The Degree of Global-State Awareness in Self-Organizing Systems

Christopher Auer, Patrick Wüchner, and Hermann de Meer

Faculty of Informatics and Mathematics,
University of Passau
94032 Passau, Germany
{auerc,wuechner,demeer}@fim.uni-passau.de

Abstract. Since the entities composing self-organizing systems have direct access only to information provided by their vicinity, it is a non-trivial task for them to determine properties of the global system state. However, this ability appears to be mandatory for certain self-organizing systems in order to achieve an intended functionality.
Based on Shannon's information entropy, we introduce a formal measure that allows to determine the entities' degree of global-state awareness. Using this measure, self-organizing systems and suitable system settings can be identified that provide the necessary information to the entities for achieving the intended system functionality.
Hence, the proposed degree supports the evaluation of functional properties during the design and management of self-organizing systems. We show this by applying the measure exemplarily to a self-organizing sensor network designed for intrusion detection. This allows us to find preferable system parameter settings.

Keywords: Self-organizing systems, Mathematical modeling, Quantitative evaluation, Information theory, System design, Sensor networks

1 Introduction

Self-organization is foreseen to enable efficient, scalable, and robust large-scale distributed systems, like the future Internet. However, the design of self-organizing systems (SOSs) that fulfill a certain intended functionality is a difficult task. Three different design approaches are sketched in [1]: the trial-and-error approach, the bio-inspired design, and the design by learning from an omniscient entity (see [2]).

In SOSs, the entities can only observe events that happen in their immediate vicinity. This makes it a non-trivial task to design the entities such that they foster the desired functionality of the SOS. Entity design can be simplified if the entities have access to information on the SOS' global state. This, however, presumes that the system entities are provided with the necessary global state information to foster the preferred system behavior. A quantitative characterization of the entities' ability to derive such global state information could help

to identify such SOSs that can then be further investigated (e.g., by the method in [2]) towards how the entities can foster the intended functionality. In this paper, we are proposing such a quantitative characterization.

Several formal measures have been proposed for describing certain properties of SOSs, like autonomy [3, 4], emergence [3–5], adaptivity, homogeneity, and resilience [6]. However, these measures do not evaluate to which extent the entities can derive the necessary global-state information from the information provided by their vicinity.

In this paper, we propose a novel measure of the entities' degree of global-state awareness. By evaluating this degree, systems can be identified in which, over time, the necessary information is communicated to the entities, i.e., the entities become *aware* of important properties of a former global system state. These system candidates can then be further investigated, e.g., by the method proposed in [2], towards how the entities can use the provided information in a purposeful, target-oriented manner. The resulting system can finally be evaluated by using the measure of target orientation proposed in [6].

Hence, the main contribution of this paper is proposing an answer to the question *if*, and to which degree, the SOS under investigation is able, in principle, to provide the necessary information to the system's entities. For instance, in an SOS where global consensus has to be found in a completely decentralized manner, a low degree of global-state awareness indicates that the system entities are not provided with sufficient information. Hence, the system is not suitable to fulfill the task and should be redesigned.

We derive the measure of the degree of the entities' global-state awareness utilizing Shannon's information entropy (see [7]) and are hence in line with the previously proposed measures of autonomy [3, 4], emergence [3–5], and homogeneity [6].

As an illustrative example, we apply the proposed measure to find suitable system parameters for a sensor network (adopted from [8]) which is designed for intrusion detection. It can be shown that the simulated sensor network indeed is able to fulfill the desired functionality in a self-organizing manner if the system parameters are chosen such that the degree of global-state awareness is maximized. The sensors are then able to reach a global consensus on whether an alarm triggered by a subset of the sensors was a false positive.

The remainder of the paper is organized as follows: In Sec. 2, we provide a reminder of Shannon's information entropy, which is essential to the understanding of following sections. We also introduce the model representing the SOS. As the main contribution of this paper, we introduce the measure of the entities' degree of global-state awareness in Sec. 3. In Sec. 4, we apply the proposed measure to discuss a sensor network for intrusion detection as an illustrative example. How the degree of global-state awareness can assist during the design of SOSs is sketched in Sec. 5. A conclusion and directions for future work are given in Sec. 6.

2 Background

This section briefly recapitulates the basic concepts of information theory and introduces the system model.

2.1 Entropy of Information

In this paper, a random variable X is denoted by an upper-case letter and the realization of X is denoted by the lower-case letter x. For the range of X we use the bold-face character $\mathbf{X}$. In our case, the range of any random variable is finite, i.e., $\#\mathbf{X} < \infty$.

As a measure of uncertainty, a useful tool is Shannon's information entropy as defined in [7]: given some discrete random variable X with finite range $\mathbf{X}$, the entropy of X is defined as:

$$\mathrm{H}[X] = -\sum_{x \in \mathbf{X}} P[X = x] \cdot \log_2 P[X = x]\,. \tag{1}$$

The entropy is 0 iff X almost surely takes a value $x \in \mathbf{X}$, i.e., $P[X = x] = 1$ for a single $x \in \mathbf{X}$. $\mathrm{H}[X]$ takes its maximum iff X is uniformly distributed. Generally, the lower the value of the entropy, the more certain the outcome of a random variable can be predicted.

If X, Y are two discrete random variables with finite ranges $\mathbf{X}, \mathbf{Y}$, then knowing the outcome of Y might reduce the uncertainty of the outcome of X. When the outcome of Y is known, the remaining entropy of X is measured by the conditional entropy $\mathrm{H}[X|Y] = \mathrm{H}[X, Y] - \mathrm{H}[Y]$ (cf. [7, 9]). It can be shown (cf. [9]) that $\mathrm{H}[X|Y] = \mathrm{H}[X]$ iff X and Y are independent and $\mathrm{H}[X|Y] = 0$ iff $X = f(Y)$, where f is a non-stochastic function.

2.2 System Model

In this paper, we focus on the class of technical SOSs that can be modeled as discrete-event systems consisting of a finite set $\mathbf{N}$ of entities. An example of such a system is depicted in Figure 1.

At each time step $t \in \mathbb{N}_0$, each entity $n \in \mathbf{N}$ receives input $i_{t,n}$ from its vicinity that contains neighboring entities and possibly comprises parts of the environment. Entity n then produces the output $o_{t,n}$ which is received in the next time step $t + 1$ by entity n's neighbors and possibly also influences the environment. For example, in Figure 1, $o_{t,n_{2,2}} = i_{t+1,n_{2,1}}$.

We assume that the SOS' entities can be modeled as deterministic finite-state Mealy automatons: The finite state space of entity n is denoted by $\mathcal{S}_n$ and its transition function by ζ_n that maps the input $i_{t,n}$ and current state $s_{t,n} \in \mathcal{S}_n$ of entity n to its output $o_{t,n}$ and its successor state $s_{t+1,n} \in \mathcal{S}_n$.

The tuple $\gamma_{t,n} = (i_{t,n}, s_{t,n})$ is called *local configuration* of entity n at time step t: At each time step, each entity can only observe its local configuration, i.e., its own state and the input provided by its vicinity. Up to some time step t, the

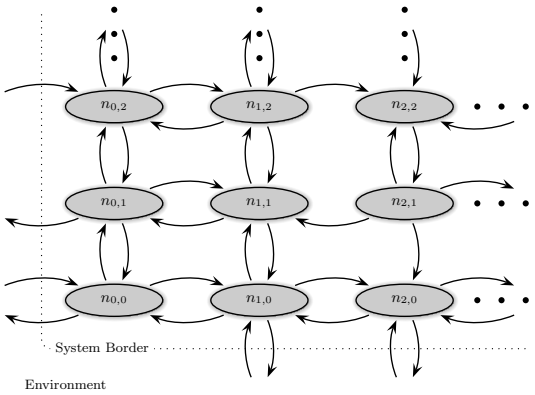

Fig. 1. Model of a self-organizing system: The system consists of several entities (nodes) and can be distinguished from its environment (system border). The entities may exchange information (arrows) with their vicinity that might contain parts of the environment.

sequence of local configurations $(\gamma_{t',n})_{t'=0\ldots t}$ constitutes the *local history* $\overleftarrow{\gamma}_{t,n}$ of entity n. The state of the whole SOS at time step t can be fully described by the local configurations of all entities: We call $\gamma_t = (\gamma_{t,n})_{n\in\mathbf{N}}$ the *configuration* at time step t. The set of all possible configurations is called *state space* of the SOS and is denoted by $\mathbf{\Gamma}$.

The configuration γ_0 at time step 0 is called *initial configuration (IC)*. The IC is modeled by the random variable Γ_0 with range $\mathbf{\Gamma}_0 \subseteq \mathbf{\Gamma}$, where $\mathbf{\Gamma}_0$ is the set of all possible ICs which all have a non-zero probability. Since the IC is chosen randomly, all subsequent configurations can also be described by random variables even if there is no random input from the environment. We denote the random variable of the SOS' configuration γ_t by the capital letter Γ_t. Accordingly, we denote with $\Gamma_{t,n}$ and $\overleftarrow{\Gamma}_{t,n}$ the random variables of entity n's local configuration $\gamma_{t,n}$ and of the local history $\overleftarrow{\gamma}_{t,n}$, respectively.

In principle, our system model assumptions can be relaxed by generalizing the model using the modeling approaches presented in [3] (asynchronous communication and stochastic transition functions) and [4] (continuous time and continuous state space). However, due to space limitations and to preserve clarity, we refrain from applying these generalizations here.

3 The Degree of Global-State Awareness

In this section, we present the main contribution of this paper. Our main goal is to evaluate an entity's ability to obtain global-state information when only the information provided by the entity's vicinity is directly available.

3.1 Classification Problem

To define the measure of the entities' degree of global-state awareness as general as possible, we introduce the notion of classification problems. In general, it is unsuitable, unnecessary, and often even impossible to distribute the full information on the system's global state to each single entity.

Hence, we aggregate global states that share a common property of interest to the same state class. For an entity it is then sufficient to figure out the state class, i.e., to solve the classification problem by only taking the information into account that is provided by its vicinity.

We assume that the entities need to determine the state class of the SOS at some time step t. Without loss of generality, we refer to this time step t as the initial time step $t = 0$. Therefore, we aggregate initial system states of the set of possible ICs $\mathbf{\Gamma}_0$ to state classes to define the *classification problem* as follows:

Definition 1 (Classification Problem). *Let* $\mathbf{L}$ *be a partition of* $\mathbf{\Gamma}_0$*, i.e.:*

$$\bigcup_{l \in \mathbf{L}} l = \mathbf{\Gamma}_0 \wedge \forall l, l' \in \mathbf{L}, l \neq l' : l \cap l' = \emptyset \wedge \emptyset \notin \mathbf{L}\,.$$

$\vartheta : \mathbf{\Gamma}_0 \to \mathbf{L}$ *is a function which maps any initial configuration to its corresponding state class in* $\mathbf{L}$*:* $\forall \gamma_0 \in \mathbf{\Gamma}_0 : \forall l \in \mathbf{L} : \vartheta(\gamma_0) = l \iff \gamma_0 \in l\,.$

Applying function ϑ *to random variable* $\mathit{\Gamma}_0$ *produces the random variable* $L = \vartheta(\mathit{\Gamma}_0)$*, which is the random variable of the state class of* $\mathit{\Gamma}_0$*. The realization of* L *is denoted by* l*.* L *naturally inherits the probability distribution from* $\mathit{\Gamma}_0$*.*

The problem of determining the state class of the initial configuration $l \in \mathbf{L}$ *(i.e., the outcome of* L*) given only the local history of an entity is called* classification problem $\mathbf{L}$.

In the following, we assume that a classification problem is non-trivial in the sense that $\mathbf{L}$ contains at least two state classes.

The mapping of the IC γ_0 to the property of interest can be represented in form of a function $\psi : \mathbf{\Gamma}_0 \to \mathbf{P}$, i.e., γ_0 has the property $\psi(\gamma_0)$, where $\mathbf{P}$ is the set of global-state properties of interest that depend on the specific application scenario. This implies the partition $\mathbf{L}_\psi$ in which all ICs are aggregated that map to the same value of ψ. Hence, the system entities can obtain the value of $\psi(\gamma_0)$ after determining the corresponding state class of $\mathbf{L}_\psi$, i.e., after solving the classification problem $\mathbf{L}_\psi$.

3.2 Defining the Degree of Global-State Awareness

As a reminder, $L = \vartheta(\mathit{\Gamma}_0)$ denotes the random variable of the state class of the random IC $\mathit{\Gamma}_0$. To solve a classification problem $\mathbf{L}$, the entities of an SOS have to decrease their uncertainty about the random variable L, given their local history. In order to measure the uncertainty about L when the local history is given, we use Shannon's information entropy (see Sec. 2.1):

Definition 2 (Degree of Global-State Awareness). *The* degree of global-state awareness $\omega_{t,n}(\mathbf{L})$ *observable by entity* $n \in \mathbf{N}$ *at time step* t *is defined by:*

$$\omega_{t,n}(\mathbf{L}) = 1 - \frac{\mathrm{H}[L|\overleftarrow{\Gamma}_{t,n}]}{\mathrm{H}[L]}\,. \tag{2}$$

The system's overall degree of global-state awareness *is then defined as the limiting value for* $t \to \infty$*, averaged over all entities:*

$$\omega(\mathbf{L}) = \lim_{t\to\infty} \frac{1}{\#\mathbf{N}} \sum_{n\in\mathbf{N}} \omega_{t,n}(\mathbf{L})\,. \tag{3}$$

Note that $\mathrm{H}[L]$ measures the uncertainty of predicting L when no additional information is given. Since $\mathrm{H}[L|\overleftarrow{\Gamma}_{t,n}] \leq \mathrm{H}[L]$, we have $\omega_{t,n}(\mathbf{L}) \in [0,1]$ and $\omega(\mathbf{L}) \in [0,1]$.

By the definition of the entropy (Eq. (1)), the denominator in Eq. (2) is strictly greater than 0 since $\mathbf{L}$ consists of at least two state classes that have a non-zero probability.

Intuitively, $\omega_{t,n}(\mathbf{L})$ measures how certain entity n can determine the outcome of L at time step t by taking its local history $\overleftarrow{\Gamma}_{t,n}$ into account. If $\omega_{t,n}(\mathbf{L}) \approx 1$, then $\mathrm{H}[L|\overleftarrow{\Gamma}_{t,n}]$ is small compared to $\mathrm{H}[L]$ and, hence, n can use its local history to determine the outcome of L with a high certainty. As a matter of fact, $\omega_{t,n}(\mathbf{L}) = 1$ iff there exists a non-stochastic function f with $L = f(\overleftarrow{\Gamma}_{t,n})$ (see Sec. 2.1). At the other extreme, if $\omega_{t,n}(\mathbf{L}) \approx 0$, then $\mathrm{H}[L|\overleftarrow{\Gamma}_{t,n}] \approx \mathrm{H}[L]$, which implies that the local history contributes almost no information about the outcome of L. Again, $\omega_{t,n}(\mathbf{L}) = 0$ iff L and $\overleftarrow{\Gamma}_{t,n}$ are independent random variables (see Sec. 2.1).

The system's overall degree of global-state awareness $\omega(\mathbf{L})$ (Eq. (3)) measures to which extent the information about L is distributed within the SOS in the long-term. Note that a high overall degree of global-state awareness indicates the distribution of the information about the IC's state class L among most entities since all entities equally contribute to $\omega(\mathbf{L})$.

4 Intrusion Detection in Sensor Networks

To show the applicability of the proposed measure, we now utilize it to find suitable parameter settings for a sensor network designed for intrusion detection. It can be shown that the sensor network indeed is able to fulfill the desired functionality in a self-organizing manner as long as suitable system parameters are chosen. The degree of global-state awareness helps to find these system parameters.

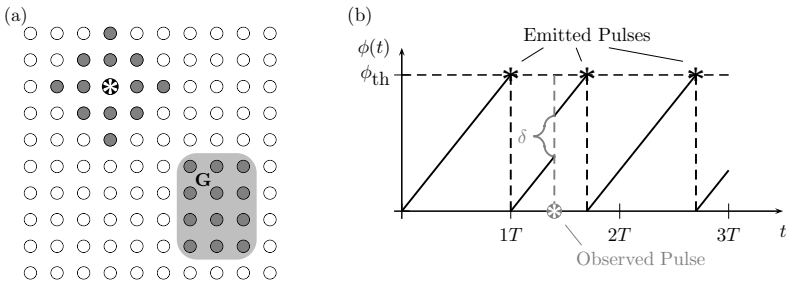

Fig. 2. Sensor network (a) and phase function $\phi(t)$ of pulse-coupled oscillator (b)

4.1 Scenario and Problem Description

Consider the distributed sensor network sketched in Figure 2(a). Suppose that the group **G** of sensors detects an intrusion. In absence of an omniscient central entity and with a non-zero probability of single sensors giving a false alarm, a consensus needs to be achieved in a completely decentralized and self-organized manner whether the alarm was false positive. Due to hard resource constraints, this should also be accomplished by exchanging a minimum amount of information. In the following, we use the degree of global-state awareness to find system candidates for which the sensors are able to find a global consensus on the detected intrusion.

In [8], a mechanism is described how a sensor network can reach global consensus on an intrusion detected by a sensor group **G**. Each sensor periodically emits pulses (illustrated by black asterisk in Fig. 2(a)) with a fixed identical period T but varying phases. These pulses can be observed by all sensors in its neighborhood (gray dots) only.

In the following, we neglect transmission delays which are discussed in, e.g., [10]. With the help of [10], the following discussion can similarly be applied to more realistic scenarios where delays cannot be neglected, e.g., in the internet.

Before intrusions can be detected, all sensors have to synchronize their phases. According to [11], which is based on the theoretical model of pulse-coupled oscillators presented in [12], synchronization can be achieved using the following mechanism: Each entity calculates a local phase ϕ following the phase function illustrated in Figure 2(b). The phase is increased linearly over time. If the entity receives no pulse, it periodically emits a pulse with a period of T. If an entity observes a pulse from another entity, it additionally increases its current phase by $\delta = (\alpha - 1)\phi + \beta$, where $\alpha > 1$ and $\beta > 0$ are constant system parameters (see also [11, 12] for details). If the phase reaches a threshold value ϕ_{th}, the entity emits a pulse and resets its phase to $\phi = 0$. It is proved in [12] that using this mechanism, synchronization can be reached almost surely if an entity's pulse can be directly observed by all other entities. Moreover, it is shown in [13] that pulse-coupled oscillators can also synchronize if each entity communicates only with its nearest neighbors.

After synchronization, the sensor network is ready for detecting intrusions using the method described in [8]: On detection of an intruder, the sensors of group $\mathbf{G}$ shift their phase by a predefined amount of $\Delta\phi$. This results in a partitioning of the network into two groups where intra-group synchronization is still given. Inter-group synchronization, however, is no longer given. The groups automatically start to resynchronize until, at some time instant, all sensors are again synchronized and all phases have been shifted by some $\Delta\Theta$ compared to the phase before the intrusion was detected. According to [8], if $\Delta\phi$ is chosen appropriately and if the sensor network is fully connected, i.e., every sensor receives pulses from all other sensors, then each sensor is able to infer from $\Delta\Theta$ whether group $\mathbf{G}$ was large enough to exclude false alarm. However, in [8], no concrete advice is given how to find appropriate system parameters (i.e., α, β, $\Delta\phi$, and ϕ_{th}) for this mechanism, despite focusing on fully connected networks. Additionally, [8] neglects that each sensor can observe more than just $\Delta\Theta$: Each sensor might use the whole history of its local observations to infer whether a false alarm has occurred. In Section 4.2, we now derive suitable system parameters by evaluating the degree of global-state awareness of a sensor network that is not even fully connected.

4.2 Application of Proposed Measure to Sensor Network

Without loss of generality, we assume that each entity $n \in \mathbf{N}$ ($\mathbf{N} = 1, \ldots, 100$) of the sensor network should conclude that an intrusion has in fact occurred, if more than one fifth of all sensors detected the intrusion, i.e., $\#\mathbf{G} > \#\mathbf{N}/5$.

We model the sensor network at discrete time steps that are defined at all time instants where at least one sensor emits a pulse. At these time steps, each sensor can be modelled as a finite state automaton: As input from its neighboring sensors, each sensor either perceives a pulse or not. Additionally, at time step 0, the detection of an intrusion by a sensor is modelled as an input from the sensor's environment. To keep the model simple, we assume that after time step 0 no further intrusions are detected. The state $s_{t,n}$ of sensor n at time step t is given by the sensors' phase value $\phi_n(t)$. As output, the sensor may or may not emit a pulse to its neighbors. In order for the state space of the sensor network to be discrete, we assume, without loss of generality, that at time step 0 all sensors are synchronized and have a phase value of 0 except for the sensor group $\mathbf{G}$ that have an identically shifted phase value of $\Delta\phi$. The resulting configuration of the sensor network is referred to as IC γ_0. This implies a discrete set of possible ICs $\mathbf{\Gamma}_0$ and, since the sensor network is investigated at discrete time steps, also the set of possible system configurations $\mathbf{\Gamma}$ and sensor states $\mathcal{S}_n$ are discrete.

To keep the model concise, we assume that intrusions are always detected by compact rectangular groups of sensor nodes, i.e., $\mathbf{G}$ forms a rectangular subset with width w and height h within the sensor grid of size 10×10 nodes (see Fig. 2(a)). The width and height of $\mathbf{G}$ are described by independent random variables W and H, respectively. Both random variables have the range $\{0, \ldots, 10\}$, where $W = 0$ or $H = 0$ implies that no sensor has triggered an alarm. Given the

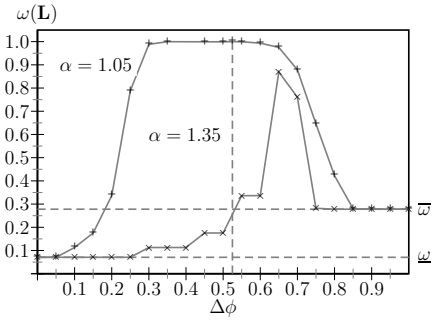

Fig. 3. The degree of global-state awareness $\omega(\mathbf{L})$ (y-axis) depending on the phase shift $\Delta\phi \in [0, 0.05, \ldots, 1.0]$ (x-axis) for $\alpha = 1.05$ and $\alpha = 1.35$.

width w and height h, the position of $\mathbf{G}$ is then chosen according to a uniform distribution on all possible positions within the sensor grid.

According to the notation introduced in Section 3.1, there are two equivalence classes on the IC $\gamma_0 \in \mathbf{\Gamma}_0$ of interest. Class $l_{>1/5}$ contains all ICs where $\#\mathbf{G} > \#\mathbf{N}/5$ and intrusion should be recognized by all sensors. Class $l_{\leq 1/5}$ contains all ICs where $\#\mathbf{G} \leq \#\mathbf{N}/5$ and should be treated as false alarm. It hence has to be checked for which system parameters each sensor is able to solve the classification problem $\mathbf{L} = \{l_{>1/5}, l_{\leq 1/5}\}$.

By simulating the sensor network for all possible $\mathbf{G}$ and weighing the outcomes with the according probabilities, exact values for the degree of global-state awareness $\omega(\mathbf{L})$ can be obtained by applying Eq. (2). In particular, we investigate the influence of the system parameters α and $\Delta\phi$ on the system's ability to solve the classification problem $\mathbf{L}$. For β we choose a value of 0.01 and $\phi_{th} = 1$, which are reasonable values according to [11].

Figure 3 shows the influence of $\Delta\phi$ on the degree of global-state awareness $\omega(\mathbf{L})$ (y-axis) for $\alpha = 1.05$, $\alpha = 1.35$, and $\Delta\phi \in [0, 0.05, 0.1, \ldots, 1]$ (x-axis). It can be seen that with increasing $\Delta\phi$, both curves initially increase, reach a maximum, and then drop. For both extremes, $\Delta\phi$ close to 0 and 1, $\omega(\mathbf{L})$ is small which implies that the sensors cannot decide certainly whether $L = l_{\leq 1/5}$ or $L = l_{>1/5}$. This is because the difference between the phases of $\mathbf{G}$ and $\mathbf{N} \setminus \mathbf{G}$ is small. Hence, the two sensor groups almost immediately resynchronize and the information about L has no time to spread within the network.

For $\Delta\phi \to 0$, $\omega(\mathbf{L})$ tends to the limiting value $\underline{\omega}$. In this case, the sensors in $\mathbf{G}$ detect an intrusion but do not shift their phase. Each sensor in $n \in \mathbf{G}$ knows that at least one sensor, namely n itself, has detected an intrusion. This reduces the uncertainty about L for n and leads to the limiting value $\underline{\omega}$. At the other extreme, for $\Delta\phi \to 1$, the sensors in $\mathbf{G}$ detect an intrusion, do not shift their phase, but emit a pulse which is also perceived by the sensors neighboring $\mathbf{G}$. All these sensors can then reduce their uncertainty about L, resulting in the limiting value $\overline{\omega}$ which is greater than $\underline{\omega}$.

For $\alpha = 1.05$, the degree of global-state awareness is close to 1 for $\Delta\phi \in [0.3, 0.6]$. Indeed, for $\Delta\phi = 0.525$ (vertical dashed line in Fig. 3), $\omega(\mathbf{L}) = 1$, which has an important implication. It follows that $\mathrm{H}[L|\overleftarrow{\Gamma}_{t,n}] = 0$ for a sufficiently large t and all $n \in \mathbf{N}$. From information theory (see Sec. 2.1), it is then known that there exists a non-stochastic function $f_n : \overleftarrow{\mathbf{\Gamma}}_{t,n} \rightarrow \mathbf{L}$ for every sensor $n \in \mathbf{N}$ such that $f_n(\overleftarrow{\Gamma}_{t,n}) = L$, almost surely. Each sensor $n \in \mathbf{N}$ can then apply f_n to obtain L. Such a mapping can, for instance, be obtained by the simulation process we used to calculate the degree of global-state awareness. Another approach to implement f_n is discussed in Section 5.

It can also be seen in Fig. 3 that the curve for $\alpha = 1.35$ remains well below the curve for $\alpha = 1.05$. In [12, 13] it is shown that the time until the sensors reach synchronization is inversely proportional to α. Intuitively, a reduced convergence time prevents the information about state class L from spreading within the whole network since the sensors within the vicinity of $\mathbf{G}$ resynchronize too quickly. Hence, suitable choices for α and $\Delta\phi$ maximize the degree of global-state awareness while minimizing the convergence time of the network to assure that the information about an intrusion is spread quickly within the network.

5 Effects on Entity Design

Remember that, by using the degree of global-state awareness, it is possible to characterize *if* it is, in principle, possible for the entities to obtain the desired global-state information. However, the degree does not indicate *how* the system entities can use the information provided by their vicinity to derive the desired global-state information and *how* this information can be used to foster the desired functionality of the SOS. In this section, we sketch how suitable local interaction strategies can be found.

A degree of global-state awareness of 1 implies that there exists a non-stochastic function that maps any local history to the corresponding global state class. In principle, such a mapping could be obtained by a simulation process similar to the one we used to produce the results shown in Sec. 4. However, the resulting table that maps a local history to the corresponding state class would be very large and, hence, cumbersome, if not impossible, to implement in devices with limited memory and computation power, e.g., wireless sensors. Furthermore, in a real-world scenario, events may happen that are not encountered by the simulation process.This could lead to an undefined input value for the obtained mapping.

In [2], we presented a method to derive local interactions strategies for entities of an SOS by learning from an omniscient entity, called the *Laplace's Daemon* (LD). We now discuss the application of LDs to obtain a mapping from the local history to the corresponding state class of the IC. In a simulation environment, the LD of each entity is equipped with information about the IC's state-class. At each time step, as input, each LD receives the local configuration of the corresponding entity and outputs the current state-class of the IC. The sequence of local inputs to and outputs from each LD generated during the simulation is

investigated by a time series analysing algorithm (CSSR algorithm; cf. [14, 15]) to obtain a minimal Markov chain description of each LD. The obtained Markov chains can be effectively implemented in the entities and are minimal in the sense that they only use the relevant information from the local history to predict the global-state class. Furthermore, the derived mappings even produce reasonably reliable results when they encounter situations that were not faced during the simulation. This information about the global-state class can then be used by the entities to foster the desired functionality of the SOS. A high degree of global-state awareness indicates that it is worthwhile to apply the approach presented in [2] to optimize the given SOS.

6 Conclusion

In order for the entities of a self-organizing system to optimize the overall system performance, it is useful for them to know certain aspects of the system's global state. To model such aspects of the system's global state, we introduced the classification problem where system global states that share common properties of interest are aggregated to state classes. As the main contribution of this paper, the degree of global-state awareness was introduced that evaluates to which extent the entities of a self-organizing systems are able to find out the respective state class while using only the local information provided by the entities' vicinity. By using our measure, it is possible to find preferable system parameters that enable the entities to adjust their behavior for an optimized overall system performance.

As an illustrative example, a sensor network for intrusion detection was used in this paper. By applying the measure of the entities' degree of global-state awareness, we were able to find system parameter settings that allow all sensors to determine whether an intrusion was detected by a significant number of sensors in a completely decentralized manner. The sensors can use this global-state information to exclude false alarms.

By using the found system parameter settings, it is then possible to derive local interaction strategies that can rely on this global information. We also described how to obtain such interaction strategies in general, i.e., by learning from an omniscient entity, an approach we presented in [2].

In near future, we intend to apply our measure of global-state awareness to other application scenarios. We also plan to investigate further properties of self-organizing systems, e.g., decentralization, while also providing suitable and generally applicable formal measures of these properties.

Acknowledgements

The authors thank Dr. Richard Holzer for fruitful discussions. This research is partially supported by the AutoI project (STREP, FP7 Call 1, ICT-2007-1-216404), by the ResumeNet project (STREP, FP7 Call 2, ICT-2007-2-224619), and by the Network of Excellence EuroNF (FP7, IST 216366).

References

1. Elmenreich, W., De Meer, H.: Self-organizing networked systems for technical applications: A discussion on open issues. In Hummel, K.A., Sterbenz, J.P.G., eds.: Proc. of the 3rd International Workshop on Self-Organizing Systems (IWSOS '08), Vienna, Austria. Volume 5343 of LNCS., Vienna, Austria, Springer Verlag (December 2008) 1–9
2. Auer, C., Wüchner, P., De Meer, H.: A method to derive local interaction strategies for improving cooperation in self-organizing systems. In Hummel, K.A., Sterbenz, J.P.G., eds.: Proc. of the 3rd International Workshop on Self-Organizing Systems (IWSOS '08), Vienna, Austria. Volume 5343 of LNCS., Vienna, Austria, Springer Verlag (December 2008) 170–181
3. Holzer, R., de Meer, H., Bettstetter, C.: On autonomy and emergence in self-organizing systems. In: Proc. of the 3rd International Workshop on Self-Organizing Systems (IWSOS '08), Vienna, Austria. Lecture Notes in Computer Science, Springer Verlag, Heidelberg (December 2008) 157–169
4. Holzer, R., de Meer, H.: On modeling of self-organizing systems. In: Proc. of Autonomics 2008, Turin, Italy. (September 2008)
5. Mnif, M., Müller-Schloer, C.: Quantitative emergence. In: Proc. of the 2006 IEEE Mountain Workshop on Adaptive and Learning Systems (SMCals 2006), Piscataway, NJ, USA, IEEE (July 2006) 78–84
6. Holzer, R., de Meer, H.: Quantitative modeling of self-organizing properties. In Plattner, B., Spyropoulos, T., Hummel, K.A., eds.: Proc. of the 4th International Workshop of Self-Organizing Systems (IWSOS '09). LNCS, Zurich, Switzerland, Springer Verlag (December 2009)
7. Shannon, C.E.: A mathematical theory of communication. In: Bell System Technical Journal. Volume 27. (1948) 379–423
8. Hong, Y., Scaglione, A.: Distributed change detection in large scale sensor networks through the synchronization of pulse-coupled oscillators. In: Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), Montreal, Canada. Volume 3. (May 2004) 869–872
9. Cover, T.M., Thomas, J.A.: Elements of information theory. Wiley-Interscience (1991)
10. Ernst, U., Pawelzik, K., Geisel, T.: Synchronization induced by temporal delays in pulse-coupled oscillators. Phys. Rev. Lett. **74**(9) (Feb 1995) 1570–1573
11. Tyrrell, A., Auer, G., Bettstetter, C. In: Biologically Inspired Synchronization for Wireless Networks. Volume 69/2007 of Studies in Computational Intelligence. Springer Verlag, Heidelberg (2007) 47–62
12. Mirollo, R.E., Strogatz, S.H.: Synchronization of pulse-coupled biological oscillators. SIAM Journal on Applied Mathematics **50**(6) (1990) 1645–1662
13. Lucarelli, D., Wang, I.: Decentralized synchronization protocols with nearest neighbor communication. In: Proc. of the 2nd International Conference on Embedded Networked Sensor Systems, Baltimore, MD, USA, ACM (2004) 62–68
14. Shalizi, C.R.: Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata. PhD thesis, University of Wisconsin (2001) Supervisor: Martin Olsson.
15. Shalizi, C.R., Shalizi, K.L.: Blind construction of optimal nonlinear recursive predictors for discrete sequences. In: AUAI '04: Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, Arlington, Virginia, United States, AUAI Press (2004) 504–511