Moving Objects: Combining Gradual Rules and Spatio-Temporal Patterns

Phan Nhat Hai[1,2], Pascal Poncelet[1], Maguelonne Teisseire[2]

[1] *LIRMM Lab., University of Montpellier 2*

[2] *TETIS Lab. 500 rue Jean-François Breton, F-34093 Montpellier, France*

Corresponding author: teisseire@teledetection.fr

***Abstract*— Mining gradual patterns plays a crucial role in many real world applications where very large and complex numerical data must be handled, e.g., biological databases, survey databases, data streams or sensor readings. Gradual rules highlight complex order correlations of the form *"The more/less X, then the more/less Y"*. Such rules have been studied for a long time and recently scalable algorithms have been proposed to address this issue. However, mining gradual patterns remains challenging in mobile object applications. In the other hand, mining frequent moving objects patterns is also very useful in many applications such as traffic management, mobile commerce, animals tracking. Those two techniques are very efficient to discover interesting rules and patterns; however, in some aspect, each individual technique could not help us to fully understand and discover interesting items and patterns. In this paper, we present a novel concept in that gradual pattern and spatio-temporal pattern are combined together to extract gradual-spatio-temporal rules. We also propose a novel algorithm, named *GSTD*, to extract such rules. Conducted experiments on a real dataset show that new kinds of patterns can be extracted.**

***Keywords*— Gradual rule, graduality, spatio-temporal pattern, moving objects, gradual-spatio-temporal rule.**

I. INTRODUCTION

Nowadays, many electronic devices are used to deal with real world applications. Telemetry attached on wildlife, GPS set on cars, sensor networks, and mobile phones carried by people have enabled tracking of almost any kind of data and lead to an increasing large amount of data containing moving objects and numerical data. Therefore, analysis on such data to find interesting patterns draws increasing attention in many real world applications such as gradual itemsets, movement patterns, animal behaviors, route planning and vehicle control.

One of the data analysis task for extracting patterns from such data is to find moving object clusters, i.e. group of moving objects that are traveling together. A moving object cluster can be defined in both spatial and temporal dimensions: (1) a group of moving objects should be geometrically close to each other, (2) they should be together for at least some minimum numbers of certain timestamps. In this context, many recent studies are interested in mining moving object clusters including moving clusters [4], flocks [6, 7, 13], convoys [1, 2, 12] and trajectory [5, 11, 3, 14].

Another data analysis technique is mining gradual pattern. Mining gradual patterns plays a crucial role in many real world applications where very large and complex numerical data must be handled, e.g., biological databases, survey databases, data streams or sensor readings. Gradual rules highlight complex order correlations of the form "*The more/less X, then the more/less Y*". Such rules have been studied and many algorithms have been proposed to address this issue [8, 10]. However, mining gradual patterns remains challenging in mobile object applications.

These two techniques are very efficient to discover interesting rules and patterns. Nevertheless, some new kind of interesting patterns may be extracted by combining these approaches.

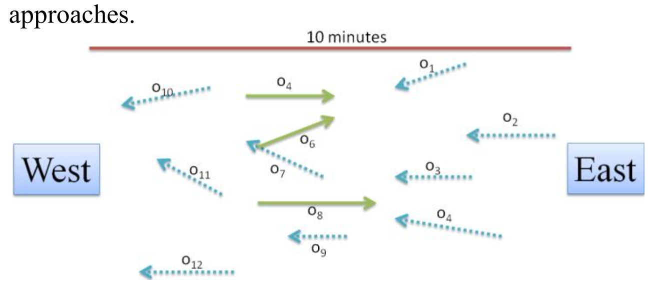


Fig. 1 Loss of interesting gradual-moving patterns

For instance, let take a look at Fig 1, there are 12 objects moving independently in 10 minutes. Individual moving objects clusters techniques including moving cluster [4], flock [6, 7], convoy [1, 2, 12] and trajectory [5, 11, 3] detecting are not efficient in this situation. In fact, these techniques require objects to be together for at least some minimum numbers of certain timestamps; therefore, they are not adapted to extract some interesting patterns in this context. Similarly, gradual rule techniques are also not adapted since they are commonly applied on numerical data while, in our context, we consider moving objects.

In this paper, we present a novel algorithm in that gradual rule and spatio-temporal pattern are combined together to extract *gradual-spatio-temporal* rules. From Fig 1, an interesting rule could be: "*The more time is going on, the more objects are moving from east to west*". More precisely, we present gradual-spatio-temporal patterns with the associated support and confidence definitions. After defining gradual-spatio-temporal rule, we also propose a novel algorithm, named *GSTD*, to extract interesting patterns. Experimental results show that we can propose new kinds of patterns.

The remaining of the paper is organized as follows. Section 2 discusses the related work. The gradual-spatio-temporal rule concept is given in Section 3. We introduce the *G*radual-*S*patio-*T*emporal rules *D*iscovering method, called *GSTD* algorithm in Section 4. Conducted experiments on real dataset are shown in Section 5. Finally, we conclude in Section 6.

II. Related Work

Related work on moving object clustering can be categorized into two main approaches: moving object cluster discovery and trajectory clustering. The former focuses on individual moving objects and tries to find clusters of objects with similar moving patterns or behaviors. For instance, the flock notion, firstly introduced in [13] and further studied in [6, 7], is defined as a group of moving objects moving together in a disc of a fixed radius for *k* consecutive timestamps. A recent study by Jeung [1, 2] proposes the convoy concept, an extension of flock, where spatial clustering is based on density. In [4], authors propose to extract moving clusters, i.e. a group of moving objects which have considerably portion of overlap at any two consecutive timestamps.

The second kind of approaches aims at finding trajectory clusters which reveal the common paths for a group of moving objects. The first and most difficult challenge for trajectory clustering is to give a good definition of similarity between two trajectories. Many methods have been proposed [5, 11, 14]. As pointed out in Lee et al. [14], distance measure established on whole trajectories may miss interesting common paths in sub-trajectories. To find clusters based on sub-trajectories, Lee et al. [14] proposed a partition-and-group framework.

The common part of mentioned work above is that moving object must move together in *k* consecutive timestamps. However, object movements are very complex and include a lot of noise (time gaps and space gaps). To address this issue, Florian Verhein [5] introduces the approach in that objects should move together and noise is reduced by solving time gaps and space gaps problem. Recently, the innovation movement model has also been proposed by Zhenhui Li [3] where moving objects are not required to be together in consecutive time timestamps. They only need to be together at least some minimum discrete timestamps. Comparing with all these definitions and movement models, gradual-spatio-temporal rule is more general since objects are not required to move together.

Gradual rules are usually applied on data sets with m attributes $(X_1, X_2, \dots, X_m)$ defined on numeric domains $dom(X_i)$. A dataset $\mathcal{D}$ is a set of rows (m-uplets) of $dom(X_1) \times, \dots, dom(X_m)$. In this scope, a gradual item is defined as a pair of an attribute and a variation $* \in \{\leq, \geq\}$. Let A be an attribute then the gradual item $A^{\geq}$ means that the attribute A is increasing. For example, $Salary^{\geq}$ is a gradual item meaning the "Salary" is increasing. Gradual items can be grouped together to define gradual rules. Such a gradual rule could be $(Age^{\geq}, Salary^{\geq})$ meaning that the older an employee is, the higher the salary the employee will be.

Another definition of support and confidence of a gradual rule was proposed in [8]. In fact, the support of a gradual rule $A_1^{*_1}, \dots, A_p^{*_p}$, is defined as the maximal number of rows $\{r_1, \dots, r_l\}$ for which there exists a permutation π such that $\forall j \in [1, l-1], \forall k \in [1, p]$, it holds $A_k\left(r_{\pi_j}\right) *_k A_k\left(r_{\pi_{j+1}}\right)$. More formally, denoting $\mathcal{L}$ the set of all such sets of rows, the support of a gradual rule is defined as follows. Assume that $s = A_1^{*_1}, \dots, A_p^{*_p}$ be a gradual rule, we have:

$$supp(s) = \frac{max_{L_i \in \mathcal{L}} |L_i|}{|\mathcal{D}|}$$

The authors propose a heuristic to evaluate this support for gradual itemsets, in a level-wise process that considers itemsets of increasing lengths. In this method, called *GRITE*, the data is presented through a graph where nodes are defined as the objects in data, and vertices express the precedence relationships derived from the considered attributes.

Unfortunately, to the best of our knowledge, mining gradual patterns in mobile object applications have not been studied properly. In this paper, we introduce the novel concept of *gradual-spatio-temporal* rule in which graduality will be embedded into *spatio-temporal* pattern (mobile object applications). Additionally, we also propose an effective algorithm, named *GSTD*, to discover meaningful gradual-spatio-temporal rules.

III. Problem Statement

A. Rule Definition

Prior to providing the definition of a rule, we introduce the notion of positive and negative moving object which are defined as follows:

Definition 1. Let $\Delta = \{\Delta_{EW}, \Delta_{WE}, \Delta_{NS}, \Delta_{SN}\}$. In this study, we consider four different directions sequentially being "E→W", "W→E", "N→S", "S → N" with E, W, N, S are abbreviations of east, west, north and south. Let Δ_q be a direction with $q \in \{EW, WS, NS, SN\}$. A positive moving object o is an object whose the movement is the same as Δ_q in the time duration $[t_i, t_j]$. In contrast, negative moving object o is an object whose the movement is different from Δ_q in the time duration $[t_i, t_j]$.

For instance (See Fig 1), the objects having the direction represented by the dotted arrows, for the east to west direction, are positive moving objects and the other ones (i.e. o_4, o_6 *and* o_8) are negative moving objects.

Assume that we have a set of moving objects $O = \{o_1, o_2, \dots, o_z\}$, a set of timestamps $T = \{t_1, t_2, \dots, t_n\}$, a variation $* \in \{+, -\}$ and a set of directions $\Delta = \{\Delta_{EW}, \Delta_{WE}, \Delta_{NS}, \Delta_{SN}\}$. The mentioned gradual-spatio-temporal rule could be regularly formed into $< T^*, O^* \rightarrow \Delta_q >$ with $q \in \{EW, WS, NS, SN\}$.

For instance, the rule r is "the more time is going on, the more objects are moving from east to west" will be represented as $< T^+, O^+ \rightarrow \Delta_{EW} >$ where T^+ means the time is going on, O^+ means the number of objects are increasing, Δ_{EW} means the objects are moving from east to west. We can recognize that there could be many different rules based-on the variation of $*$ and Δ.

B. Supporting Time Segment and Supporting Time Pattern

In our study, one of the most crucial issues is to define gradual-spatio-temporal pattern such that graduality could be smoothly expressed. Let take a look to the rule "*The more time is going on, the more objects are moving from east to west*", we have three different data dimensions being time, number of objects, moving area. On the entire rule's point of view, graduality is taken into account of all the data dimensions. Therefore, the graduality of gradual-spatio-temporal pattern definition should be expressed in terms of time, number of objects and space.

In this section, we proposed the novel concept of supporting segment to reflect the graduality.

Definition 2. Given a time segment $[t_i, t_{i+1}]$, a set of moving objects $O = \{o_1, o_2, \dots, o_z\}$, the rule $r: < T^*, O^* \rightarrow \Delta_q >$ with $q \in \{EW, WS, NS, SN\}$. The *support value of a rule* in $[t_i, t_{i+1}]$, denoted $\sigma(r)_{[t_i,t_{i+1}]}$, is evaluated as follows:

$$\sigma(r)_{[t_i,t_{i+1}]} = \frac{p_{moving,r_{[t_i,t_{i+1}]}}}{p_{moving,r_{[t_i,t_{i+1}]}} + n_{moving,r_{[t_i,t_{i+1}]}}} \quad (1)$$

$$\bar{\sigma}(r)_{[t_i,t_{i+1}]} = \frac{n_{moving,r_{[t_i,t_{i+1}]}}}{p_{moving,r_{[t_i,t_{i+1}]}} + n_{moving,r_{[t_i,t_{i+1}]}}} \quad (2)$$

$$\sigma(r)_{[t_i,t_{i+1}]} + \bar{\sigma}(r)_{[t_i,t_{i+1}]} = 1$$

where $p_{moving,r_{[t_i,t_{i+1}]}}$ and $n_{moving,r_{[t_i,t_{i+1}]}}$ stand for the number of positive and negative moving objects during the time duration $[t_i, t_{i+1}]$. Note that $\bar{\sigma}(r)_{[t_i,t_{i+1}]}$ is the complement of $\sigma(r)_{[t_i,t_{i+1}]}$.

As previously seen in the previous section, too many rules can be detected. We can also notice that each interesting rule has different starting timestamp. For instance (See Fig 2), the rule $r_1: < T^+, O^+ \rightarrow \Delta_{EW} >$ is interesting from timestamp $[t_1, t_2]$. Its value is 1 because of $p_{moving,r_{1[t_1,t_2]}} = 3$ (three objects respect the rule, i.e. o_1, o_2, o_3) and $n_{moving,r_{1[t_1,t_2]}} = 0$. While the rule $r_2: < T^+, O^+ \rightarrow \Delta_{WE} >$ is interesting from timestamp $[t_2, t_3]$ (the support value is 3/7 ($p_{moving,r_{2[t_2,t_3]}} = 3$, i.e. o_3, o_4, o_5) and $n_{moving,r_{2[t_2,t_3]}} = 4$, i.e. o_1, o_2, o_6, o_7). Therefore, a minimum support threshold σ_0 for the starting timestamp is proposed to hold this issue. Additionally, uninteresting rules, whose support values are appropriately low in all timestamps in T, could be limited by using this minimum support threshold σ_0.

Definition 3. Given a set of timestamps $T = \{t_1, t_2, \dots, t_n\}$, a time segment $[t_{st}, t_{st+1}]$, the rule $r: < T^*, O^* \rightarrow \Delta_q >$. $[t_{st}, t_{st+1}]$ is the *starting timestamp* of the rule r if it is the first timestamp (from timestamp t_1) whose support value is larger than the minimum support threshold σ_0.

$$\begin{cases} \sigma(r)_{[t_{st},t_{st+1}]} \geq \sigma_0 \\ \forall u \ (1 \leq u \leq st-1): \begin{cases} \sigma(r)_{[t_{st},t_{st+1}]} > \sigma(r)_{[t_u,t_{u+1}]} \\ \sigma(r)_{[t_u,t_{u+1}]} < \sigma_0 \end{cases} \end{cases}$$

After detecting the starting timestamp t_{st}, the rule r will be considered as an interesting rule candidate.

Definition 4. Given a time segment $[t_i, t_{i+1}]$, a rule candidate $r: < T^*, O^* \rightarrow \Delta_q >$ with $n-1 \geq i \geq t_{st}$ (t_{st} is the starting timestamp of *r*). $[t_i, t_{i+1}]$ is a *supporting time segment* if and only if the object movements satisfy the rule in $[t_i, t_{i+1}]$. It means that: 1) if $O^* = O^+$ (resp. $O^* = O^-$), the number of positive moving objects in $[t_i, t_{i+1}]$ is larger than or equal (resp. smaller than or equal) the number of positive moving objects in $[t_{i-1}, t_i]$; 2) if $O^* = O^+$ (resp. $O^* = O^-$), the support value, denoted $\sigma(r)_{[t_i,t_{i+1}]}$, is larger than or equal (resp. smaller than or equal) the support value in $[t_{i-1}, t_i]$, $\sigma(r)_{[t_{i-1},t_i]}$; 3) the support value is larger than or equal to the minimum support threshold σ_0.

$$if\ O^* = O^+ then \begin{cases} p_{moving,r_{[t_i,t_{i+1}]}} \geq p_{moving,r_{[t_{i-1},t_i]}} & : condition(1) \\ \sigma(r)_{[t_i,t_{i+1}]} \geq \sigma(r)_{[t_{i-1},t_i]} & : condition(2) \\ \sigma(r)_{[t_i,t_{i+1}]} \geq \sigma_0 & : condition(3) \end{cases}$$

$$if\ O^* = O^- then \begin{cases} p_{moving,r_{[t_i,t_{i+1}]}} \leq p_{moving,r_{[t_{i-1},t_i]}} & : condition\ (1) \\ \sigma(r)_{[t_i,t_{i+1}]} \leq \sigma(r)_{[t_{i-1},t_i]} & : condition(2) \\ \sigma(r)_{[t_i,t_{i+1}]} \geq \sigma_0 & : condition(3) \end{cases}$$

For instance, assume that $\sigma_0 = 0.3$, let consider the rule $r: < T^+, O^+ \rightarrow \Delta_{EW} >$. From Fig 2, we have $p_{moving,r_{[t_3,t_4]}} = 5$ (i.e. o_7, o_1, o_2, o_3, o_8), $n_{moving,r_{[t_3,t_4]}} = 2$ (i.e. o_4, o_5) and therefore $\sigma(r)_{[t_3,t_4]} = 5/7$ (formula 1), $\bar{\sigma}(r)_{[t_3,t_4]} = 2/7$ (formula 2). Similarly, $\sigma(r)_{[t_2,t_3]} = 4/7$, so

TABLE 1. NOTATIONS DESCRIPTION.

Notation	Description
$O = \{o_1, o_2, \dots, o_z\}$	Objects set
$T = \{t_1, t_2, \dots, t_n\}$	Set of timestamps
$* \in \{+, -\}$	Variation
Δ	Set of directions
r	Gradual-spatio-temporal rule
θ	Threshold for neutral time segment defining
$\sigma(r)_{[t_i,t_{i+1}]}$	Support of rule r in time segment $[t_i, t_{i+1}]$
$\bar{\sigma}(r)_{[t_i,t_{i+1}]}$	Complement of $\sigma(r)_{[t_i,t_{i+1}]}$
$\mathcal{D}$	Set of time patterns
$\sigma(r)_T$	Support of rule r in set of timestamps T
$c(r)_T$	Confidence of rule r in set of timestamps T
$p_{moving,r_{[t_i,t_{i+1}]}}$	Number of positive moving objects
$n_{moving,r_{[t_i,t_{i+1}]}}$	Number of negative moving objects
p_s	Supporting time pattern
p_{ns}	Non-supporting time pattern
p_{nt}	Neutral time pattern

$\sigma(r)_{[t_3,t_4]} > \sigma(r)_{[t_2,t_3]}$ (condition 2). Additionally, $p_{moving,r_{[t_3,t_4]}} > p_{moving,r_{[t_2,t_3]}} = 4$ (condition 1) and $\sigma(r)_{[t_3,t_4]} > \sigma_0$ (condition 3). So, $[t_3, t_4]$ is a supporting time segment.

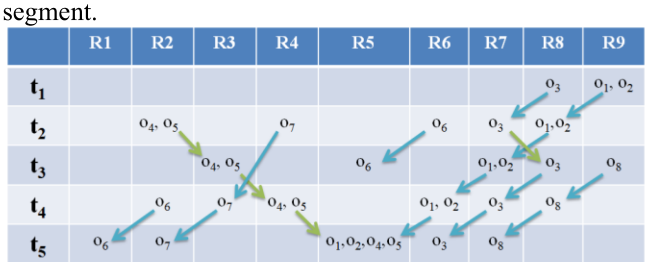


Fig. 2 Example of moving objects

At the moment, we already proposed the definition of the support for a time segment. In further, we also have non-supporting time segment and neutral time segment. Non-supporting time segment is a time segment in that there are many moving objects but not enough to satisfy the rule. In the other hand, neutral time segment $[t_i, t_{i+1}]$ is a time segment in that there are not enough moving objects. To distinguish neutral time segment from other kinds of time segment, we defined a minimum moving objects threshold θ so that if the total number of moving objects in $[t_i, t_{i+1}]$ is smaller than θ then $[t_i, t_{i+1}]$ is a neutral time segment.

$$p_{moving,r_{[t_i,t_{i+1}]}} + n_{moving,r_{[t_i,t_{i+1}]}} < \theta$$

where θ is a predefined threshold.

Now, we propose the *k*-supporting time pattern definition (denoted p_s), *k*-non-supporting time pattern definition (denoted p_{ns}) and *k*-neutral time pattern (denoted p_{nt}).

Definition 5. Let $p_s = (t_i, t_{i+k}), |p_s| = k\ (k \geq 1)$, p_s is a *k-supporting time pattern* if and only if $\forall j\ (0 \leq j < k)$, $[t_{i+j}, t_{i+j+1}]$ is a *supporting time segment*. In the opposite, $p_{ns} = (t_i, t_{i+k})$ is a *k-non-supporting time pattern* if and only if $\forall j\ (0 \leq j < k), [t_{i+j}, t_{i+j+1}]$ is a *non-supporting time segment*. Finally, $p_{nt} = (t_i, t_{i+k})$ is a *k-neutral time pattern* if and only if $\forall j\ (0 \leq j < k), [t_{i+j}, t_{i+j+1}]$ is a *neutral time segment*.

After defining all the *k*-supporting time patterns, *k*-non-supporting time patterns and *k*-neutral time patterns, we have a set of patterns $\mathcal{D} = (\{p_{s,1}, p_{s,2}, \ldots, p_{s,n}\}; \{p_{ns,1}, p_{ns,2}, \ldots, p_{ns,m}\}; \{p_{nt,1}, p_{nt,2}, \ldots, p_{nt,l}\})$ including all the patterns. With $p_{s,i}$ is the supporting time pattern i^{th}, $p_{ns,j}$ is the non-supporting time pattern j^{th} and $p_{nt,p}$ is the neutral time segment p^{th}.

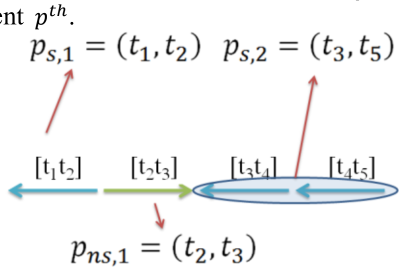


Fig. 3 Supporting and non-supporting time patterns in Fig 2

For instance (See Fig. 3), we have two supporting time patterns. $p_{s,1} = (t_1, t_2)$ is a 1-supporting time pattern, $p_{s,2} = (t_3, t_5)$ is a 2-supporting time pattern. $p_{ns,1} = (t_2, t_3)$ a 1-non-supporting pattern.

C. Support and Confidence of Rules in a Set of Timestamps

Usually, the support of a rule is sequentially defined as the number of items that satisfy the rule in gradual rule discovering and the number of objects that travel together in moving objects clusters discovering. While, there is a little bit difference in our context: 1) items are moving objects, while they usually are numerical in gradual itemsets discovering techniques, 2) the objects are not necessary to be together, while they are required to travel together in certain timestamps in moving objects clusters discovering studies.

Recall the main issue of our study, gradual rule definition and spatio-temporal pattern definition are combined together (called gradual-spatio-temporal rule). For instances: "*The more time is going on, the more objects are moving from east to west*" or "*The more time is going on, the more objects are collected together*". Let consider the above rules, we could recognize that objects movements could satisfy the rule in some time duration and do not support the rule in other time duration. In further, the much supporting time is, the larger support value will be. It means the support of a rule and the supporting time closely correlate to each other and therefore support should be defined based on the supporting time. Similarly, confidence of a rule is also correlated to supporting time.

In the previous section, we already proposed the definition of supporting time pattern. Based on it, the support value of the rule in all time duration T (denoted $\sigma(r)_T$) and the confidence value of the rule (denoted $c(r)_T$) are defined as follows:

In a set of patterns $\mathcal{D} = (\{p_{s,1}, p_{s,2}, \ldots, p_{s,n}\}; \{p_{ns,1}, p_{ns,2}, \ldots, p_{ns,m}\}; \{p_{nt,1}, p_{nt,2}, \ldots, p_{nt,l}\})$ of the rule $r: < T^*, O^* \rightarrow \Delta_q >$ with $q \in \{EW, WS, NS, SN\}$, we have:

$$\begin{cases} \{p_{s,1}, p_{s,2}, \ldots, p_{s,n}\}\ supports\ the\ rule\ r \\ \{p_{ns,1}, p_{ns,2}, \ldots, p_{ns,m}\}\ does\ not\ support\ the\ rule\ r \\ \{p_{nt,1}, p_{nt,2}, \ldots, p_{nt,l}\}\ does\ not\ support\ the\ rule\ r \end{cases}$$

We can recognize that supporting time patterns, $\{p_{s,1}, p_{s,2}, \ldots, p_{s,n}\}$, support the rule while other patterns do not support the rule. Therefore, support $\sigma(r)_T$ of a rule should be the total number of supporting time segments and confidence $c(r)_T$ of rule should be computed as the number of supporting time segments to all time segments.

Definition 6. Given a set of timestamps $T = \{t_1, t_2, \ldots, t_n\}$ and a set of time patterns $\mathcal{D} = (\{p_{s,1}, p_{s,2}, \ldots, p_{s,n}\}; \{p_{ns,1}, p_{ns,2}, \ldots, p_{ns,m}\}; \{p_{nt,1}, p_{nt,2}, \ldots, p_{nt,l}\})$. The *support of a rule*, denoted $\sigma(r)_T$, is the total number of supporting time segments:

$$\sigma(r)_T = \sum_{i=1}^{n} |p_{s,i}|$$

The *confidence of a rule*, denoted $c(r)_T$, is used to measure how well we believe that an arbitrary set of moving objects could satisfy the rule in all time duration *T*. It is evaluated as the number of supporting time segments to all time segments:

$$c(r)_T = \frac{\sum_{i=1}^{n} |p_{s,i}|}{\sum_{i=1}^{n} |p_{s,i}| + \sum_{j=1}^{m} |p_{ns,j}| + \sum_{k=1}^{l} |p_{nt,k}|}$$

For instance (See Fig 3), we have:

$$\sigma(r)_T = \sum_{i=1}^{2} |p_{s,i}| = 3$$

$$c(r)_T = \frac{\sum_{i=1}^{2} |p_{s,i}|}{\sum_{i=1}^{2} |p_{s,i}| + \sum_{j=1}^{1} |p_{ns,j}|} = \frac{3}{4} = 0.75$$

IV. Gradual-Spatio-Temporal Rule Discovering Algorithm

In this section, we present the *G*radual-*S*patio-*T*emporal rule *D*iscovering (GSTD) algorithm defined for extracting rules.

Algorithm **GSTD:** *G*radual-*S*patio-*T*emporal rule *D*iscovering
Input: a set of objects O, a set of timestamps $T = \{t_1, t_2, \dots, t_n\}$, a set of directions $\Delta = \{\Delta_{EW}, \Delta_{WE}, \Delta_{NS}, \Delta_{SN}\}$, a minimum support threshold σ_0, a minimum moving objects threshold θ.
Output: a set of rules r.

$candidateSet = \emptyset$;
for each timestamp $t_i \in T$
  for each direction $\Psi \in \Delta$
    $r < T^+, O^+ \rightarrow \Psi >.p_{[t_i,t_{i+1}]} \leftarrow ComputePositiveMovingObject(O)$;
    $r < T^+, O^+ \rightarrow \Psi >.n_{[t_i,t_{i+1}]} \leftarrow ComputeNegativeMovingObject(O)$;
    $r < T^+, O^+ \rightarrow \Psi >.\sigma_{[t_i,t_{i+1}]} \leftarrow \frac{p}{p+n}$;
    if $r \notin candidateSet$ **then**
      if $r < T^+, O^+ \rightarrow \Psi >.\sigma_{[t_i,t_{i+1}]} \geq \sigma_0$ **then**
        $r < T^+, O^+ \rightarrow \Psi >.t_{st} \leftarrow t_i$; // t_{st} is the starting point
        $candidateSet \leftarrow candidateSet \cup r < T^+, O^+ \rightarrow \Psi >$;
      end if
    end if
    for each rule $r \in candidateSet$ // $r < T^+, O^+ \rightarrow \Psi >$
      if $p_{[t_i,t_{i+1}]} \geq p_{[t_{i-1},t_i]}$ *and* $\sigma(r)_{[t_i,t_{i+1}]} \geq \sigma(r)_{[t_{i-1},t_i]}$
          and $\sigma(r)_{[t_i,t_{i+1}]} \geq \sigma_0$ *and* $(p_{[t_i,t_{i+1}]} + n_{[t_i,t_{i+1}]}) \geq \theta$ **then**
        $r.sp \leftarrow r.sp + 1$; //$r.sp$ is the supporting time
      end if
    end for
  end for
end for
for each rule $r \in candidateSet$ // $r < T^+, O^+ \rightarrow \Psi >$
  $\sigma(r)_T \leftarrow r.sp$;
  $c(r)_T \leftarrow \frac{r.sp}{|T|}$;
  return r;
end for

For each timestamp t_i, for each direction Ψ, we compute the number of positive moving objects and the number of negative moving objects and the supporting value, $\sigma_{[t_i,t_{i+1}]}$, for the rule $r < T^+, O^+ \rightarrow \Psi >$. Next, we check whether the current rule r holds the minimum support threshold condition to be inserted into the set of candidates. Next, each rule r in $candidateSet$ will be checked to verify that it is satisfied of not. If it is satisfied, we increase the supporting time, $r.sp$, by one and continue scanning the next timestamp. After all the timestamp t_i is scanned, we return all the rules with appropriate confidence and support.

V. Experimental Results

In order to evaluate the GSTD algorithm, we use the Swainsoni dataset[1] [3] including 43 objects evolving over time and 4225 different timestamps. The dataset is gathered from July 1995 to June 1998. GSTD algorithm is applied to discover gradual-spatio-temporal rules in this dataset. By setting $\sigma_0 = 0.3$ and $\theta = 2$ we can find four different rules with corresponding support and confidence (See Table 2).

Table 2. Support and Confidence for each Rule.

Rule	$\sigma(r)_T$	$c(r)_T$
$< T^+, O^+ \rightarrow \Delta_{NS} >$	2653	0.6416
$< T^+, O^+ \rightarrow \Delta_{SN} >$	1944	0.4617
$< T^+, O^+ \rightarrow \Delta_{WE} >$	2725	0.6472
$< T^+, O^+ \rightarrow \Delta_{EW} >$	2060	0.4903

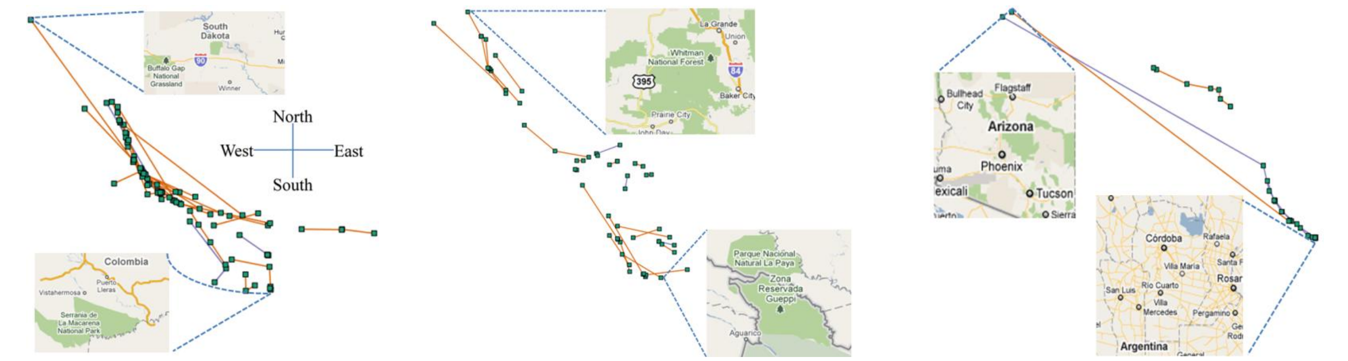
In Table 2, we can notice that the supports and confidences of rules $< T^+, O^+ \rightarrow \Delta_{NS} >$ and $< T^+, O^+ \rightarrow \Delta_{WE} >$ are higher than other rules. Intuitively, we could conclude that the more objects are moving from North to South and from West to East.

To highlight the difference between our patterns and the convoy or closed swarm, we extracted them by using the closed swarms and convoy algorithms[2][15]. As too many patterns can be extracted from the dataset, we only present the most typical ones in Fig 4. Let take a look at the Fig 4 (a)(b)(c)(d), they are the longest patterns of different gradual-spatio-temporal rules which are discovered from the dataset. The orange connections are the moving paths of positive moving objects and the blue ones are the moving paths of negative moving objects. Fig 4 (e), (f) present the typical convoy and closed swarm being extracted from the dataset.

Patterns extracted from gradual-spatio-temporal rules are quite different compare with the ones from convoy and closed swarm. In fact, patterns from convoys or closed swarms could be expressed as "*several objects move together from a specific location to other specific location*". For instance (See Fig 4(e)(f)), from 1996-11-17 to 1997-02-28, two objects move together from Carrasco national park (Bolivia) to Cordoba city (Argentina) while from 1996-08-29 to 1997-06-17, two objects move together from Challis national forest (USA) to Mar Del Plata city (Argentina). In the other hand, the extracted patterns from the gradual-spatio-temporal rules could be described as "*the more time is going on, the more objects are moving from a specific location to other specific location*". For instance (See Fig 4(a)), from 1996-10-07 to 1996-10-30, the more time is going on, the more objects are moving from South Dakota (USA) to "Serrainia de La Macarena" national park (Colombia).

---

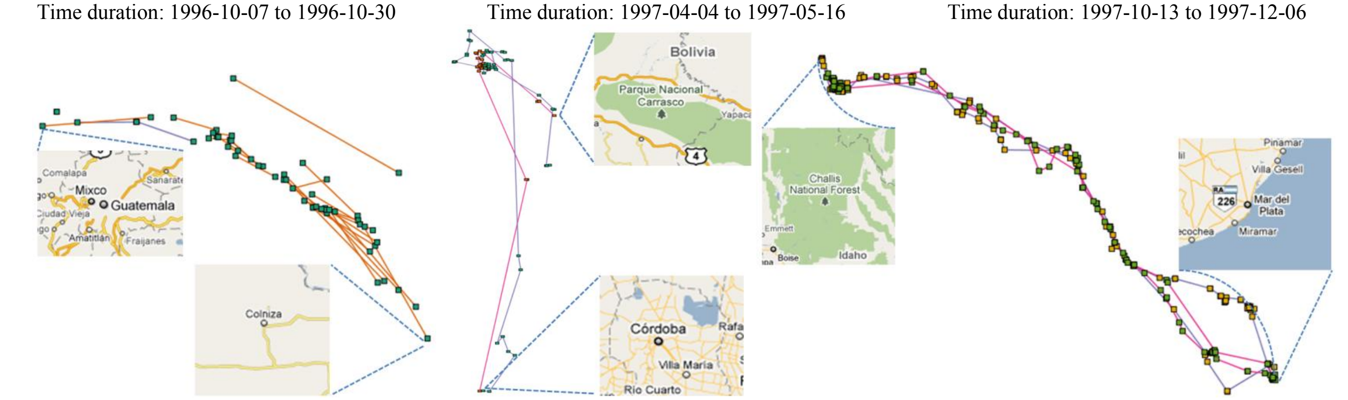
[1] http://www.movebank.org
[2] http://dm.cs.uiuc.edu/movemine/



(a)-Longest pattern of $r\!<T^{+},O^{+}\rightarrow\Delta_{NS}\!>$
Time duration: 1996-10-07 to 1996-10-30

(b)-Longest pattern of $r\!<T^{+},O^{+}\rightarrow\Delta_{SN}\!>$
Time duration: 1997-04-04 to 1997-05-16

(c)-Longest pattern of $r\!<T^{+},O^{+}\rightarrow\Delta_{WE}\!>$
Time duration: 1997-10-13 to 1997-12-06



(d)-Longest pattern of $r\!<T^{+},O^{+}\rightarrow\Delta_{EW}\!>$
Time duration: 1997-03-10 to 1997-03-29

(e)-Convoy
Time duration: 1996-11-17 to 1997-02-28

(f)-Closed swarm
Time duration: 1996-08-29 to 1997-06-17

Fig. 4 The difference between gradual-spatio-temporal rule patterns and convoy, closed swarm

Additionally, in term of object movements, the patterns expressed in gradual-spatio-temporal rules comparing to convoy or closed swarm are quite different. There are many different objects being moving independently in the gradual-spatio-temporal rules (See Fig 4(a)(b)(c)(d)). While, the objects travel together in convoy and closed swarm (consecutive or non-consecutive timestamps). In fact, the objects belong to the convoy and the closed swarm (See Fig 4 (e)(f)) always travel together in consecutive timestamps.

Finally, we can conclude that gradual-spatio-temporal rule's patterns are interesting and quite different with previous work in terms of presented knowledge and object movements.

VI. Conclusion And Future Work

In this paper, we proposed the concept of gradual-spatio-temporal rule in that the graduality is embedded into spatio-temporal patterns. This concept is quite novel and different from the previous work and enables the discovery of new interesting moving object rules. In further research, GSTD will be improved to automatically discover all the interesting gradual-spatio-temporal rules.

References

[1] H. Jeung, M. L. Yiu, X. Zhou, C. S. Jensen, H. T. Shen. Discovery of Convoys in Trajectory Databases. VLDB '08, August, 2008, Auckland, New Zealand.

[2] H. Jeung, X. Zhou, H. T. Shen. Convoy Queries in Spatio-Temporal Databases. ICDE 2008, April, 2008, Cancun, Mexico.

[3] Z. Li, B. Ding, J. Han, R. Kays. Swarm: Mining Relaxed Temporal Moving Object Clusters. VLDB 2010, September, 2010, Singapore.

[4] P. Kalnis, N. Mamoulis, S. Bakiras. On Discovering Moving Clusters in Spatio-temporal Data. SSTD 2005, Angra dos Reis, Brazil, August, 2005. LNCS 3633.

[5] F. Verhein. Mining Complex Spatio-Temporal Sequence Patterns. SDM'09. April, 2009, John Ascuaga's Nuggest-Sparks, Nevada.

[6] J. Gudmundsson, M. J. van Kreveld, and B. Speckmann. Efficient detection of motion patterns in spatio-temporal data sets. In GIS, 2004.

[7] J. Gudmundsson and M. van Kreveld. Computing Longest Duration Flocks in Trajectory Data. In GIS, 2006.

[8] L. Di Jorio, A. Laurent and M. Teisseire. Mining Frequent Gradual Itemsets from Large Databases. In Proceedings of the 8th International Symposium on Intelligent Data Analysis Lyon France, pp. 297-308, September 2009.

[9] F. Berzal, J. Cubero, D. Sanchez, M. Vila, and J. Serrano. An Alternative Approach to Discover Gradual Dependencies. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15(5):559-570, 2007.

[10] S. Ayouni, A. Laurent, S. B. Yahia, and P. Poncelet. Mining Closed Gradual Patterns. ICAISC 2010. Zakopane, June, 2010.

[11] J. Yang, M. Hu. TrajPattern: Mining Sequential Patterns from Imprecise Trajectories of Mobile Object. EDBT 2006, Munich, Germany, March, 2006. LNCS 3896, pp.664-681.

[12] G. Al-Naymat, S. Chawla, and J. Gudmundsson. Dimensionality reduction of long duration and complex spatio-temporal queries. ACM SAC 2007, Seoul, Korea, March, 2007.

[13] P. Laube and S. Imfeld. Analyzing relative motion within groups of trackable moving point objects. In GIS, 2002.

[14] J.-G. Lee, J. Han, and K.-Y. Whang. Trajectory clustering: A partition-and-group framework. In SIGMOD, 2007.

[15] Z. Li, M. Ji, J.-G. Lee, L. Tang, Y. Yu, J. Han, and R. Kays. Movemine: Mining moving object databases. In *SIGMOD*, 2010.