Advantages of hierarchical organization: from pigeon flocks to optimal network structures

Anna Zafeiris
MTA-ELTE Statistical and Biological Physics Research Group, Pázmány Péter stny. 1A, Budapest, Hungary, H-1117

Tamás Vicsek (vicsek@hal.elte.hu)
Department of Biological Physics, Eötvös University - Pázmány Péter stny. 1A, Budapest, Hungary, H-1117
MTA-ELTE Statistical and Biological Physics Research Group, Pázmány Péter stny. 1A, Budapest, Hungary, H-1117

Abstract

Hierarchical organization is abundant in both natural and artificial social structures, among humans just as among animal species. However, despite of this abundance and the vast literature on related fields, quantitative results are rare, and mostly considering two-level hierarchies in which individuals are either "leaders" or "followers". In this paper we overview some of our latest quantitative results, including biological observations and computational models. These studies reveal multi-level hierarchical structures as opposed to the two-level organization. We also review a measure that can serve as a common base for the quantitative determination of the level of hierarchy of complex networks.

Keywords: Hierarchy, multi-level hierarchy, hierarchy measure.

Introduction

Motivation

In a recent study ultra-light GPS devices had been installed onto homing pigeons in order to record their flight trajectories on their way home with high precision (see Fig. 1 *a*). The pairwise velocity correlation analysis, applied on these GPS logs, revealed that during collective decision making (i.e., navigating home as a single group), pigeons choose their common direction of flight as a result of dynamically changing hierarchical leadership-followership interactions (Nagy et al., 2010). The question whether this hierarchy is somehow related to the pecking order (the dominance structure birds known to live in), naturally arises.

In order to get insight into the above problem, the social dominance and the in-flight leader-follower relationships had been compared in a flock of homing pigeons, consisting of ten individuals (Nagy et al., 2013). Surprisingly enough, dominance and leadership hierarchies have been found to be completely independent of each other (See Fig. 1 *B* and *C*). Dominance is known to be correlated with aggression and access to food, based on some individual characteristics such as physical strength, in order to strangulate the violence to a minimum level within the group. At the same time, these results imply that the stable leadership hierarchies arising in the air must be the resultant of a different set of individual competences. But how do these hierarchies emerge? What are the main features optimized by hierarchy: Flow of information? More efficient production? Controllability? Better decision making process? And what are the main signatures of hierarchy? Despite the abundant literature on hierarchy (Celko, 2012; Chisholm, 1992; Saaty, 2012; Kipfer, 2001; Williamson, 1983; Thompson et al., 1991), widely accepted quantitative interpretation of the origin and emergence of multi-level hierarchies do not exist, leaving all the above questions open.

In this paper we attempt to give an overview of some of the first results that try to approach these questions in a quantitative way.

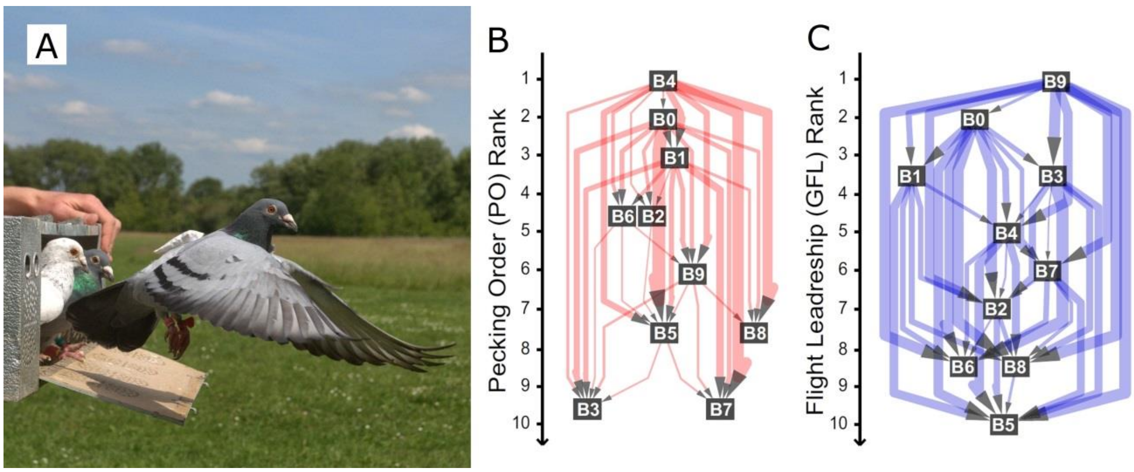

Figure 1 – ***Dominance and leadership networks in a flock of homing pigeons, based on data collected by recently developed technologies, such as high-precision GPS devices.*** ***(a)*** *Releasing pre-trained homing pigeons from the loft with small, ultra-light, high-precision GPS devices on their back (the device is in the small white bag on the back of the bird). (Courtesy of Zs. Ákos).* ***(b)*** *The pecking order and* ***(c)*** *the flight leadership network. Directed edges point from the dominant (or leader) towards the subordinate (or follower) with edge widths corresponding to interaction strength. Nodes are ordered vertically according to rank, with dominants on the top. The two hierarchies are obviously very different. From (Nagy et al., 2013).*

Hierarchy measure for complex networks

In absence of a widely accepted definition of 'hierarchy', finding a well-usable quantity (measure) seems to be a promising first step towards a quantitative approach. Although measures have been proposed, they have various undesirable properties, such as involving free parameters that are often unknown (Carmel et al., 2002) or being defined only to some specific graph types, like fully directed or fully undirected graphs (Trusina et al., 2004), preventing them from becoming universally accepted. According to our expectations, an ideal quantity should satisfy the following conditions:

1. Absence of free parameters and *a priori* metrics in the definition.
2. The definition should be for unweighted directed graphs (digraphs) and it should be easily extendable to both weighted and undirected graphs.
3. The hierarchy measure should be helpful for generating a layout of the graph.

We distinguish three different types of hierarchies: "flow", "nested" and "order". In case of an *order hierarchy*, the network behind the system is neglected, and the system is described basically as an ordered set, based on an arbitrary characteristic of the elements. An example for this is the grades in a school, where student are organised into grades according to their age, and are referred to as 'first graders', 'second graders', etc. In case of a *nested hierarchy*, higher level elements consist of lower level ones, like in the biological organizational hierarchy: molecules form cells, cells form tissues, tissues form organs, etc. When a network is structured in a *flow hierarchy*, the system can be described by a directed graph in which two nodes are connected if one influences the other. In case of armies, the nodes can be the ranks with edges indicating the flow of orders. Since order and nested hierarchy types can be converted into flow hierarchy, the ideal quantity is defined for flow hierarchy.

In this section we overview a measure proposed by (Mones et al., 2012) which satisfies all the above expectations and has been developed for flow hierarchy. Accordingly, we believe that this quantity can be widely accepted and used to measure the level of hierarchy. The central idea is that the rank of a node within a network should be related to its 'impact' on the whole system, and 'impact' can be measured by the ratio of nodes that are reachable from the given node. Such a quantification can be done via the concept of local reaching centrality:

Local reaching centrality, $C_R(i)$, of node *i* in an unweighted directed graph, *G*, is the proportion of all nodes in the graph that can be reached from node *i* via outgoing edges, that is, the maximum possible number of nodes reachable from node *i* divided by $N - 1$. In short, *hierarchy is the heterogeneous distribution of the local reaching centralities of the nodes* (see Fig. 2).

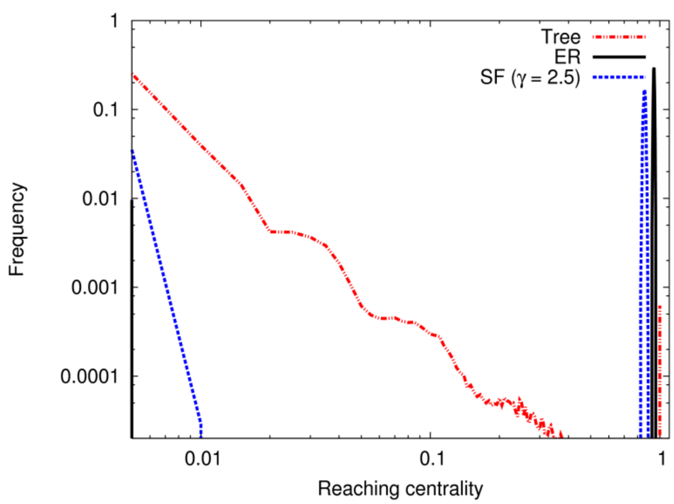

Figure 2 – ***Distributions of the local reaching centralities for the different network types: Tree, Erdős-Rényi(ER) and scale-free(SF).*** *All the curves are averages of 1000 graphs with N=2000, of the appropriate graph type. From (Mones et al., 2012).*

In order to demonstrate the above idea (namely that the distributions of the local reaching centralities reveal the hierarchical nature of a network), three different graph types are compared in Figure 1: Erdős-Rényi (ER), hierarchical ("Tree"), and Scale-free (SF). As it can be seen, the three distributions are markedly different: The directed tree is the most heterogeneous, following a power-law which is distorted due to the random branching numbers, with very few nodes having $C_R(i)$ close to 1. Accordingly, this one is the most hierarchical network. Since a well-usable measure does not return a distribution, but much more a number, the proposed definition is an expression that grasps the heterogenity of the distribution as follows. Let C_R^{max} denote the highest local reaching centrality. Then, *Global Reaching Centrality, GRC*, is defined as next:

$$GRC = \frac{\sum_{i \in V} [C_R^{\max} - C_R(i)]}{N - 1} \tag{1}$$

The GRC values for our three example graphs are the following:

Tree: 0.997 ± 0.001, which is the highest.

Scale-free: 0.127 ± 0.008, that is, SF networks are slightly hierarchical,

Erdős-Rényi: 0.058 ± 0.005, that is, these are not hierarchical at all.

These values (the means and variances) have been calculated for an ensemble of 1000 graphs, and they demonstrate nicely that the measure returns values that are close to our ‘intuitions’.

The above definition applies to unweighted directed graphs. The generalizations to undirected and/or weighted networks are based on the modification of the definition of the local reaching centrality. Regarding the visualization abilities, Figure 6 depicts two hierarchical networks plotted using the present approach.

Group performance is maximized by hierarchical competence distribution

Hereunder, in the present section and in the next one, we shall overview two frameworks investigating the emergence of hierarchy within groups consisting of individuals with diverse abilities.

The first framework was motivated by the observations made on pigeon flocks (and shortly discussed in the Introduction), namely that individuals contribute unequally in solving a common problem (navigating home). Among others, it turned out clearly that the flight leadership network is much more complicated than a two-level hierarchy in which one or a few leaders would lead the rest of the group. Accordingly, the network structure depicted on Fig 1 *C* is probably based on the different navigational skills of the pigeons, but is this only an assumption that “seems reasonable” or is this really the case? And if so, then how are these skills distributed exactly within the flock?

In order to answer these questions we have defined four models, all describing a problem that had to be solved by a group of interconnected agents: (i) a minimal model, in which the group had to find the correct answer choosing from two options (yes/no, -1/1, etc.), called “*Voting model*”, (ii), a very general model, in which a number sequence had to be estimated, called “*Sequence guessing model*”, (iii) a model for a specific problem, “*Direction finding*”, and finally (iv), the above mention “*Flocking model*”, in which a group of boids (units moving in a 2D surface) had to find a pre-defined target. The models are defined in a way that many real-life problems can be mapped on them.

For more details see (Zafeiris and Vicsek, 2013).

In each model, the quality of the groups' performance, *Pe*, is quantifiable and characterized by a parameter with values in the [0, 1] interval. The capability of the group members, Co_i, also varied between 0 (complete ignorance) and 1 (perfect knowledge). Each model consisted of the following iterative steps. (i) The behavior $Be_i^{(t+1)}$ of agent (member) *i* at time step *t*+1, depended both on its own estimation $f(Co_i)$ regarding the correct solution, and on the (observable) average behavior of its neighbors $j(\epsilon R)$ in the previous step *t*, $<Be^t_j>_{j\epsilon R}$:

$$Be_i^{(t+1)} = (1-\lambda_i)\, f(Co_i) \oplus \lambda_i \langle Be_j^t \rangle_{j\in R} \qquad (2)$$

$\oplus$ denotes "behaviour-dependent summation", where "behaviour" refers to various actions, such as estimating a value, casting a vote or turning into a direction, etc. λ_i defines the *pliancy* (disposition to follow others) of unit *i*, also by a parameter taking values on the [0, 1] interval. Some kind of noise, explicitly or implicitly, was incorporated into all models.

The aim of the study was to identify the optimal distribution of the competence values (Co_i*) and pliancy values (*λ_i*) within the group.* The reason behind the phenomenon that during collective decision making processes not all members have perfect knowledge is that learning (gaining knowledge and information) is a costly process (and unnecessary at the same time, since group members can learn and copy from each other). Accordingly, our question can be put as follows: *If the 'available amount' of competence is limited, then how to distribute it among the group members, if the group is to perform as efficient as possible?* Apart from the scientific value of this question, potential applications include choosing the best composition for a team, where "best" means better performance using the smallest possible amount of resources ("competence costs money").

Regarding the communication network (who shares information with who) we assumed four different graph types: (i) Erdős-Rényi, (2) Small-world, (3) hierarchical and (4) a real-life one describing the friendship relations in a high school, referred to as 'Friendship'. Figure 3 *c* summarizes our results for our minimal ('Voting') model on the Friendship graph (Fig. 3 *a*, *b*). In the Flocking model – in contrast with the other models – those boids communicated with each other which were closer than a given distance called 'Range of Interaction' (ROI), as depicted on Fig. 3 *d*.

In all cases, we find that the optimal distribution of competences is a highly skewed function with a structured fat tail, as opposed to the often-assumed 'bimodal' distribution in which individuals are either 'informed' or 'not-informed'. We believe that this more continuous nature of the optimal distribution is due to a phenomenon that we call "information spreading or mixing", which can be summarized as next: *Multi-level hierarchical interactions make the spreading (mixing) of the information between the individuals much more efficiently than in a two-level" system.*

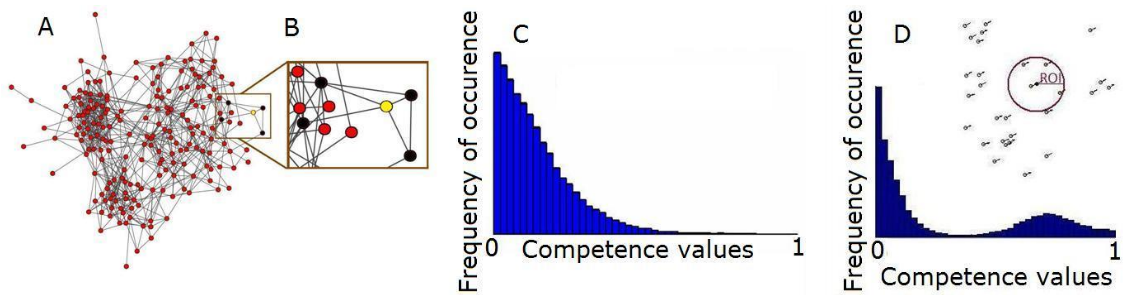

Figure 3 **– *Optimal competence distribution for The Voting model and for the Flocking model.*** ***a,*** *The "Friendship graph", a real-world social network reflecting the amity relations in a high school among 204 students.* ***b,*** *An enlarged portion of the network showing the influential relations from the viewpoint of the node coloured yellow.* ***c,*** *The optimal competence distribution for the Voting model: a highly skewed function with a fat tail.* ***d,*** *The optimal competence distribution for the Flocking model: the distribution of the competence values is a highly skewed function in this case too, with a structured tale. From (Zafeiris and Vicsek, 2013).*

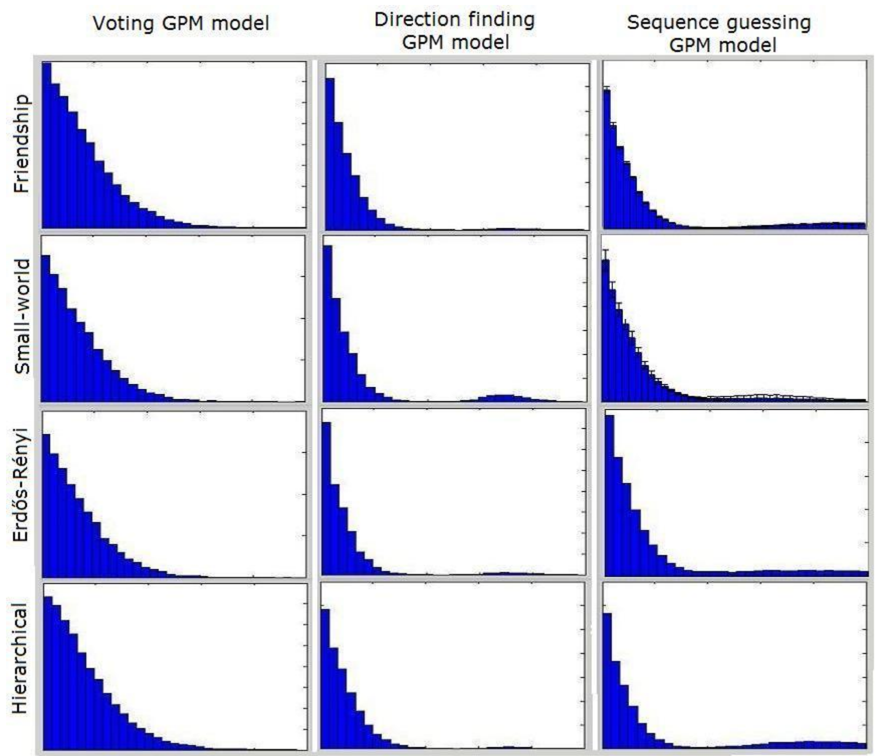

Figure 4.- ***The optimal competence distributions for the three models and four network types.*** *The optimal distributions are hierarchically ordered, highly skewed functions often with a structured tail. The size of the "Friendship" network is N=204, while the other graphs contain N=200 nodes. In the last column (belonging to the Sequence guessing model) we have marked the error bars for the Friendship and small world networks (the error bars for the rest of the plots fall into the same range). "GPM model" is the abbreviation of "group performance maximization model". From (Zafeiris and Vicsek, 2013).*

Emergence of hierarchical cooperation among selfish individuals

In the previous section we have seen that if a problem is to be solved by interconnected agents, the *group* performs optimal if the members are structured into a multilevel hierarchical leadership network. But what kind of structure emerges – if any – if individuals keep in mind only their own interest?

(Nepusz and Vicsek, 2013) have designed and investigated a simple model in which selfish individuals are trying to optimize their own success in a continually changing environment, and found that such a setup leads to a hierarchical network-type organization. The obtained structures possess the two, perhaps most important features of complex systems: a simultaneous presence of adaptability and stability. By such a model, which due to its simplicity is applicable to a wide range of actual situations, we may take a significant progress towards getting a deeper insight into the hierarchy producing mechanism.

The main assumptions of the model are: i) groups of individuals are typically embedded into a changing environment and better adaptation (finding out about the new conditions as soon as possible) is one of the core advantages an individual (or the group) can have. Importantly, ii) the abilities of the actors to gain advantage from their environment on their own is obviously diverse, thus, iii) individuals are trying to follow the decisions of their group mates (learn from them) in proportion with the degree they trust the judgment of the others as compared to their own level of competence. iv) Maintaining a decision-making connection with a group mate has a cost (effort). When these common and natural assumptions are integrated into a game theoretical-like stochastic model, the process results in the emergence of a collaboration structure in which the leader-follower relationships manifest themselves in the form of a multi-level, directed hierarchical network. Neglecting any of the above four points leads to loosing the emergence of a multi-level hierarchical structure.

The main steps of the decision making process in the model are:

1. First the changing environment is defined (the state of which the individuals have to guess to gain benefit) in a very simple but unpredictable way. The state of the environment is chosen to have a value of 1 or 0 with a probability *p.* Such a definition corresponds to a random walk with a characteristic time of changing its direction proportional to $1/p$
2. Each of the individuals has a pre-defined ability (according to a given distribution with values between 0 and 1, see Fig. 5) to make a proper guess of the state of the environment. The guess of each actor in each turn is based on its interactions with the agents he/she trusts the most by making a weighted average of its own estimation and that of the most trusted $k=1, 2, \ldots, m$ actors (typically from 2 to 7).
3. After all agents completed a round of making their guesses, the actual state of the environment is revealed, letting the units learn which ones have made the correct estimate.
4. The above information allows the construction of an updated ‘trust matrix’ in which the elements correspond to the degree agent *i* trusts agent *j*. This trust is proportional to the number of times agent *i* made use of the estimate of agent *j* in such a way that the guess by *j* contributed positively to the correct guess of agent *i*. Accordingly, individuals are trusted on the basis of their prior performance.

More trusted agents are „listened to” more frequently. Naturally, the trust matrix is updated as the collective decision making process progresses.

The above steps iterate during which the system typically converges to a trust matrix in which the values depend on a non-trivial way on the original abilities of the agents. A typical run starts with a uniform (except the diagonal) trust matrix which then evolves in time in such a way that after some time, the values of the permanently changing matrix more or less suddenly jump into a state which optimizes the overall performance of the group to a much higher degree than a random matrix (See Fig. 5).

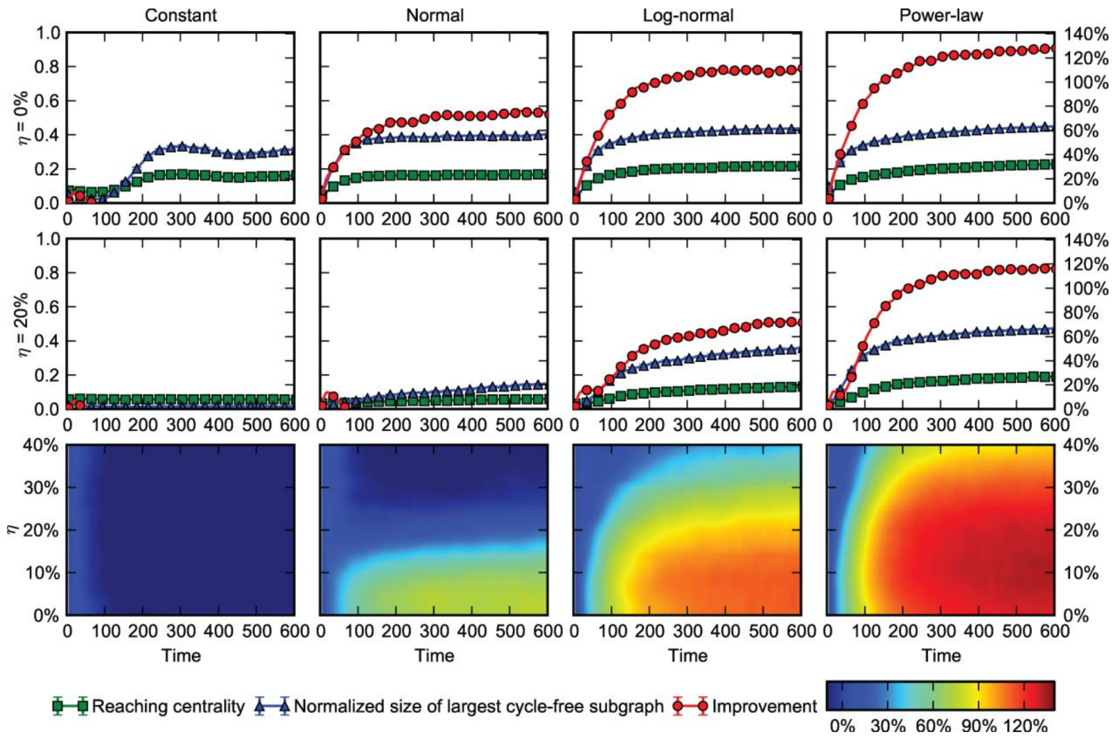

Figure 5 - ***Behaviour and performance of the model as a function of time and noise for various ability distributions****. The columns correspond to constant, normal, log-normal and power-law fitness (or ability) distributions with a mean value of 0.25 and a variance of 1/48. The upper row corresponds to the case of no noise; the middle row corresponds to 20% relative noise. The green and blue lines correspond to two hierarchy measures (fraction of forward arcs and global reaching centrality, expressed as numbers between 0 (no hierarchy) to 1 (maximal hierarchy that is theoretically possible). Red lines indicate the improvement of the overall performance of the individuals, expressed as percentages on the right axis. The heat maps in the bottom line represent the improvement as a function of time and relative noise level. Each data point is averaged from 500 trials with N=256 individuals; error bars represent the standard error of the mean but they are smaller than the corresponding markers on the plot. From (Nepusz and Vicsek, 2013).*

The evolved trust matrix possesses all the information about the network that has emerged during the run. It is straightforward to derive the graph from this matrix: each agent is a node and the weight of the edge between nodes *i* and *j* is the *(i,j)* element of the matrix. Those edges are taken into account in the graph which have the highest weights –

which are the strongest ties. These networks are found to be hierarchically ordered with multiple levels in them. Figure 5 depicts how hierarchy emerges as a function of time by depicting the values of two complementary measures: the ‘global reaching centrality’, GRC, proposed by (Mones et al., 2012) and discussed in the section entitled ‘Hierarchy measure for complex networks’, and the normalized fraction of forward arcs, defined by (Eades et al., 1993). As it can be seen, the individuals show a strong tendency to structure themselves into a multilevel hierarchical organization which, importantly, seems to be the case in real-life social interactions as well. In a recent experiment, called “Liskaland”, the way by which the leadership-followership relations are being built up has been studied in a group of 86 people. In this experiment, human actors were participating in a process in which they were interested in gaining advantage (in the form of making a larger amount of virtual money) during a camp organized along the economic theory of Liska (Liska, 2007). The participants were interested in getting good advice from others, and the information about their tendencies to follow others had been recorded using an online questionnaire (for more details see Mones et al). Fig. 6 shows the outcome of the experimentally registered network of directed interactions between the participants built up during their one week long interactions as compared to one of the typical networks the model reviewed in the present section predicts, for the same parameters, i.e., 73 nodes (13 participants remained segregated) and 142 edges. The good qualitative and quantitative agreement between the experimental and the model network provides a promising evidence in favour of the approach.

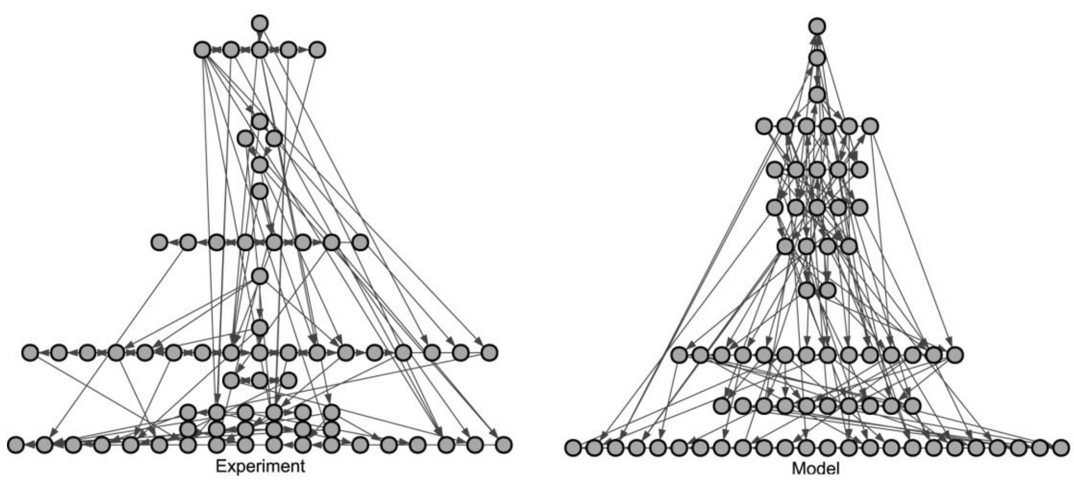

*Figure 6**: Qualitative comparison of the experimental and the modelling results.** The network on the right hand side had been generated by the approach reviewed in the present section showing features similar to those obtained during the Liskaland experiment. The data were plotted using the method introduced in section "Hierarchy measure for complex networks". From (Nepusz and Vicsek, 2013).*

Conclusion

Animal (and human) groups have been emerging as a result of evolutionary processes. Accordingly, it can be assumed that they are highly optimized systems regarding the most essential aspects. Since biological observations show that for complex societies multi-level hierarchical structures are abundant, this type of organization is most probably more advantageous than the other ones, including the 'egalitarian' type and the 'two-level' structure. We believe/conjecture that one of the advantages lie in the phenomenon called "information spreading or mixing theory" which claims that information is spreading more effectively (more quickly) in a multi-level hierarchical structure than it spreads in other kinds of networks.

Furthermore, members of the same group can maintain not only one, but more hierarchies in parallel, always arranging themselves according to the one that is optimal to the actual task and conditions. An example for this phenomenon is the pecking order contra the flight leadership network within a flock of homing pigeons, depicted on Figure 1 *B* and *C*, which networks apparently differ from each other.

Importantly, such multilevel hierarchies appear spontaneously as well, when selfish individuals are trying to optimize their own success, regardless of any kind of group interest.

We consider the studies overviewed in the present paper as the first steps towards a quantitative interpretation of the origin and emergence of multi-level hierarchies.

References

Carmel, L., Haren, D. and Koren, Y. (2001) *Drawing Directed Graphs Using One-Dimensional Optimization*. Springer-Verlag, Heidelberg, pp. 193–206.

Celko, J. (2012), *Joe Celko's Trees and Hierarchies in SQL for Smarties*, Morgan Kaufmann, Burlington.

Chisholm, D. (1992), *Coordination Without Hierarchy*, University of California Press, Oakland.

Eades, P., Lin, X., and Smyth, W.F. (1993) "A fast and effective heuristic for the feedback arc set problem" *Inform Process Lett*, Vol. 47, No. 6, pp. 319-323.

Kipfer, B. A. (2001) *The Order of Things: How Everything in the World is Organized into Hierarchies, Structures, and Pecking Orders*, Random House Reference, New York.

Mones, E., Vicsek, L. and Vicsek, T. (2012), "Hierarchy Measure for Complex Networks", *PLoS ONE*, Vol. 7, No. 3, doi:10.1371/journal.pone.0033799.

Mones, E et al *Experimental results on the spontaneous emergence of hierarchical networks of leadershiprelations in groups of humans optimizing their individual benefits*, to be published.

Nagy, M., Ákos, Zs., Bíró, D. and Vicsek, T. (2010), "Hierarchical group dynamics in pigeon flocks", *Nature*, 464, pp. 890-893.

Nagy, M., Vásárhelyi, G., Pettit, B., Roberts-Mariani, I., Vicsek, T. and Bíró, D. (2013), "Context-dependent hierarchies in pigeons", *PNAS*, doi:10.1073/pnas.1305552110.

Nepusz, T. and Vicsek, T. (2013), "Hierarchical self-organization of non-cooperating individuals", *PLoS ONE*, Vol. 8, No. 12, doi: 10.1371/journal.pone.0081449.

Saaty, T. L. (2012), *Decision Making for Leaders: The Analytic Hierarchy Process for Decision in a Complex World*, RWS Publications, Pittsburgh.

Thompson, G., Frances, J., Levacic, R. and Mitchell, J. C. (eds.)(1991), *Markets, Hierarchies and Networks: The Coordination of Social Life*, SAGE Publications Ltd,

Trusina, A., Maslov, S., Minnhagen, P. and Sneppen K (2004), "Hierarchi measures in complex networks." *Phys Rev Lett*, Vol. 92, pp. 178702.

Williamson, O. E. (1983) *Markets and Hierarchies: Analysis and Antitrust Implications*, Free Press, New York.

Zafeiris, A. and Vicsek, T. (2013), "Group performance is maximized by hierarchical competence distribution", *Nature Communications*, Vol. 4, doi:10.1038/ncomms3484