Modelling the dynamics of cluster formation

Filip De Smet

Supervisors: Dirk Aeyels, Gert De Cooman

Abstract—**We introduce a dynamical model of mutually attracting agents with the long term behavior consisting of agents organized into several groups or clusters. The cluster structure is completely characterized by means of a set of inequalities in the parameters of the model and transitions between different cluster structures take place when the intensity of the attraction is varied. We illustrate the relation with the Kuramoto model on interconnected oscillators and we discuss an application on compartmental systems.**

Keywords—**Clustering, Kuramoto model, compartmental systems.**

I. Introduction and Motivation

The clustering phenomenon is observed in fields ranging from the exact sciences to social and life sciences; consider e.g. swarm behavior of animals or social insects [1], opinion formation [2] or the clusters in the frequency space for synchronized coupled oscillators [3] as a model for heart cells. We present a model that captures this phenomenon and at the same time allows a mathematical analysis. We formulate necessary and sufficient conditions for the occurrence of a given cluster structure. The model, or elaborated versions, may be used as a tool for explaining some of the phenomena mentioned above, for systematizing arguments, for prediction and for control.

II. The Dynamics

In general interacting agents are bound to generate unpredictable unstructured behavior. We propose a mathematical model with a particular type of interaction such that through self-organization a structure emerges where sets of clustered agents find themselves in balance.

We present a simple model admitting a succinct formulation of the conditions governing the cluster configuration at which the system settles. The differential equations for the model consisting of N agents ($N > 1$) are

$$\dot{x}_i(t) = b_i + \frac{1}{N}\sum_{j=1}^{N} f(x_j(t) - x_i(t)), \qquad (1)$$

$\forall\, t \in \mathbb{R}, \forall\, i \in \{1,\ldots,N\}$, with $x_i(t), b_i \in \mathbb{R}$; $x_i(t)$ describes the state of agent i at time t, $\dot{x}_i(t)$ the time derivative, and b_i represents the autonomous component in each agent's behavior. The summation term represents the attraction exerted by the other agents on each agent. The function $f : \mathbb{R} \to \mathbb{R}$ is odd and non-decreasing with respect to distances between agents. This implies a *symmetric attraction* between any two agents. We assume that the interaction intensifies with separation up to a certain saturation level:

$$\exists\, d > 0 : f(x) = F, \quad \forall\, x \geq d.$$

Author and supervisors are with the SYSTeMS Research Group, Dept. of Electrical Energy, Systems and Automation, Ghent University, Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium. E-mail: Filip.DeSmet@UGent.be,Dirk.Aeyels@UGent.be,Gert.DeCooman@UGent.be.

A Lipschitz condition on f is introduced for technical reasons; it guarantees a unique solution to the differential equation with respect to a set of initial conditions.

For convenience we will present our results for the model (1), but an extension to the following generalization is possible [4]:

$$\dot{x}_i(t) = b_i + A_i \sum_{j=1}^{N} K_{ij}\gamma_j f\left(x_j(t) - x_i(t)\right), \qquad (2)$$

$\forall\, t \in \mathbb{R}$, $\forall\, i \in \{1,\ldots,N\}$. The function f has the same characteristics as before. The interpretation of b_i remains unaltered. The parameters A_i and γ_i are all positive. The matrix K is symmetric and irreducible with $K_{ij} \geq 0$. It is important to notice that the interaction structure of (2) is completely arbitrary, while model (1) involves all-to-all coupling. The elements of matrix K represent levels of attraction between agent pairs (e.g. no attraction between agents i and j if $K_{ij} = 0$); the extent to which each individual agent j tends to attract other agents is denoted by γ_j. The parameter A_i reflects the sensitivity of agent i to interactions with other agents.

III. Analysis and Results

Assume that, for a particular solution of (1), the behavior of the agents can be characterized as follows by an ordered set of *clusters* $(G_1,\ldots,G_M)$ defining a partition of $\{1,\ldots,N\}$:

- The distances between agents in the same cluster remain bounded (i.e. $|x_i(t) - x_j(t)|$ is bounded for all $i, j \in G_k$, for any $k \in \{1,\ldots,M\}$, for $t \geq 0$).
- After some positive time T, the distances between agents in different clusters are at least d and grow unbounded with time.
- The agents are ordered by their membership to a cluster: $k < l \Rightarrow x_i(t) < x_j(t), \forall\, i \in G_k, \forall\, j \in G_l, \forall\, t \geq T$.

We will refer to this behavior as *clustering behavior*.

For any non-empty set $G_0 \subset \{1,\ldots,N\}$, with the number of elements denoted by $|G_0|$, we introduce the notation $\langle b\rangle_{G_0}$ for the average value of b_i over G_0:

$$\langle b\rangle_{G_0} \triangleq \frac{1}{|G_0|}\sum_{i\in G_0} b_i.$$

In [4] we derive the following set of *necessary and sufficient* conditions for clustering behavior of *all* solutions of the system (1), with the cluster structure $(G_1,\ldots,G_M)$ independent of the initial condition:

$$\langle b\rangle_{G_{k+1}} - \langle b\rangle_{G_k} > \frac{F}{N}\left(|G_{k+1}| + |G_k|\right), \quad \forall\, k \in \{1, \ldots, M-1\}; \tag{3}$$

$$\langle b\rangle_{G_{k,2}} - \langle b\rangle_{G_{k,1}} \le \frac{F}{N}|G_k|, \quad \forall\, G_{k,1}, G_{k,2} \subsetneq G_k, \text{ with } G_{k,2} = G_k \setminus G_{k,1}, \quad \forall\, k \in \{1, \ldots, M\}. \tag{4}$$

The characteristics of the interaction play a key role in the proof. Since f is odd all internal interactions (i.e. interactions between agents in the same cluster) cancel when calculating the velocity of the 'center of mass' of a cluster, similar to the cancellation of internal interactions in mechanics. The saturation of f implies that the interactions between agents from different clusters reduce to F/N or $-F/N$ whenever agents from different clusters are separated over at least a distance d. The monotonicity of f will guarantee that the resulting clustering behavior is independent of the initial condition.

Under the assumption of clustering behavior and taking into account the previous considerations, the ordering of the agents and distances growing unbounded with time for agents in different clusters will lead to the condition (3). Similarly, since distances between agents from the same cluster remain bounded the condition (4) can be derived. This implies the necessity of the inequalities (3) and (4) for the existence of a solution of (1) satisfying clustering behavior. Next we give an outline of the proof of sufficiency.

The main idea is to pick an initial condition for which agents from different G_k will always (i.e. for all $t \ge 0$) be separated over at least a distance d, with their interaction saturated as a consequence. Invoking the condition (4) it can then be shown that the differences in $x(t)$-values will be bounded for agents in the same G_k. From this boundedness together with the condition (3), it will follow that the differences in $x(t)$-values for agents from different G_k will grow unbounded. The solution of (1) corresponding to this particular initial condition will exhibit clustering behavior (with $T = 0$, and the clusters equal to the G_k). Any other solution $\hat{x}$ of (1) will exhibit the same clustering behavior (i.e. identical clusters, possibly a different value for T). This follows by observing that the distance in the state space $\mathbb{R}^N$ between x and $\hat{x}$ is a non-increasing function of time: $\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{N}(x_i(t) - \hat{x}_i(t))^2\right) \le 0$.

We also indicate that for every given set of parameters b_i and F there exists a *unique* ordered partition $(G_1, \ldots, G_M)$ of clusters satisfying (3) and (4) implying a unique clustering behavior. In general there exist $N-1$ *bifurcation* values for the intensity of attraction F, defining N intervals for F; each interval corresponds to a particular cluster configuration, and transitions to new cluster configurations take place at these bifurcation points. We refer to [4] for full details.

IV. Applications

A. The Kuramoto model

The Kuramoto model [5] is a mathematical model describing systems of coupled oscillators. The oscillators are characterized by an individual frequency, which determines their behavior if there is no interaction. When the coupling strength is increased oscillators tend to form clusters, with all members of a cluster moving at the same long term average frequency. Simulations indicate that the clustering behavior is independent of the initial condition, as in the model (1), and also the transitions between the different clusters for varying coupling strength are similar.

Although analytical results for the Kuramoto model exist, a complete analytical description — as we have given for the model (1) — is not available, and therefore the results for systems (1) and (2) may be useful in the investigation of the Kuramoto model or in the investigation of synchronization of coupled oscillators in general.

B. Compartmental systems

Consider N different basins connected by horizontal pipes, each basin furthermore subject to either a constant external inflow or outflow of a fluid, e.g. water. We assume that the pipes have a maximal throughput, which is independent of the direction of the flow, and denoted by K_{ij} for the pipe connecting basins i and j. Representing the water height of basin i by x_i, the pressure difference between basins i and j will be proportional to $x_j - x_i$, and thus the volume flow rate through the connecting pipe can be represented by $K_{ij} f(x_j - x_i)$ where f relates throughput through a pipe — normalized to one — to the pressure difference — expressed in difference in water level height. Defining the appropriate parameters one easily derives the model (2).

The objective consists in checking whether a network of connected basins is prone to flooding. Assuming that the total external inflow equals the total external outflow, the desired behavior corresponds to one cluster at zero velocity. This will be fulfilled as long as, for all partitions of the set of basins into two nonempty subsets, the interconnections have the capacity to transport a net external inflow rate from one part of the network to the other part with the same net external outflow rate. This can be expressed by conditions analogous to (4). If these are not satisfied, a set of basins will overflow.

Acknowledgments

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

Filip De Smet is a Research Assistant of the Research Foundation - Flanders (FWO - Vlaanderen).

References

[1] Craig W. Reynolds, "Flocks, herds, and schools: A distributed behavioral model," *Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987, (SIGGRAPH '87 Conference Proceedings).

[2] Rainer Hegselmann and Ulrich Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation.," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002.

[3] Steven H. Strogatz, *SYNC: The Emerging Science of Spontaneous Order*, Hyperion Press, 2003.

[4] Dirk Aeyels and Filip De Smet, "Clustering in a network of non-identical and attractive agents," In preparation.

[5] Y. Kuramoto, "Cooperative dynamics of oscillator community," *Prog. Theoret. Phys. Suppl.*, vol. 79, pp. 223–240, 1984.