

Research Article

Identification of V-Formations and Circular and Doughnut Formations in a Set of Moving Entities with Outliers

**Francisco Javier Moreno Arboleda, Jaime Alberto Guzmán Luna,
and Sebastian Alonso Gomez Arias**

Universidad Nacional de Colombia, Sede Medellín, Bloque M8A, Medellín, Colombia

Correspondence should be addressed to Jaime Alberto Guzmán Luna; jaguzman@unal.edu.co

Received 1 November 2013; Revised 8 February 2014; Accepted 24 February 2014; Published 10 April 2014

Academic Editor: J.-C. Cortés

Copyright © 2014 Francisco Javier Moreno Arboleda et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Diverse movement patterns may be identified when we study a set of moving entities. One of these patterns is known as a V-formation for it is shaped like the letter V. Informally, a set of entities shows a V-formation if the entities are located on one of their two characteristic lines. These lines meet in a position where there is just one entity considered the leader of the formation. Another movement pattern is known as a circular formation for it is shaped like a circle. Informally, circular formations present a set of entities grouped around a center in which the distance from these entities to the center is less than a given threshold. In this paper we present a model to identify V-formations and circular formations with outliers. An outlier is an entity which is part of a formation but is away from it. We also present a model to identify doughnut formations, which are an extension of circular formations. We present formal rules for our models and an algorithm for detecting outliers. The model was validated with NetLogo, a programming and modeling environment for the simulation of natural and social phenomena.

1. Introduction

Diverse movement patterns may be identified when we study a set of moving entities, for example, a flock of birds [1] and a school of fish [2]. One of these patterns is known as a V-formation for it is shaped like the letter V; see Figure 1. Another movement pattern is known as a circular formation for it is shaped like a circle; see Figure 2.

Informally, a set of entities shows a V-formation if the entities are located on one of their two characteristic lines. The lines meet in a position where there is just one entity considered the leader of the formation [3]. Several authors have analyzed V-formations. In [4, 5], there is an attempt to explain from a physical point of view the reasons why certain species of birds, such as Canadian geese (*Branta canadensis*), red knots (*Calidris canutus*), and plovers (*Calidris alpina*), tend to fly this way.

Other authors try to simulate V-formations at a computational level. For instance, Nathan and Barbosa [6] propose a model based on rules that allows us to generate V-formations depending on specific parameters. The authors validated their

model using NetLogo [7], a programming and modeling environment to simulate natural and social phenomena.

On the other hand, a circular formation is a set of entities grouped around a common center and where the entities' distance to the center is less than a given threshold. Regarding related works with circular formations, we identified the following.

In [8] the authors experimented with a set of data referring to the movement of different animal species. It was found that despite being in different ecosystems, species follow similar behavioral patterns. The authors also tried to model general grouping behaviors of fish, birds, insects, and even people. One of these behaviors is circular formation in which they identified physical forces: attraction, repulsion, alignment, and frontal interaction.

On the other hand, researchers in the field of robotics and in control theory, inspired by social grouping phenomena and by the patterns of birds and fish, have developed applications to coordinate the movement of multivehicle systems. Among these patterns are circular [9] and V-formations; see Figure 3.



FIGURE 1: V-formation in birds. Source: [9].



FIGURE 2: Circular formation in a fish bank. Source: [21].

On the other hand, regardless of the type of formation, Reynolds [2, 10] proposes a computer model of coordinated animal motion for bird flocks and fish schools. His model considers three simple rules that act upon individuals: (i) separation: steer to avoid crowding local flockmates, (ii) alignment: steer towards the average heading of local flockmates, and (iii) cohesion: steer to move toward the average position of local flockmates.

Although the previous works allow the simulation of a set of moving entities, they are not aimed at the explicit identification of V-formations and circular formations. The identification of these types of formations may be useful in fields as zoology, to analyze the movement of birds [3] and fish [11], and in the military and videogames where squadrons of combat planes, ships, and robots usually assume these types of formations [3, 12]. V-formations also usually appear in stock markets (stock prices) [13]. In this direction the model of Andersson et al. [14], although it is not aimed at the explicit identification of formations, identifies an entity leader in a set of moving entities. We use their model as the basis for identifying V-formations and circular formations, as we will show in our paper.

This paper is organized as follows. In Section 2, we present our models for V-formations and circular and doughnut formations, which are an extension of circular formations. In Section 3, we present support for outliers in our models. In Section 4, we present experiments. In Section 5, we conclude the paper and propose future works.

2. V-Formations and Circular and Doughnut Formations

2.1. Andersson's Model. Next, we present the essential elements of Andersson's model [14].

Consider a set E of n entities $\{e_1, e_2, \dots, e_n\}$ that move in a space, usually in a region during a time interval $[t_1, t_f]$. This space is represented by the Euclidian plane. Time is represented continuously. Tp denotes an infinite set of time-points: $\{t \mid t \in [t_1, t_f]\}$. On the other hand, Ts denotes a finite set of discrete time-points $\{t_1, t_2, \dots, t_f\}$. Every $t_i \in Ts$ represents a time-step and corresponds to a time-point in which the position of a moving entity (and perhaps other data) was recorded. A unit-time-interval is an open interval between two consecutive time-steps, that is, (t_{i-1}, t_i) , for all $t_i \in Ts, i \neq 1$.

The coordinates of an entity at a time-point are given by a pair of functions $xpos$ and $ypos$, both having the signature $E \times Tp \rightarrow \text{Real}$. Andersson assumes that between two consecutive time-steps an entity moves along a straight line with constant direction and constant velocity. This assumption has been adopted in other works (see [15–17] for a discussion). Therefore, in such a model in order to be practical enough, the recording frequency of the positions of moving objects should be adapted to the nature of each domain. For example, when detecting bird formations the elapsed time between two consecutive time-steps should not be longer than 30 seconds.

Consider an entity e_i at a time-point t , where $t_{x-1} < t < t_x, t_x \in Ts, x \neq 1$; then the direction of e_i at t , denoted by $d(e_i, t)$, is the angle specified by the straight line segment that goes from $(xpos(e_i, t_{x-1}), ypos(e_i, t_{x-1}))$ to $(xpos(e_i, t_x), ypos(e_i, t_x))$; see Figure 4. The angle $d(e_i, t)$ is between $[0, 2\pi)$. Andersson declares the direction of an entity at a time-step t_x to be *undefined*, because at a time-step an entity might change its direction (for more details see [14]).

The front region of an entity e_i is a region associated with e_i at a time-point t , where $t_{x-1} < t < t_x, t_x \in Ts, x \neq 1$, which represents the perception region of an entity, for example, a visual or an auditive range. The front region of e_i is defined as follows: consider three straight line segments s_0, s_1 , and s_2 , each of length r . Each segment has an endpoint at $(xpos(e_i, t), ypos(e_i, t))$. The direction of the segment s_0 is $d(e_i, t)$. Segments s_1 and s_2 form angles of $\alpha/2$ and $-\alpha/2$ ($0 \leq \alpha < 2\pi$) with regard to segment s_0 , respectively. The circular sector (a wedge-shaped region) with a radius r ($r \geq 0$), delimited by s_1 and s_2 , makes up the front region of e_i at t and it is denoted by $front(e_i, t)$; see Figure 5.

An entity e_j is in front of an entity e_i , at a time-point t ($t \in Tp, t \notin Ts$), where $t_{x-1} < t < t_x, t_x \in Ts, x \neq 1$, if e_j is in the front region of e_i , that is, if $(xpos(e_j, t), ypos(e_j, t))$ is inside $front(e_i, t)$; this is denoted $e_j \in front(e_i, t)$ and we say that e_i follows e_j (e_i is a follower of e_j) at t . We consider an additional restriction to establish that one entity follows another: let $\beta \in [0, \pi]$; then entity e_i follows entity e_j at t if (i) $e_j \in front(e_i, t)$ and (ii) $\|d(e_i, t) - d(e_j, t)\| \leq \beta$; see Figure 6. An entity e_i is said to follow an entity e_j during a time interval $[t_a, t_b]$, where t_a and t_b are time-points in Tp , if and only if e_i follows e_j at t , for all $t \in [t_a, t_b], t \notin Ts$.

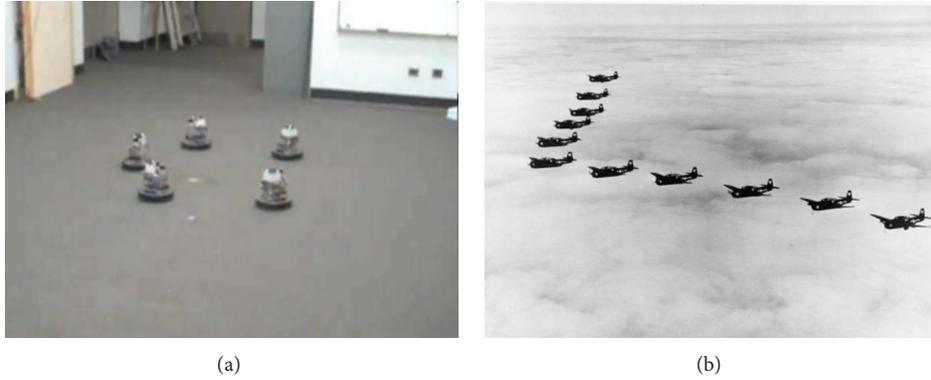


FIGURE 3: Robots in circular formation and planes in V-formation. Sources: [11, 12].

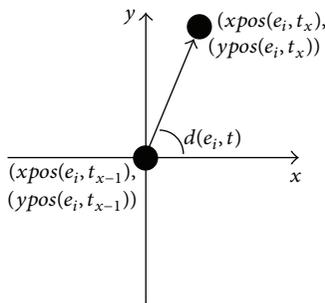


FIGURE 4: $d(e_i, t)$: angle of entity e_i at a time-point t .

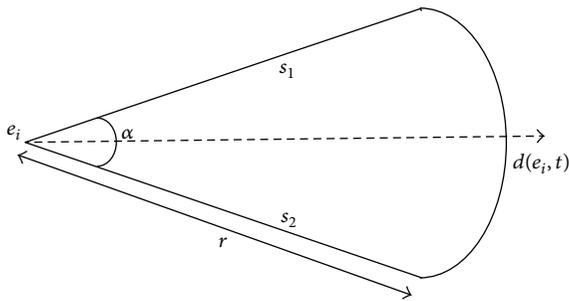


FIGURE 5: Front region of an entity e_i at a time-point t .

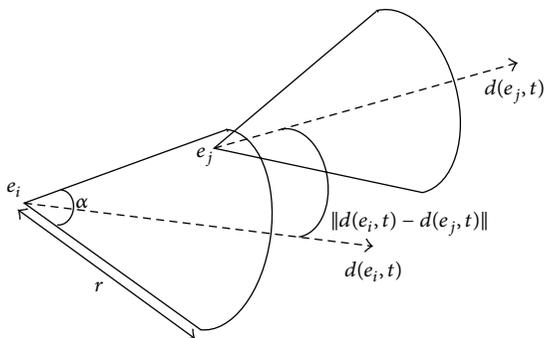


FIGURE 6: Entity e_i follows entity e_j at a time-point t .

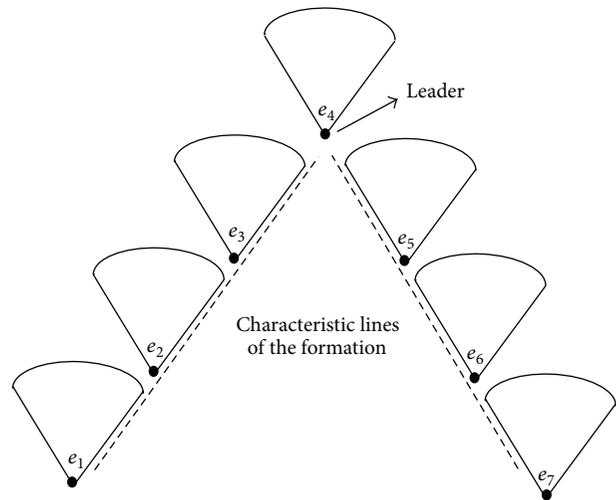


FIGURE 7: An example of a V-formation at a time-point t .

An entity e_i is said to be a *leader* of a formation during a time interval $I = [t_a, t_b]$, where t_a and t_b are time-points in Tp , if (i) e_i does not follow any entity during I and (ii) e_i is at least followed by m entities at each time-point $t \in I, t \notin Ts$. It is said that there is a *leadership pattern* if e_i is a leader of at least m entities for at least k unit-time-intervals.

2.2. V-Formations. Consider the formation in Figure 7 where entities displayed a V-formation at a time-point t , and e_4 is the leader entity. Note that if we applied Andersson's model [14] to this formation, we would observe that no entity contains e_4 in its front region. Therefore, this example shows that Andersson's model [14] is inadequate to identify leadership patterns in these types of formations, unless we consider a wide front region for each entity, as shown in Figure 8. Thus, the entities on the extremes, e_1 and e_7 , will require a wide front region to be able to contain the leader entity e_4 ; this is not a realistic assumption in V-formations since an entity's visual field is not usually that wide [4].

In the following section, we extend Andersson's model to identify leadership patterns in V-formations.

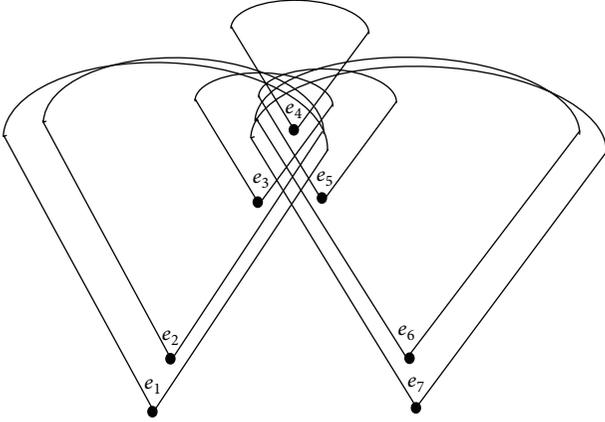


FIGURE 8: A V-formation: entities e_1 and e_7 would require a wide front region in order to contain the entity leader e_4 .

2.3. *V-Formation Model.* Let $F = \{e_1, e_2, \dots, e_n\}$ be a set of moving entities at a time-point t , $\text{card}(F) \geq 3$. F is a V-formation at t if the following hold.

- (i) There exist nonempty subsets A and B of F such that $\text{card}(A) \geq 2$, $\text{card}(B) \geq 2$, $A \cup B = F$, and $A \cap B = \{e_k\}$, where e_k is said to be the leader entity of the formation.
- (ii) Entities in A tend to form a straight line l_1 at t .
- (iii) Entities in B tend to form a straight line l_2 at t .
- (iv) Straight lines l_1 and l_2 meet at position $(x_{\text{pos}}(e_k, t), y_{\text{pos}}(e_k, t))$.
- (v) $\angle ap_t > 0$ (the smallest angle defined by straight lines l_1 and l_2 at t).

Regarding conditions (ii) and (iii), to establish if a set of entities tend to form a straight line, we use Pearson's correlation coefficient r [18]. Thus, given a set of points $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, r indicates how much they adjust to a straight line (linearity degree). $r \in (-1, 1)$; if $|r| \approx 1$, then points tend to form a straight line. r is calculated as shown in (1). A threshold μ_p may be specified by the user to indicate the linearity degree required for the lines of the formation; that is, $|r| \geq \mu_p$:

$$r = \frac{n * (\sum_{i=1}^n x_i * y_i) - \sum_{i=1}^n x_i * \sum_{i=1}^n y_i}{\sqrt{[n * \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2] * [n * \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}} \quad (1)$$

To obtain the equation for each straight line ($y = mx + b$) characteristic of the formation (l_1 and l_2) we may apply the equations that correspond to the straight line which most suits a set of points $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$; see (2). Table 1 presents the results for the formation in Figure 9:

$$m = \frac{n * (\sum_{i=1}^n x_i * y_i) - \sum_{i=1}^n x_i * \sum_{i=1}^n y_i}{n * \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (2)$$

$$b = \frac{\sum_{i=1}^n y_i - m * \sum_{i=1}^n x_i}{n}.$$

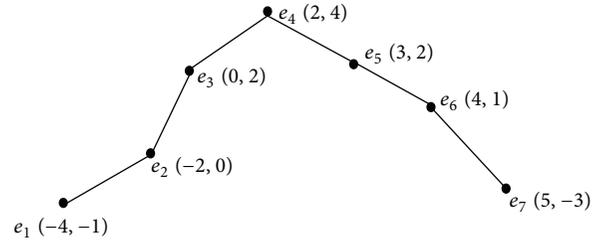


FIGURE 9: A formation of entities at a time-point t .

Regarding condition (v), $\angle ap_t$ is calculated as follows: we get straight lines l_1 and l_2 from the formation and find the smallest angle between them as follows.

a is the positive angle in $[0, 2\pi)$ that is specified by the line segment that goes from p_k to p_i , where p_k is the position of e_k (leader entity) according to equation for l_1 and p_i is the position of $e_i \in A$, $e_i \neq e_k$ according to equation for l_1 .

b is the positive angle in $[0, 2\pi)$ that is specified by the line segment that goes from p_k to p_j , where p_k is the position of e_k (leader entity) according to equation for l_2 and p_j is the position of $e_j \in B$, $e_j \neq e_k$, according to equation for l_2 .

Let $w = |a - b|$; then $\angle ap_t = w$ if $w \leq \pi$, and $\angle ap_t = 2\pi - w$, otherwise. For example, in the formation in Figure 9, $a = 0.7$ rad (40.36°), $b = 5.14$ rad (294.44°), and $w = 4.43$ rad (254.08°); thus, $\angle ap_t = 1.84$ rad (105.91°).

In an analogous way to leadership patterns, we say that there is a *V-formation pattern* if the set F of moving entities shows a V-formation for at least k unit-time-intervals.

2.4. *Circular and Doughnut Formations.* In this section, we present a formal model to identify circular formations. We also present doughnut formations, which are an extension of circular formations.

2.4.1. *Circular Formation Model.* Let $F = \{e_1, e_2, \dots, e_n\}$ be a set of moving entities at a time-point t . Let C be the *minimum circle* that encloses the entities in F and let C_r be its radius and let (C_x, C_y) be its center (Welzl [19] and Megiddo [20] have showed that this circle can be found in linear time). F is a circular formation at t if

- (i) $C_r \leq R$ (a user-given radius with center at (C_x, C_y)),
- (ii) the minimum number of members in the formation is N_{\min} .

In Figure 10, we show a circular formation with 9 members.

2.4.2. *Doughnut Formation Model.* Let $F = \{e_1, e_2, \dots, e_n\}$ be a set of moving entities at a time-point t . Let C be the minimum circle that encloses the entities in F and let C_r be its radius and (C_x, C_y) its center. F is a doughnut formation at t if

TABLE 1: Applying our model to the formation in Figure 9.

Characteristic line	Entities that make up the characteristic line	Characteristic line equation	Pearson's coefficient	Entity coordinates: original and calculated using the characteristic line equation
l_1	$A = \{e_1, e_2, e_3, e_4\}$	$y = 2.1 + 0.85x$	0.99	e_1 : (-4, -1) and (-4, -1.3) e_2 : (-2, -0) and (-2, 0.4) e_3 : (0, 2) and (0, 2.1) e_4 : (2, 4) and (2, 3.8)
l_2	$B = \{e_4, e_5, e_6, e_7\}$	$y = 8.7 - 2.2x$	0.97	e_4 : (2, 4) and (2, 4.3) e_5 : (3, 2) and (3, 2.1) e_6 : (4, 1) and (4, 0.1) e_7 : (5, -3) and (5, -2.3)

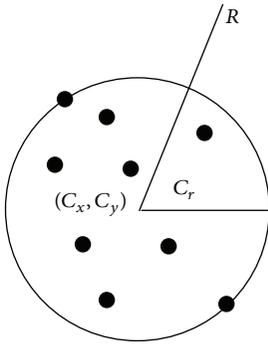


FIGURE 10: A circular formation.

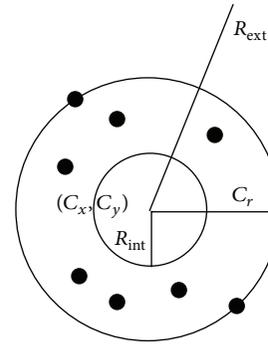


FIGURE 11: A doughnut formation.

- (i) $C_r \leq R_{ext}$ (external radius, a user-given radius with center at (C_x, C_y)),
- (ii) for each entity of the formation its distance d regarding (C_x, C_y) is greater than R_{int} (internal radius, a user-given radius with center at (C_x, C_y) , $R_{int} \leq C_r$),
- (iii) the minimum number of members in the formation is N_{min} .

Note that a circular formation is a particular case of a doughnut formation, where $R_{int} = 0$. In Figure 11, we show a doughnut formation with 8 members.

Circular and doughnut patterns are also defined in an analogous way to a V-formation pattern.

3. Outliers

Informally, for V-formation an outlier is an entity which is away from its characteristic lines, and for circular formations it is an entity found beyond the radius of the formation (R).

3.1. Outliers in V-Formations. There are sets of entities which tend to display a V-formation; they may have at a time-point t entities which are away from their characteristic lines and which, therefore, affect Pearson's coefficient. These entities are called outliers [22, 23].

There are numerous methods to detect outliers in different domains [24]. Listing 1 presents an algorithm that receives an array of m entities ($lineMembers$) which form

a characteristic line of a formation at a time-point t . The algorithm determines if, after removing a maximum number of entities on the given array, the Pearson's coefficient surpasses a given threshold μ_p . For example, if it is permitted to remove a maximum of two entities from a characteristic line of the formation, it is considered that the entities in Figure 12 display a V-formation having two outliers on each characteristic line. Then, the algorithm receives the minimum value of Pearson's coefficient μ_p which should be met and a maximum percentage of entities ($percentageOutliers$) which are permitted to be removed from the array of entities. This percentage is calculated with regard to the total number of entities m .

Example 1. Consider the set of entities in Figure 13 which would form a straight line if entities e_2 and e_4 were not considered; that is, e_2 and e_3 are outliers.

Input algorithm parameters:

$lineMembers = [e_1, e_2, \dots, e_8]$, $\mu_p = 0.99$ and $percentageOutliers = 20\%$.

Step 1. Equation of l : $y = -0.754x + 2.01$.

Step 2. $nbrOutliers = \lceil 8 * 0.2 \rceil = 2$.

Step 3. e_2 and e_4 .

Step 4. $lineMembers = [e_1, e_3, e_5, e_6, e_7, e_8]$.

ALGORITHM: Outlier detection on a characteristic line**INPUT:** $lineMembers = [e_1, e_2, \dots, e_m]$ // Array of m entities μ_p // Threshold for Pearson's coefficient $percentageOutliers$ // Maximum percentage of outliers permitted on $lineMembers$ **OUTPUT:** $outliers = []$ // Array of outliers**BEGIN**

- (1) Obtain the equation of the line l which most suits the positions (x, y) of the entities in $lineMembers$, see (2).
- (2) Get the maximum number of outliers permitted on $lineMembers$: $nbrOutliers = \lceil m * percentageOutliers/100 \rceil$.
- (3) Find the $nbrOutliers$ entities in $lineMembers$ which have the maximum distance to l .
- (4) Remove from $lineMembers$ the entities found in Step 3.
- (5) Calculate Pearson's coefficient using positions (x, y) of each entity in $lineMembers$:
 $Pearson = PearsonCoefficient(lineMembers)$.
- (6) If $Pearson \geq \mu_p$ then return in $outliers$ the entities found in Step 3.

END

LISTING 1: Outlier detection algorithm.

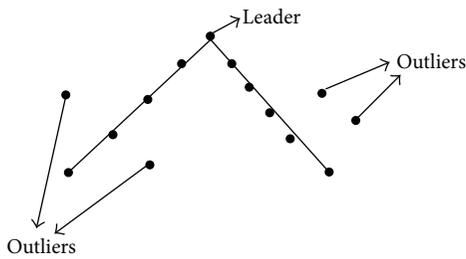


FIGURE 12: A V-formation with two outliers.

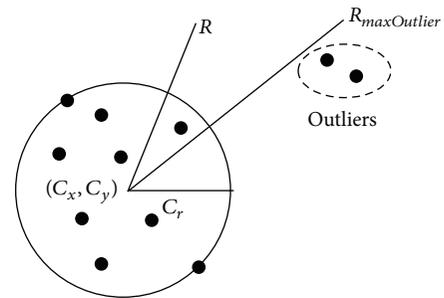


FIGURE 14: A circular formation with two outliers.

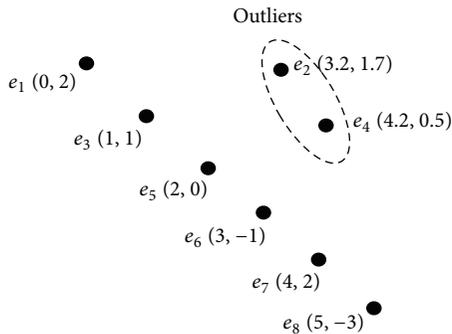


FIGURE 13: Characteristic line of a V-formation with two outliers.

Step 5. $Pearson = PearsonCoefficient(lineMembers) = 1$.

Step 6. $1 \geq 0.99$; then return $outliers = [e_2, e_4]$.

The algorithm determines that entities e_2 and e_4 are outliers. Then, we conclude that entities e_1, e_3, e_5, e_6, e_7 , and e_8 tend to form a straight line with Pearson's coefficient greater than 0.99.

3.2. Outliers in Circular Formations. Let $F = \{e_1, e_2, \dots, e_n\}$ be a set of moving entities at a time-point t . Consider an entity $e_i \in F$ such that $F - \{e_i\}$ is a circular formation and

the distance from e_i to (C_x, C_y) (center of the minimum circle that encloses the entities in $F - \{e_i\}$) is greater than R at t . This entity may be considered as a member of the formation which is temporally away from it; that is, e_i is an outlier. To identify this type of entities, we introduce a parameter $R_{maxOutlier}$, where $R_{maxOutlier} > R$. An entity is considered an outlier at a time-point t if its distance d to (C_x, C_y) is greater than R and less than $R_{maxOutlier}$; see Figure 14.

Since the distancing of the outlier entity from the formation is temporal, an analyst may introduce a second parameter $T_{maxTimeSeparation}$ to control the maximum continuous time of permitted distancing. That is, if an entity separates from a circular formation at a time-point t , then, to be considered an outlier, it will have to reincorporate ($d \leq R$) to the formation before $t + T_{maxTimeSeparation}$. This same aspect can also be considered for outliers in V-formations.

An analyst can also specify a maximum permitted number of outliers $maxNumberOutliers$ in the formation. This value can be calculated from a percentage ($percentageOutliers$) with regard to the total number of entities in the formation.

3.3. Outliers in Doughnut Formations. Let $F = \{e_1, e_2, \dots, e_n\}$ be a set of moving entities at a time-point t . Consider an entity

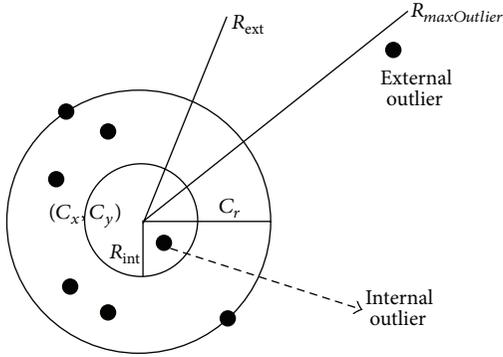


FIGURE 15: Outliers in doughnut formations.

TABLE 2: Parameters to generate V-formations in NetLogo using Nathan's model.

Number of entities	15
Vision parameters	
Vision distance	9 patches
Vision cone	1.79 rad (103°)
Obstruction cone	0.75 rad (43°)
Movement parameters	
Base speed	0.2
Speed change factor	0.15
Updraft distance	9 patches
Too close	3.1
Max. turn	0.14 rad (8°)

$e_i \in F$ such that $F - \{e_i\}$ is a doughnut formation. For doughnut formations we consider two types of outliers: external and internal. An entity e_i is considered an external outlier at t if its distance d to (C_x, C_y) (center of the minimum circle that encloses the entities in $F - \{e_i\}$) is greater than R_{ext} and less than $R_{maxOutlier}$, where $R_{maxOutlier} > R_{ext}$. On the other hand, e_i is considered an internal outlier at t if d is less than R_{int} ; see Figure 15.

Analogously to circular formations, $T_{maxTimeSeparation}$ and $maxNumberOutliers$ values may be specified, not only for external outliers but also for internal outliers.

4. Experiments and Discussion

4.1. Experiment for V-Formations. For our experiment, we used Nathan's model [6] implemented in NetLogo in order to generate V-formations. NetLogo is a programmable modeling environment for simulating natural and social phenomena that have been used in previous works cite9, 26.

We worked with a population of 15 entities. The dimensions of the Euclidian plane where the entities move were $x \in [-35, 35]$ and $y \in [-35, 35]$, where the unit of measure for distance in NetLogo is a *patch*. To generate V-formations, we used the parameters in Table 2. These parameters allow us to define aspects related to vision and movement of the entities in Nathan's model.

For outlier detection, we considered for our model the following parameters: $\mu_p = 0.90$ and $percentageOutliers = 20\%$. Our experiment began from tick 40199 and ended at tick 40400, because in Nathan's model the generation of V-formations requires a stabilization time (tick 40199 in our experiment, a *tick* is the unit of measure for time in NetLogo). Time-steps were taken in intervals of 20 ticks, for if we consider a smaller step, a change in the position and direction of an entity would be practically imperceptible. Thus, the total number of steps was 10. Figure 16 shows the position of entities at ticks 40200, 40220, 40380, and 40400 and Table 3 summarizes our results.

4.1.1. Analysis of V-Formation Results. According to the results shown in Table 3, our model classified several sets of entities as V-formations. These results agree with what was expected since Nathan's model had already stabilized and V-formations were being generated effectively for those times. Moreover, it was necessary to apply the outlier detection algorithm just in two ticks (at ticks 40380 and 40400, see entities in brackets in Table 3, second column). On the other hand, we found that formations showed a high linearity degree since Pearson's coefficients in all cases were greater than 0.92.

We also applied Andersson's model. We considered $m = 2$ and for the front region of each entity the parameters in Table 2; that is, $s_0 = 9$ patches (vision distance) and $\alpha = 1.79$ rad (103°) (vision cone). Andersson's model only identified formations of two and three entities. This is reasonable as explained in Section 2.2 (see Figure 8). In addition, his model does not give information about the shape of the formation that it identifies; even if we considered a wider front region, his model would not indicate that a formation is, for example, V-shaped.

4.2. Experiments for Circular and Doughnut Formations. For experiments with circular and doughnut formations, we also worked in NetLogo and used Wilensky's model [26] (a model inspired in Reynolds' work [10]), which generates randomized formations of entities in NetLogo.

We worked with a population of 102 entities. The dimensions of the Euclidian plane where the entities move were $x \in [-35, 35]$ and $y \in [-35, 35]$. To generate formations, we used the parameters in Table 4.

The model was executed on 1200 consecutive ticks (one run) and we conducted an analysis of circular formations and doughnut formations every 400 ticks. We considered that a set of entities showed a circular/doughnut pattern if the set showed a circular/doughnut formation during all the run (1200 ticks). A total of 100 runs were conducted. The parameters used to detect circular and doughnut formations are shown in Tables 5 and 6. Table 7 summarizes our results for the 100 runs. Figure 17 shows the results for one of the runs.

4.2.1. Analysis of Experiments for Circular and Doughnut Formations. According to the results shown in Table 7,

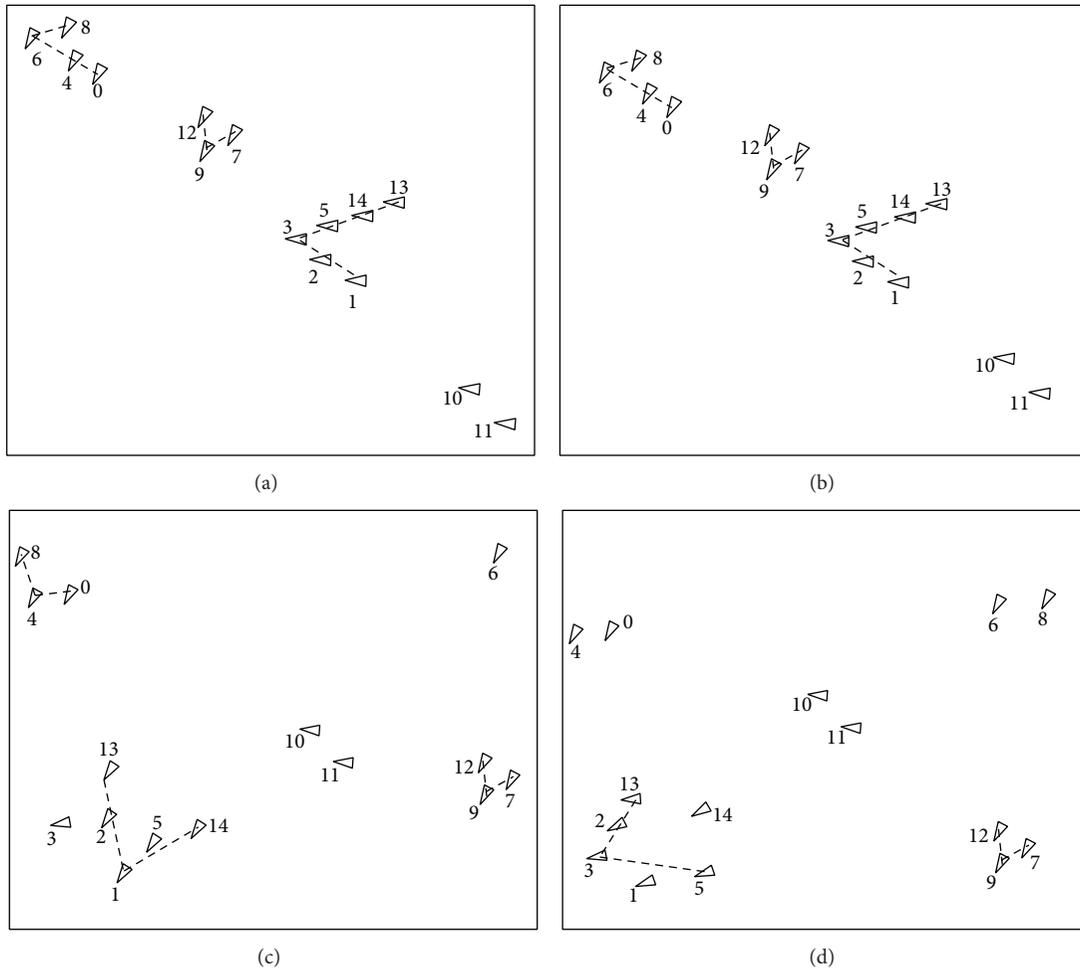


FIGURE 16: V-formations. Position of entities at ticks (a) 40200, (b) 40220, (c) 40380, and (d) 40400.

the model classified several sets of entities as circular formations and around three remained throughout the 1200 ticks of a run. Regarding doughnut formations, the model detected around two formations that remained during the 1200 ticks of a run. These results agree with what was expected since Wilensky's model had already stabilized in NetLogo.

With respect to Andersson's model, we considered $m = 3$ and the parameters in Table 2. His model identified 752 formations during the 100 runs, a value that is greater than the total number of circular and doughnut formations identified during the 100 runs. This is reasonable because with $m = 3$ (note that the average number of entities in his formations was four) it is expected that any set of at least three entities that are close to each other (considering $s_0 = 9$ patches) will be identified as a formation. However, as our previous experiments with V-formations, his model is unable to indicate the shape of a formation that it identifies.

5. Conclusions and Future Work

In this paper, we propose two models:

- (I) a model to identify V-formations with outliers. The model considers the location of entities to determine if they form this type of formation during a time interval;
- (II) a model to identify circular formations with outliers. The model considers the location of entities to determine if they form this type of formation during a time interval. In addition, we proposed an extension to identify doughnut formations with outliers.

The rules for model (I) are flexible since they allow V-formations which are not necessarily perfectly aligned, in accordance with the real world. Furthermore, we consider outliers in V-, circular, and doughnut formations, that is, members of the formation that could be temporarily away from it. Our experimental results in NetLogo showed that our models identified these types of formations in such simulation environment where they were generated. We also showed that when we applied Andersson's model, his model did not give information about the shape of the formation that it identified; that is, his model suffers from lack of semantic information about shape formations.

Regarding future work, we plan to conduct a series of experiments in the stock market where V-formations usually

TABLE 3: V-formations: results of the experiment in NetLogo.

Tick	Formations identified by our model (leader bolded, outliers in brackets)	Pearson's coefficient straight line 1	Pearson's coefficient straight line 2	Opening angle	V-formation	Formations identified by Andersson's model (leader bolded)
40200	{1, 2, 3 , 5, 13, 14}	{1, 2, 3} → 0.99	{3, 5, 13, 14} → 0.99	1.5 rad (86.49°)	Yes	{2, 3, 5} {13, 14 }
	{7, 9 , 12}	{7, 9} → 1	{9, 12} → 1	1.33 rad (75.92°)	Yes	{7, 9 , 12}
	{0, 4, 6 , 8}	{0, 4, 6} → 0.98	{6, 8} → 1	0.52 rad (29.97°)	Yes	{0, 4 } { 6 , 8}
	{10, 11}	NA	{10, 11} → 1	NA	No	{ 10 , 11}
40220	{1, 2, 3 , 5, 13, 14}	{1, 2, 3} → 0.99	{3, 5, 13, 14} → 0.99	1.51 rad (86.49°)	Yes	{2, 3, 5} {13, 14 }
	{7, 9 , 12}	{7, 9} → 1	{9, 12} → 1	1.33 rad (75.92°)	Yes	{7, 9 , 12}
	{0, 4, 6 , 8}	{0, 4, 6} → 0.98	{6, 8} → 1	0.52 rad (29.97°)	Yes	{0, 4 } { 6 , 8}
	{10, 11}	NA	{10, 11} → 1	NA	No	{ 10 , 11}
⋮						
40380	{2, {3}, 13, 1 , 5, 14}	{2, 13, 1} → 1	{1, 5, 14} → 0.95	0.79 rad (45.13°)	Yes	{ 1 , 5}
	{7, 9 , 12}	{7, 9} → 1	{9, 12} → 1	1.33 rad (75.92°)	Yes	{7, 9 , 12}
	{0, 4 , 8}	{0, 4} → 1	{4, 8} → 1	2.68 rad (153.74°)	Yes	No
	{10, 11}	NA	{10, 11} → 1	NA	No	{ 10 , 11}
	{6}	NA	NA	NA	NA	NA
40400	{{1}, 5, 3 , 2, 13, {15}}	{5, 3} → 0.93	{3, 2, 13} → 0.98	0.86 rad (49.5°)	Yes	{2, 3 } { 1 , 5}
	{7, 9 , 12}	{7, 9} → 1	{9, 12} → 1	1.33 rad (75.92°)	Yes	{7, 9 , 12}
	{0, 4}	NA	{0, 4} → 1	NA	No	No
	{10, 11}	NA	{10, 11} → 1	NA	No	{ 10 , 11}
	{6, 8}	NA	{6, 8} → 1	NA	No	No

TABLE 4: Parameters to generate formations in NetLogo using Wilensky's model.

Number of entities	102
Vision parameters	
Vision distance	3 patches
Minimum Separation	1 patch
Movement parameters	
Maximum angle of rotation	0.08 rad (4.75°)
Maximum following angle	0.04 rad (2.50°)
Maximum angle of separation	0.06 rad (3.5°)

TABLE 5: Parameters for circular formation detection in NetLogo.

Minimum number of entities (N_{min})	5
Maximum distance (R) from the center of the minimum enclosing circle to the entities	15 patches
Maximum percentage of outliers permitted (<i>percentageOutliers</i>)	30%
$R_{maxOutlier}$	30 patches
$T_{maxTimeSeparation}$	400 ticks

TABLE 6: Parameters for doughnut formation detection in NetLogo.

Minimum number of entities (N_{min})	5
Maximum distance (R_{ext}) from the center of the minimum enclosing circle to the entities	30 patches
Maximum distance (R_{int}) from the center of the minimum enclosing circle to the entities	10 patches
Maximum percentage of outliers permitted	30%
$R_{maxOutlier}$	50 patches
$T_{maxTimeSeparation}$	400 ticks

appear [13]. Moreover, we plan to extend our models to identify *isolated entities*, that is, entities that even if they are considered members of a formation, they follow their own path and do not influence the path of other entities [27]. We also plan to propose models to identify patterns such as *convergence*, that is, a set of entities that converge or approach a place; *divergence*, that is, a set of entities that disperse or move away from a place [28], and *self-organization*, that is, a set of entities that move as a formation without there being a leader or an entity guiding the rest [29] or that this leader is unknown to the members [30–32].

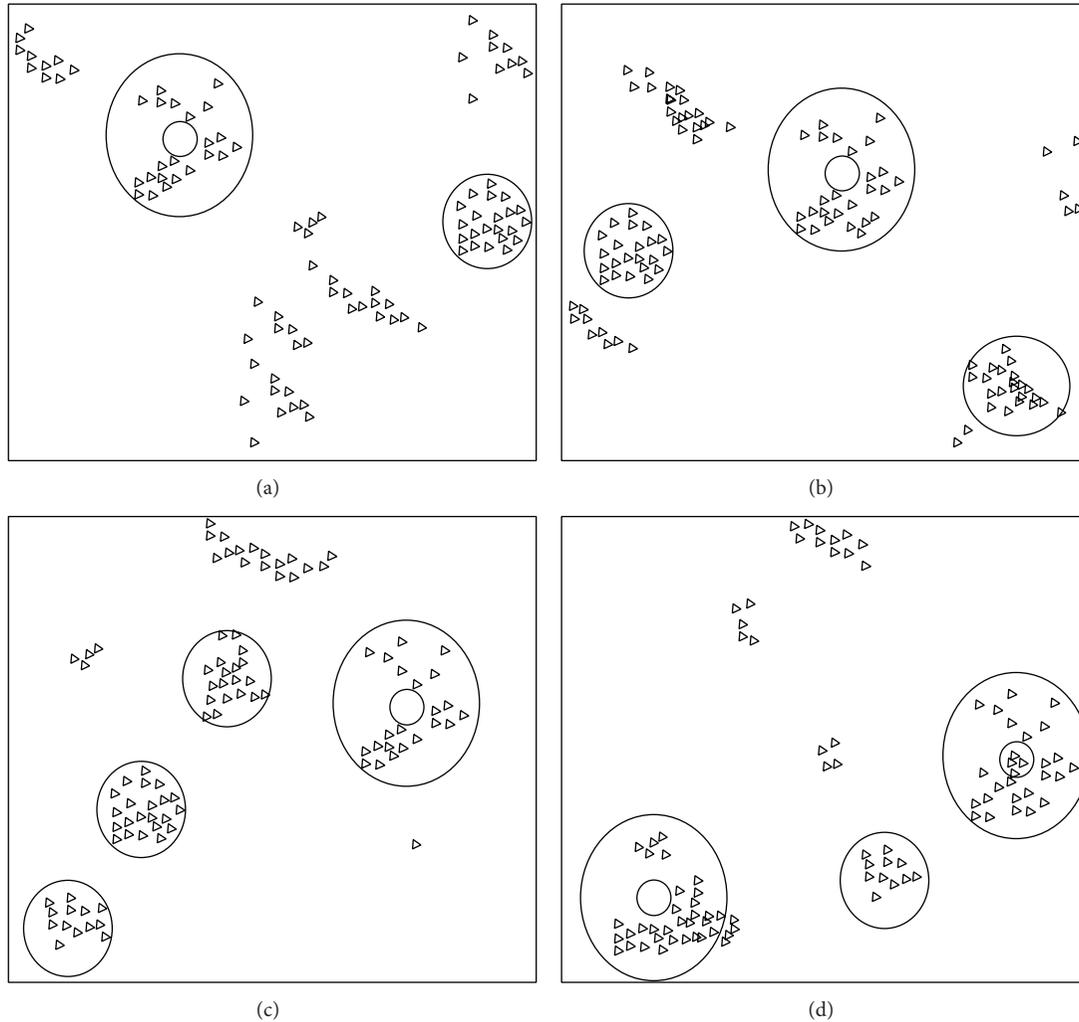


FIGURE 17: Circular and doughnut formations. Position of entities at ticks (a) 0, (b) 400, (c) 800, and (d) 1200.

TABLE 7: Circular and doughnut formations: results of the experiments in NetLogo.

	Circular	Doughnut	Andersson's model
Total number of patterns identified during the 100 runs	332	198	752
Average number of patterns identified in each run (1200 ticks)	3	2	8
Average number of entities in each pattern	13	22	4
Average number of outliers	2	Internal: 3 External: 2	NA

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This paper presents preliminary results of the project “Apoyo al Grupo de Sistemas Inteligentes Web-SINTELWEB” with Quipú code 205010011129, developed at the Universidad Nacional de Colombia, Sede Medellín.

References

- [1] S. Dodge, R. Weibel, and A. K. Lautenschütz, “Towards a taxonomy of movement patterns,” *Information Visualization*, vol. 7, pp. 240–252, 2008.
- [2] C. W. Reynolds, “Flocks, herds, and schools: a distributed behavioral model,” *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.
- [3] F. Cattivelli and A. H. Sayed, “Self-organization in bird flight formations using diffusion adaptation,” in *Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '09)*, pp. 49–52, December 2009.
- [4] P. Seiler, A. Pant, and K. Hedrick, “Analysis of bird formations,” in *Proceedings of the 41st IEEE Conference on Decision and Control (CDC '02)*, vol. 1, pp. 118–123, December 2002.

- [5] C. M. Sewatkar, A. Sharma, and A. Agrawal, "A first attempt to numerically compute forces on birds in v formation," *Artificial Life*, vol. 16, no. 3, pp. 245–258, 2010.
- [6] A. Nathan and V. C. Barbosa, "V-like formations in flocks of artificial birds," *Artificial Life*, vol. 14, no. 2, pp. 179–188, 2008.
- [7] U. Wilensky and W. Rand, "Making models match: replicating an agent-based model," *Journal of Artificial Societies and Social Simulation*, vol. 10, no. 4, p. 2, 2007.
- [8] R. Lukeman, Y.-X. Li, and L. Edelstein-Keshet, "Inferring individual rules from collective behavior," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 28, pp. 12576–12580, 2010.
- [9] "Jets 'should copy birds to save energy'—telegraph" 2013, <http://www.telegraph.co.uk/news/uknews/8846224/Jets-should-copy-birds-to-save-energy.html>.
- [10] C. Reynolds, *Boids: Background and Update*, 2001, <http://www.red3d.com/cwr/boids>.
- [11] W. L. Romey, "Individual differences make a difference in the trajectories of simulated schools of fish," *Ecological Modelling*, vol. 92, no. 1, pp. 65–77, 1996.
- [12] N. Moshtagh, N. Michael, A. Jadbabaie, and K. Daniilidis, "Bearing-only control laws for balanced circular formations of ground robots," in *Proceedings of Robotics: Science and Systems IV*, MIT Press, 2008.
- [13] K. Nesbitt, "Interacting with stock market data in a virtual environment," in *Data Visualization*, pp. 273–282, Springer, New York, NY, USA, 2001.
- [14] M. Andersson, J. Gudmundsson, P. Laube, and T. Wolle, "Reporting leadership patterns among trajectories," in *Proceedings of the ACM Symposium on Applied Computing (SAC '07)*, pp. 3–7, New York, NY, USA, March 2007.
- [15] R. H. Güting and M. Schneider, *Moving Objects Databases*, Elsevier, San Diego, Calif, USA, 2005.
- [16] J. Gudmundsson and M. Van Kreveld, "Computing longest duration flocks in trajectory data," in *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS '06)*, pp. 35–42, November 2006.
- [17] M. Benkert, J. Gudmundsson, F. Hübner, and T. Wolle, "Reporting flock patterns," *Computational Geometry*, vol. 41, no. 3, pp. 111–125, 2008.
- [18] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.
- [19] E. Welzl, "Smallest enclosing disks (balls and ellipsoids)," in *New Results and New Trends in Computer Science*, vol. 555 of *Lecture Notes in Computer Science*, pp. 359–370, Springer, Berlin, Germany, 1991.
- [20] N. Megiddo, "Linear-time algorithms for linear programming in R^3 and related problems," in *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science (SFCS '08)*, pp. 329–338, IEEE, New York, NY, USA, 1982.
- [21] S. Wilde, *Thoughts Forming in a Fish*, 2013, <http://www.stuartwilde.com/2013/02/thoughts-forming-in-a-fish/>.
- [22] D. M. Hawkins, D. Bradu, and G. V. Kass, "Location of several outliers in multiple-regression data using elemental sets," *Technometrics*, vol. 26, no. 3, pp. 197–208, 1984.
- [23] I. Ben-Gal, "Outlier detection," in *Data Mining and Knowledge Discovery Handbook*, pp. 131–146, Springer, New York, NY, USA, 2005.
- [24] S. Papadimitriou, H. Kitagawa, P. B. Gibbons, and C. Faloutsos, "LOCI: fast outlier detection using the local correlation integral," in *Proceedings of the 19th International Conference on Data Engineering*, pp. 315–326, March 2003.
- [25] B. W. Miller, I. Breckheimer, A. L. McCleary et al., "Using stylized agent-based models for population-environment research: a case study from the Galápagos Islands," *Population and Environment*, vol. 31, no. 6, pp. 401–426, 2010.
- [26] U. Wilensky, *NetLogo Flocking Model*, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, Ill, USA, 1998.
- [27] P. Laube and S. Imfeld, "Analyzing relative motion within groups of trackable moving point objects," in *Geographic Information Science*, pp. 132–144, Springer, New York, NY, USA, 2002.
- [28] J. Gudmundsson, M. Van Kreyeld, and B. Speckmann, "Efficient detection of motion patterns in spatio-temporal data sets," in *Proceedings of the 12th ACM International Symposium on Advances in Geographic Information Systems (GIS '04)*, pp. 250–257, November 2004.
- [29] J. A. Canizo, J. A. Carrillo, and J. Rosado, *Collective Behavior of Animals: Swarming and Complex Patterns*, Departament de Matemàtiques, Universitat Autònoma de Barcelona, Barcelona, Spain, 2009.
- [30] Z. Wang and D. Gu, "Distributed cohesion control for leader-follower flocking," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '07)*, July 2007.
- [31] H. Su, N. Zhang, M. Z. Q. Chen, H. Wang, and X. Wang, "Adaptive flocking with a virtual leader of multiple agents governed by locally Lipschitz nonlinearity," *Nonlinear Analysis B: Real World Applications*, vol. 14, no. 1, pp. 798–806, 2013.
- [32] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 293–307, 2009.