# The Emergence Of Groups In The Evolution Of Friendship Networks

Zeggelink, Evelien P.H.; Stokman, Frans N.; Bunt, Gerhard G. van der

*Journal of Mathematical Sociology*
1996, Vol. 21(1–2), pp. 29–55
Reprints available directly from the publisher
Photocopying permitted by license only



# THE EMERGENCE OF GROUPS IN THE EVOLUTION OF FRIENDSHIP NETWORKS

EVELIEN P. H. ZEGGELINK, FRANS N. STOKMAN and
GERHARD G. VAN DE BUNT*

*ICS, University of Groningen, Gr. Rozenstraat 31, 9712 TG Groningen,
The Netherlands*

Friendship networks usually show a certain degree of segmentation: subgroups of friends. The explanation of the emergence of such groups from initially dyadic pair friendships is a difficult but important problem. In this paper we attempt to provide a first contribution to the explanation of subgroup formation in friendship networks by using the LS set as a definition for a friendship group. We construct a dynamic individual oriented model of friendship formation and provide preliminary simulation results that give an idea of how to continue the process of explaining group formation.

## 1. INTRODUCTION

It is general knowledge that *friendship groups* are very important aggregations from the view of many of its members. The group of which an individual is a member influences his behavior and attitudes.[1] How such groups *emerge*, from individual behavior and attitudes on the other hand, is a completely different, but at least as important, process. It is a difficult process however.

In a population of individuals who have to interact with each other for a longer period of time within a certain context, it is not solely dyadic friendships that are developed. Once individuals deepen and strengthen their dyadic friendship relationships, they influence each other's personal lives, thoughts and actions. As a result of restricted time, effort and from emotional motives, they bring their friends together because when one's friends know one another, it is easier to relate closely and frequently with each of them (Feld, 1981). It is clear that the larger the number of friends that two individuals have in common, the higher the probability that these two will be introduced to each other and the more encouragement (conscious or unconscious) there will be for them to become friends (Hammer, 1979, 1980; Salzinger, 1982). Chances increase that these individuals get to like each other because they have *common friends* and consequently have a higher probability to have common interests. This positively valued interaction through common friends and frequent interaction leads to an increasing degree of overlap in friendships of

*We are indebted to Tom Snijders, Tom Fararo, and Patrick Doreian for fruitful comments on earlier versions of this manuscript.

[1]In the following, whenever we use the male form to refer to an individual, we mean the female reference as well.

the individuals concerned (among others, Freeman, 1992; Homans, 1950; Ridgeway, 1983; Romney and Faust, 1982). As such, the 'friends' go through a process from mainly togetherness to a higher level of 'groupness' (Sherif and Sherif, 1964). If the size of network is sufficiently large, recognizable groups of friends can then be distinguished in friendship networks (Bernard and Killworth, 1973). In general, these friendship groups are very *homogeneous*. The most important source of this is the initial selection of friends that is based on similarity (see also Leenders, 1996). Group homogeneity is further increased by conformity pressures and homophilic selection of new members (Cohen, 1977).

A group is however more than just an arbitrary, but homogeneous, aggregate of friendships. In the first place there is an *observable difference* between the *density* of friendships within the group as compared to the sparse number of friendships between group members and non-group members (Reitz, 1988). As such, it is important to realize that a three-person group of individuals $i$, $j$, and $k$ has a different sociological meaning than a triad consisting of the dyadic pairs $i$ and $j$ $i$ and $k$, and $j$ and $k$, who all interact with each other but never interact simultaneously (Wilson, 1982). The former 'real' groups structure may however, and usually does, result from the latter.

Moreover, groups are of *limited size*, and there is '*more*' than just structure and homogeneity. This 'more' results from the fact that more internal friendships (greater density within the group) lead to a higher degree of *closeness*, i.e. the overall extent to which 'deep' feelings exists between group members. Members are more committed, care more about the group, and develop specialized group cultures, inside jokes, rituals, and norms to regulate important activities frequently engaged in (Ridgeway, 1983). Norms are also established to set bounds for treatment of members and maintenance of the group as a unit. This relates to the fact that the group is more discernable, the less external friendships. This character as a unit is also recognized explicitly in the sense of "we" and "they", "the club in the back".

Although much global facts and intuitive notions are known about friendship groups in general, it may have become clear that a *definition of friendship group* is extremely difficult to give. There seems to be no generally acknowledged definition of a group. One of the most enduring concerns of social network scholars has been the attempt to discover the subgroups into which a network can be divided. Different concepts have been developed to define these subgroups as there are cliques, clusters, clubs, clans, cores, circles, and components (among others, Alba, 1973; Alba and Moore, 1978; Luce and Perry, 1949; Mokken, 1979; Seidman, 1983). All definitions have their specific advantages and disadvantages depending on the research purpose. Researchers in social psychology also apply many different (but less mathematical and strict) definitions of groups. Nevertheless, there seems to be some overlap in the use of one or more of the following elements (Shaw, 1983): Within groups there are interactions between the group members, members perceive the others (the members are aware of the existence of the group), members develop shared perceptions, affective ties and (consciously or unconsciously) organize the group with regard to roles, statuses, and norms.

The definition of a group is already difficult, let alone the attempt to explain the emergence of such groups. It is a fundamental problem for which no satisfying detailed explanation and models are present (see e.g. Fararo and Doreian, 1996): Research has rather been focused on the maintenance of existing group structures, (the emergence of) behavioral patterns, like norms and values, within groups, than on the first emergence of them (Mullen and Goethals, 1987; Shaw, 1983; Ridgeway, 1983). In this paper we attempt to provide a first contribution to the explanation of emergence of groups by the construction of an individual oriented model of friendship network evolution by focusing on the structural aspects of groups only. This means that we will pay neither attention to homogeneity aspects of groups nor pay attention to aspects related to norm emergence and so on (e.g. Flache and Macy, 1996). Before doing so, we both need an appropriate group definition and some knowledge of the sparse sociological and social psychological literature on group emergence and dynamics. We focus on the development of a friendship network in a relatively small and closed population of individuals who are initially mutual strangers and who will interact with each other for a certain time in a specific context.



In Section 2 we present a summary of the most common theories on group emergence and dynamics. In Section 3, the idea of dynamic individual oriented models of network evolution is introduced. A definition of group is presented in Section 4 followed by its incorporation in the individual oriented 'group' model in Section 5. Section 6 deals with the simulation results of this model and in Section 7 we conclude and provide some suggestions for future developments.

## 2. THEORIES OF GROUP FORMATION

Most theories on group formation can be found in the social psychological literature on group dynamics. However, usually these theories are concerned with *task oriented* groups: small sets of people that come together to solve a problem, take a decision, are a therapeutic group and so on. The group dynamics research usually is concerned with the evolution of relationships, and the organization that develops within such a single small group (Mullen and Goethals, 1987; Shaw, 1983; Ridgeway, 1983). These processes differ from the ones of our interest. We want to address the process of friendship group formation within a larger population (friendship networks). Such groups may define a group goal after a while (and will establish group norms), but the pure coming into existence of such groups is the intriguing phenomenon. Therefore theories on why people would attempt to get into a group, why people feel attracted to a group, are only of secondary importance, because such groups first have to be established. There has been hardly any study that examines such processes.

We can however use the following. Initiating the first friendship with an individual outside the group can be quite risky for a group member when his group has strong norms regarding 'out group behavior'. Therefore, an individual takes the judgments of his group subconsciously into consideration when deciding whether to become friends with a non-group member. If group members do succeed in establishing new friendships outside of the group, these new friends will not self-evidently become members of the group, but will be more easily accessible potential candidates for friendship for other group members and later on group membership. Newcomers to a 'nearly saturated' group are accepted only if they were already friends with a group member in some other context (Freeman, Freeman, and Michaelson, 1988). It is also imaginable that the internal 'outseeker' is isolated from the group.

In the literature, groups are often referred to as self-regulating entities. The view and the behavior of the individual with respect to his group or other groups is usually of minor importance. We argue that it is the individual who decides whether he wants to belong to a group, individuals build and constitute the group, and the individuals have internal and external friendships. In short, a group does not behave, an individual does. A member of one group might, however, behave differently from a member of another group, and probably even more differently from an individual who does not belong to any group at all. To illustrate some important elements we give the following findings: New relations tend to develop within the existing group and access to different groups is limited (Granovetter, 1982). Group members make more choices within the group than out of it and as a result, new friendships that group members develop remain within the group (Salzinger, 1982). Groups influence and constrain individual behavior but the reverse direction is at least as important: Individual behavior has an impact on group structure and composition. This proposition again underlines the importance of micro-macro approaches (Doreian et al., 1996, Leenders, 1995) and the focus on the link between local and global properties (Skvoretz et al., 1996).

## 3. INDIVIDUAL ORIENTED MODELS OF NETWORK EVOLUTION

The question addressed in this paper is how to explain the emergence of groups in friendship networks. The attempt to answer this question is part of an extended goal to predict *structure* in friendship networks in general. Within the scope of this larger goal, previous models have already been developed (Zeggelink, 1993, 1994, 1995). The general approach constructing these models is that networks are conceived as the macro level and individuals and their behavior as the micro level. The intriguing aspect is to predict network structure from individual behavior regarding relationship formation and dissolution. This is particularly important when relationships are not formed in view of some goal at the macro level (or are not predetermined), but emerge from individual choice: individuals initiate, build, maintain, and break up friendships and thereby determine the overall structure of the friendship network. This emphasis on the mechanisms underlying the existence of groups in larger social structures resembles the reflections on the generative approach (dynamics or microanalysis) of tripartite structures and the idea of action structure as the content of social structure (Fararo and Doreian, 1984).

The models that have been developed and the one to be developed here all "start from scratch" in the sense that the initial situation of the network evolution is a population of unrelated individuals as in Skvoretz et al. (1996). The individuals are mutual strangers. Also we consider closed populations: no individual enters or leaves the population.[2] As such we try to capture the main determining mechanisms of network evolution. Thus, in the case of the goal of the model developed here, no

[2]Allowing individuals to leave or to enter the population would lead to models in the direction of the most ambitious dynamic theories, according to Doreian et al. (1996).



subgroups are present in the beginning of the process. This situation, as well as the situation of an 'open' population, can however be incorporated easily with the approach we take.

The models are based on the approach of methodological individualism. We capture the basic principles of this approach, in our framework of evolving friendship networks, in three components. The *first* component concerns the rational behavior of an individual with respect to initiating, maintaining, and breaking off friendships. Simple individual behavioral rules can be extracted from the literature: goals of the individual can be defined in terms of the need for social contact (basic model: Zeggelink, 1993, 1994), the preference for similarity (similarity model: Zeggelink, 1993, 1995), and the structure of his friendship relationships (present group model). The latter refers to preferences with regard to within group orientation versus out group friendships and so on. The individual is however hampered by several aspects like his amount of information and his capability of imagination. This leads to bounded rationality: the individuals use simple heuristics to choose such that they expect that their goals will be approaches as close as possible (Snijders, 1996).

The *second* component in the prediction involves the constraints from the network on the individual orientations: the network as it is constructed influences the individual choices. It is clear that the composition of the population already determines the availability of desired friends, but the structure of the network also determines, among other things, the possible availability of individuals as friends, their positions in the network, number and sizes of friendship groups (and thereby availability).

The *third* component, the transformation rule, subsequently shows how individual choices interact and jointly determine the structure of the emergent friendship network. The appropriate transition is not necessarily the simple aggregation of individual preferences, but is a complex combination of interdependent individual behaviors.

These basic principles of methodological individualism have been implemented in dynamic individual oriented models in an *object oriented programming* environment (among others, Stokman and Van Oosten, 1994; Stokman and Zeggelink, 1996; Rumbaugh et al., 1991). Object oriented programming or modelling is a natural way to approach and model phenomena by basing the model on communicating objects that act and react to each other. By considering individuals (but also relationships and networks) as objects, such models take into account the differences in attributes and behavior of individuals and furthermore explicitly consider the influence of the network on the individual. Moreover, the dependence between relationships of different individuals and the fact that individuals may act simultaneously are taken into account (see also Skvoretz et al., 1996). Details of this kind of modelling are presented in Section 5 when we introduce the group aspect in individuals' behavior. First we need an appropriate group definition.

## 4. DEFINITION OF A GROUP

It is difficult to strictly define the concept of friendship group verbally (Section 1), let alone in a technical way. One may therefore argue that it is useless to define a group in a strict sense when it is not even possible to define it verbally. Nonetheless, an explicit group definition which best fits the properties described in Section 1 is needed for constructing the individual oriented model. In this view of the individual oriented model, the intuitive definition should represent how people observe groups. This implies that it should capture the contrast between the structural character of the group itself and the structure between itself and other groups or individuals. It appears that the rule people use is approximately the following. An individual is assigned to group $J$ if he interacts, on the average, more with others in $J$ than with others not in $J$. It is the *large number of friendships within* as well as *the small number of friendships outside the group* that is distinctive (Freeman, Freeman and Michaelson, 1988, 1989; Sailer and Gaulin, 1984).

Numerous representations of groups (subgroups, subnetworks, or subsets) exist that attempt to represent the characteristics mentioned in Section 1. Before we present some of these definitions, we introduce some graph theoretical definitions.

Let $G = (V,E)$ be an undirected graph with vertex set $V$ and edge set $E$. $G' = (V',E')$ is a *subgraph* of $G$ if its vertex set $V'$ is a subset of $V$, and $E'$ consists of all edges of $E$ that are incident with vertices in $V'$. $G$ is called a supergraph of $G'$. The number of vertices in a subset $V'$ of $V$ is denoted $|V'|$ and is called its *size.* A subgraph $G'$ of $G$ is *maximal* with respect to some property if that property holds for the subgraph $G'$ but is not valid anymore for any other subgraph of $G$ that is a supergraph of $G'$.

For the translation to sociological terms, graph corresponds with (friendship) network, edge corresponds with friendship, and vertex with individual. Subnetwork is the sociological counterpart of subgraph.

A *clique* is a maximal complete subnetwork of three or more individuals: every individual in the clique is friends with every other individual in it (Harary, Norman, and Cartwright, 1965; Luce and Perry, 1949). The clique definition is very strict. Friendships between the individuals in the clique are relevant in the first place. External friendships are considered implicitly because the definition of clique involves the aspect of maximality. A drawback of the clique definition is that one individual may be contained in many different cliques.

Alba and Moore (1978) propose a combination process for cliques to identify larger, less restrictive, subnetworks in the network: *social circles.* Since, in general, many small overlapping cliques exist in a friendship network, they suggest to merge these cliques into large subnetworks when they overlap sufficiently according to some threshold.

Various other technical generalizations of the clique exist which are less restrictive and are all concerned only with friendships within the groups. An *n-clique* is a maximal subset of individuals such that the largest geodesic[3] between any two individuals in it is not longer than $n$ (Alba, 1973; Luce, 1950). Limitations of the $n$-clique are that its diameter[4] may be larger than $n$ because the geodesics do not necessarily have to be contained in the subnetwork. As an extreme instance, an $n$-clique can be disconnected.[5]

The definitions of $n$-cliques, $n$-clubs and $n$-clans focus only on distances. An *n-clan* is an $n$-clique in which the distance between all individuals in the subnetwork is not larger than $n$ for paths within the subnetwork. An *n-club* is a maximal subnetwork of diameter $n$. The distance between all individuals in the subgraph is less than or equal to $n$ (Mokken, 1979).

Definitions of $k$-degree sets, $k$-cores, and $k$-plexes focus on numbers of friendships (Seidman, 1983; Seidman and Foster, 1978). The definition of *k-degree set* requires that members of the subset have at least $k$ friends within the subset. $k$-degree sets are numerous and can be overlapping A *k-core* is a maximal $k$-degree set. $k$-cores do not partially overlap and form a hierarchical ordering of disjoint subsets of individuals. A *k-plex* is a maximal subnetwork $V'$ in which every individual is friends with at least $|V'| - k$ individuals in the subnetwork. A $k$-plex is robust to the departure of individuals (in contrast to an $n$-clique), and furthermore has limited diameter.

All these definitions of groups focus on different properties like the minimum number of internal friendships, the maximum number of external friendships, the maximum distance between members of the group, the relative number of friendships in a group, et cetera. Some definitions are related to each other by being more or less restrictive than the other. The main shortcoming in these definitions is that they all lack a simultaneous consideration of both the internal and external friendships of the group members. The definition of LS set does however take both into account (Luccio and Sami, 1969; Seidman, 1983; Borgatti et al., 1990).

DEFINITION 1 *A cutset $S$ of a connected graph is a set of edges of $G$ such that its removal from $G$ disconnects $G : G - S$ is disconnected. The number of edges in a cutset that has the minimum number of edges, is called the edge connectivity $\lambda(G)$ of $G$.*

*The edge connectivity $\lambda(v,w)$ of the pair of vertices $v$, $w$ is the minimum number of edges that must be removed to disconnect them. $\lambda(V') = \min\{\{\lambda(v,w) \mid v,w \in V'\}\}$ is the edge connectivity of $V'$, where $V'$ is a subset of $V$.*

DEFINITION 2 *The number of edges joining two subsets $V_1$ and $V_2$ of $V$ is denoted $\alpha(V_1,V_2)$. The external edges of a subset $V'$ of $V$ are the edges joining vertices of $V'$ with vertices of $V - V'$. The number of external edges of a subset $V'$ is denoted $\alpha(V') = \alpha(V',V - V')$.*

DEFINITION 3 *A subset $L$ of the vertex set $V$ of a graph $G$ is an LS set if and only if for any proper subset $K$ of $L$, $\alpha(K,L - K) > \alpha(K,V - L)$.*

Thus, an LS set is a set of vertices $L$ in which each proper subset has more edges to its complement within $L$ than to the outside of $L$. An alternative definition is:

DEFINITION 4 *A subset $L$ of $V$ of a graph $G$ is an LS set if for any proper subset $K$ of $L$, $\alpha(K) > \alpha(L)$.*

[3]A geodesic is a shortest path between two vertices, if connected, in a graph.
[4]The diameter of a graph is the length of any longest geodesic.

[5]A graph is connected if every single vertex in it can be reached (by a path) from every other vertex. As soon as two vertices are not reachable from each other, the graph is called disconnected.

According to Definition 4, an LS set with $\lambda = 1$ can exist only if it is disconnected from the rest of the graph.[6] At least one vertex of such an LS set can become disconnected if just one edge would be removed. Therefore we consider only LS sets with $\lambda > 1$ for an LS set to have any meaning as a 'group'.

Every individual vertex and $V$ itself are trivial LS sets. All other LS sets are called nontrivial LS sets. For what follows, we pay special attention to the following properties of LS sets.

THEOREM 1 *Let L and M be LS subsets of* $V(G)$. *Then if* $L \cap M \neq \emptyset$, *either* $M \supset L$ *or* $L \supset M$.[7]

Thus, LS sets cannot, in contrast to most other definitions of subnetworks, partially overlap but they can contain each other. The fact that LS sets cannot overlap is not problematic because we consider prespecified closed contexts in which the individuals interact. Usually in such environments, relatively small, informal groups based on positive close personal relationships, seldom overlap (Freeman, 1992). Another important property is described in Proposition 1.

PROPOSITION 1 *Suppose L is an LS set of a graph* $G = (V, E)$. *Then for all* $u, v, w \in L$, *and* $x \in V - L$, $\lambda(u, v) > \lambda(w, x)$.

Thus in order to disconnect two vertices, the number of edges to be removed for vertices in an LS set is always larger than that for a vertex within the LS set and a vertex outside the LS set. LS sets are difficult to disconnect by removing edges: there is relatively high density within the LS set as compared to the loose connectedness to the outside.[8]

In sociological terms, the relatively high value of $\lambda$ represents the close internal structure of the group: members have many direct or indirect friends (at distance two) in common which contributes to the reinforcement of every friendship. (Since we assumed $\lambda > 1$, group members have at least two friends within the group). Consequently, the dissolution of a friendship will be more negative (than outside a group) because two individuals who used to be friends will, at least indirectly, remain confronted with each other. Thus a high value of $\lambda$ does not only mean that more friendships should dissolve for the group to fall apart, but it is also more difficult to dissolve a friendship. Thus the higher the value of $\lambda$ in a group, the stronger the 'group feelings' of its members. In general, LS sets cannot be characterized by

[6]Note that we do not say that all LS sets disconnected from the rest of the graph have edge connectivity $\lambda = 1$.

[7]Consequently, LS sets can be partitioned in a nested way. The minimal LS sets have high edge connectivity. Every pair of vertices within the LS set is connected by a relatively large number of independent paths. Independent paths have no edges in common, so when one edge is removed the two vertices are always connected through at least one other path. The higher the edge connectivity, the more difficult it is to disconnect two vertices in the LS set and the more robust it is with regard to the removal of edges. The maximal LS sets have lower edge connectivity and are less connected.

[8]Although LS sets are quite insensitive to the removal of edges, they can be sensitive to the removal of vertices. In our sociological translation, these vertices are individual group members. Since we assume closed populations, this is not too big a problem. Moreover, it is known from reality that a group may fall apart as one of its members leaves.



edge connectivity alone. LS sets with equal edge connectivity can have different sizes and different structures when of equal size.

The low value of $\alpha$ represents that an LS set does not contain any individual or subgroup with more 'linkage' outside the group than within the group. LS sets can therefore be assumed to be stable. Moreover, the value of $\alpha$ is a measure of distinctiveness: the smaller $\alpha$ the more distinct the group. These and other aspects are illustrated in Figure 1. We present all possible LS sets of size 3 and size 4, and examples of no LS sets, in graphs of size $g \geq 6$. We present only graphs of size $g = 6$, but any number of vertices can be added, while the LS set remains an LS set, as long as there are no edges between these added vertices and the vertices in the LS set. Different values of $\alpha$ and $\lambda$ are possible. Black vertices represent LS set members. Different individuals in the same LS set can have different positions, e.g. central or peripheral. For central individuals, the ratio of internal friendships to external friendships exceeds that of peripheral individuals. Figure 1 shows how these different situations can be distinguished according to the individual that has external friendships. E.g. in Figure 1(f), the individual with relatively many internal friendships also has the external friendship, whereas in Figure 1(g), the individual with relatively few internal friendships has the external friendship.

We now briefly examine the presence and detection of LS sets in a larger network. Consider graph $G$ with $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}, v_{14}\}$ in Figure 2. $G$ contains 2 LS sets: $L = \{v_1, v_2, v_3\}$ and $M = \{v_{11}, v_{12}, v_{13}, v_{14}\}$. Table 1 shows why $M$ is an LS set (see Definition 4). $\{v_4, v_5, v_6\}$ is not an LS set because e.g. $\alpha(\{v_5\}) = \alpha(\{v_4, v_5, v_6\}) = 2$, while it should be $\alpha(\{v_5\}) > \alpha(\{v_4, v_5, v_6\})$ for an LS set. $\{v_7, v_8, v_9, v_{10}\}$ neither is an LS set because $\alpha(\{v_9\}) = \alpha(\{v_7, v_8, v_9, v_{10}\})$.

With the choice of the LS set as a definition for a social group, we implicitly assume that the structural difference between internal and external friendships is sufficient to detect 'real' friendship groups. Three individuals may therefore constitute an LS set while they never interact simultaneously. Since the very nature of groups depends in the first place on the friendships between the individuals comprising them (Breiger, 1974), the definition seems reasonable. A disadvantage of LS sets however is that they do not or rarely appear in empirical data sets because they cover just so many aspects of the intuitive notion of a sociological group. Nevertheless, it is the best definition to be used in our individual oriented models, because it is so relevant for the aspect how individuals observe groups and because no better definition is available.

Another important aspect that deserves more attention is that individuals cannot belong to several groups simultaneously because LS sets cannot partially overlap. As a first start, it is plausible to assume that individuals do not belong to more than one group if the only relationship of interest is the friendship relationship in small closed populations. However, in the future we would want to be able to consider the situation of individuals belonging to several groups, and as a consequence, occupying different roles in different groups. Bipartite and tripartite structural analyses become relevant (Breiger, 1974; Wilson, 1982; Fararo and Doreian, 1984) to examine the duality between individuals and groups: groups are defined by its members, and at the same time, an individual's individuality is determined by the different groups he belongs to.

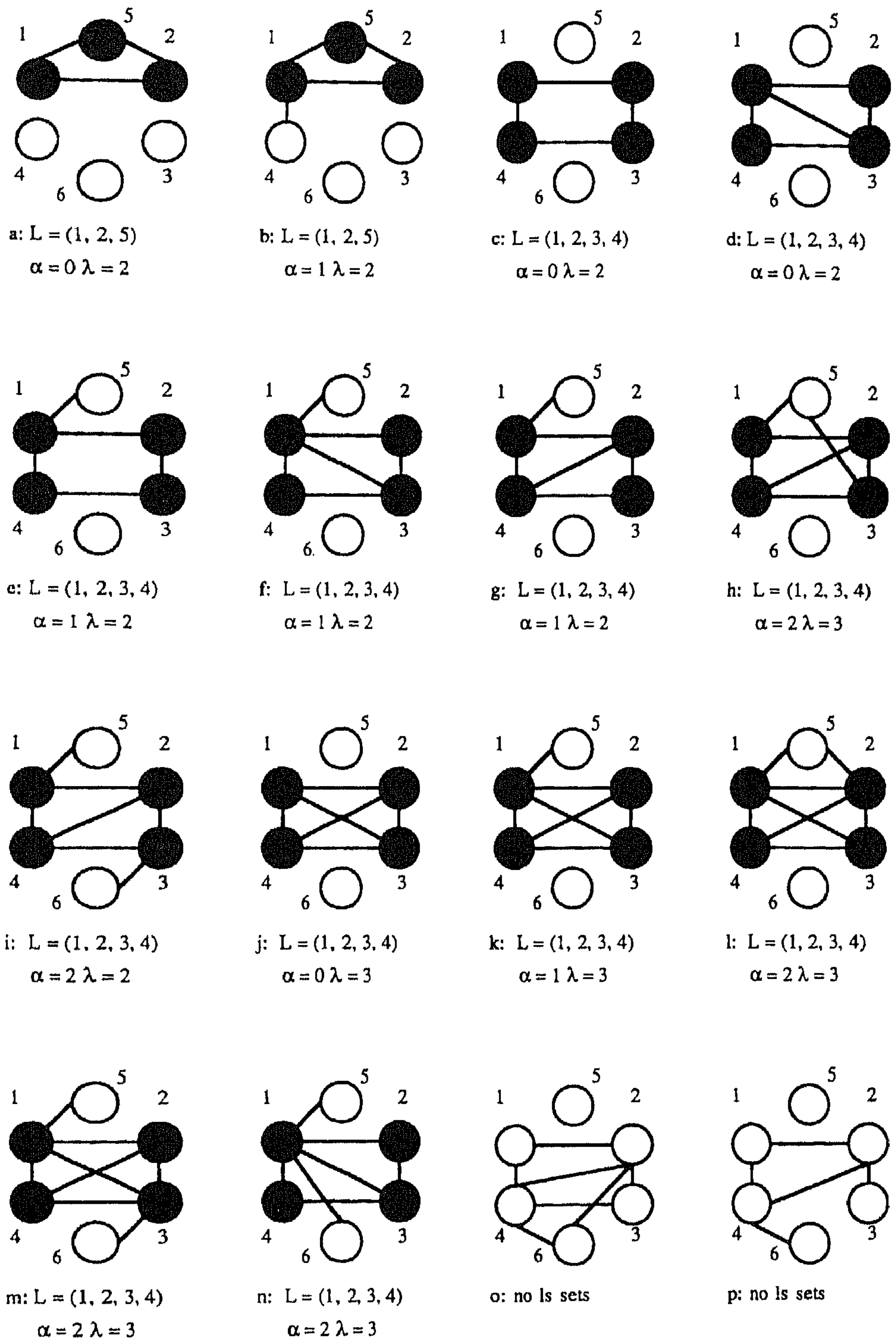



FIGURE 1. All possible LS sets of size $|L| = 3$ and $|L| = 4$ in graphs of size $g \geq 6$.



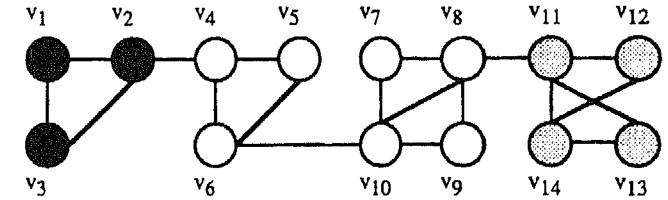



FIGURE 2. Two different LS sets $L = \{v_1, v_2, v_3\}$ and $M = \{v_{11}, v_{12}, v_{13}, v_{14}\}$.

TABLE 1
All Proper Subsets of LS Set $M = \{v_{11}, v_{12}, v_{13}, v_{14}\}$ in Figure 2 and Corresponding Values of $\alpha$

| Subset $K$ | $M - K$ | $\alpha(K, M - K)$ | $\alpha(K, V - M)$ |
|---|---|---|---|
| {11} | {12 13 14} | 3 | 1 |
| {12} | {11 13 14} | 2 | 0 |
| {13} | {11 12 14} | 2 | 0 |
| {14} | {11 12 13} | 3 | 0 |
| {11 12} | {13 14} | 3 | 1 |
| {11 13} | {12 14} | 3 | 1 |
| {11 14} | {12 13} | 4 | 1 |
| {12 13} | {11 14} | 4 | 0 |
| {12 14} | {11 13} | 3 | 0 |
| {13 14} | {11 12} | 3 | 0 |
| {11 12 13} | {14} | 3 | 1 |
| {11 12 14} | {13} | 2 | 1 |
| {11 13 14} | {12} | 2 | 1 |
| {12 13 14} | {11} | 3 | 0 |

## 5. MODEL OF GROUP FORMATION

Individuals can be seen as goal directed in pursuing friendships (Miell and Duck, 1986). Therefore, behavioral rules of individuals are based on *tension minimization* with respect to so-called issues. An issue being any kind of dimension with respect to friendships one has an opinion about, and one thinks is changeable by one's own actions (Hoede, 1990). Depending on the number and kind of issues, the interrelatedness between these issues and the set of allowable individual actions, different models of individual behavior can be specified. The general behavior is concerned with establishing friendships. An individual's state with regard to the presence and configuration of friendships is summarized in his tension. Every individual aims to achieve a tension 0 (ideal states on all issues), and consequently always tries to *reduce* his *tension* with respect to the issues.

DEFINITION 5 *Let $z$ be the number of issues, let $\Delta_{ij}(t)$ be the ith actor's tension with respect to the jth issue at time $t$, and let $w_{ij}$ be the importance of the jth issue to the ith actor. Then the general form of the multidimensional tension function for the ith actor at time $t$ is given by:*

$$\Delta_i(t) = \sum_{j=1}^{z} w_{ij} \Delta_{ij}(t). \tag{1}$$

*where $w_{ij}$ of each issue and the $\Delta_{ij}$ functions are chosen to be nonnegative. Usually tension $\Delta_{ij}$ is given by some function of the difference between the evaluation of an ideal (preferred) state and the evaluation of the current state, according to actor i, on dimension or issue j.*

This attempt to reduce tension motivates the individual's behavior. How he attempts this depends on the following three aspects (see also Snijders' approach (1996)).

1. The *'capability of imagination'* refers to several elements. The first element is that he applies so-called myopic behavioral rules, meaning that he can only imagine what happens at the next point in time as a result of his own actions in attempting to minimize tension. This capability of imagination is, however, further limited. The actual result of his own actions depends on the unknown actions of the other individuals. The individual therefore assumes that the actions of the others are such that his own actions will lead to minimally attainable tension values.

2. The *set of allowable actions* consists in the first place of *'extending'* and *withdrawing'* friendship choices towards other individuals. In principle, only *positive choices*, representing the willingness to establish a friendship with the individual to whom the choice is directed, are considered. An individual may make as many choices as he would like to make in accordance with the issues that are relevant. An individual is allowed to send *negative* messages to represent that he does not want to be friends with an individual that chooses him. This does not necessarily mean that he has a negative attitude towards that individual, but represents that he is not open to form another friendship.[9]

3. The *amount of information* of the individual is limited. We assume that every individual at least perceives all positive choices and negative messages that are directed towards him. He also knows the total number of individuals in the population and is able to observe whether he is a group member or not. If so, he knows who are group members and who their friends are.

The 'group' model is based on two issues. The first issue is the only issue from the basic model (Zeggelink, 1993) and represents every individual's need for social contact (the number of friends):

$$\Delta_{i1}(t) = |df_i - f_i(t)|, \tag{2}$$

where $df_i \leq g - 1$ is $i$'s desired number of friends and $f_i(t)$ is $i$'s actual number of friends at time $t$. The individuals meet as strangers, thus $f_i(0) = 0$ for all $i$ and

[9]One extra element needs introduction here. Individuals do not wait infinitely long for a reciprocated choice from someone they want to be friends with, i.e., asymmetric choices tend to be withdrawn. This behavior can be modelled by giving all individuals a so-called 'waiting period': the maximum amount of time that they will not withdraw an unreciprocated choice. To avoid the problem of interpersonal comparisons of waiting periods, we introduce a so-called *'waiting equilibrium'*. This is the situation (configuration of states) that all individuals wait for reactions of other individuals in the form of reciprocated friendship choices or negative messages, but no individual does make such a choice or send such a negative message because he has no impetus to do so. We assume that the minimal waiting period of all individuals is larger than the time the process needs to reach this waiting equilibrium. To keep the process running, one randomly chosen individual (the most impatient), will withdraw (randomly one of) his unreciprocated choice(s). We perceive the act of exceeding individual $i$'s waiting period by individual $j$ as similar to the act of individual $j$ sending a negative to message to $i$. Both represent no urgent desire of $j$ to become friends with $i$.



$\Delta_{i1}(0) = df_i$. Since tension increases if $f_i(t)$ increases beyond $df_i$ and since an individual himself can always decide whether he wants to be friends with another individual or not, it can be assumed that $f_i(t) \leq df_i$:

$$\Delta_{i1}(t) = df_i - f_i(t), \qquad 0 \leq f_i(t) \leq df_i. \tag{3}$$

The specification of tension with respect to the second issue, the group element, is more complicated. We assume that groups arise by accident. This is not too crude an assumption because in real friendship formations, individuals cannot observe when a group can readily be formed. It is only as a side effect of having friends in common that groups get a chance to develop. Once an individual is a group member, he tries to guarantee the future of that group but also takes care of his non-group goal (his number of friends). As such, the model remains to be an individual oriented model in which the structure of the friendship relations within one's group may become relevant in the 'calculation' of one's individual behavior, but in which groups themselves do not have explicit preferences, and thereby do not show any particular behavior. However, the fact that every individual group member tries to assure that the group structures 'strengthens' may be considered as some form of striving for group interest.

Moreover, we assume that non-group members do not explicitly seek group membership. Such individuals might enter a group, but the proposal to the group is always by accident. Later we will show why this assumption does not make a large difference for the evolution of the friendship networks as it follows from the model as specified hereafter.

We make the following distinction. Let $L_i(t)$ be the smallest, if possible, non-trivial LS set to which $i$ belongs at time $t$ (an individual cannot belong to two or more disjoint LS sets simultaneously). If a non-trivial LS set exists, $i$ is called a group member. If no non-trivial LS set exists, $L_i(t)$ is defined as the trivial singleton set containing just $i$, and $i$ is a non-group member. Let $\lambda(L_i(t))$ be the edge connectivity of $L_i(t)$, and let $\alpha(L_i(t))$ be the number of its external edges. For non-trivial LS sets, the minimum group size $|L_i(t)|_{\min}$ is 3, because $\lambda(L_i(t))$ is assumed to be larger than 1 (see comments following Definition 4). The maximum nontrivial group size $|L_i(t)|_{\max}$ is $g - 1$. If no confusion can arise, we will omit the subscript $i$ and the arguments $t$ and $L_i(t)$.

If an individual is a group member, his group sense is stronger the smaller his and the group's total number of external friendships, and the larger his and the group's total number of internal friendships. The simplest definition of tension is concerned with the total number of external friendships of all group members $\alpha(L_i(t))$, and attributes this value to every single group member. Similarly, $\lambda(L_i(t))$ is an overall representative of tension with regard to the internal structure.[10] Tension should then increase with increasing $\alpha$ or decreasing $\lambda$. Figure 1 clarifies this derivation: the larger $\lambda$, the more outstanding the density of structure within the LS set. The smaller

[10]Thus, for the sake of simplicity we assume that each group member has the same tension with respect to the group element, regardless of his position in the friendship group. Since every individual in the group also has an individual specific component in his tension function, the behavior for group members does not have to be the same.

$\alpha$, the more outstanding the distinctive appearance of the LS set. However, LS sets of different size can have equal values of $\alpha$ and $\lambda$. If $\alpha(L) = \alpha(M)$ and $|L| > |M|$, $L$ might be considered as the group with less tension because for the smaller group $M$ chances are higher that the group will no longer be an LS set if the same extra number of external friendships is established than for the larger group $L$.

For equal $\lambda$, it is more difficult to compare different group sizes because edge connectivity is not defined only with respect to friendships within the group. Since the precise functional shape of the tension component is not so relevant for the derivation of behavioral rules, the only specified properties of the tension component are that it increases with increasing $\alpha$, with decreasing $\lambda$, and with decreasing $|L|$. Moreover, the total tension should be larger for nongroup members than for group members. Therefore, we define

$$\Delta_{i2}(t) = (g-1)^2 - \frac{\lambda(L_i(t))|L_i(t)-1|}{\alpha(L_i(t))+1}, \tag{4}$$

where $L_i(t)$ is the smallest, if possible non-trivial, LS set to which $i$ belongs. Groups arise by chance, and not until then, as $|L_i(t)| > 1$, is this component relevant for group members. The value $(g-1)^2$ assures non-negative tension values and can be considered as the desired value (in combination with $\alpha = 0$) for the group dimension. Tension values do not have to differ for different values of density within LS sets because $\alpha$, $\lambda$, and $|L|$ might be equal. Compare for example the LS set in Figure 1(e) with that in Figure 1(f). Despite these restrictions, for a first attempt to include group formation in a simple manner, (4) is an appropriate definition of tension.

Now, let $w_{i1}$ and $w_{i2}$ be the importances for individual $i$ of the number of friends and the group, respectively. Since we do not compare tensions between individuals, we can assume that $w_{i1} \equiv 1$. For simplicity we assume $\gamma = w_{i1}/w_{i2}$ ($\gamma > 0$) to be equal for all individuals. Accordingly, the tension function becomes:

$$\Delta_i(t) = df_i - f_i(t) + \gamma\left((g-1)^2 - \frac{\lambda(L_i(t))|L_i(t)-1|}{\alpha(L_i(t))+1}\right). \tag{5}$$

The behavior that follows from the group model is different for individuals who do and individuals who do not belong to groups ($L_i(t) = \{i\}$). The general idea is that tension will reduce if a friendship is established. For group members it is reduced more if $\lambda$ or $|L|$ increases or $\alpha$ decreases. The behavior of non-group members is therefore equal to the behavior that would follow from the basic model, in which only the first issue is relevant. The derivation of behavioral rules for group members is difficult because it is not straightforward how $\lambda$, $\alpha$ or $|L|$ will be changed as a result of one's own actions. These values heavily depend on actions of the other individuals. Therefore, we use only the aspect that *group members* attempt to achieve smaller values of $\alpha$ and larger values of $\lambda$ in the process of establishing the desired number of friends. We assume that for every individual the number of friends is so important ($\gamma$ is so small) that he will never establish more friendships than the desired number of friends.

The general situation of an individual $i$ now is as follows. At a certain moment in time, he has a number of friends (reciprocated friendship choices, $f_i(t)$), makes a



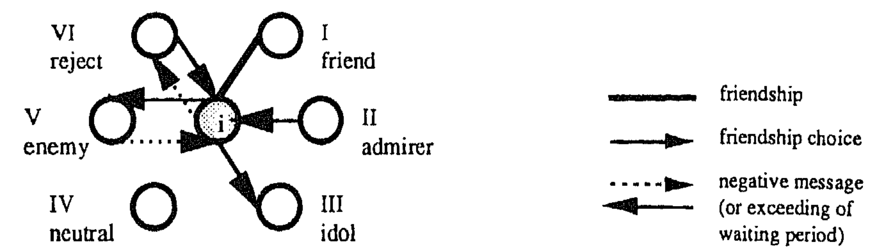



FIGURE 3. Classification of all other individuals according to $i$.

number of unreciprocated friendship choices, receives a number of unreciprocated friendship choices, and receives a number of negative messages. Accordingly, $i$ divides the other individuals $j$ into six mutually exclusive classes. Except for the class 'friends', all other class names are only for convenience when referring to them. They do not have any relevant sociological meaning. The division is illustrated in Figure 3. A mutual friendship choice is represented by a solid line, an unreciprocated friendship choice is represented by a thin arrow, and a negative message (or an exceeding of waiting period) is represented by a broken arrow. Since the latter can occur only in the presence of a positive choice in the opposite direction, this positive choice is presented too.

Individual $i$ 'ego' divides all other individuals $j$ into six classes:

I *friends* $= \{j \mid j$ sends a positive choice towards $i$ and receives a positive choice from $i\}$

II *admirers* $= \{j \mid j$ sends a positive choice to $i$ and $i$ does not reciprocate (yet)$\}$

III *idols* $= \{j \mid j$ receives a positive choice from $i$ and $j$ does not reciprocate (yet)$\}$

IV *neutrals* $= \{j \mid j$ does not choose $i$, and $i$ does not choose $j\}$

V *enemies* $= \{j \mid j$ has sent/sends a negative message to $i$ (or has waited too long to react on $i$)$\}$

VI *rejects* $= \{j \mid j$ has received/receives a negative message from $i$ (or $i$ exceeded $j$'s waiting period)$\}$.

For all $i$, $\Delta_i(0) = df_i$. Depending on $i$'s tension value and the configuration of these classes, $i$ will undertake action to reduce tension as much as possible. Given the capability of imagination, the set of allowable actions, and the amount of information, $i$ assumes that all other individuals behave 'optimally' in the view of his own tension reduction such that when he behaves 'optimally', his tension will maximally reduce in the next step. Consequently his total number of friendship choices (including the reciprocated ones) is smaller than $df_i$, he will add choices. In one step, more than one choice may be added.

- Rejects and enemies are not potential candidates for friendship. The class of rejects is empty until $i$ has the desired number of friends or once he has been the least patient individual in the waiting equilibrium. In the former case tension is zero, in the latter case, rejects are not considered as candidates because they once withdrew a choice towards $i$ or because $i$ waited too long to react.[11] Enemies have shown, in one way or the other, not to be willing to establish a friendship. Choices to enemies are always withdrawn and never remade because they reacted negatively or waited too long to react to a friendship attempt.

- The attempt to maximally minimize tension thus leads to a preference ordering of potential friends in the remaining classes II, III, and IV. For a *non-group member*, tension is reduced with certainty (and maximally) if he chooses an admirer.[12] At first sight, it seems to make sense to replace existing choices to idols with choices to admirers in order to obtain a reciprocated choice with certainty. However, if every individual does so, this effect is lost because in the next step an admirer may have become a neutral. Thus, this preference for admirers holds only when new choices have to be added. Next in the preference order are neutrals. Choices towards idols and friends already exist. Consequently, as a result of the limited availability of individuals in classes II and IV, at a certain moment in time, the number of choices may be smaller than $df_i$. If the difference between actual and desired number of choices is larger than the number of individuals in the category that the individual wants to choose from, he chooses randomly from this category.
- Both group- and non-group members, will never withdraw reciprocated choices because tension would increase. Consequently, once established, friendships are maintained. If $f_i(t) = df_i$, $i$'s tension is minimal, and he has no impetus to act. However, other individuals observe that $i$ does not need any more friendships, therefore it is justified to assume that $i$ will send a negative message to those individuals that still try to initiate a friendship with him. In other words, individuals who send a superfluous friendship choice will be rebuffed.
- In a similar way, if $i$ receives a negative message from $j$, he knows that his tension will never be reduced by keeping his choice extended to $j$. He withdraws the choice extended to $j$ and places $j$ in class $V$.
- If $i$ is a group member, he makes an extra subdivision within the classes of potential friends. They are distinguished into *group members*, *indirect external friends* and *non-group members*. Indirect external friends are individuals who are not $i$'s friends but friends with one or more of his fellow group members. Those are the individuals that determine the value of $\alpha$. Within class II an individual distinguishes admirers that are group members (IIa), admirers that are indirect external friends (IIb), and admirers that are non-group members (IIc). A same distinction is made in classes III and IV. Possible candidates for friendship are in classes II and IV. A *group member* also takes into account that $\lambda$ might increase when he establishes a friendship with a group member who was not his friend yet. This leads to a preference for individuals in IIa over individuals in IIb and IIc. Analogously, individuals in IVa are preferred over those in IVb and IVc.
- In contrast to behavior of non-group members, a group member $i$ will also *replace* choices to idols according to the above preference order, i.e. choices towards idols in classes IIIb or IIIc will, if possible be replaced by choices towards individuals in IIa or IVa, and in this preference order (IIa > IVa > IIIb, IIIc). For the purpose of increasing $\lambda$, we also assume $i$ prefers individuals in class IVa over those in IIb and IIc (IVa > IIb, IIc).
- Another way to reduce tension for a group member is by reducing $\alpha$. If $i$ himself would have an external friendship, he could dissolve this friendship (withdraw the friendship choice) and try to establish a new one with a group member. Since this seems a rather rude way of behaving, and to maintain the largest analogy with the basic model, $i$ does not withdraw friendship choices to friends. This limits the way to reduce $\alpha$. $\alpha$ might be reduced (not with certainty) with a similar increase in group size $|L|$ by the establishment of a friendship with an indirect external friend ($b$-classes). So, these individuals are next in the preference order, and a group member prefers them in order of the number of group members they are friends with. The higher this number, the larger the probability that $\alpha$ will decrease with a simultaneous increase in group size. Again, a group member will also replace choices towards individuals in class IIIc to indirect external friends (IIb > IVb > IIIc).
- For the purpose of decreasing $\alpha$, we assume that individuals in class IVb are preferred over those in IIc.
- Summarized, the following preference order is obtained when $i$ is a group member: IIIa > IIa > IVa > IIIb > IIb > IVb > IIIc > IIc > IVc.

[11]This is an assumption that may not mimic reality but suffices as a first rule of behavior.

[12]This preference is justified from the fact that people are attracted to those who like them.



In actual friendship formations, non-group members probably attempt to become group members by making friendship choices towards group members. The specification of behavior for group members shows that such non-group members will have very small probabilities of becoming friends with group members if they are not already linked to another group member.

In the implementation of the model, and in its illustration (Figure 4), different actions occur in different steps. Adding and removing (one or more) positive choices take place in one step. In another step, negative messages are sent. This distinction in alternating steps is not meant to have any correspondence with reality.

In Figure 4 we present a possible network development for this model. We consider a set of $g = 8$ individuals who all want to establish 3 friendships. The description of the process is as follows:

$t = 0$: The initial situation where the tension of every individual is maximal.

$t = 1$: Every individual randomly makes as many choices as his desired number of friends. 8 establishes the desired number of friendships. Other individuals also establish friendships. Consequently, 1, 2, 3 and 8 constitute an LS set with $\lambda = 2$ and $\alpha = 1$ (dotted vertices).

$t = 2$: 8 sends a negative message to 5 because his incoming choice is superfluous.

$t = 3$: 5 removes the corresponding choice and makes a new choice towards his admirer 3. 1, 2, and 3 observe that they are group members and replace choices to non-group members (IIIc) by choices to group members (IVa). As a result, 1 and 3 have the desired number of friends. Notice that 5 preferred 3 because 3 was an admirer in the previous time step. At the same time however, 3 removes this choice because he prefers a group member above 5.

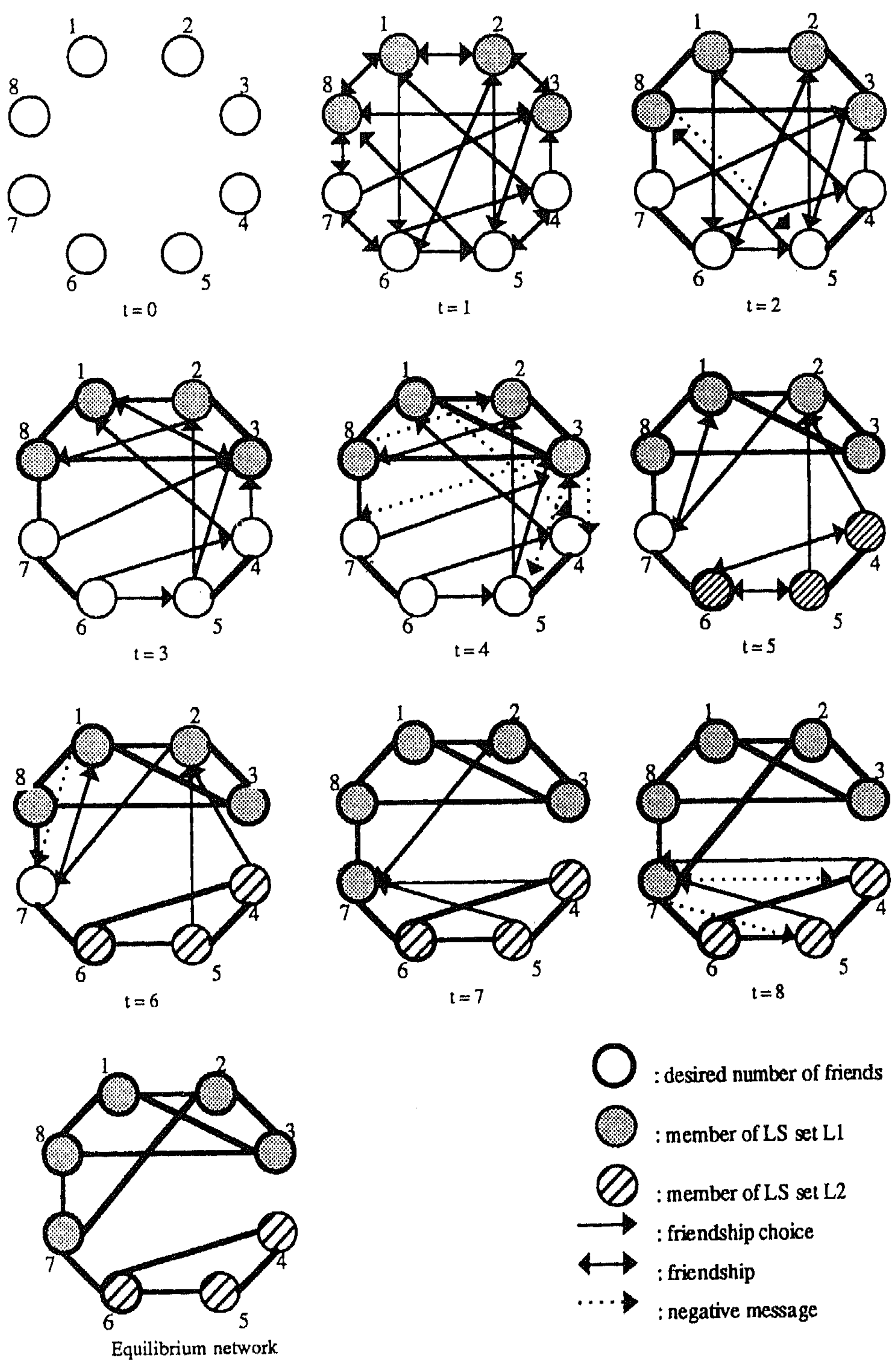



FIGURE 4. Evolution of a friendship network.



$t = 4$: 8 sends a negative message to 2 (although 2 is a group member, 8's number of friends is more important). 3 sends negative messages to 4, 5, and 7. 1 sends a negative message to 4.

$t = 5$: 4, 5, and 7 remove the corresponding choices. 7 has no admirers and randomly chooses 1. 4 and 5 both received an incoming choice from 6 and both choose him. As a result, an LS set (striped vertices) emerges containing 4, 5, and 6 with $\lambda = 2$ and $\alpha = 1$. 2 could not choose any more group members and moved to 7 because he is an indirect external friend (IVb).

$t = 6$: 1 sends a negative message to 7.

$t = 7$: 7 removes his choice to 1 and prefers admirer 2. 7 and 2 have the desired number of friends. A larger LS set emerges including individual 7 also: $\lambda$ and $\alpha$ do not change. 4 and 5 observe that they are group members, they cannot replace choices to non-group members by choices to group members. Thus, they replace them by choices to the indirect external friend 7 (IVb).

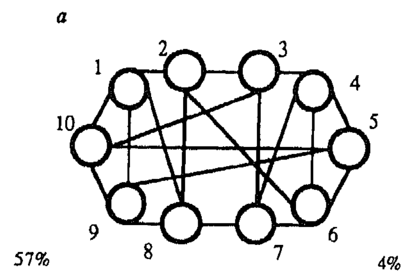
$t = 8$: 7 sends negative messages to individual 4 and 5. 4 and 5 are the only individuals who have tension with respect to the number of friends, each attempts to establish friendships with others who have not rejected them yet. However, all alternative individuals have the desired number of friends, send negative messages and the network reaches equilibrium. This network consists of two LS sets: $L = \{1,2,3,7,8\}$, $\alpha(L) = 1$, $\lambda(L) = 2$, and $M = \{4,5,6\}$, $\alpha(M) = 1$, $\lambda(M) = 2$.

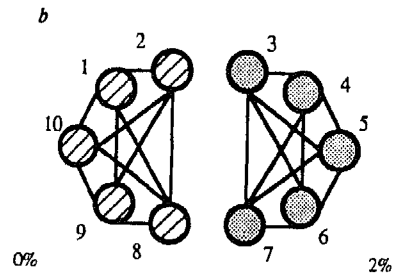
## 6. RESULTS

In order to extract the most important tendencies that arise from the group issue, we consider only population sizes $g = 10$ and $g = 15$, $df_i = 4$ or $df_i = 6$ for all $i$. With all initial situations we run 100 simulations. We investigate the emergence and presence of LS sets in equilibrium in particular. If LS sets are present, we are interested in their number and characteristics. Other results concerning the overall structure of the friendship network and effects of mean desired number of friends and population size are given in Zeggelink (1993).

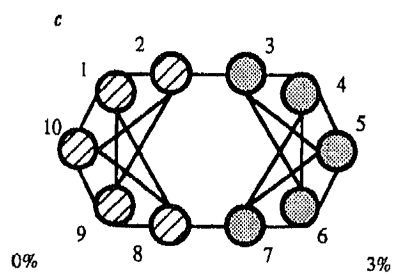
Possible equilibrium networks for $g = 10$ and $d = 4$ are presented in Figure 5(a) to 5(m). A summarized overview of these figures is presented in Table 2. For the sake of comparison we also present results for the basic model, when the group issue is irrelevant (as if every individual would behave like a non-group member). In the first column we present the figure number of the network. The second column contains the mean established number of friends in equilibrium. In the third column the number of LS sets in the network is presented. The fourth column presents characteristics of all LS sets that are present in the network. Columns five and six show the percentage of occurrences in the basic model and the group model respectively.

The number of LS sets in equilibrium varies between 0 and 2. Sizes of LS sets vary between 3 and 9. Values of $\lambda$ vary between 2 and 4, values of $\alpha$ vary between 0 and 2. A closer look at Figures 5(a) to 5(m) sugests that the small values of mean number of friends in equilibrium resulting from the group model (in comparison with the values for the basic model) are caused by the fact that individuals start to choose group members once they are LS set members. Consequently, non-group

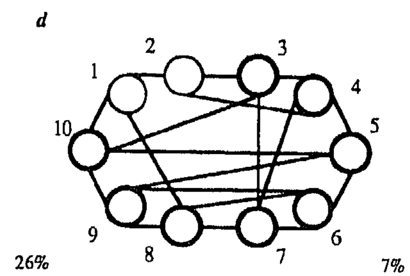


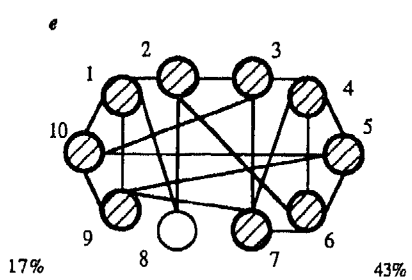




L = {1, 2, 8, 9, 10} $\alpha = 0$ $\lambda = 4$
M = {3, 4, 5, 6, 7} $\alpha = 0$ $\lambda = 4$

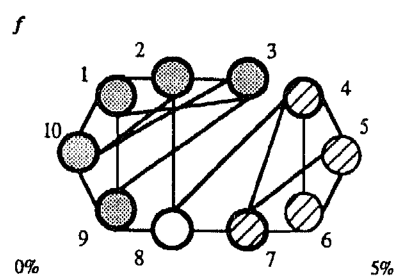



L = {1, 2, 3, 4, 5} $\alpha = 0$ $\lambda = 4$
M = {6, 7, 8, 9, 10} $\alpha = 0$ $\lambda = 4$




L = {1, 2, 3, 4, 5, 6, 7, 9, 10} $\alpha = 2$ $\lambda = 4$




L = {1, 2, 3, 9, 10} $\alpha = 2$ $\lambda = 4$
M = {4, 5, 6, 7} $\alpha = 2$ $\lambda = 3$

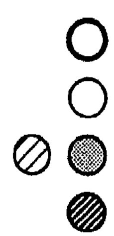
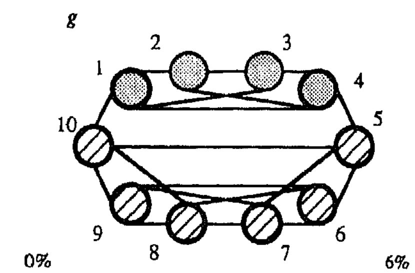
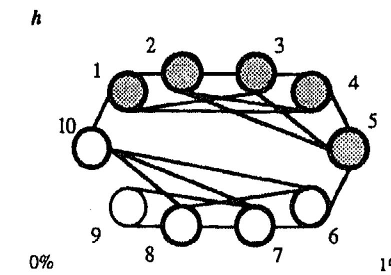



FIGURE 5. LS sets in equilibrium for $g = 10$. If LS sets are present, we present their members and their values of $\alpha$, and $\lambda$. We also present the percentages of occurrence in the basic model (left) and the group model (right).

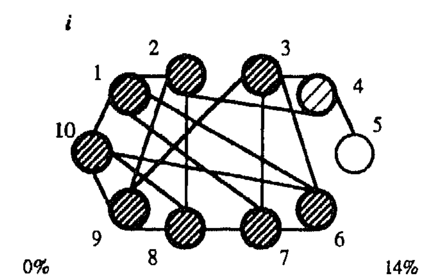




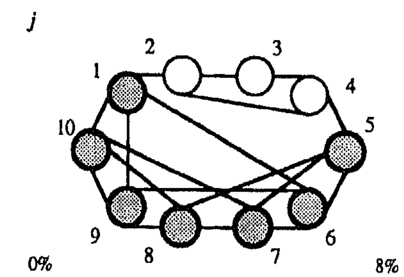

L = {5, 6, 7, 8, 9, 10} $\alpha = 2$ $\lambda = 4$
M = {1, 2, 3, 4} $\alpha = 2$ $\lambda = 3$



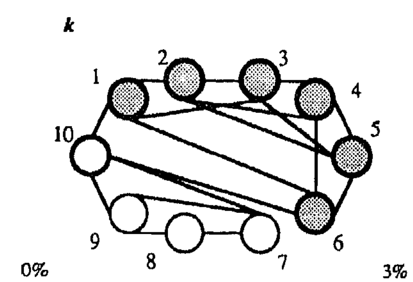

L = {1, 2, 3, 4, 5} $\alpha = 2$ $\lambda = 4$



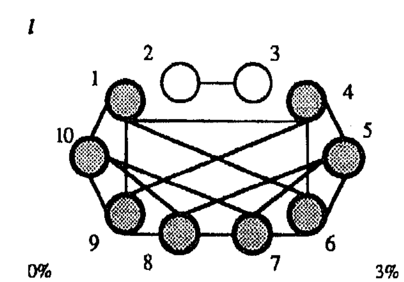

L = {1, 2, 3, 6, 7, 8, 9, 10} $\alpha = 2$ $\lambda = 4$
M = {1, 2, 3, 4, 6, 7, 8, 9, 10} $\alpha = 1$ $\lambda = 2$



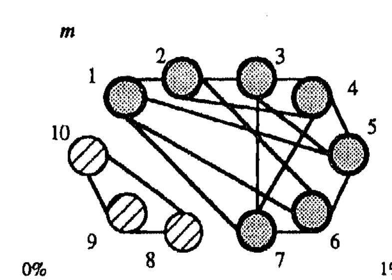

L = {1, 5, 6, 7, 8, 9, 10} $\alpha = 2$ $\lambda = 4$




L = {1, 2, 3, 4, 5, 6} $\alpha = 2$ $\lambda = 4$




L = {1, 4, 5, 6, 7, 8, 9, 10} $\alpha = 0$ $\lambda = 4$




L = {1, 2, 3, 4, 5, 6, 7} $\alpha = 0$ $\lambda = 4$
M = {8, 9, 10} $\alpha = 0$ $\lambda = 2$

FIGURE 5. *(Continued.)*

TABLE 2
Characteristics of Networks in Terms of LS Sets for Basic Model and Group Model
Population Size $g$ = 10 (100 runs)

| Fig. 5 | Mean # Friends in Equilibrium | # LS Sets | Characteristics LS Sets Size $\|L\|$, Number of External Friendships $\alpha$, Edge Connectivity $\lambda$ | % of Occurrences in Basic Model | % of Occurrences in Group Model |
|---|---|---|---|---|---|
| *a* | **4.0** | 0 | | 57 | 4 |
| *b* | 4.0 | 2 | $\|L\| = 5 \quad \alpha = 0 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 5 \quad \alpha = 0 \quad \lambda = 4$ | | |
| *c* | 4.0 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 3 |
| | | | $\|M\| = 5 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *d* | **3.8** | 0 | | 26 | 7 |
| *e* | 3.8 | 1 | $\|L\| = 9 \quad \alpha = 2 \quad \lambda = 4$ | 17 | 43 |
| *f* | 3.8 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 5 |
| | | | $\|M\| = 4 \quad \alpha = 2 \quad \lambda = 3$ | | |
| *g* | 3.8 | 2 | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 6 |
| | | | $\|M\| = 4 \quad \alpha = 2 \quad \lambda = 3$ | | |
| *h* | 3.8 | 1 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 1 |
| *i* | **3.6** | 2* | $\|L\| = 8 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 14 |
| | | | $\|M\| = 9 \quad \alpha = 1 \quad \lambda = 2$ | | |
| *j* | 3.6 | 1 | $\|L\| = 7 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 8 |
| *k* | 3.6 | 1 | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 3 |
| *l* | **3.4** | 1 | $\|L\| = 8 \quad \alpha = 0 \quad \lambda = 4$ | 0 | 3 |
| *m* | 3.4 | 2 | $\|L\| = 7 \quad \alpha = 0 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 3 \quad \alpha = 0 \quad \lambda = 2$ | | |
| | | | | 100 | 100 |

*The smaller LS set is a subset of the larger LS set, there is one individual that does not belong to the smaller group, but does belong to the larger group.

members have more difficulties finding the desired number of friends. The *individuals who do not have the desired number of friends, more often do not belong to a group than individuals who have the desired number of friends.* Nevertheless, individuals having less friends than desired *can* belong to LS sets (see Figures 5(f) and 5(g)). Moreover, when LS sets are present, usually all individuals with the desired number of friends belong to LS sets (exceptions in Figures 5(f), 5(h), 5(k)). If they do not belong to LS sets, they *usually occupy a bridge function between groups and non-group members, or simply between groups.*

The equilibrium structures of the group model do not self-evidently contain LS sets. In 89% of the runs, LS sets appear in equilibrium (in contrast to 17% for the basic model). However, many of these LS sets are relatively 'large'. In 60% of the simulations, LS sets of size $|L| \geq 8$ appear. This result that 1 or 2 individuals do not belong to an LS set occurs rarely with the basic model. The question is whether these 8 or 9 individuals can be considered to form a group. Omitting these equilibrium structures, only 29% of the runs remains to be investigated. This low percentage can be explained by several reasons. First, *the probability that a (small) LS set emerges based on the initial random choices of the individuals is small.* Second, *once emerged LS sets may disappear during the process because for every individual, the number of friends is more important than the age issue. As a consequence, more external friendships can be established than 'allowable' for the LS set to remain an LS set.*



TABLE 3
Characteristics of Networks in Terms of LS Sets for Basic Model and Group Model
Population Size $g$ = 15 (100 runs)

| Ref | Mean # Friends in Equilibrium | # LS Sets | Characteristics LS Sets Size $\|L\|$, Number of External Friendships $\alpha$, Edge Connectivity $\lambda$ | % of Occurrences in Basic Model | % of Occurrences in Group Model |
|---|---|---|---|---|---|
| *a* | **4.0** | 0 | | 68 | 4 |
| *b* | 4.0 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 9 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *c* | 4.0 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 11 |
| | | | $\|M\| = 10 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *d* | 4.0 | 2 | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 8 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *e* | 4.0 | 2 | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 9 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *f* | 4.0 | 2 | $\|L\| = 7 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 7 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *g* | 4.0 | 2 | $\|L\| = 7 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 2 |
| | | | $\|M\| = 8 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *h* | **3.8** | 0 | | 18 | 2 |
| *i* | 3.8 | 1 | $\|L\| = 14 \quad \alpha = 2 \quad \lambda = 4$ | 13 | 41 |
| *j* | 3.8 | 1 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 1 |
| *k* | 3.8 | 1 | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 1 |
| *l* | 3.8 | 2 | $\|L\| = 4 \quad \alpha = 2 \quad \lambda = 3;$ | 1 | 3 |
| | | | $\|M\| = 11 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *m* | 3.8 | 3† | $\|L\| = 6 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 7 \quad \alpha = 0 \quad \lambda = 2;$ | | |
| | | | $\|N\| = 8 \quad \alpha = 0 \quad \lambda = 4$ | | |
| *n* | 3.8 | 3† | $\|L\| = 5 \quad \alpha = 0 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 9 \quad \alpha = 2 \quad \lambda = 4;$ | | |
| | | | $\|N\| = 10 \quad \alpha = 0 \quad \lambda = 2$ | | |
| *o* | **3.6** | 1 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 1 |
| *p* | 3.6 | 1 | $\|L\| = 11 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 1 |
| *q* | 3.6 | 1 | $\|L\| = 12 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 6 |
| *r* | 3.6 | 1 | $\|L\| = 13 \quad \alpha = 2 \quad \lambda = 4$ | 0 | 2 |
| *s* | 3.6 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 7 \quad \alpha = 2 \quad \lambda = 4$ | | |
| *t* | 3.6 | 2 | $\|L\| = 5 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 14 \quad \alpha = 1 \quad \lambda = 2$ | | |
| *u* | 3.6 | 2 | $\|L\| = 13 \quad \alpha = 2 \quad \lambda = 4;$ | 0 | 11 |
| | | | $\|M\| = 14 \quad \alpha = 1 \quad \lambda = 2$ | | |
| *v* | 3.6 | 3† | $\|L\| = 5 \quad \alpha = 0 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 9 \quad \alpha = 2 \quad \lambda = 4;$ | | |
| | | | $\|N\| = 10 \quad \alpha = 0 \quad \lambda = 2$ | | |
| *w* | 3.6 | 3† | $\|L\| = 5 \quad \alpha = 0 \quad \lambda = 4;$ | 0 | 1 |
| | | | $\|M\| = 8 \quad \alpha = 2 \quad \lambda = 4;$ | | |
| | | | $\|N\| = 9 \quad \alpha = 0 \quad \lambda = 2$ | | |
| | | | | 100 | 100 |

†The smaller LS set is a subset of the larger LS set, there is one individual that does not belong to the smaller group, but does belong to the larger group.

This also explains why $\lambda = 4$ for most larger LS sets: every individual seeks to establish 4 friendships.

Results for $g$ = 15 and $d$ = 4 are summarized in Table 3 (similar to Table 2). For $g$ = 15, LS sets appear more often than for $g$ = 10. Omitting LS sets that contain smaller LS sets $(m,n,v,w)$, the number of groups varies between 0 and 2. Sizes of LS sets vary between 4 and 14. Values of $\lambda$ vary between 2 and 4, values of $\alpha$ vary between 0 and 2.

For the basic model only 14% of all equilibrium networks contains LS sets. If they are present, they are in almost all cases trivial ($|L| = 14$). In the group model, LS sets emerge in 94% of all equilibrium networks. For $g = 10$, we did not consider LS sets for which $|L| \geq 0.8g$ For $g = 15$, we therefore do not consider LS sets with $|L| \geq 12$, and 34% of the equilibrium networks remains (29% for $g = 10$). Since equilibrium networks hardly emerged with LS sets of size $8 \leq |L| \leq 12$, this percentage of 34% is almost equal to the case when we omit equilibrium structures with LS sets only of size $|L| \geq 8$. Apparently, the probability that small LS sets exist in equilibrium for $d = 4$ does not significantly differ for $g = 10$ and $g = 15$. Intuitively one could argue that the probability that a 'small' LS set emerges by accident is smaller in a larger population. However, once an LS set does emerge, *it can develop further more easily in a larger population because then every individual has more options to choose additional friends without causing his group to fall apart as a result of too many external friendships* ($\alpha$) *by a simultaneous increase in group size.*

For $g = 15$ and $d = 6$, equilibrium networks containing LS sets emerge in 99% of the runs. However, only in 14% of the runs are these LS sets of size $|L| < 0.8$ $g = 12$. Again, the reason is the relatively high desired mean degree.

## 7. DISCUSSION

The presented group model is a preliminary step in the direction of a model that captures subgroup formation. In this model, groups emerge 'accidentally', and once they emerged, members try to guarantee the future of their group with possibly negative consequences for non-group members who might not succeed in establishing the desired number of friends. In some cases, these initially non-group members also succeed in 'establishing a group'.

More extensive discussions of results with this preliminary 'group' model are presented in Zeggelink (1993). There we also compare the predictions of the model with those of the basic model and the 'similarity' model in terms of generally applied structural characteristics that describe friendship networks.

In Zeggelink (1993) and Van dė Bunt and Zeggelink (1993), a limited confrontation of the models with 'real life' data can be found. 'Optimal' testing of the models requires knowledge about a closed set of initially mutual strangers, their need for social contact and the friendship network among these individuals (preferable at consecutive points in time). The number of available data sets that meet these requirements was very limited. The most appropriate data are those collected by Hallinan and colleagues that were used in several publications (among others: Hallinan, 1979; Hallinan and Kubitschek, 1990; Hallinan and Sörensen, 1985; Hallinan and Williams, 1987). The longitudinal data describe friendship reports of children in 11 different classrooms from grades 4 to 7 in the United States. The data were collected at six points in time at six week intervals. Unfortunately, the first point in time is not the moment at which the individuals are mutual strangers. Other important drawbacks of the data are that the children were allowed to mention both best friends and friends, there are no data directly referring to a variable expressing 'need for social contact', and there is a large number of missing cases. Thus, the data are not very suitable. Therefore, at the moment, we are gathering data among first year sociology students at the University of Groningen.



With the data on classrooms, only weak and strong aspects of the models were examined. We tested our models by comparing their predictions about several structural parameters.

No large difference exists between the predictions of the basic model and the group model for these empirical data. Of all three models, the 'similarity' model, not presented here, performed best. The main behavioral rule in this model is that individuals prefer to be friends with those who are similar to them. For the classes considered here, similarity on gender was an important constraint for friendship choice.

The fact that LS sets are rare in empirical data was not too big a problem for these comparisons because we could use other structural characteristics of the friendship networks to assess the strength of predictions. However, more fruitful comparisons between theoretically predicted and empirical networks would be possible if a means could be developed to quantify the degree to which a network structure departs from the presence of LS sets. Or, as Borgatti et al. (1990) suggest, to come up with a measure of the extent to which individuals depart from belonging to an LS set. These kind of developments are of higher concerns once adaptations have been applied to the model presented here. On the basis of simulation results on artificial populations, and confrontations with empirical data, some shortcomings of the present 'group' model were derived. There is too much emphasis on the need for social contact. An individual always attempts to establish the desired number of friends and never establishes more friendships than desired. Another element concerning this number of friends is the impossibility of friendship dissolution. As a result, especially the group aspect does not live up to its promise because the number of friends for an individual is always much more important than his group 'state'. As a result, groups that emerged may disappear again because individuals who lack friends continue looking for friends disregarding the possible collapse of the group.

The need for social contact becomes less important by adapting appropriately the weights of the issues in the tension functions. It may also be more appropriate to define a minimum and a maximum value for the desired number of friends. The group aspect becomes much more interesting when individuals may dissolve friendships to avoid the break up of the group, or to strengthen the group's 'state'.

The accidental arising of groups from common friendships is not too crude an assumption when friendships are established on the basis of similarity. However, in the present model of group formation, no such similarity aspect was taken into account. Groups had to emerge from purely random choices. A joint model of the similarity and group model is promising when the previous adaptations concerning, among other things, the importance of the number of friends relative to the 'characteristics of friends' and 'group', and the possible friendship dissolution, have been elaborated (Zeggelink, 1996).

No distinction of behavior was made among different group members and between group members and non-group members. If different tension values are defined for individuals in different positions in one group, behavior with respect to the group may differ among members of one group. As a result, the inclusion of the aspect that non-group members purposely seek group membership becomes meaningful.

Another aspect concerns the improvement of empirical testing and model developments that result from its conclusions.[13] The possibility of testing whether the possible adaptations of the models improve predictions of network structure depends on the availability of empirical data. In the present models, the need for social contact played a very important role. It is however a concept that is very difficult to operationalize, particularly with secondary data. Since in adapted versions of the model the exact determination of the need for social contact would be less relevant, it may be less of a problem. However, it is preferable to gather data specifically in view of the testing of these models.

## REFERENCES


Alba, R. (1973) A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology* **3:** 113–126.

Alba, R., and Moore, G. (1978) Elite social circles. *Sociological Methods and Research* **7:** 167–187.

Bernard, H., and Killworth, P. D. (1973) On the social structure of an ocean-going research vessel and other important things. *Social Science Research* **2:** 145–184.

Borgatti, S., Everett, M., and Shirey, P. (1990) LS sets, lambda sets and other cohesive subsets. *Social Networks* **12:** 337–357.

Breiger, R. (1974) The duality of persons and groups. *Social Forces* **53:** 181–190.

Cohen, J. (1977) Sources of peer group homogeneity. *Sociology of Education* **50:** 227–241.

Doreian, P., Kapuscinski, R., Krackhardt, D., and Szczypula, J. (1996) A brief history of balance through time. *Journal of Mathematical Sociology* **21:** 113–131.

Fararo, T., and Doreian, P. (1984) Tripartite structural analysis: Generalizing the Breiger-Wilson formalism. *Social Networks* **6:** 141–175.

Fararo, T., and Doreian, P. (1996) The theory of solidarity: An agenda of problems for mathematical sociology. To appear in *Journal of Mathematical Sociology.*

Feld, S. (1981) The focussed organization of social ties. *American Journal of Sociology* **86:** 1015–1035.

Flache, A., and Macy, W. M. (1996) The weakness of strong ties: Collective action failure in a highly cohesive group. *Journal of Mathematical Sociology* **21:** 3–28.

Freeman, L. (1992) The sociological concept of 'group': An empirical test of two models. *American Journal of Sociology* **98:** 152–166.

Freeman, L., Freeman, S., and Michaelson, A. (1988) On human social intelligence. *Journal of Social and Biological Structure* **11:** 415–425.

Freeman, L., Freeman, S., and Michaelson, A. (1989) How humans see social groups: A test of the Sailer–Gaulin models. *Journal of Quantitative Anthropology* **1:** 229–238.

Granovetter, M. (1982) The strength of weak ties: A network theory revisited. In P. Marsden and N. Lin (Eds.), *Social Structure and Network Analysis*, Sage, Beverly Hills, 105–130.

Hallinan, M. (1979) The process of friendship formation. *Social Networks* **1:** 193–210.

Hallinan, M., and Kubitschek, W. (1990) The formation of intransitive triads. *Social Forces* **69:** 505–519.

Hallinan, M., and Sörensen, A. (1985) Class size, ability group size, and student achievement. *American Journal of Education*, 71–89.

Hallinan, M., and Williams, R. (1987) The stability of student's interracial friendships. *American Sociological Review* **42:** 653–664.

Hammer, M. (1979) Predictability of social connections over time. *Social Networks* **2:** 165–180.

Hammer, M. (1980) Social access and the clustering of personal connections. *Social Networks* **2:** 305–325.

Harary, F., Norman, R., and Cartwright, D. (1965) *Structural Models*, Wiley, New York.

Hoede, C. (1990) Social atoms: Kinetics. In J. Weesie and H. Flap (Eds.), *Social Networks Through Time*, ISOR, Utrecht, 45–63.


[13]See Snijders (1996) for a more direct link between theoretical models and empirical data.




Homans, G. (1950) *The Human Group*, Harcourt Brace, New York.

Leenders, R. Th. A. J. (1996) Dynamics of friendship and best friendship choices. *Journal of Mathematical Sociology* **21:** 133–148.

Leenders, R. Th. A. J. (1995) Longitudinal behavior of network structure and actor attributes: Sources and consequences of misspecification. In P. Doreian and F. N. Stokman (Eds.), *The Evolution of Social Networks*, New York: Gordon and Breach.

Luccio, F., and Sami, M. (1969) On the decomposition of networks into minimally interconnected networks. *IEEE Transactions on Circuit Theory* **16:** 184–188.

Luce, R. (1950) Connectivity and generalized cliques in sociometric group structure. *Psychometrika* **15:** 169–190.

Luce, R., and Perry, A. (1949) A method of matrix analysis of group structure. *Psychometrika* **14:** 94–116.

Miell, D., and Duck, S. (1986) Strategies in developing friendships. In V. Derlega and B. Winstead (Eds.), *Friendship and Social Interaction*, New York: Springer Verlag.

Mokken, R. (1979) Cliques, clubs, and clans. *Quality and quantity* **13:** 161–173.

Mullen, B., and Goethals, G. (Eds.) (1987) *Theories of Group Behavior*, New York: Springer Verlag.

Reitz, K. (1988) Social groups in a monastery. *Social Networks* **10:** 343–357.

Ridgeway, C. (1983) *The Dynamics of Small Groups*, New York: Sint Martens Press.

Romney, A., and Faust, K. (1982) Predicting the structure of a communication network from recalled data. *Social Networks* **4:** 285–304.

Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., and Lorensen, W. (1991) *Object Oriented Modeling and Design*, Englewood Cliffs: Prentice Hall.

Sailer, D., and Gaulin, S. (1984) Proximity, sociality and observation: The definition of social groups. *American Anthropologist* **86:** 91–98.

Salzinger, L. (1982) The ties that bind: The effect of clustering on dyadic relationships. *Social Networks* **4:** 117–145.

Seidman, S. (1983) LS sets and cohesive subsets of graphs and hypergraphs. *Social Networks* **5:** 92–96.

Seidman, S., and Foster, B. (1978) A graph theoretic generalization of the clique concept. *Journal of Mathematical Sociology* **6:** 139–154.

Shaw, M. (1983) *Group Dynamics: The Psychology of Small Group Behavior*, New York· McGraw-Hill.

Sherif, M., and Sherif, C. (1964) *Reference Groups: Exploration into Conformity and Deviation of Adolescents*, New York: Harper and Row.

Skvoretz, J., Faust, K., and Fararo, T. (1996) Social structure, networks, and E-state structuralism models. *Journal of Mathematical Sociology* **21:** 57–76.

Snijders, T. A. B. (1996) Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology* **21:** 149–172.

Stokman, F., and Van Oosten, R. (1994) The exchange of voting positions: An object oriented model of policy networks. In B. Bueno de Mesquita and F. N. Stokman (Eds.), *Decision Making in the European Community: Models, Applications, and Comparisons*, New Haven: Yale University Press.

Stokman, F., and Zeggelink, E. (1996) Is politics power or policy oriented? A comparative analysis of dynamic access models in policy networks. *Journal of Mathematical Sociology* **21:** 77–111.

Van de Bunt, G., and Zeggelink, E. (1993) Can we predict structure in friendship networks? A nempirical test of an individual oriented dynamic simulation model. Paper presented at the *Third European Conference on Social Network Analysis*, Munich, Germany, June 10–13.

Wilson, T. (1982) Relational networks: An extension of sociometric concepts. *Social Networks* **4:** 105–116.

Zeggelink, E. (1993) *Strangers Into Friends: The Evolution of Friendship Networks Using an Individual Oriented Approach*, Amsterdam: Thesis Publishers.

Zeggelink, E. (1994) Dynamics of structure: An individual oriented approach. *Social Networks* **16:** 295–333.

Zeggelink, E. (1996) Evolving friendship networks: An individual oriented approach implementing similarity. *Social Networks* **17:** 83–110.

Zeggelink, E. (1996) Group formation in friendship networks. To appear in *Journal of Mathematical Sociology.*