2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

# Critical density for highest cohesion in dynamical models of pigeon flocks

Graciano Dieck Kattas and Michael Small

Department of Electronic and Information Engineering
The Hong Kong Polytechnic University
Email: g.dieck@polyu.edu.hk, ensmall@polyu.edu.hk

**Abstract**—A computer algorithm is used to build nonlinear dynamical models of pigeon flocks from 3D positional data. Temporal separation measures are used to quantify the dynamics of the model in a global scale over several simulations. The model is capable of exhibiting realistic collective behavior that is highly dependent on speed and density. The simulations show that for fixed initial speeds, there is a critical density value for the population which gives the best flock cohesion.

## 1. Introduction

The study of the collective interactions which affect the dynamics of a population is an emerging research field of interest for scientists and engineers. The simple local actions taken by every individual are influenced by others, and they affect the global behavior of the whole population as if they were a single entity. Many different phenomena in nature, especially animals, have been observed to follow such behavioral rules. The cohesive movement of animals aggregating together is commonly refered as swarming. In particular, bird flocking is a well studied case that shows traces of synchronized movement and alignment by the interacting birds. Recently, accurate positional data of homing pigeon flights has been obtained from GPS devices and it lead to convincing conclusions of the existence of leaders in flocks [1]. Photographs have also advanced the understanding of the interactions between birds [2]. Essentially, the advances of measurement technology have improved the available data to find new discoveries about real collective dynamics.

Dynamical systems with collective behavior have usually been described using nonlinear mathematical models. Initial efforts considered partial differential equation models with diffusion components to generalize animal swarming [3]. Later models, such as the well known “Boids”, considered simple dynamics with attraction, alignment, and short-range repulsion, as the main mechanisms for collective motion [4]. The Vicsek model showed how simple orientation alignment between particles can produce a variety of swarming behaviors [5]. The common approaches on mathematically modeling dynamical systems with collective behavior have been based on proposing mathematical functions based on physical laws, which are later verified with either manual observations of the simulations [4, 5] or parameter tuning from data [6, 7]. An alternate way of constructing models is to use a computer algorithm that receives time series data as input to fit a model that can emulate the behavior as well as possible. This data-driven philosophy, sometimes referred as “system identification”, has enjoyed recent success with ordinary differential equation modeling of biological systems [8, 9], as well as with earlier applications using discrete difference equations to model chaotic behavior [10, 11]. As a first step to test our approach, we used simulated data from the Vicsek model [5] to build a basic swarming model using a modeling algorithm based on radial basis functions [10, 11]. Our results confirmed that the method is adequate for modeling basic swarming dynamics [12], and as a consequence, in this paper we extend our methodology to address the modeling of real experimental data from pigeon flights.

## 2. Input data: pigeon flight data from GPS

Recent studies in pigeon flocking retrieved accurate 3D positional data of homing pigeon trajectories using GPS devices [1]. The data sets consist of eleven free flights with the pigeons moving around near their nests, and four homing flights that involve the flock moving together to reach a destination. In this study, only the datasets of the homing flights will be considered, since their dynamics are much more concrete and manageable. The data from the homing flights was further sampled to provide a single large dataset consisting of data from all the flights which is easier to handle for a computer algorithm, and from which a single general model can be built. With this objective, sampling rates of two seconds were used on the original data to reduce the size. In addition to this, the flights were cut to delete segments with no significant movement of the birds. Far-away pigeons that were not interacting were also removed in order to have a dataset that resembles a fully interactive population as well as possible. All of this was done by manual visualization of the trajectories.

## 3. Modeling scheme

Now that the input data has been defined, the modeling scheme that we selected to characterize the flight dynamics will be presented. We decided to consider a single general model to be followed by all the birds in our simulations, just like in the classical Boids and Vicsek models [5, 4], so that the essential flocking behavior that is followed by all the pigeons can be emphasized in the model. A black-box modeling philosophy was considered, in order to focus on obtaining complex functions which can closer emulate the nonlinear behavior through simulations, instead of giving a good symbolic representation. The selected modeling method and its extension towards collective pigeon flights shall be outlined in the next two subsections.

### 3.1. Radial basis functions

We selected a radial basis function method originally presented in [10, 11] as the modeling structure and algorithm to use, since it has verified capabilities of modeling and emulating nonlinear dynamics. In general, the method receives a scalar time series $y(t)$ as input and attempts to build the best model of the form:

$$y(t+1) = f(\mathbf{z}(t)) + \epsilon(t) \tag{1}$$

where $\mathbf{z}(t) = [y(t), y(t-1), ..., y(t-d)]$ is the embedding of the system and $\epsilon(t)$ is the model prediction error. The embedding consists of the past values from the time series data $y(t)$ that will be considered for calculating the prediction for time $t+1$. The samples used for model retrieval are built from time series $y(t)$ using the embedding $\mathbf{z}(t)$. The function $f$ to build has a general structure that consists of a linear term equivalent to AR models, and a nonlinear term of sums of radial basis functions. The computational method that optimizes the model essentially consists in generating random candidate radial basis functions, choosing the ones which best follow the data, and estimating parameters; minimizing the Minimum Description length [13] to prevent data overfitting. Specific details about the structure and algorithm can be found in the original papers [10, 11].

### 3.2. Relative position modeling for general flocking model

We are interested in building a single model using relative positioning, from the data of the homing flights. The idea is to use the same general model for each bird of the system, just like the Vicsek or Boids models. Even if the data is 3D, we decided to exclude the height component, in order to consider the two coordinates not influenced by gravity. Therefore we require two functions for a single model (to predict each coordinate) and thus the model can be defined as:

$$\mathbf{f}[\mathbf{z}(t)] = \begin{pmatrix} f_1(\mathbf{z}(t)) \\ f_2(\mathbf{z}(t)) \end{pmatrix} \tag{2}$$

To consider relative positional data in our model, the general form in (1) is extended to predict the two-dimensional relative change in position $\Delta\mathbf{x}_i(t+1) = \mathbf{x}_i(t+1) - \mathbf{x}_i(t)$:

$$\Delta\mathbf{x}_i(t+1) = \mathbf{f}[\mathbf{z}_i(t)] + \mathbf{e}_i(t) \tag{3}$$

where $\mathbf{e}(t)$ is an array with the model prediction errors for each coordinate. The embedding $\mathbf{z}_i(t)$ of a bird $i$ must contain enough information from the data to guarantee an adequate prediction. We propose embedding:

$$\mathbf{z}_i(t) = \begin{pmatrix} \Delta\mathbf{x}_i(t) \\ \langle\Delta\mathbf{x}_i(t)\rangle_M \\ \mathbf{x}_i(t) - \langle\mathbf{x}_i(t)\rangle_M \end{pmatrix} \tag{4}$$

where the first component represents the positional change of bird $i$ at time $t$: $\Delta\mathbf{x}_i(t) = \mathbf{x}_i(t) - \mathbf{x}_i(t-1)$, the second component is the average positional change of the $M$ nearest neighbors, and the third component is the averaged positional difference between $i$ and its nearest neighbors. At a contrast to the Boids and Vicsek model, we decided to consider topological interactions (fixed number of nearest neighbors) instead of metric interactions (all neighbors within a distance). This was inspired from a recent study which concluded that birds interact with a fixed number of neighbors [2]. In simple terms, our embedding uses the movement of bird $i$, the averaged movement of its $M$ nearest neighbors, and its directional separation to the averaged position of its neighbors, to predict the positional change of $i$.

### 3.3. Unbiasing the data

It is important to note that the input data introduced in section 2 is based upon the absolute positions of birds, and that there is an obvious embedded navigational bias. This can be understood by visualizing that if the data consists of mostly movement towards a roughly southwest direction, then most data segments will follow that direction. Since it is of our interest to build a general flocking model using relative positioning, the bias was attenuated by transforming the data. We performed uniform 2D rotation transformations to produce the final data points on which to apply prediction (4) and embedding (5) for model building. To remove the navigational bias, each rotation angle for an instance of (5) is calculated so that the orientations of the prediction, $\Delta\mathbf{x}_i(t+1)$, span a full circle ($2\pi$) in the whole dataset. By uniformly spanning a full circle in the navigational direction of the samples, the bias is reduced.

## 4. Measuring flock dynamics

Although velocity correlations have been used previously to quantify collective behavior in animals [3, 1], we introduce a positional measure to characterize swarming dynamics with a single time-depdendent variable by defining the average separation of individuals to the mean position of the population at a given time $t$ as:

$$\delta_g(t) = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i(t) - \langle\mathbf{x}(t)\rangle_N\| \tag{5}$$

Our goal is to use this statistic to analyze the dynamical properties of a bird population from its trajectory data. This measure can be useful to perform qualitative analysis of a single dataset, with simple visualization of transient and steady state properties of the swarm. The average separation over all time intervals $T$ is a more compact statistic, and will be defined as $\overline{\delta}_g = \langle \delta_g(t) \rangle_T$.

## 5. Methodology

The algorithmic process to build a single model will be outlined:

Input: Time series matrix containing positional data from multiple homing flights : $\mathbf{x}(t)$

For each positional coordinate $j$:

1. Build samples for the function $f_j$ using as output $\Delta x_i^{(j)}(t+1)$ with its respective embedding $\mathbf{z}_i(t)$, using data from all particles/pigeons.
2. Run the radial basis modeling algorithm.
3. Set the retrieved function as $f_j$.

The radial basis modeling algorithm [10, 11] optimizes over a random set of generated functions to build the model, which makes it necessary to run the algorithm several times to retrieve different models for the same dataset and find the most appropriate one. Details on the model selection criteria will be described in the results section.

## 6. Results

The data from the homing flights that was described in section 2 was used to build five models for each structure with $M = 1$ to $M = 4$. The selection of the "best" model was done by using ten different sets of initial conditions that resemble the properties of one of the homing flights to simulate flights of $N = 9$ particles, and using $\overline{\delta}_g$ measures to compare between the cohesion of the models. The flight with lowest average separation $\overline{\delta}_g$, was selected as the "best model" due to its higher flocking behavior; and it was found to be a model with $M = 4$. This model was used for all the analyses and it will be referred as "the model" in general terms.

With different populations sizes and parameter values, our model was able to simulate many realistic behaviors, ranging from global to local swarming. In order to gather information about the qualitative dynamics using $\delta_g(t)$, we decided to limit our simulations by only changing the initial densities and speeds of the birds, using a population size ($N$) and number of nearest neighbors ($M$) fixed at 300 and 4 respectively. The initial density of the population was varied according to $r = 250r_c$, where $r$ is the radius (in meters) of the circle in which the initial positions of the particles are distributed with a uniform random distribution, and $r_c$ is the actual coefficient that is varied for each case. The initial speeds (the magnitudes of $\Delta\mathbf{x}_i(1)$) were varied according to $v_i = 30v_c$, by modifying $v_c = [0, 1]$, while the initial orientations of $\Delta\mathbf{x}_i(1)$ were obtained from a random uniform distribution $U(0, 2\pi)$.

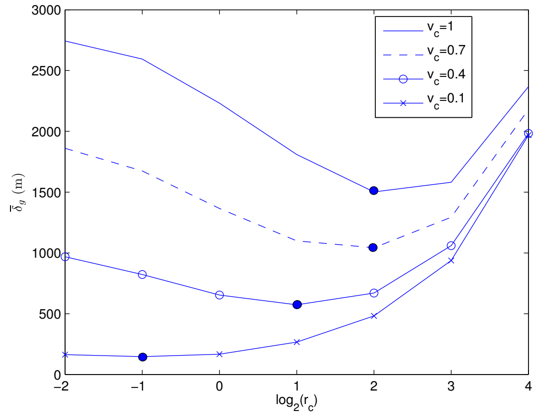


Figure 1: For fixed initial speeds, average separation $\overline{\delta}_g$ as a function of density (radius $r_c$)

One of the most interesting conclusions we could get from the simulation data is that for a fixed initial speed $v_c$, there is a critical density value which gives best cohesion in the flock. This means that high density scenarios do not guarantee that the flock will stay together. Figure 1, shows $\delta_g(t)$ values averaged over time for different initial density values. It clearly shows how for a fixed initial speed, by decreasing the density (increasing $r_c$) the separation of the flock decreases until reaching an in-between critical value, and then increases for lower densities. All this implies that a critical density distribution facilitates flock cohesion when starting from a disorganized initial state. Also of importance, the figure shows how in general, lower speeds give lower separations, which is consistent with the fact that birds can reach their neighbors much more easily if the overall movement is slower.

To better illustrate this behavior, figures 2 and 3 show simulations of two cases with different initial density values but the same initial speeds ($v_c = 0.5$). For an early time interval, figure 2 shows how the individuals in the case of higher density ($r_c = 0.5$), even though densely concentrated, move almost uniformly outward in all directions from the center, while the lower density distribution ($r_c = 2$) has significant clusters of individuals aggregating. After some time, in figure 3 we can see how the low density simulation clearly has a more cohesive and aligned group than the more dispersed high density population.

## 7. Conclusions

By verifying the model simulations, our results confirm that both the separations (densities) and speeds of birds in pigeon flocks are important parameters that characterize the dynamics of the system, in the form of flock cohesion and behavior. We discovered that a critical separation

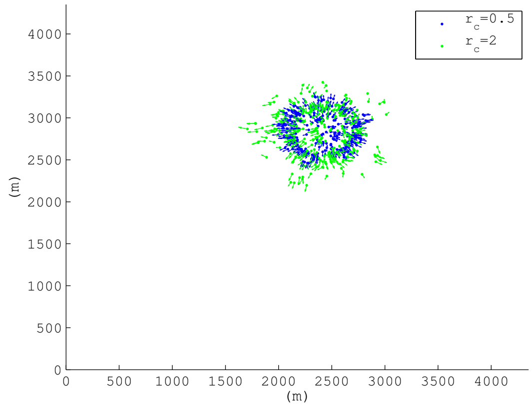


Figure 2: Comparison of simulations with different initial densities and $v_c = 0.5$ at $t = 50$

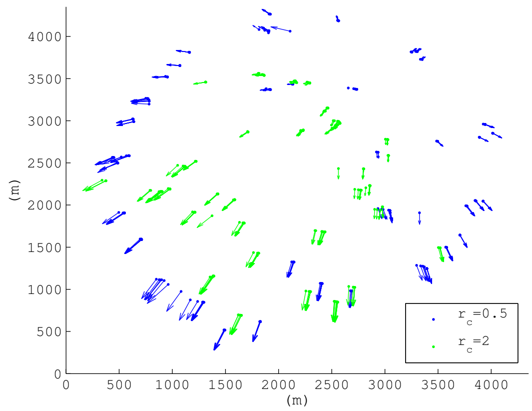


Figure 3: Comparison of simulations with different initial densities and $v_c = 0.5$ at $t = 200$

between individuals shall achieve the best overall aggregation, which also implies that a mechanism similar to the alignment, short-range repulsion, and attraction behaviors from classic models [3, 4] exists in pigeon flight. In general, we have demonstrated that using experimental data to fit black-box nonlinear dynamical system models, it is possible to simulate realistic dynamics featuring collective behavior, and to get important conclusions about the source system.

**Acknowledgments**

The authors would like to thank M.Nagy for providing the pigeon flight data. GDK is currently supported by the Hong Kong PHD Fellowship Scheme (HKPFS) from the Research Grants Council (RGC) of Hong Kong

## References

[1] M. Nagy, Z. Akos, D. Biro, and T. Vicsek, "Hierarchical group dynamics in pigeon flocks," *Nature*, vol. 464, pp. 890–893, Apr. 2010.

[2] M. Ballerini, et al., "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 1232–1237, Jan. 2008.

[3] A. Okubo and S. A. Levin, *Diffusion and Ecological Problems: Modern Perspectives*. Springer Verlag, 2001.

[4] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 25–34, 1987.

[5] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, pp. 1226–, Aug. 1995.

[6] R. Lukeman, Y.-X. Li, and L. Edelstein-Keshet, "Inferring individual rules from collective behavior," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 12576–12580, July 2010.

[7] A. Eriksson, M. Nilsson Jacobi, J. Nystrom, and K. Tunstrom, "Determining interaction rules in animal swarms," *Behavioral Ecology*, vol. 21, pp. 1106–1111, Sept. 2010.

[8] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 9943–9948, June 2007.

[9] P. Gennemark and D. Wedelin, "Efficient algorithms for ordinary differential equation model identification of biological systems," *Systems Biology, IET DOI - 10.1049/iet-syb:20050098*, vol. 1, no. 2, pp. 120–129, 2007.

[10] K. Judd and A. Mees, "On selecting models for nonlinear time series," *Physica D: Nonlinear Phenomena*, vol. 82, pp. 426–444, May 1995.

[11] M. Small and K. Judd, "Comparisons of new nonlinear modeling techniques with applications to infant respiration," *Physica D: Nonlinear Phenomena*, vol. 117, pp. 283–298, June 1998.

[12] G. Dieck Kattas and M. Small, "Building computational models of swarms from simulated positional data," *Lecture Notes in Computer Science*, vol. (to appear), 2011.

[13] J. Rissanen, *Stochastic Complexity in Statistical Inquiry Theory*. World Scientific Publishing Co., Inc., 1989.