direct contacts becomes larger than one; a process which emerges in form of negative externality. However, as we mentioned in Sect. 3.1, SaaS service networks do not have this limitation. The payoffs of objects in our network formation model would not be affected by the connectivity degree of their direct contacts.

As a future extension of the current work, we aim to work on the similarity measure of profiles of SaaS services. Furthermore, alternatively to the focus of this paper on the collaboration between SaaS services, association rule mining of observed usage patterns of different SaaS services by service developers may also lead to interesting results in the future.

References

- Barabasi A, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439): 509–512
- Buechel B (2008) Network formation with closeness incentives. *Networks, Topology and Dynamics*. Springer

- Carayol N, Roux P (2005) Collective innovation in a model of network formation with preferential meeting. *Nonlinear dynamics and heterogeneous interacting agents*, Springer Berlin Heidelberg, chapter 10:139–153
- Carayol, N., Roux, P. (2009) Knowledge flows and the geography of networks: a strategic model of small world formation. *J Econ Behav Organ* 71:414–427
- Erdős P, Rényi A (1959) On random graphs. *Publicationes Mathematicae* 6:290–297
- Erdős P, Rényi A (1960) On the evolution of random graphs. *Publ Math Inst Hung Acad Sci* 5:17–61
- Erdős P, Rényi A (1961) On the strength of connectedness of a random graph. *Acta Mathematica-Scientia Hung* 12:261–267
- Galeotti A, Goyal S, Jackson MO, Vega-Redondo F, Yariv L (2010) Network games. *Rev Econ Stud* 77(1): 218–244
- Jackson MO (2008) *Social and economic networks*. Princeton University Press, New Jersey
- Jackson MO, Rogers BW (2005) The economics of small worlds. *J Eur Econ Assoc* 3:617–627
- Jackson MO, Wolinsky A (1996) A strategic model of social and economic networks. *J Econ Theor* 71:44–74
- Johnson C, Gilles R (2000) Spatial social networks. *Rev Econ Des* 5:273–299
- Kim K, Altmann J (2012) A complex network analysis of the weighted graph of the Web 2.0 service network. *Advances in Collective Intelligence 2011, AISC* 113:79–90
- Kim K, Altmann J (2013) Evolution of the software-as-a-Service innovation system through collective intelligence. *International Journal of Cooperative Information Systems*, 22(3):61–72
- Kim K, Altmann J, Hwang J (2011) Measuring and analyzing the openness of the web 2.0 service network for improving the innovation capacity of the Web2.0 system through collective intelligence. *On Collective Intelligence*, Springer Berlin Heidelberg, 93–105
- Koohborfardhaghighi S, Altmann J. How placing limitations on the size of personal networks changes the structural properties of complex networks. *Proceedings of the companion publication of the 23rd international conference on World wide web companion*. International World Wide Web Conferences Steering Committee, 2014:1017–1022.
- Möhlmeier P, Agnieszka R, Emily T (2013) A degree-distance-based connections model with negative and positive externalities. *Working Papers* 479, Bielefeld University, Center for Mathematical Economics
- Morrill T (2011) Network formation under negative degree-based externalities. *Int J Game Theor* 40:367–385