**3. Swarm Clustering**

**3.1 Introduction**

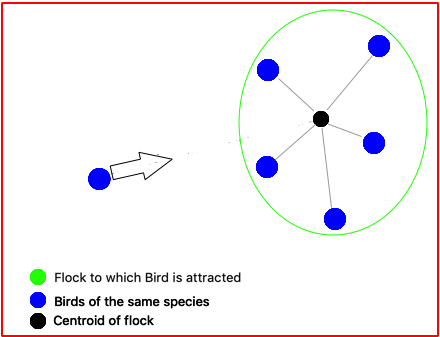
**‘**Birds of the same feather flock together.**’**

We use the idea cited by [9]Flock by Leader - Abdelghani Bellaachia, Anasse Bari

We simulate nature by assuming each stock to be a bird. We follow the three rules - Cohesion, Alignment, Separation.

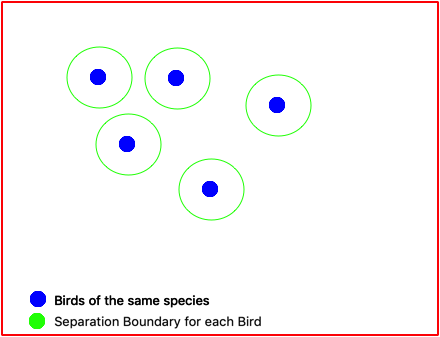
**3.1.a Cohesion**

Attraction to Birds of the same species.



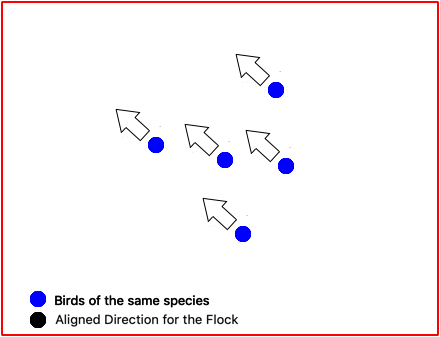
**3.1.b Separation**

Maintain a minimum distance of separation between all birds.



**3.1.c Alignment**

The entire flock always has the same aligned direction.



**3.2 Algorithm**

**Algorithm :** Leaderless Swarm clustering

**Input:** Algorithm Parameters, Similarity values for particles

**Output:** Final Clusters

Initialize the population randomly.

**While** *~StopCondition()* **do**

**For** each particle in Space

Find same species

Update Direction

Update Velocity

**If** within Constraints()

Move()

**Else**

Wait()

**Output** Space

**3.3 Approach and Implementation**

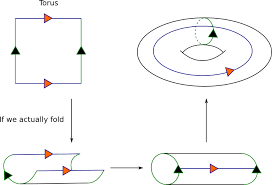
**The approach has been explained in the following sections.**

* **Coordinate Space**
* **Similarity**
* **Modelling a Stock as a Bird**
* **Time Period**
* **Attributes**

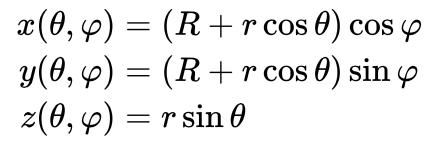
**3.3.a Coordinate Space**

The Swarm Clustering algorithm involves having individual entities moving around in a space. To simulate a boundaryless space we use a novel approach of transforming a 2-Dimensional NxN Grid into a continuous toroidal space. This is required as flocks should never meet a boundary as it will obstruct free flocking behaviour. An alternate approach could be to model the space on a  sphere or a graph.

We transform the 2-Dimensional space by merging both the boundaries along the x-axis and respectively for the y-axis. This is visually represented as below.



A Torus can be represented parametrically as:



Where

*θ*, *φ* are angles which make a full circle, so that their values start and end at the same point,

*R* is the distance from the center of the tube to the center of the torus,

*r* is the radius of the tube.

**3.3.b Similarity**

To find Birds of the same species we use a scale from 0 to 1 and  a threshold value t. All birds which have a similarity value greater than t is considered to be a fellow of the same species.

We use percentage increase in the Closing price of two stocks on a day to day basis to simulate on time period. Percentage increase is calculated after normalising to a value of 1. 20% increase corresponds to 1 while 0 corresponds to a 20% drop in price. 20% is the maximum allowed change also known a sa Trading Curb.

**3.3.c Modelling a Stock as a Bird**

We consider a toroidal surface to model our swarms on. This is shown in the Coordinate space section.

The stocks here are represented by birds which are initially randomly spawned all across the map.

We model a bird with the following attributes:

* Position : (x,y) coordinates
* Direction : The direction the bird is moving towards - Calculated as a vector sum of the attraction vector and alignment vector.
* Speed : Speed of a bird is proportional to multiple factors including number of birds of the same species in it’s neighbourhood.

Every stock is assigned a separate bird and each bird is aware of only the history of it’s stock and knows similar birds at any given stage of time.

We keep track of the birds history(where all it has been) and finally the clusters they form after a definite amount of time.

**3.3.d Time Period**

A time period is the smallest unit of time in which an observable action can be taken in our model. These actions involve the movement of a bird in the space.

In a given time period every bird has 9 options. To move in one of the 8 directions or stay in it’s position due to constraints which include separation and model design.

**3.3.e Attributes**

Speed

We set the speed proportional to the distance to centroid of the flock, taking into account a threshold distance so all the birds ner to the centroid have the same speed.

We have used a variation of the Acceptance-rejection Sampling to enforce speed in a 2D grid where in a  given time period a bird can only move to the adjacent box or wait.

**If** dist to centroid() < threshold distance :

   speed=0.5

**Else**:

   speed=0.5 + dist to centroid() / N (Grid Length)

R <= Random number

**If** (R<speed):

   return d

**Else**:

   return (0,0)

Direction

A vector sum of the Attraction vector to flock and Alignment vector of flock. The Attraction vector is along the the line joining bird and centroid of the flock. The alignment vector is the mean direction of the flock.

We have reduced any arbitrary direction in polar coordinates to one of the 4 fundamental directions in a 2D space by randomly selecting from a weighted distribution proportional to the direction.

(u,v) <= Direction vector

R <= Random number between 0 and 1

**If** R<u/(u+v):

   d=(u/(abs(u)),0)

**Else**:

   d=(0,v/(abs(v)))

Threshold

The Threshold value has been set to 0.7

Separation Distance

Each square in the grid is scaled to the minimum separation distance and hence the separation rule is enforced by simply constraining at most one bird to a box in the grid.

Line of Sight Radius

Every bird is set with a variable parameter line of sight, which is the maximum distance upto which it can recognise birds of the same species. So the bird exhibits flocking behaviour with respect to birds present in this radius. This has been set to N/3 where N is the Grid Length. So a bird has access to a an area Π/9 th of the entire space.

Total simulation duration

We have found a common duration for all stocks between the years 2013-2018. Every day of the Stock Market being open represents one Time Period as defined in Section 3.3.c.