

# Quantifying Organization by Means of Entropy

Jia-Liang Lu, Fabrice Valois, Mischa Dohler and Dominique Barthel

**Abstract**—With the aim to reduce disorder and improve efficiency, nodes in an ad hoc network run a self-organization scheme to cooperatively organize the network. Although metrics such as complexity or self-stability are commonly used for evaluation, to the best of our knowledge, none of them quantifies the efficiency to build and maintain an organization (order). We henceforth apply the notion of entropy to ad hoc type wireless networks facilitating a quantification of the internal organizational state generated by different self-organization schemes. Invoking node and link failure probabilities, we expose the dependency of the organizational state on the self-organization protocol of choice.

**Index Terms**—self-organization, entropy, ad hoc networks

## I. INTRODUCTION

The trend towards ubiquitous wireless communication by means of WSN/MANET requires the use of self-organization paradigms in order to minimize the need of configuration, to facilitate networking and to enable quick deployment [1]. Self-organization refers to non-centralized processes facilitating the emergence of an *order* in the network. Self-organization schemes hence facilitate the provision of a logical organization in the network based on the physical topology with the aim to support applications and services [1]. The use of self-organization schemes should increase the order in the network.

We believe that the organization exhibits a microscopic level and a macroscopic one. At microscopic level, the organization is based on logical links, representing local interactions between nodes: a node can communicate with one neighbor (existence of a logical link) but not with another one (non-existence of a logical link) according to a self-organization scheme. At macroscopic level, all these local interactions form a virtual structure which is flexible and adaptive.

Several metrics [2], [3] are used to evaluate the self-organization strategies such as the protocol cost (e.g. control message overhead), the complexity of the algorithm, and the self-stabilization properties (e.g. convergence time). However, these metrics are not suitable to describe the degree of the order in the network.

To this end, entropy is a well established notion that had been introduced by Clausius [4] as a unique measure for the process of a reversible change in thermal energy with respect to the absolute temperature; he focused on the macroscopic behavior of microscopic chemical reactions and proposed the thermodynamic entropy. Based on this, Boltzmann defined

the statistical entropy [5] from combining microstates (giving rise to the different configurations of a system) as  $S = -k_B \sum_i p_i \ln p_i$ , where  $p_i$  is the probability that a microstate  $i$  occurs during the system's fluctuations and  $k_B$  being the Boltzmann constant; it is applicable to characterize the order of the system and how the system self-organizes among different entities.

Later, Shannon introduced the concept of information entropy  $H$  (today known as Shannon entropy) [6]; Shannon's entropy relates to the representation but not contents of information by quantifying its uncertainty. This notion was used by Shannon himself to quantify the capacity of a transmission channel and has also been extended to other problems and domains: the Shannon entropy has been used in complex systems and social problems [7]; also, in [8], a social entropy has been introduced to measure the robot group diversity (self-organization in a robotic team system); etc. Above and related approaches, however, use information entropy as a local metric which helps guiding self-organization to achieve optimal local decisions w.r.t. to some metrics [9], [10].

Contributions extending these concepts to describe emergent system behaviors are scarce. In [11], the authors metaphorically describe the conditions under which coordination can emerge from individually autonomous actions in a multi-agent system; i.e., emergent coordination can arise through coupling the macro level (a concomitant decrease in entropy) to an entropy-increasing process at a micro level such as pheromone evaporation.

In contrast to above, we adopt the definition of *entropy* used in statistical thermodynamics to quantify the order of a network according to the node distribution and channel properties under different organizational schemes. In our opinion, it is a key metric because it exhibits the susceptibility of self-organization protocols to network inherent parameters, such as link and node reliability.

## II. EXTENDED DEFINITION OF ENTROPY

Entropy is a macroscopic description of a system taking into account microscopic interactions. In ad hoc type wireless networks, the use of self-organization schemes restricts these interactions. For example, when a local minimum spanning tree (LMST) organization is applied, a node does not communicate with some of its neighbors; the interactions among these nodes are eliminated. In order to quantify the impact of self-organization schemes on the internal organization of a network, we apply the concept of entropy to the links connecting the nodes. Let us assume that a link exists between nodes  $u$  and  $v$  with probability  $p(u, v)$ ; then the state of organization, i.e. the entropy of this link is given as  $E = -p(u, v) \log p(u, v)$ . High values of  $E$  indicate high disorders; low values of  $E$  high orders, i.e. a better organization.

Manuscript received May 14, 2007; revised August 16, November 11 and December 3, 2007. This work is partially funded by France Telecom R&D under CRE No 46130157 and French ARESA RNRT Project ANR-05-RNRT-01703.

J. LU and F. Valois are with ARES INRIA / CITI Lab, INSA-Lyon, F-69621, France (Email: {firstname.lastname}@insa-lyon.fr)

M. Dohler and D. Barthel are with France Telecom R&D 28 Chemin du Vieux Chene, 38243 Meylan Cedex, France (Email: {firstname.lastname}@orange-ftgroup.com)

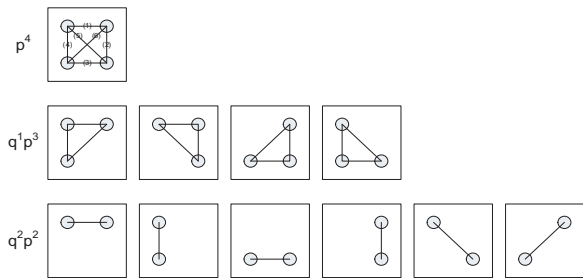


Fig. 1. Flat organization.

For  $p(u, v)$  being close to zero, the entropy is close to zero; this indicates that this link is well organized even though virtually absent; however, it is almost deterministically absent. For  $p(u, v)$  being close to one, the entropy is also close to zero; this also indicates that this link is well organized because virtually deterministically present. Intermediate values of  $p(u, v)$  yield larger disorder because the link may or may not be present.

We assume the physical structure of an ad hoc type wireless network as a random geometric graph  $G(X; r)$  which indicates the locations of nodes in set  $X$  and the physical radio link between nodes within the radio vicinity  $r$ . By summing the entropy of each link, the global entropy of the network is:

$$E = \sum_{u, v \in X} -p(u, v) \log p(u, v). \quad (1)$$

Let us now apply this concept to the network. To this end, we observe that the existence of a link is the result of:

- two independent random effects, i.e. the existence of the nodes on either end of the link, and the availability of the wireless channel leading to a random network state, and
- one deterministic effect, i.e. the way the protocol organizes the network for a given realization of the random network state.

Let us assume that a node exists with probability  $p$  and hence that it does not exist with probability  $q = 1 - p$ . Furthermore, the wireless channel is assumed to be in a good state with probability  $c$ . Let us now calculate the entropy per link and, then, per system for a flat and LMST organization scheme, assuming a simple illustrative case as per Figures 1 and 2. This will then be extended in Section III to more general topologies.

#### A. Flat Organization Scheme

As per Figure 1, the probability that link (1) exists is  $p_1 = c \cdot (p^4 + 2qp^3 + q^2p^2) = c \cdot p^2$ . The remaining five link probabilities ( $p_i$ ) can similarly be calculated, all of which lead to the same above-given expression. The state of organization (entropy) per link is hence  $-p_1 \log p_1$ , which gives the entire system entropy as  $E = -\sum_{i=1}^6 p_i \log p_i = -6 \cdot p_1 \log p_1$ .

Let us now assume that  $m$  links exist in a flat network composed of  $N$  nodes. The value of  $m$  depends on the communication radius and the geographical distribution of nodes. Given a density of the network  $\rho$ , we obtain  $m$  as

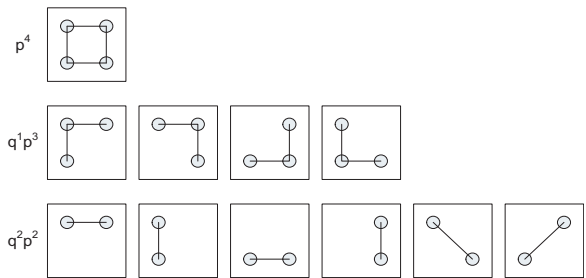


Fig. 2. LMST organization.

$m = \rho \cdot N/2$ . The entropy of a network using flat organization is the sum of the entropy of all links; we hence have  $E = -m \cdot cp^2 \cdot \log(cp^2)$ .

#### B. LMST Self-Organization Scheme

LMST [12] is a link pruning based self-organization scheme which provides a logical structure with fewer links than those in the flat physical structure. All nodes are kept in the resulting structure. Each node locally computes its two-hop minimum spanning tree (MST). For two neighborhood nodes  $u$  and  $v$ , if  $u$  is in the MST of  $v$  and  $v$  is in the MST of  $u$ , then the link  $(u, v)$  is in the LMST: it is kept active in the network, whereas the rest of links are dropped in the organization.

As per Figure 2, the probability that links (1)-(4) exist is  $p_1 = c \cdot (p^4 + 2qp^3 + q^2p^2)$  and the probability that links (5) and (6) exist is  $p_5 = c \cdot q^2p^2$ . The entire system entropy is hence  $E = -[4 \cdot p_1 \log p_1 + 2 \cdot p_5 \log p_5]$ .

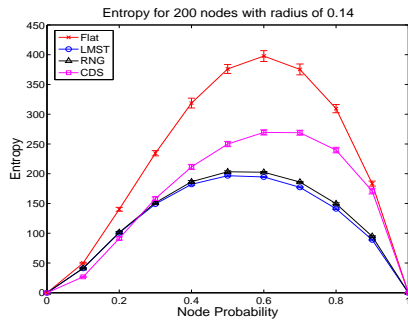
To obtain these probabilities, and hence the total entropy, for a general network topology is analytically infeasible. Furthermore, an exhaustive summation over all possible link formations is NP-complete. This required us to use simulations, the results of which are described in subsequent section.

### III. ENTROPY EVALUATION

Subsequent results are obtained for four different self-organization protocols, namely flat, LMST, Relative Neighborhood Graph (RNG) [13] and Connected Dominating Set (CDS) [2]. For the latter, the following rules are used for the links selection: *i*) all the links between a dominant and its dominated are kept active, *ii*) we keep only one active link to connect a dominated to only one dominant.

To numerically obtain the entropies, we generate a network with  $N$  nodes and let them appear with probability  $p$  in the network. Then we apply the probability  $c$  on the existing links. On this resulting graph, we apply given organizational schemes and get a subgraph, say  $G^{LMST}(X, r, p, c)$  in the case of LMST. We record the existing links for the particular network realization. This is then re-run for the same topology. After sufficient runs, we obtain the average probability that a link exists in the organization. According to (1), the entropy for this topology is calculated. Finally, a set of above procedures is repeated to obtain an entropy value which is statistically significant for the given self-organization protocol.

Without depicting the analytical entropy results of example topologies given in Figures 1 and 2, the following can be observed. LMST has a lower entropy over all state probabilities

Fig. 3. Entropy for varying  $p$ ;  $c = 1$ .

and hence a higher degree of organization. Furthermore, for low ( $p \rightarrow 0$ ) or high ( $p \rightarrow 1$ ) node existence probabilities, the difference between both schemes is negligible. For more uncertain configurations, it is better to use LMST. It can further be observed that for nodes occurring with increasing likelihood ( $p \rightarrow 1$ ), the channel plays an increasingly important role.

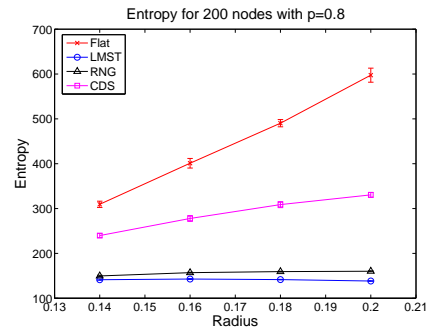
Figure 3 compares entropy results for flat, LMST, RNG and CDS schemes for  $N = 200$  nodes and a normalized communication radius of  $r = 0.14$ . The use of LMST/RNG/CDS decreases the entropy and hence increases the degree of organization in the network. This in turn requires less links to be reconfigured and hence less energy/overhead spent in the case of network changes. LMST and RNG provide a higher order than CDS, i.e. have a higher level of determinism in their local algorithms and – from an organization point of view – are thus the better choice. This is due to LMST and RNG keeping fewer possible interactions in their resulting logical structure, whereas all interactions between dominating nodes in CDS are kept. We also observe that LMST and RNG are less susceptible to changes of link existence probabilities, i.e. a change in  $p$  will cause less change in entropy than for flat and CDS topologies. Finally, as exemplified by a horizontal line in the figure, the entropy and hence the energy/overhead needed for reconfiguration is the same at low and high  $p$  despite the systems exhibiting entirely different communication behaviors.

Figure 4 shows the evolution of entropy with different radii for the same number of nodes and constant probabilities. The entropy of flat organization increases almost in a linear fashion; however, the entropy of LMST and RNG are almost stable with different radii. This can be explained as following:

- In the flat organization, with the increasing radius, nodes have more neighbors. The number of microscopic states of local interactions hence increases. This makes the global network entropy increase as well.
- In self-organization schemes, the local interactions are constrained by the algorithm running on every node. Fewer interactions are encouraged in the network. It is known that the density of LMST and RNG structure is small and almost constant; therefore, despite varying radii, almost a constant entropy is observed.

#### IV. CONCLUSIONS

Self-organization has received much attention lately, mainly driven by the increasing network size and complexity. In this

Fig. 4. Entropy for varying  $r$ ;  $p = 0.8$ .

paper, we extended the metric based on the notion of entropy to evaluate the organizational state of a network which has then been applied to various known self-organization schemes.

This approach facilitated to gain the following quantitative and qualitative insights into the behavior and design of self-organization protocols: First, any of the chosen protocols yields a higher organizational state than a flat topology. Second, for an increasing likelihood of a node being available, the channel commences to play a dominant role in the organizational state of the network. Third, LMST and RNG yield the same level of organization, despite being based on fairly different rules. Finally, whilst flat and CDS strongly depend on the communication radius, LMST and RNG exhibit a high degree of organizational stability.

#### REFERENCES

- [1] C. Prehofer and C. Bettstetter. Self-Organization in Communication Networks: Principles and Design Paradigms. *IEEE Communication Magazine*, pages 78–85, July 2005.
- [2] J. Wu and H. Li. On calculating Connected Dominating Set for efficient routing in ad hoc wireless networks. *Proc. of the Third Int' Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications, Seattle, USA*, pages 7–14, August 1999.
- [3] J. Lu, F. Valois, and D. Barthel. Low-Energy Self-organization Scheme for Wireless Ad Hoc Sensor Networks. *4<sup>th</sup> WONS, Obergurgl, Austria*, 2007.
- [4] R. Clausius. *Annalen der Physik*, (79):368397, 500524, 1850.
- [5] L. Boltzmann. Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Warmegleichgewicht. *Wien. Ber.*76, 373–435, Band II, pages 164–223, 1877.
- [6] C.E. Shannon. The mathematical theory of communication. *University of Illinois Press*, 1949.
- [7] K. Sneppen, A. Trusina, and M. Rosvall. Measuring information networks. *Pramana - Journal of Physics*, 64:1121–1125, 2005.
- [8] T. Balch. Hierarchic social entropy: an information theoretic measure of robot group diversity. *Autonomous Robots*, 8:209–237, 2000.
- [9] T.H. Labella and F. Dressler. A bio-inspired architecture for division of labour in SANETS. *IEEE/ACM BIONETICS 2006, Cavalese, Italy*, 2006.
- [10] K. Sneppen, A. Trusina, and M. Rosvall. Hide-and-seek on complex networks. *Europhysics Letters*, 2005.
- [11] V. Parunak and S. Brueckner. Entropy and self-organization in multi-agent systems. *Proc. of the fifth international conference on Autonomous agents*, pages 124–130, 2001.
- [12] N. Li, C. Hou, and L. Sha. Design and analysis of an MST-based topology control algorithm. *INFOCOM, San Francisco, USA*, 3:1702–1712, July 2003.
- [13] G. Toussaint. The relative neighborhood graph of finite planar set. *Pattern Recognition*, 1980.