



**HAL**  
open science

# Modeling a Surveillance Network Based on Unit Disk Graph Technique - Application for Monitoring the Invasion of Insects in Mekong Delta Region

Viet Truong Xuan, Xuan-Hiep Huynh, Minh Le Ngoc, Alexis Drogoul

► **To cite this version:**

Viet Truong Xuan, Xuan-Hiep Huynh, Minh Le Ngoc, Alexis Drogoul. Modeling a Surveillance Network Based on Unit Disk Graph Technique - Application for Monitoring the Invasion of Insects in Mekong Delta Region. The 15th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA-2012), Sep 2012, Malaysia. pp.228-242. hal-00855152

**HAL Id: hal-00855152**

**<https://hal.science/hal-00855152>**

Submitted on 18 Sep 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Modeling a Surveillance Network based on Unit Disk Graph technique - Application for monitoring the invasion of insects in Mekong Delta Region

Viet Xuan Truong<sup>1,3</sup>, Hiep Xuan Huynh<sup>2</sup>, Minh Ngoc Le<sup>3</sup>, and Alexis Drogoul<sup>1,2</sup>

<sup>1</sup> UMI 209 UMMISCO-IRD/UPMC, Bondy, France

txviet@ctu.edu.vn, alexis.drogoul@gmail.com

<sup>2</sup> DREAM Team/UMI 209 UMMISCO-IRD, Can Tho University, Vietnam

hxhiep@ctu.edu.vn

<sup>3</sup> Faculty of Computer Science & Engineering, HCMUT, Ho Chi Minh city, Vietnam

minhle@cse.hcmut.edu.vn

**Abstract.** Applying the computer science in environment and ecology management becomes an important trend in recent years. Modeling an environmental and ecological system, normally, needs three main components: (1) a dynamic system model for the environment and ecology reality, (2) a Surveillance Network Model (SNM), and (3) an evaluation module helping to validate the accuracy of the whole system. This paper aims at introducing an abstract model of SNM based on Unit Disk Graph (UDG) technique, the second component of the system. Furthermore, we also introduce the Brown plant hopper Surveillance Network Model (BSNM), a concrete case of SNM which works as an agent-based model. The BSNM helps to model the light-trap network (one kind of spatial sampling devices); this model is used to measure and to evaluate the invasion of Brown Plant Hoppers (BPHs), one kind of dangerous insects attacking rice in a large area. The simulation is applied in three provinces of Mekong delta region, Vietnam.

**Keywords:** Surveillance network, sensor network, estimation, system dynamics, Agent-Based Model (ABM), sampling, Brown Plant Hopper (BPH), Unit Disk Graph (UDG), spatial data mining, graph mining

## 1 Introduction

Finding dynamic laws of an ecological and environmental system is one of the interesting problems involved by multiple applied domains. Before, this work is only concerned by ecologists or biologists; the commonly used solutions can be mathemat-

ical models, statistical models or the GIS-based spatial analysis methods (GIS - Geographical Information System). However, with the need to understand more deeply about multiple complex relations between multiple objects and multiple interactions, the traditional methods become exhausted in some situations. In recent years, the complex system approach, with the supports of modeling and simulation techniques, allows the scientists to get more conveniences from computer sciences; some high level algorithms can be applied to retrieve the hidden emergence, like spatial or graph data mining.

In this paper, we will present in details a modeling process for a Surveillance Network Model (SNM) - a model, which can measure and evaluate the situation of environment. The model must appropriate with any existing specific surveillance system, and tries to get more information from this. We try to build an abstract model for all types of surveillance network; the measurement devices can be the sensors or simply the sampling points. The observed objects of these surveillance networks are the ecological and environmental systems. Some kinds of SNM can be the meteorological or hydrological stations, the sampling network used in the forest or insect invasion surveys. Finding the hidden spatial and temporal relations between the spatial measuring devices is the main objective of each SNM.

There are three important hypotheses needed to be improved in a surveillance network:

- Hypothesis 1: The collected data on the same device in a short period is correlated.
- Hypothesis 2: The collected data on different close devices (by geographical distance), in the same period, is also correlated.
- Hypothesis 3: A set of devices can give more information than one single device inside this set.

We conduct to verify the first hypothesis on the data of brown plant hoppers (BPH) at 26 light-traps in Dong Thap province, and the result is very encouraging. By comparing the collected data between two consecutive days, all popular correlation ranks [8] (like Kendall, Spearman or Pearson) always return a high rank of correlation average (greater 65%); and even the Spearman rank returns a correlation average greater 80% (See Fig. 7 in Section 5 for more details). However, it's not easy to verify the two last hypotheses if without a mechanism to determine the neighboring relationships in space. And the problem is how to build a mechanism helping to verify effectively all offered hypotheses (also for further hypotheses).

To solve the problem, we will try to model the SNM as a spatio-temporal simulation model, and we notice that the Unit Disk Graph is a good suggestion for determining the neighboring relationships between spatial measuring devices. UDG is a kind of the intersection graphs [1][2][3] allowing to build the connection from one node to its neighbors in a fixed radius (in the standardized case, this radius is equal to one specific unit). UDG can be directional or unidirectional, depending on the specific problem and the dynamic laws of the observed objects.

In general, in this paper, we visually present an abstract model and the implementation process for a particular Surveillance Network Model. This model has their cen-

tral elements being the spatial measuring devices and uses UDG as a solution for setting up the internal connections. SNM allows to measure and to evaluate the status of the observed objects in a specific spatial region. Based on this abstract model, a Brown Plant Hoppers Surveillance Network Model (BSNM) is also implemented and used to measure and evaluate the BPH invasion in rice; the simulation is applied in three provinces (Soc Trang, Hau Giang and Bac Lieu) - the typical provinces in the Mekong Delta region of Vietnam.

The next sections are arranged as the following. Some important related works are introduced in Section 2. Section 3 introduces the abstract description of SNM for an ecological and environmental system dynamics. Section 4 explains some abstract classes inside the abstract model of SNM: Node, Edge, Graph and Competency. The implementation of the BSMN is introduced in Section 5; this model is an Agent-Based Model (ABM) and works on GAMA platform [7][19][20]. And we summarize the contribution and suggest some research directions in Section 6. Some acknowledges for the investors are honorably written in the last section.

## 2 Related works

Spatial sampling technique is always the first investment of any environmental or ecological surveillance system, such as agricultural monitoring [9], or forest management [10][11]. The simplest technique, but often used, is representative because of its simplicity in aggregating the sampling results.

Finding the emergence from the sampling data is always a difficult problem. There are many approaches, where the modeling and the simulation are one of important trends in recent years. Some general modeling and simulation techniques are introduced in [16]; also some models help to simulate the propagation process of BPHs in the Mekong Delta region, see [6][17][21] in GAMA platform [7][19][20]. These models perform a successful integration between the ABM model and the geographical information systems (GIS). Some examples in [16][18] present some robotic models. The modeling techniques focus on the main properties and the essential behaviors of the research objects and always try to simplify the complexity of the real system.

The simulation of system dynamics is also an interesting investment. Some of them can be found in [12][13][14][15]. These studies try to find out the dynamics laws of the ecological and environmental system and to simulate all of them; some prediction models can be built by this way.

### Unit disk graph

The UDG technique is firstly introduced in 1990 by Clark [1], and is used widely in ad-hoc communication. In the wireless sensor network, UDG is used to solve the problem of optimizing the localization, some research for this trend can be found in [2] and [3]. Besides, UDG is known as a solution for some NP-Hard problems in mathematics such as vertex coloring, maximum independent set or maximum clique [1][2].

However, there are not many investments of UDG in managing an ecological and environmental system. An interesting research allows to use the UDG as a bridge between an Individual-Based Model (IBM) and an Equation-based model (EBM), proposed by MSI/IRD research group [4]. We also introduced some UDG-based research for estimating the BPH density and modeling the surveillance network in [5] and [6]. Based on these first successes, we have to generalize all process of modeling and implementing an ecological and environmental surveillance network based on UDG. In the proposed model, UDG is used as a useful tool helping to determine the connection topology of the surveillance network.

### 3 Abstract description of Surveillance Network Model for an ecological and environmental system dynamics

In an ecological environment management system, there are two main participants as the following:

- **Observed objects:** The objects observed by the managers. These objects can be the plants in forest ecology, the fishes in marine ecosystem ecology or the insects in ecological agricultural systems.
- **Surveillance Network Model:** A closed model allows to monitor and to evaluate the situation of observed objects on particular spatial and temporal scales. This model works as a consistent system where each internal component has a specific relation with the others. These relations can be one of the following:
  - Spatial relation: Monitoring the observed objects based on the adjacent relation of the spatial measuring nodes.
  - Timing relation (time series process): Monitoring the dynamics of the observed objects via the time series data at multiple measuring nodes.
  - Aggregation/combination relation: These relations consider the capacity of retrieving the emergence based on a set of elements inside the network. They can contain all relations related with the statistical or data mining techniques.

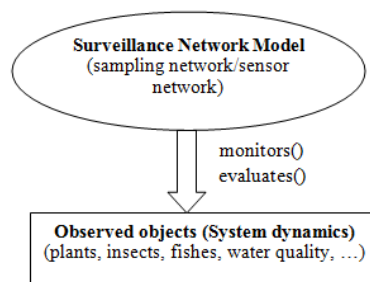
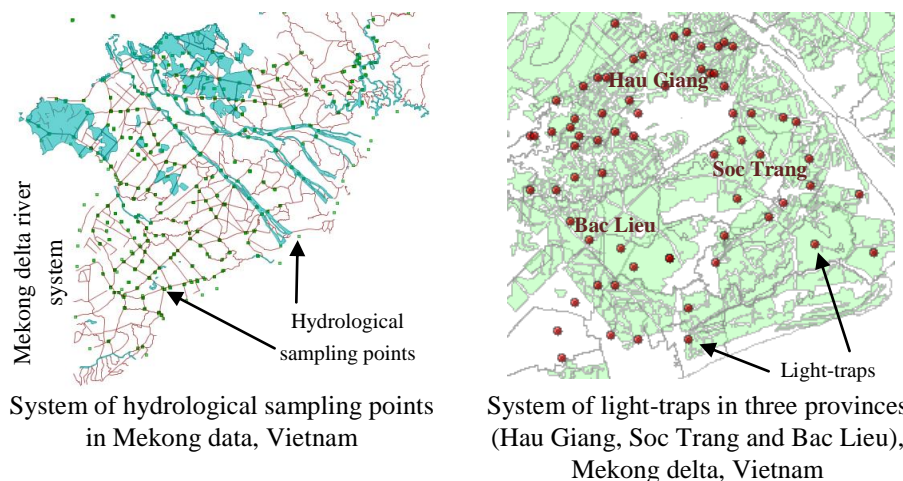


Fig. 1. Abstract view of the Surveillance Network Model.

So, the question is how to model a surveillance network which satisfies all above relations. The UDG-based approach can answer thoroughly all needs related. In the UDG, the graph node and edge contain a particular weight (the data type of the weight can be a complex structure), which describes the spatio-temporal relations of the observed objects. Also in the graph structure, we consider the relation between the different measuring devices as the view of clustering (a technique in data mining [24] [25]), to observe the relation based on some aggregation or combination techniques. The final goal of SNM is providing more useful information about the observed objects for the managers; this model will operate as a decision-support system.

#### 4 Modeling the Surveillance Network based on Unit Disk Graph

In this section, we introduce in details the implementing process for a UDG-based surveillance network. We try to build a model as abstractive as possible (while the next section will introduce a concrete case for the rice-hoppers surveillance network). With a particular system, we can apply this abstract model to build a concrete model depending on specific conditions.



**Fig 2. Environmental sampling systems.**

Fig. 2 introduces two examples for the environmental sampling systems. These systems have a lot of sampling points to measure the specific objects. For the hydrological system the observed objects can be the water levels and flows; while the light-traps system concentrates on the density of multiple types of rice-hoppers.

## 4.1 Unit Disk Graph

### 4.1.1 Definition

In graph theory, a Unit Disk Graph  $G = \langle V, E \rangle$  is a graph building from a set of vertices  $V = \{v_1, v_2, \dots, v_n\}$  and a set of edges  $E = \{e_1, e_2, \dots, e_m\}$ ; where if any pairs of vertices  $\{v_i, v_j\}$  satisfy the condition of  $d(v_i, v_j) \leq r$ , the pair  $\{v_i, v_j\}$  will be added into  $E$ . The function  $d(v_i, v_j)$  is called the distance function of  $G$  and  $r$  is called the connected radius. In this graph, each vertex  $v_i$  is considered as the centered point and the zone determined by the connected radius  $r$  around  $v_i$  forms a circle shape (or disk shape), then we call  $G$  the disk graph. When  $r$  is unit and is fixed for the whole graph, we have a unit disk graph.

In extended case, the disk shape can be changed depending on a specific use-case, for example an ellipse shape. Figure 2 introduces two different cases of UDG with (a) a circle disk and (b) an ellipse disk.

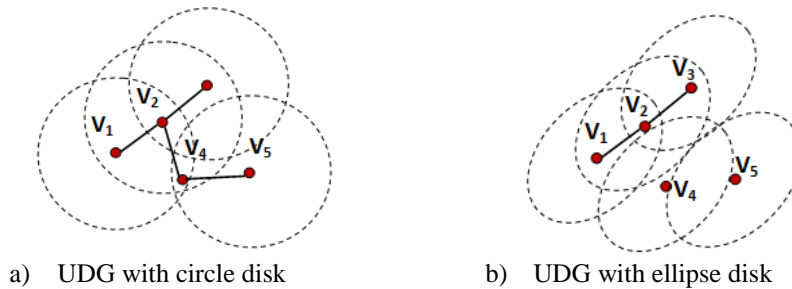


Fig. 3. Two different kinds of UDG.

In case of ellipse distance, the centered point is replaced by the foci of the ellipse which compose two special points  $F_1$  and  $F_2$  on the ellipse's major axis and are equidistant from the centered point. The radius  $r$  is replaced by  $r_e$  with two values  $[a_e, b_e]$  where  $a_e$  and  $b_e$  are called the major and minor radius relatively. The connected condition for this case is determined as  $d_e(v_i, v_j) = d(v_i, F_{1j}) + d(v_i, F_{2j}) \leq 2a_e$ .

In Figure 2, the UDG in (a) has  $V = \{v_1, v_2, v_3, v_4, v_5\}$  and  $E_1 = \{(v_1, v_2), (v_2, v_3), (v_2, v_4), (v_4, v_5)\}$  while in (b), the graph has the same set of vertices but the set of edges  $E_2$  has only two elements with  $E_2 = \{(v_1, v_2), (v_2, v_3)\}$ . The second choice of ellipse disk is an extended suggestion and we can define more kinds of disk in case of necessary.

### 4.1.2 Connection analysis

In the simple case, the connected condition can be defined only by the spatial distance. But in the real context, perhaps there are some special cases:

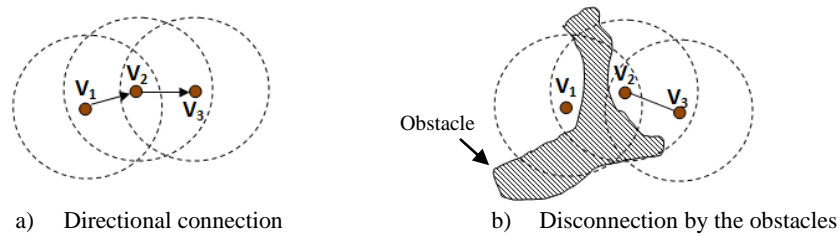
- (1) Two vertices having the "Cause-Effect" relation: can be considered as "directional connection".

This relation depends on the dynamic laws of the observed objects. If the movement is directional, the surveillance graph is necessary to be a directional graph for some research conveniences. For example, in the hydrological sys-

tem, the water flow direction can also be considered as the direction of the edges. An example for this case is illustrated in Fig. 4.

- (2) Two vertices having the topological disconnection can be considered as “not connected”.

This relation appears when there is an obstacle between two nodes; in this case, although these two nodes are very close but the necessary relation for the research doesn't exist, then there is no edge connecting these two nodes. In observing the environment, in this case we can find out a break of spatial topology in a local space around two nodes. For example, if two hydrological sampling points have a close distance but they lie on two independent channels (and two these channels have the intersection in a very far distance), these nodes will be disconnected. An example for this case is introduced in Fig. 4.



**Fig 4.** Different cases for the connection inside UDG.

Suppose that  $Obstacle(v_i, v_j)$  is the relation in which exists an topological obstacle between two nodes  $v_i, v_j$ , we can determine the existence of edge  $(v_i, v_j)$  as the following:

$$\exists(v_i, v_j) \rightarrow d(v_i, v_j) \leq r_{disk} \text{ and } \nexists Obstacle(v_i, v_j)$$

Denote the scalar and direction relation between two nodes as  $(v_i, v_j)$  and  $(\overline{v_i}, \overline{v_j})$ . The existence of a scalar edge  $(v_i, v_j)$  is determined by the existence of two opposite directional edges, we have:

$$\exists(v_i, v_j) \rightarrow \exists(\overline{v_i}, \overline{v_j}) \text{ and } \exists(\overline{v_j}, \overline{v_i})$$

#### 4.1.3 Disk radius and the graph topology

The topology of the disk graph is significantly changed as changing the disk radius. From being all disconnected this topology can be all connected (complete graph) as increasing the disk radius from *zero* (0) to unlimited value ( $+\infty$ ).

$R_{min}$  and  $R_{max}$  is called the smallest and largest Euclidean distance existing between any pair of nodes. We found that when  $r < R_{min}$ , the graph does not exist any edge ( $E = \emptyset$ ); when  $r \geq R_{max}$ , the graph becomes a complete graph (i.e. exists edge  $(v_i, v_j) \forall i, j$ ).

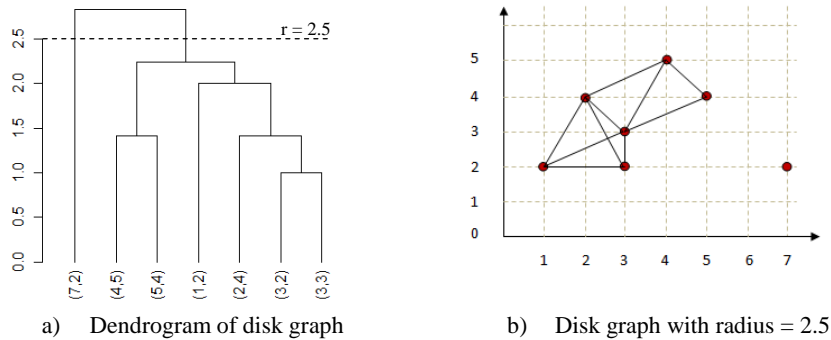
Consider a graph with 5 nodes  $V = \{(1, 2), (2, 4), (2, 2), (3, 3), (4, 5), (5, 4), (7, 2)\}$ , we have a distance matrix as follows:



Nodes	(1, 2)	(2, 4)	(3, 2)	(3, 3)	(4, 5)	(5, 4)	(7, 2)
(1, 2)	0	$\sqrt{5}$	2	$\sqrt{5}$	$3\sqrt{2}$	$2\sqrt{5}$	$6 (R_{max})$
(2, 4)		0	$\sqrt{5}$	$\sqrt{2}$	$\sqrt{5}$	3	$\sqrt{29}$
(3, 2)			0	$1 (R_{min})$	$\sqrt{10}$	$2\sqrt{2}$	4
(3, 3)				0	$\sqrt{5}$	$\sqrt{5}$	$\sqrt{17}$
(4, 5)					0	$\sqrt{2}$	$3\sqrt{2}$
(5, 4)						0	$2\sqrt{2}$
(7, 2)							0

**Table 1.** Distance matrix.

Based on the distance matrix, we easily found the value of  $R_{max} = 6$  and  $R_{min} = 1$ . Therefore, the graph topology is only changed when moving the disk radius in the range  $[1, 6]$ . The dendrogram for the above example is shown in Fig. 5b.



**Fig 5.** Dendrogram and graph topology.

Dendrogram in Fig. 5a can help to observe the graph topology, and also the number of sub-graph with a specific value of radius  $r$ . For example, in case  $r = 2.5$ , there are two sub-graphs; in which a graph  $G_1$  with  $V_1 = \{(1, 2), (2, 4), (2, 2), (3, 3), (4, 5), (5, 4)\}$  and  $G_2$  with  $V_2 = \{(7, 2)\}$ .

#### 4.2 Surveillance Network Model (SNM)

In this part, we introduce the process of modeling the surveillance network based on the UDG technique. The SNM will be implemented as an object-oriented model; and the UDG technique becomes the tool to build the monitoring graph – the core element of the model.

Here, we consider of four major classes of SNM: *Node*, *Edge*, *Graph* and *Competency*. First two components will constitute the *Graph* class, the representative class for the whole surveillance system. The *Competency* is a special class which contains the evaluation information of remaining classes.

#### 4.2.1 Node class:

The *Node* class is used to model a sampling device. The spatial surveillance problem is more interested in coordinates of the sampling devices; sampling location plays a crucial role in the process of estimating and evaluating the density of the observed objects. Some related attributes of device can also be considered: its dominated zone in practice, the measuring threshold (or measuring capacity), the working status (in some situations, sampling devices are disabled due to some objective factors, such as in the rain).

General specification for the *Node* class is as follows:

##### Properties

- `nodeID` (string): Identifier of sampling device.
- `x_coordinate` (float): X coordinate of sampling device.
- `y_coordinate`: Y coordinate of sampling device.
- `nodeName`: Node name.
- `currentDensity`: Density of the last period.
- `historicalDensities[]`: Densities of the recent periods (the duration depends on the specific needs).
- `capacity`: The measuring threshold on which the sampling device can support.
- `workingStatus`: Working: ON/Non-working: OFF.
- `competency`: Type of *Competency* class, which contains the evaluation information for the node.

##### Actions

- `measures()`: Measuring the current population density of the observed objects.
- `correlateByTimeMoving ()`: Calculate the correlation coefficient [8] by all pair of measured values between two last consecutive periods.
- `getAdjacentInfo(deep)`: Getting information from all neighbors of this node, the *deep* parameter is used to decide the scope of adjacent relation from this node.

**Algorithm 1.** Calculating the correlation coefficient by time moving (Implemented in *Node.correlateByTimeMoving()*).

##### Input:

`historicalDensities[]`: Array contains the historical data of the node.  
`duration`: Timing duration of the observation.  
`correlationType`: PEARSON/SPEARMAN/KENDALL.

##### Output:

`correlationCoef`: Correlation coefficient.

**Complexity:**  $\theta(n \log n)$ .

##### Implementation:

```
1. Begin
2.   i = 0;
3.   For t from 1 to duration
4.     Begin
5.       vectorX[i] = historicalDensities[t];
```

```

6.     vectorY[i] = historicalDensities[t];
7.     i++;
8.     End
9.     correlationCoef = cor(vector, vectorY, correlationType);
10.    Return correlationCoef;
11.    End

```

#### 4.2.2 Edge class:

Determining the existence of edges in the graph is a difficult problem. We need to determine the multiple relations between different sampling devices while these relations don't exist explicitly. Because the dominated zone of sampling devices isn't concerned (it's too small in comparing with the sampling scope), their characteristics are completely different from a wireless telecommunication network (or a radar network). Meanwhile, we are interested in some potential relationships to establish the edge connections. For example, in the BPHs surveillance network, we pay attention to their migration behaviors along the wind.

General specification for the Edge class is as follows:

##### Properties

- edgeID: Identifier of the edge.
- sourceNodeID: Identifier of the source node.
- destinationNodeID: Identifier of the destination node.
- deviationAmount: Defined as the density difference between the source and destination nodes.
- length: Euclidean distance from the source node to the destination node.
- competency: Type of *Competency* class, which contains the evaluation information for the edge.

##### Actions

- getInternalCorrelation(CorrelationType): Get the correlation between the sources and the destinations (all adjacent edges).

#### 4.2.3 Graph class

This class is the representative for the Surveillance Network. *Graph* class keeps a central position in controlling all tasks of measuring and evaluating in the system. Due to this reason the evaluation results retrieved from this class is always the final conclusion. There are multiple actions implemented in this class, almost of them are the aggregation and combination tasks.

General specification for the Graph class is as follows:

##### Properties

- nodesList: List of nodes existing in the graph.
- edgesList: List of edges existing in the graph.
- competency: Type of *Competency* class, which contains the evaluation information for the graph.

### Actions

- `resetEdges()`: Rebuilding the new network from the current context.
- `estimateDensity` (Method): Helping to estimate the density for the whole region in the GRID model, some methods can be the IDW [22], Kriging [23], UDG-based [5]
- `getCorrelationByPeriods()`: Calculating the correlation coefficient in data of all nodes between two consecutive periods: current and previous periods.
- `getCorrelationByEdges()`: Calculating the correlation coefficient in data between all pairs of nodes existing in the `edgesList`.

**Algorithm 2.** Implementing the edge connections (Implemented in `Graph.resetEdges()`).

**Input:**

nodesList: list of nodes  
radius: radius of unit disk graph

**Output:**

edgesList: list of edges

**Complexity:**  $\theta(n \log n)$ .

**Implementation:**

```
1. Begin
2.   For i from 0 to (nodesList.length - 2)
3.   For j from i + 1 to (nodesList.length - 1)
4.     If (nodesList[i].workingStatus == TRUE)
5.       and (nodesList[j].workingStatus == TRUE)
6.       Then If (distance(nodesList[i], nodesList[j]) ≤ radius)
7.         Then If (canConnect(nodesList[i], nodesList[j]))
8.           Then
9.             newEdge = createNewEdge(nodesList[i], nodesList[j]);
10.            edgesList.add(newEdge);
11.          End
12.        End
13.      End
14.    End
15.  Return edgesList;
16. End
```

**Algorithm 3.** Calculating the correlation coefficient in data of all nodes between two consecutive periods current and previous periods (Implemented in `Graph.getCorrelationByPeriods()`).

**Input:**

nodesList: list of nodes  
historicalDensities[]: list of nodes  
correlationType: PEARSON/SPEARMAN/KENDALL

**Output:**

correlationCoef: Correlation coefficient.

**Complexity:**  $\theta(n)$ .**Implementation:**

```
1. Begin
2.   t = currentPeriod;
3.   For i from 0 to (nodesList.length - 1)
4.     Begin
5.       vectorX[i] = nodesList[i].currentDensity;
6.       vectorY[i] = nodesList[i].historicalDensities[t - 1];
7.     End
8.   correlationCoef = cor(vector, vectorY, correlationType);
9.   Return correlationCoef;
10. End
```

**Algorithm 4.** Calculating the correlation coefficient in data between all pairs of nodes (Implemented in *Graph.getCorrelationByEdges()*).

**Input:**

edgesList: list of edges.

correlationType: PEARSON/SPEARMAN/KENDALL.

**Output:**

correlationCoef: Correlation coefficient.

**Complexity:**  $\theta(n)$ .**Implementation:**

```
1. Begin
2.   For i from 0 to (edgesList.length - 1)
3.     Begin
4.       srcID = edgesList[i].sourceNodeID;
5.       dstID = edgesList[i].destinationNodeID;
6.       vectorX[i] = nodesList[srcID].currentDensity;
7.       vectorY[i] = nodesList[dstID].currentDensity;
8.     End
9.   correlationCoef = cor(vector, vectorY, correlationType);
10. Return correlationCoef;
11. End
```

#### 4.2.4 Competency class

*Competency* class is defined as an abstract class to evaluate the capacity of another class. The properties of this class are designed to evaluate the capacity of another class. This class becomes a private property of three above classes. General specification for the *Competency* class is as follows:

##### Properties

- `correlationCoefficients[]`: Containing the correlation coefficients.
- `aggregationValues[]`: Containing the aggregation values.
- `combinationValues[]`: Containing the combination values.

##### Actions

- `evaluatesCapacity()`: Evaluating the capacity from the private properties.

## 5 Brown Plant Hopper surveillance network and Agent-based simulation

In this section, we present a concrete situation of SNM for the insect surveillance network in the Mekong delta region. The observed objects are the Brown Plant Hoppers in adult stage (when the BPHs can dispatch along the wind).

To implement the model, we need to consider two factors presented in the abstract view (Fig. 1):

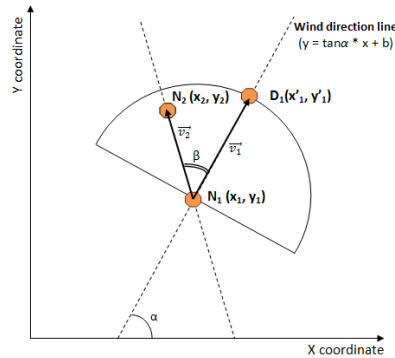
- (1) Observed objects: Brown Plant Hoppers.  
These objects must be implemented as a system dynamics. In fact, simulating the movement of these objects becomes a big challenge for any modelers. In this research, we will implement a migration model for Brown Plant Hoppers; this model will consider three natural factors: transplantation regions of rice, the river regions and the wind (velocity and direction).
- (2) Surveillance Network Model: This model has its own private name Brown Plant Hopper Surveillance Network Model (BSNM). BSNM follows the general SNM structure specified in Section 4. By the reason of migration behaviors of BPHs (depending on the wind), the graph will be directional (the edge direction is determined by the wind direction). Also, the disk radius is determined by the wind velocity. The sampling devices are the light-traps established in the plant protection management (managed by the Southern Plant Protection Research Center). The sampling stage of this sampling network is determined by one day (this stage is also considered for one simulation step).

### 5.1 Modeling the BSNM

The sampling period is determined as (by or as) one day (the real period). We considered the maximum distance on which the BPH can dispatch in each period. Suppose the  $v$  is the wind velocity,  $t$  is the dispatch time inside one period (normally  $t = 4$

hours from the evening to the night), we have  $r_{max} = v * t$ . The disk radius  $r$  of UDG should be less than  $r_{max}$ . We apply the case of model using the circle disk (Fig. 2a).

Fig. 6 analyses the connection between two light-traps depending on the velocity and the direction of wind. In this figure, we will consider the light-trap  $N_1$ , with coordinates  $(x_1, y_1)$ . Suppose that there is a light-trap  $N_2$ , with coordinates  $(x_2, y_2)$  located in the half circle lower the wind having  $r_{max} = v * t$ , we will analyze the connection between  $N_1$  and  $N_2$  (if existing).



**Fig. 6.** Connection analysis between two nodes.

In Fig. 6, call  $\alpha$  is the angle between the wind direction and the x axis. We call  $N_1$  is the source point, from where the BPH can dispatch, and a point  $D_1(x'_1, y'_1)$  is the farthest where the BPH can reach in case of moving totally along the wind direction (in one day). We see that the distance from  $N_1$  to  $D_1$  is  $r$ . We will find out the relation of  $N_1$  and  $N_2$  via  $\alpha$  and  $D_1$ .

The line connected  $N_1$  and  $D_1$  has the equation  $y = \tan \alpha * x + b$ . Then, we have:

- (1)  $|x'_1 - x_1| = r * |\cos \alpha|$
- (2)  $|y'_1 - y_1| = r * |\sin \alpha|$

Depending on the value of  $\alpha$ , we can calculate the values of  $x'_1, y'_1$  from the values of  $x_1, y_1$  and  $\alpha$ . We have the table of expressions:

$\alpha$	Expressions
$0 < \alpha \leq \pi/2$	$x'_1 = x_1 + r *  \cos \alpha ; y'_1 = y_1 + r *  \sin \alpha $
$\pi/2 < \alpha \leq \pi$	$x'_1 = x_1 - r *  \cos \alpha ; y'_1 = y_1 + r *  \sin \alpha $
$\pi < \alpha \leq 3\pi/2$	$x'_1 = x_1 - r *  \cos \alpha ; y'_1 = y_1 - r *  \sin \alpha $
$3\pi/2 < \alpha \leq 2\pi$	$x'_1 = x_1 + r *  \cos \alpha ; y'_1 = y_1 - r *  \sin \alpha $

**Table 2.** Calculation of values  $x'_1$  và  $y'_1$  following  $\alpha$  and  $x_1, y_1$ .

After determining the coordinates of  $D_1$ , we need to determine whether the  $N_2$  satisfies the condition of locating in the half circle lower the wind of  $N_1$ .

Call  $\vec{v}_1 = \overrightarrow{N_1D_1} = (x'_1 - x_1, y'_1 - y_1)$  and  $\vec{v}_2 = \overrightarrow{N_1N_2} = (x_2 - x_1, y_2 - y_1)$  are two directional vectors. We need to determine the angular  $\beta$  made by  $\vec{v}_1$  and  $\vec{v}_2$ , if  $\beta \leq \pi/2$ , then  $N_2$  satisfies the condition of locating in the half circle lower the wind of  $N_1$ ; and in the model, there is one edge connected from  $N_1$  to  $N_2$ .

The value of  $\beta$  can be determined by the expression:

$$\beta = \text{acos}\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{|\vec{v}_1| \cdot |\vec{v}_2|}\right)$$

where  $\vec{v}_1 \cdot \vec{v}_2$  is the scalar product of two vectors  $\vec{v}_1$  and  $\vec{v}_2$ , while  $|\vec{v}_1|$  and  $|\vec{v}_2|$  are respectively the length of  $\vec{v}_1$  and  $\vec{v}_2$  (by Euclidean distance).

So, the condition to establish a connection from  $N_1$  to  $N_2$  is:

$$\begin{cases} |\vec{v}_2| \leq t * v \\ |\beta| \leq \pi/2 \end{cases}$$

Where,  $t$  is the amount of time the BPH can dispatch in one day (in fact  $t = 4$  hours/day) and  $v$  is the wind velocity.

## 5.2 Experiments

The sampling data is the number of BPHs trapped by the light-trap every night, the sampling period is one day (and also considered as the simulation period). In this research, we use the light-trap data in January, 2010 of four provinces in Mekong delta region.

In this part, we introduce two main results. The first result is the verification of the laws inside the light-trap data; we applied some correlation analysis in R language [25] (introduced in Part 5.2.1). This first result suggests many ideas to develop the evaluation actions of SNM abstract model. In Part 5.2.2, we introduce the second result of this research, all outputs of the BSNM. This BSNM is built on GAMA 1.4 platform [7]; this GIS-based and ABM-based simulation platform has been developed by the research team MSI (located in the IFI, Hanoi, and part of the IRD/UPMC International Research Unit UMMISCO) since 2007.

### 5.2.1 Correlation analysis in R

Fig. 7 is the correlation analysis [8] in three popular ranks: Kendall, Pearson and Spearman in the data at 26 light-traps in Dong Thap province in January, 2010. The correlation coefficient is calculated on two vectors  $v_1$  and  $v_2$ , where  $v_1$  keeps the values of all light-traps in the day  $i$  and  $v_2$  keeps the values of all light-traps in the day  $i+1$  (two consecutive days),  $1 \leq i \leq 30$ . The analysis results demonstrate that the Spearman rank is rather high, the average  $> 0.8$ , and there is strictly the correlation between two consecutive days.



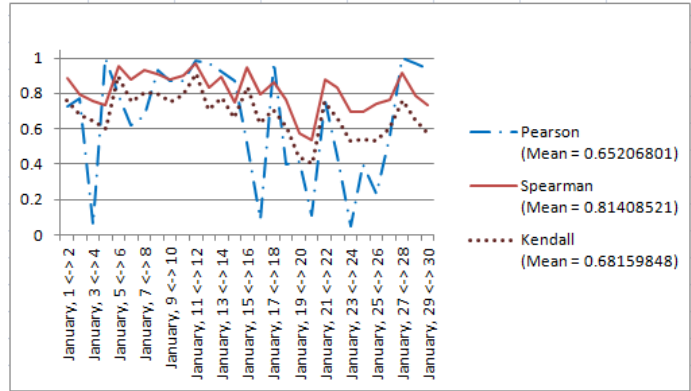


Fig. 7. Correlation analysis of light-trap data between two consecutive days in January, 2010 in Dong Thap province, Vietnam.

### 5.2.2 BSNM results

a) Directional surveillance network with different values of disk radius:

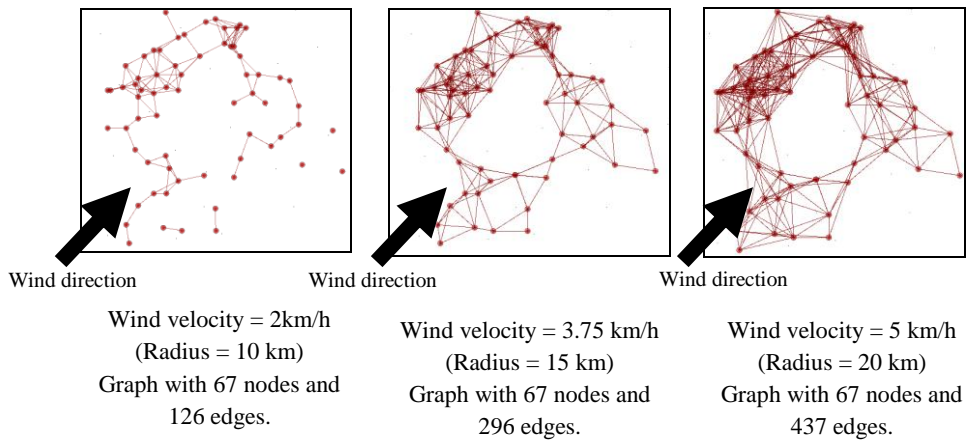
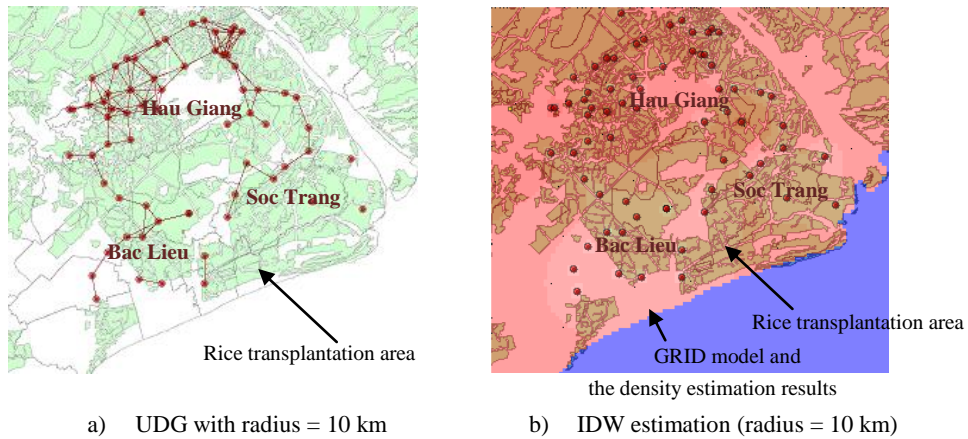


Fig. 8. Directional surveillance network

The surveillance network in Fig. 8 is built from 67 light-traps in three provinces Hau Giang, Soc Trang and Bac Lieu. In the model, all light-traps are in the running status, the wind direction is South-West ( $\alpha = \pi/4$ ).

b) Ecological and environmental model:

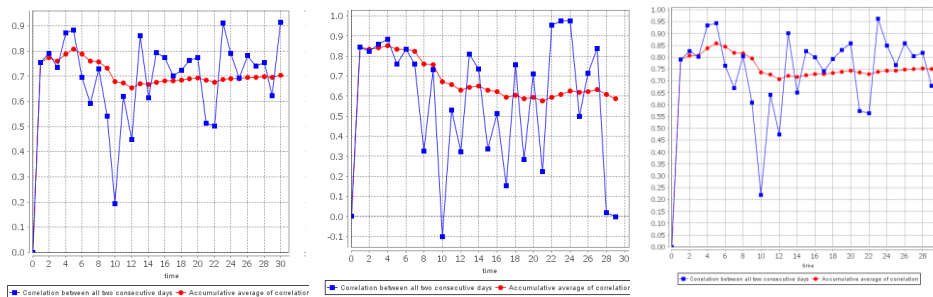
The Ecological and Environmental model is the second part of the whole system; this model is independent with the SNM and is measured and evaluated by BSNM. We don't concentrate on it in this paper (can see more details in [6]).



**Fig. 9.** Ecological and environmental model and the results of density estimation algorithm.

The ecological and environmental model is based on the grid one. This approach can easily help apply more estimation and propagation models. Fig. 9b shows the result of IDW estimation (using the *Graph.estimateDensity(Method = 'IDW')* action) from the SNM.

- c) Correlation analysis - Actions of Graph class:
  - Action: *Graph.getCorrelationByPeriods()*



Kendall correlation

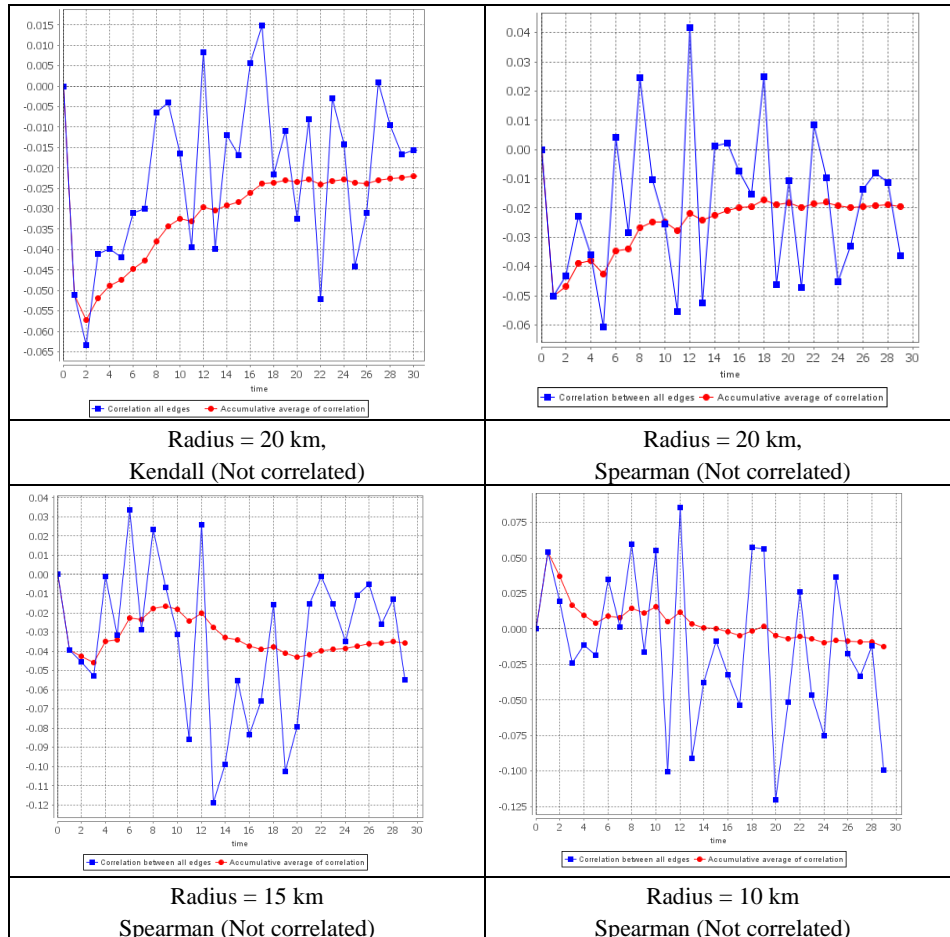
Pearson correlation

Spearman correlation

**Fig 10.** Correlation between all two consecutive days

These results is consistent with one illustrated in Section 5.2.1, but is applied in another region. Furthermore, the correlation can be automatically generated from the BSNM; we can apply this action for any kinds of data: the historical data for the analysis tasks and the propagation data for the predictive analysis.

- Action: *Graph.getCorrelationByEdges()*



**Fig 11.** Correlation between all pairs of nodes in *Graph.edgesList*.

Almost analyses of this action give a bad result in correlation rank. We can say that the Hypothesis 2 in Section 1 isn't verified. The data is nearly not correlated at all.

## 6 Conclusion

In this paper, we propose an abstract model helping to implement a surveillance network based on the UDG technique [1][2][3]. This Surveillance Network Model (SNM) can be implemented for many different ecological and environmental surveillance systems. We can consider this model as a useful tool for verifying and evaluating many scientific hypotheses related with the ecosystems. All components of SNM are object-oriented, and it's easily to adapt with many modeling methods, as well as The 15th International Conference on Principles and Practice of Multi-Agent Systems, p. 18, 2012. © Springer-Verlag Berlin Heidelberg 2012

the multi-agents model. Four main abstract classes are defined: *Node*, *Edge*, *Graph* and *Competency* and multiple statistical and data mining algorithms [8][24][25] (e.g. the correlation [8], the spatial estimation [5][22][23]) are also concerned in this abstract model.

In addition, we are successful in implementing the Brown plant hopper Surveillance Network Model (BSNM) for the light-trap network in three provinces of the Mekong delta region: Soc Trang, Hau Giang and Bac Lieu. The BSNM is agent-based [16][17][21] and contains four types of agent relatively with all main classes in the SNM. We verify some hypotheses by using the correlation ranks (Kendall, Pearson and Spearman [8]); with it, we can get more information of the observed objects. The first hypothesis is well verified while the second one needs more investments in the future research because of the bad result in correlation. Besides, the implementation of IDW estimation [22] algorithm is also given in this model.

The investment on the aggregation and the combination algorithms are the next works of this research.

## 7 Acknowledgement

This publication was made possible through support provided by the IRD-DSF. The first author was funded by the PDI program (International Doctoral Program, website: <http://www.pdimsc.upmc.fr>).

## 8 References

1. Brent N. C., Charles J. C. and David S. J. 1990. Unit Disk Graph. Holland. Elsevier Science Publishers B.V. (North-Holland). Discrete Mathematics, 86, 165-177.
2. Wu, W., et al. 2006. Minimum connected dominating sets and maximal independent sets in unit disk graphs. Elsevier B.V., Theoretical Computer Science, 352, 1-7.
3. Marathe, M. V., Breu, H., Hunt, H. B. 1995. Simple Heuristics for Unit Disk Graphs. Wiley Periodicals, Inc. Networks, 25 (2), 59–68. DOI: 10.1002/net.3230250205.
4. Nguyen, Ngoc Doanh, Phan, Ha Duong, Thi, Ngoc Anh Nguyen, Drogoul, Alexis and Zucker, Jean-Daniel. 2009. Disk Graph as a Powerful Intermediate Representation for Bridging the Gap between IBM and EBM of Population Dynamics. IEEE-RIVF International Conference on Computing and Communication Technologies Research, Innovation and Vision for the Future.
5. Truong, X.V., Huynh, X. H., Le, N.M., Drogoul, A. 2011. Estimating the density of Brown Plant Hoppers from a light-traps network based on Unit Disk Graph The 2011 International Conference on Active Media Technology (AMT 2011), Lanzhou, China.
6. Truong, X.V., Huynh, X. H., Le, N.M., Drogoul, A. 2011. Modeling the Brown Plant Hoppers Surveillance Network using Agent-Based Model – Application for the Mekong Delta Region. In The Proceedings of The Second International Symposium on Information and Communication Technology (SoICT'11), October 13–14 2011, Hanoi, Vietnam. ISBN: 978-1-4503-0880-9
7. GAMA Platform. Website: <http://code.google.com/p/gama-platform/>

8. Myers, J. L.; Well, A. D. 2003. *Research Design and Statistical Analysis* (2nd ed.). Lawrence Erlbaum. pp. 508. ISBN: 0-8058-4037-0.
9. Kozak, M. and Wieczorowski, R. 2005. PPS sampling versus stratified sampling - Comparison of efficiency in agricultural surveys. *Statistics in transaction*, 7 (1), 5-12.
10. Köhl, M., Magnussen, S., Marchett, M. 2006. *Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory*. Berlin. Springer, ISBN: 3-540-32571-9
11. Talvitie, M., Leino, O. and Holopainen, M. 2006. *Inventory of Sparse Forest Populations Using Adaptive Cluster Sampling*. The Finnish Society of Forest Science, *Silva Fennica*, 40 (1), ISSN 0037-5330.
12. E.Perrier, N. Bird, T. Rieutord, 2010, Percolation properties of 3-D multiscale pore networks: how connectivity controls soil filtration processes, *Biogeosciences Discuss.*, 7, 2997-3018, available at [www.biogeosciences-discuss.net/7/2997/2010/](http://www.biogeosciences-discuss.net/7/2997/2010/)
13. N.Bird, E.Perrier, 2009. Percolation properties of multiscale pore networks, in press, *Geoderma*, available at <http://dx.doi.org/10.1016/j.geoderma.2009.10.009>
14. H.Laurie, E.Perrier, 2010. A multifractal model for spatial variation in species richness. *Ecological Complexity*, 7(1) 2010, pp 32-35. Télécharger
15. E.Blanchart, N.Marilleau, C.Cambier, A.Drogoul, E.Perrier, JL. Chotte, 2009.. SWORM: an agent-based model to simulate the effect of earthworms on soil structure. *European Journal of Soil Science* 60 (1) pp 13-21
16. Treuil, J-P., Drogoul, A. and Zucker, J-D. 2008. *Modélisation et Simulation à base d'Agents*. Dunod Editions
17. Phan, H. C., Huynh, X. H., Drogoul, A. 2010. An agent-based approach to the simulation of Brown Plant Hopper (BPH) invasion in the Mekong Delta. *International Conference on Computing and Communication Technologies (IEEE-RIVF 2010)*, Hanoi, Vietnam, DOI:10.1109/RIVF.2010.5633134
18. Le, N. M., Nguyen, T. 2010. A Robotics Modeling, Design and Simulation Toolbox in Ptolemy II. *Chinese Control and Decision Conference*. 2297-2301. 978-1-4244-5182-1/10
19. Taillandier, P., Vo, D.-A., Amouroux, E., Drogoul, A. 2010. GAMA: a simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. *13th International Conference on Principles and Practice of Multi-Agent System*. PRIMA 1010. Kolkata, India.
20. Amouroux, E., Gaudou, B., Desvaux, S., Drogoul, A. 2010. O.D.D.: a Promising but Incomplete Formalism For Individual-Based Model Specification. *International Conference on Computing and Telecommunication Technologies*. IEEE-RIVF 2010. Hanoi, Vietnam.
21. Ouka, A. 2005. A migration analysis of the rice planthopper *Nilaparvata lugens* from the Philippines to East Asia with three-dimensional computer simulations. Tokyo : The Society of Population Ecology and Springer-Verlag Tokyo.
22. Shepard, D. 1968. A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference*. Pp. 517–524. DOI:10.1145/800186.810616.
23. Matheron, G. 1963. *Principles of geostatistics*. *Economic Geology*. Vol. 58, pp. 1246–1266.
24. Huynh, X. H., Guillet, F., Blanchard, J., Kuntz, P., Briand, H., Gras, R. 2007. A Graph-based Clustering Approach to Evaluate Interestingness Measures: A Tool and a Comparative Study. *Quality Measures in Data Mining*, Pp. 25-50
25. Hastie T., Tibshirani, R., Friedman, R. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2<sup>nd</sup> Edition). Springer Series in Statistics. Springer-Verlag New York Berlin Heidelberg. ISBN: 0-387-95284-5.
26. R language. Website: <http://www.r-project.org/>