INSTITUTT FOR FORETAKSØKONOMI
DEPARTMENT OF BUSINESS AND MANAGEMENT SCIENCE

FOR 8 2014
ISSN: 1500-4066
February 2014

Discussion paper

The Analysis of Split Graphs in Social Networks Based on the K-Cardinality Assignment Problem

BY
Ivan Belik

Norges Handelshøyskole
NORWEGIAN SCHOOL OF ECONOMICS

THE ANALYSIS OF SPLIT GRAPHS IN SOCIAL NETWORKS BASED ON THE *K*-CARDINALITY ASSIGNMENT PROBLEM

IVAN BELIK

Norwegian School of Economics (NHH), Department of Business and Management Science,
Helleveien 30, 5045 Bergen, Norway
Email address: ivan.belik@nhh.no

Abstract

In terms of social networks, split graphs correspond to the variety of interpersonal and intergroup relations. In this paper we analyse the interaction between the cliques (socially strong and trusty groups) and the independent sets (fragmented and non-connected groups of people) as the basic components of any split graph. Based on the Semi-Lagrangean relaxation for the k-cardinality assignment problem, we show the way of minimizing the socially risky interactions between the cliques and the independent sets within the social network.

Keywords: social networks, split graphs, *k*-cardinality assignment

1. Introduction

The understanding of how social networks are currently forming, what kinds of relations exist and the possible ways to formalize, predict and manipulate these networks through their internal mechanisms and structures are important in many socio-economic settings. Socially generated networks are different from any other types of networks in terms of their internal structure, interaction mechanisms and the tools employed to analyze them. Jackson & Rogers (2007) consider the large-scale social networks as dynamic models in trying to understand how random social networks are. They argue that different situations where social network structures "…*play a central role include scientific collaborations among academics, joint research ventures among firms, political alliances, trade networks, the organization of intrafirm management, the sharing of information about job opportunities…*". Trying to understand the formation of social networks, they present a dynamic model where the nodes in the social network form the relations in two ways: (1) randomly and (2) by searching locally through the structure.

The main feature of any social network is its natural intention for the formation of the internal communities and modules. The analysis of social networks' modularity is one of the most challenging problems in the area of social network analysis. According to Newman (2006), "*the problem of detecting and characterizing this community structure is one of the outstanding issues in the study of networked systems*". In this paper Newman shows the spectral algorithm for community detection as one of the efficient methods.

The analysis of split graphs, as one of the common structures in terms of social networks' modularity detection, represents the special interest for us in the current research.

2. Split Graphs n Social Networks

The split graph is a graph that can be partitioned into the disjoint union of a clique and an independent set (Merris 2003). In terms of social networks, split graphs reflect the realistic interpersonal and intergroup relations. It is very common to see social groups whose members are closely interrelated by similar ideas and interests, such as religion, research, education, level of income etc. In terms of graph theory, these social groups form the structures called cliques (Luce & Perry 1949). Finding the maximum clique, as the largest possible subgroup of closely related people in the social network, is an important step in the network's analysis. It corresponds for searching of the most powerful group of people in the network. The problem of finding the maximum clique is NP-complete, and deterministic polynomial time algorithms do not exist (Östergård 2002). However, finding the maximum clique in small graphs is not problematic.

In contrast to cliques, independent sets in graphs, representing the sets of nodes with no edges connecting them (Boppana & Halldórsson 1992), correspond to the socially fragmented and non-interrelated groups of people. Mostly, the people in the independent sets do not know each other; have no common interests or may even compete with each other. Finding the maximum independent set is an NP-hard optimization problem (Robson 1986).

3. Concept of Trust in Social Networks

Trust is the key concept in any social network. It is a basis for the formation of social groups and coalitions, identification of the most powerful nodes in the network, and it is the determinant for the information flow in the social network (Adali, Escriva, Goldberg, Hayvanovych, Magdon-Ismail, Szymanski & Williams 2010).

When a person has to decide whether to trust the other person or not, the decision about trust is influenced by a set of different factors. Following (Adali et al. 2010), there are three basic decision components: (1) personal predisposition to trust, (2) previous relationship with the person and his relatives, friends, colleagues etc. (3) the opinion about the decisions and actions previously made by a third person.

Trust measuring in social networks is a complex process, and there are many trust models in the literature. For example, Abdul-Rahman & Hailes (2000) proposed the trust model in virtual communities based on the idea of measuring trust employing the mechanism of the experience and reputation of the network members. Considering trust as the "subjective degree of belief about agents" (McKnight & Chervany 1996; Misztal 1996), Abdul-Rahman & Hailes (2000) show how to measure trust degrees and how to assign trust weights.

Another approach was invented by Aberer & Despotovic (2001). They present the algorithms of trust measuring based on the computation of the agent's reputation. The research shows a specific way of evaluating trustworthiness based on the local trust computations.

Adali et al. (2010) represent the metrics of trust based on the analysis of the dyadic relations. They describe the idea of behavioral trust measures, which are based on determining the communication behavior of agents in the social network. The represented methods of the behavioral measures of trust were tested based on the Twitter network data.

Any trust relation is basically affiliated with the risk of making a wrong decision when communicating with the other persons in the social network. Buskens (2002) in the book "Social Networks and Trust" provides some examples about the risks of trust within the "social context" of trust relations.

Trust is highly associated with the risk of interconnection with wrong, dangerous or suspicious people and it is therefore important to assess the trustworthiness of the relations in social networks. According to Aberer & Despotovic (2001), "*this allows to compute directly the expected outcome respectively risk involved in an interaction with an agent, and makes the level of trust directly dependent on the expected utility of the interaction.*" When discussing the problems of social exchange, Molm & Takahashi (2000) specifically consider risk and trust as basic aspects in terms of classical exchange theory. According to that paper, the evaluation of trustworthiness is initially based on the analysis of risk and uncertainty of the exchange.

4. Problem Description

In the given research, we are specifically interested in the analysis of social networks with the risk of trust as the basic social factor in the interpersonal communications and socio-economic exchange. We do not concentrate on the idea of how to measure the trust (or the risk of trust) in the network. This is a topic for different research, and we have described some approaches in this area in Section 3 ("Concept of Trust in Social Networks"). Our goal is to analyse the social networks in terms of the weighted split graphs, where the edge weights correspond to the risk of trust between the nodes (i.e., persons) in the social network. By finding the maximum clique and one or more independent sets, we are minimizing the risk of interconnections between the maximum clique members and the members of the independent sets solving a *k*-cardinality assignment problem (Belik & Jörnsten 2014). In other words, we consider the clique as the socially powerful group of people, and we have the independent sets of people who wish to enter the clique (to become members of the clique). In general, clique members do not wish persons from the "external world" to enter their "internal environment". From the clique's position, this is the risk that should be minimized or avoided completely. Nevertheless, complete avoidance is almost impossible, because social networks are not closed systems. In forming connections with the "external world", clique members try to minimize their risk of interrelations. By applying Semi-Lagrangean relaxation for the *k*-cardinality assignment problem, we minimize the overall risks. The brief description of the Semi-Lagrangean relaxation for the *k*-cardinality assignment problem is represented in the following sections 5-6.

The detailed explanation of the given mechanisms (including the testing) is provided in Belik & Jörnsten (2014).

5. The Integer Programming Formulation of the *k*-cardinality Assignment Problem

The integer programming formulation of the *k*-cardinality tree problem is as follows:

$$Min \sum_i \sum_j c_{ij} x_{ij} \quad (1)$$

Subject to:

$$\sum_{ij} x_{ij} = k \quad (2)$$

$$\sum_i x_{ij} \leq 1 \qquad \forall j \in J \quad (3)$$

$$\sum_j x_{ij} \leq 1 \qquad \forall i \in I \quad (4)$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in I, j \in J \quad (5)$$

6. Semi-Lagrangean Relaxation

The formalization of the Semi-Lagrangean relaxation (Beltran et al. 2006) for the *k*-cardinality problem, relaxing the single equality constraint, was described in Belik & Jörnsten (2014) and has the following representation:

Max SL(u) subject to u≥0,

where $\mathcal{L}(u)$ is defined by the following optimization problem:

$$Min \sum_i \sum_j c_{ij} x_{ij} - u(\sum_{ij} x_{ij} - k) \quad (6)$$

Subject to:

$$\sum_{ij} x_{ij} \leq k \quad (7)$$

$$\sum_i x_{ij} \leq 1 \qquad \forall j \in J \quad (8)$$

$$\sum_j x_{ij} \leq 1 \qquad \forall i \in I \quad (9)$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in I, j \in J \quad (10)$$

7. Implementation

Consider the illustrative example of a social network with 14 people. The structure of the network is represented in Fig. 1. The edges' weights correspond to the risk of trust between the persons.

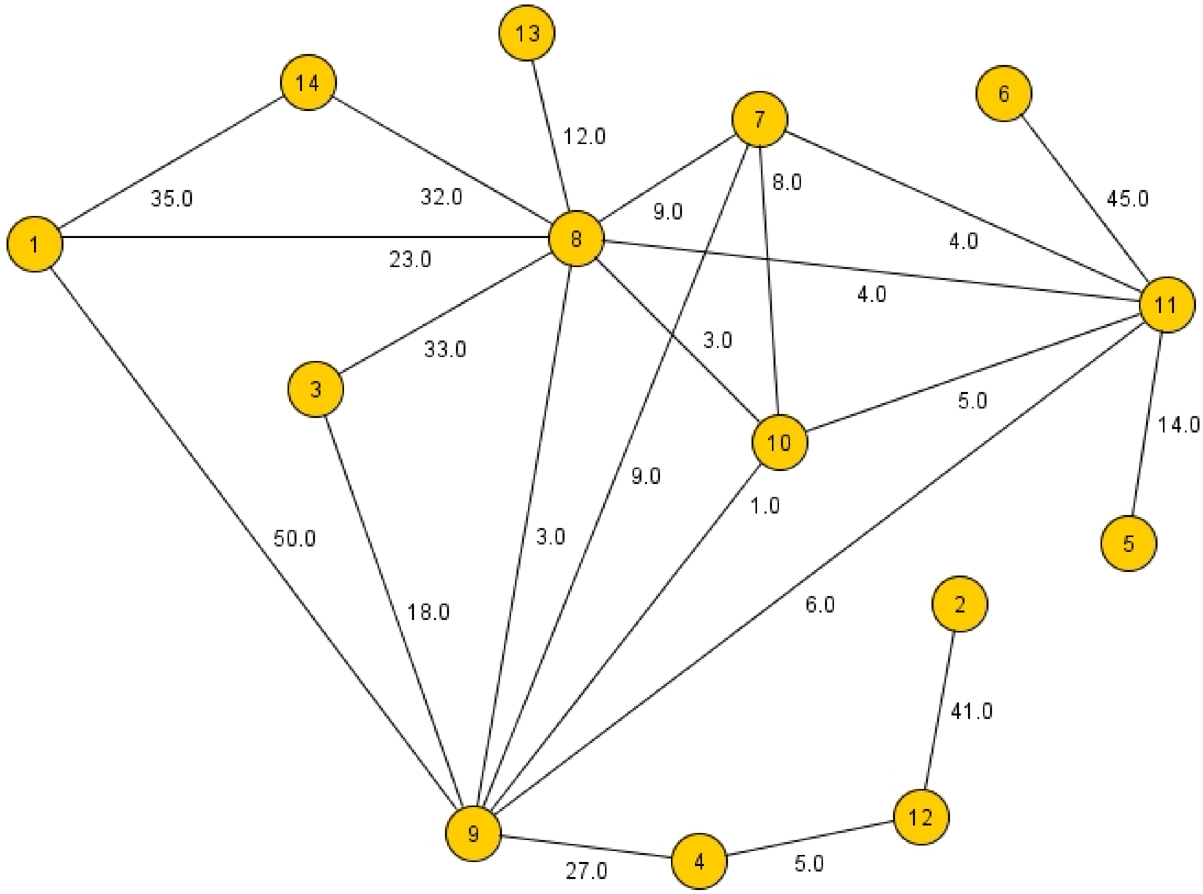

Fig. 1 Social network: illustrative example

We analyse the given graph as the split graph looking for the maximum clique and the independent sets. The results of the analysis are represented in Fig. 2.

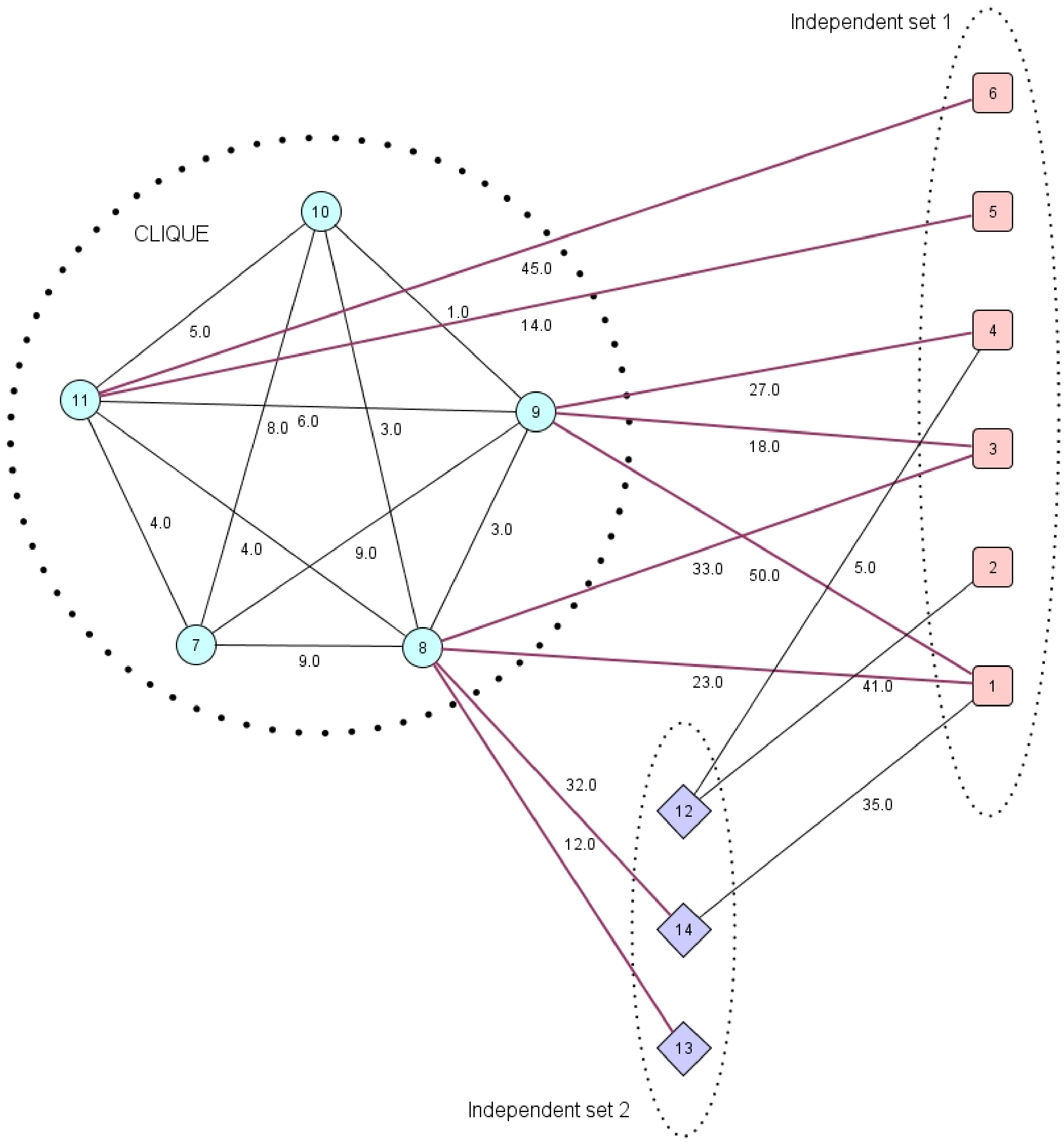

Fig. 2 Maximum clique and independent sets in the social network

According to Fig. 2 we have the clique that consists of 5 persons (i.e., 7, 8, 9, 10, and 11) and two independent sets: persons 1, 2, 3, 4, 5, 6 in the first independent set and 12, 13, 14 – in the second independent set. The clique members have to communicate with the external groups minimizing the risk of being involved in the interconnections with persons that are out of the clique.

In terms of social networks, there can be different situations. For example, the clique is the research team in the university, where some research members are interested in hiring assistants for themselves from the set of students (i.e., independent set). Each clique member is interested in hiring the right person minimizing the overall risk for the clique of hiring the wrong student from the independent set.

Another example is the company that has to choose the specialists to hire from the number of trainees. In this case, the clique is the set of company departments looking for the trainees to be hired. For instance, the IT department is looking for a programmer from one set of trainees (i.e., first independent set); the finance department might be looking for the financial analyst from another set of trainees (i.e., second independent set) etc.

Considering our split graph represented in Fig. 2 we have m=5, n=9, k=3, and the matrix of risks represented in Table 1.

Table 1. C-matrix of risks

		Indep. Set 1						Indep. Set 2		
		1	2	3	4	5	6	12	13	14
CLIQUE	7	100	100	100	100	100	100	100	100	100
	8	23	100	33	100	100	100	100	12	32
	9	50	100	18	27	100	100	100	100	100
	10	100	100	100	100	100	100	100	100	100
	11	100	100	100	100	14	45	100	100	100

In C-matrix of risks m is the number of clique members, n is the number of persons in the independent sets, k is the number of persons to be involved in the interconnection with clique members, which we assign to be equal to “3”. Weights in the matrix correspond to the level of risk with values in the range [0,100]. It is important to notice that if the edge between the member of the clique and the member of the independent set does not exist then the risk of making the interconnection is assigned to be “100” matching the highest level of risk and uncertainty.

We solve the k-cardinality assignment problem based on the Semi-Lagrangean relaxation following the procedure described in Belik & Jörnsten (2014).

The optimal Semi-Lagrangean multiplier is u^*=18. The resulting cost matrix is represented in Table 2.

Table 2. Cost matrix

		Indep. Set 1						Indep. Set 2		
		1	2	3	4	5	6	12	13	14
CLIQUE	7	X	X	X	X	X	X	X	X	X
	8	X	X	X	X	X	X	X	**-6**	X
	9	X	X	**0**	X	X	X	X	X	X
	10	X	X	X	X	X	X	X	X	X
	11	X	X	X	X	**-4**	X	X	X	X

In Table 2 all non-allowable assignments are marked by X. The solution to the problem is the "row-column" assignment 8-13, 9-3, and 11-5 of cardinality "3" with the objective function value equal to "-10". The lower bound (LBD) is 54-10=44. The feasible solution (the upper bound) is also "44". Therefore, the optimal solution has been found.

The optimal Semi-Lagrangean multiplier $u^*=18$ has a meaningful interpretation in terms of social networks. It corresponds to the highest risk of interconnection with non-clique members: the interconnection between person 9 from the clique and person 3 from the first independent set is the most risky one for the whole clique. Two other newly assigned interconnections i.e., person 11 from the clique and person 5 from the first independent set; person 8 from the clique and person 13 from the second independent set, are less risky. In terms of social networks, it helps to formulate the prospective relation strategies with the new persons. For example, person 3 might be required to be controlled more than others or might be additionally trained. In general, solving the *k*-cardinality assignment problem, we minimize the risk of the clique members' interconnections with the new persons.

6. Interpretation of the Results in Terms of Social Networks

The social meaning of the described mechanism for the minimization of risky intergroup interactions is important in terms of the trustworthiness analysis in social networks. The ability to detect the least risky relations between different types of social groups has a significant practical importance. Its central role is obvious in the research collaborations among scientists, joint ventures among companies, hiring new employees, and personnel management etc.

The case, analyzed in section 5, represents a small social network with only 14 nodes. However, the real-world social networks are characterized by large-scale structures and the multifactor complex nature of the relations. The analysis of risks in the interpersonal and intergroup communications is one important problem, but it is also important to know the ways of minimizing those risks. One of these ways has been shown in this paper.

7. Conclusion

In the given research we considered social networks in terms of split graphs. Since split graphs reflect the structure of many real-world social networks, we analyzed their internal mechanisms considering the risk of trust as the basic factor in the interpersonal and intergroup relations. In this paper we described the concepts of trust in section 3, explaining its exclusive importance.

The risk of trust between the substructures of the split graph was our main concept of analysis.

We represented the problem of the risk of trust minimization in terms of *k*-cardinality problem applying the Semi-Lagrangean relaxation to solve it. The mechanism is based on the research represented in Belik & Jörnsten (2014). The approached result shows the effective mechanism to minimize the risky interconnections that has to be established in the split graph due to the internal social requirements in the network.

References

Abdul-Rahman A, & Hailes S (2000, January) Supporting trust in virtual communities. In System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on (pp. 9-pp). IEEE.

Aberer K, & Despotovic Z (2001, October) Managing trust in a peer-2-peer information system. In Proceedings of the Tenth International Conference on Information and Knowledge Management (pp. 310-317). ACM.

Adali S, Escriva R, Goldberg M K, Hayvanovych M, Magdon-Ismail M, Szymanski B K, & Williams G (2010, May) Measuring behavioral trust in social networks. In Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on (pp. 150-152). IEEE.

Belik I, & Jörnsten K (January 15, 2014) A New Semi-Lagrangean Relaxation for the K-Cardinality Assignment Problem. NHH Dept. of Business and Management Science. Discussion Paper No. 2014/1. ISSN: 1500-4066

Beltran C, Tadonki C, & Vial J P (2006) Solving the p-median problem with a semi-Lagrangian relaxation. Computational Optimization and Applications, 35(2), 239-260.

Boppana R, & Halldórsson, M M (1992) Approximating maximum independent sets by excluding subgraphs. BIT Numerical Mathematics, 32(2), 180-196.

Buskens V (2002) Social networks and trust (Vol. 30). Springer

Jackson M O, & Rogers B W (2007) Meeting strangers and friends of friends: How random are social networks? The American Economic Review, 890-915.

Luce R D, & Perry A D (1949) A method of matrix analysis of group structure. Psychometrika, 14(2), 95-116.

McKnight D H, & Chervany N L (1996) The meanings of trust. Technical Report 94-04, Carlson School of Management, University of Minnesota, 1996.

Merris R (2003) Split graphs. European Journal of Combinatorics, 24(4), 413-430.

Misztal B A (1996) Trust in Modern Societies: The Search for the Bases of Social Order (Vol. 1). Cambridge: Polity Press.

Molm Linda D, Nobuyuki Takahashi, and Gretchen Peterson (2000) Risk and trust in social exchange: An experimental test of a classical proposition. American Journal of Sociology: 1396-1427.

Newman M E (2006) Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577-8582.

Östergård P R (2002) A fast algorithm for the maximum clique problem. Discrete Applied Mathematics, 120(1), 197-207.

Robson J M (1986) Algorithms for maximum independent sets. Journal of Algorithms, 7(3), 425-440