Trend-Based Prediction of Spatial Change

Xiaoyu Ge,[*] **Jae Hee Lee,**[†] **Jochen Renz,**[*] **Peng Zhang**[*]

[*] Research School of Computer Science, Australian National University
{xiaoyu.ge,jochen.renz,p.zhang}@anu.edu.au
[†]QCIS, FEIT, University of Technology Sydney
jaehee.lee@uts.edu.au

Abstract

The capability to predict changes of spatial regions is important for an intelligent system that interacts with the physical world. For example, in a disaster management scenario, predicting potentially endangered areas and inferring safe zones is essential for planning evacuations and countermeasures. Existing approaches usually predict such spatial changes by simulating the physical world based on specific models. Thus, these simulation-based methods will not be able to provide reliable predictions when the scenario is not similar to any of the models in use or when the input parameters are incomplete. In this paper, we present a prediction approach that overcomes the aforementioned problem by using a more general model and by analysing the trend of the spatial changes. The method is also flexible to adopt to new observations and to adapt its prediction to new situations.

1 Introduction

Spatial changes that characterise events and processes in the physical world are ubiquitous in nature. As such, spatial changes have been a fundamental topic in different research areas such as computer vision, knowledge representation and geographic information science.

In this paper, we are interested in predicting the spatial changes of evolving regions, which is an important capability for decision makers. For example in a disaster management scenario, it is essential to predict potentially endangered regions and to infer safe regions for planning evacuations and countermeasures. A more commercial example are solar power providers who need to estimate solar power output based on predicted cloud coverage.

Existing approaches to predicting changes of spatial regions typically use simulations tailored to a specific application domain, e.g., wildfire progression and urban expansion. The models behind such simulations are based on physical laws that require a detailed understanding of microscopic phenomena of the given application domain and rely on specific domain parameters that need to be known. If not all parameters are known, or if the underlying model is not fully understood, predictions based on simulations might fail.

This shortcoming is particularly problematic in disaster situations, where time is often the most critical factor. An immediate response can save lives and prevent damage and destruction. An immediate response might not leave enough time to build a domain model, to adjust it to the situation at hand, or to collect and apply the required domain parameters. What is needed is a system that provides fast and reliable predictions, that can be quickly applied to any new domain and new situation, that can adopt to new observations, and that does not require intimate domain knowledge.

We present a general prediction method with all these desired properties. Our method can be applied to any kind of evolving regions in any application domain, as long as the regions follow some basic principles of physics. Our method only requires a sequence of "snapshots" of evolving regions at different time points that allow us to extract the boundaries of the observed regions. This could be, for instance, aerial or satellite images, but also data from sensor networks, or a combination of different sources. Our method then samples random points from the region boundaries of each snapshot and infers the trend of the changes by trying to assign boundary points between snapshots. We developed a hybrid method that combines a probabilistic approach based on a Kalman filter with qualitative spatial reasoning in order to obtain accurate predictions based on the observed trend.

In the following we first present some related work before discussing the specific problem we are solving and the assumptions we are making. We then present our method for inferring the trend of changes and for predicting future changes. We evaluate our method using real world data from two completely different domains, which demonstrates the accuracy of our predictions. We conclude the paper with a discussion.

2 Related Work

Reasoning about spatial changes has been a fundamental topic in geographic information science, in knowledge representation as well as in computer vision.

In geographic information science, most prediction methods in the literature are tailored to a specific application area and are not based on trend-analysis of spatial change. For example [Alexandridis *et al.*, 2008; Ghisu *et al.*, 2015] predict wildfire progression and [Santé *et al.*, 2010; Singh *et al.*, 2015] urban expansion based on cellular automata simulations which cannot adopt sensor data to guide the prediction. Rochoux *et*

al. [2014] on the other hand use a data assimilation method (extended Kalman filter) to estimate certain environmental parameters in a wildfire simulator, the parameters including seven biomass fuel properties and the wind. Peng *et al.* [2014] track cloud movement based on a constant velocity model for forecasting short-term solar irradiance.

From the perspective of a general prediction method Junghans and Gertz [2010] propose a general method to track evolving regions in disaster scenarios which uses minimum bounding rectangles to represent the tracked regions. The method estimates the bounding rectangles of the future regions by means of linear regression.

In knowledge representation and reasoning, qualitative spatial change [Galton, 2000; 2009] has been identified as a vehicle for commonsense reasoning about space, where *predictive* reasoning plays an essential role [Bhatt, 2014]. Despite this importance, research mainly focused on detecting spatial changes [Worboys and Duckham, 2006; Jiang *et al.*, 2009; Liu and Schneider, 2011] but not on predicting them.

In computer vision, the active contour model has been used to track and predict regions with smooth and parametrisable shapes [Blake and Isard, 1998]. In this model, region boundaries are approximated by splines and their positions and shapes are tracked by means of a Kalman filter. The active contour model approach is similar to our approach in that the target objects are region boundaries and the Kalman filter is used for tracking and prediction. However, the active contour model is more focussed on tracking splines by using a default shape information of the target region and is not suitable for prediction tasks where no such default shape information is given (e.g. in wildfire or cloud movement prediction).

3 Problem Description and Approach

Our aim in this paper is to predict future changes of evolving regions by only considering information about past changes. A region R at time t in this context is defined as a set $R_t \subset \mathbb{R}^2$ that can consist of multiple connected components with holes. Each hole can recursively contain other connected components of the region. We assume that the interior of regions is homogeneous and that changes of regions are entirely captured by changes to their boundaries. We now describe the problem and the assumptions in more detail.

Input   As input we use a series of snapshots of evolving regions at different times. Each snapshot must allow us to detect the boundary of the observed regions (see Fig. 1). Hence, snapshots could be obtained, for example, from remote sensing images, sensor networks (cf. [Duckham, 2013]), or a combination of different sources. We denote the boundary of a region R_t as ∂R_t, which is a finite set of closed curves, one for each connected component and one for each hole.

Trend Analysis   Given these snapshots and the corresponding boundaries, we now try to infer the *trend* of observed changes. In order to do this, we randomly sample a set of k boundary points $z_t \in \partial R_t$ and are interested in the movement of these boundary points over time. To this end, for every pair of consecutive snapshots we assign to each boundary point
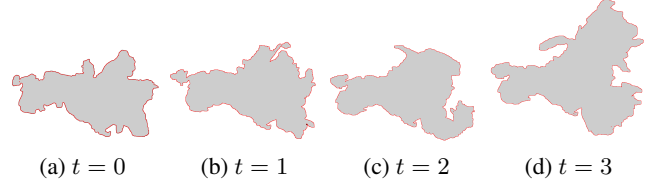

(a) $t = 0$ (b) $t = 1$ (c) $t = 2$ (d) $t = 3$

Figure 1: A wildfire progression example. The boundaries of fire regions R_t are indicated in red.

$z_t \in \partial R_t$ a boundary point $z_{t+1} \in \partial R_{t+1}$. This assignment is given by a map $f_t : \partial R_t \to \partial R_{t+1}$ that minimises the total distances[1] of the point movements, i.e.

$$f_t = \arg\min_f \sum_{z_t \in \partial R_t} \|f(z_t) - z_t\|. \qquad (1)$$

Function f_t yields a reliable assignment if each of the boundaries ∂R_t and ∂R_{t+1} consist of one closed curve as is the case in Fig. 1. However, it does not always lead to a meaningful assignment, if their topological structures are more complex. In Section 4.1 we will resolve this issue by analysing the topological structures of the boundaries and adjusting the assignment accordingly.

Once we have f_t, we can repeatedly apply f_t to ∂R_t and construct a sequence $z_0, z_1, \ldots, z_T$ of observed points from region boundaries $\partial R_0, \partial R_1, \ldots, \partial R_T$. This sequence is, however, only an approximation of the true movement of boundary points p_t and the locations of z_t are therefore subject to noise

$$z_t = p_t + b_t, \qquad (2)$$

where b_t is a zero mean Gaussian noise. We assume that this noise is random and is negatively correlated with the density D of sampled boundary points in a given area, i.e. the noise decreases with increasing D.

Prediction Model   The main assumption we make on the change of region boundaries is the following: *If there is no external cause, any point p of a region boundary $\partial R \subset \mathbb{R}$ moves with a constant speed.* We assume that there is always a noise in the model owing to the ignorance of physical parameters such as acceleration. Formally, for each $p_t \in \partial R_t$ that moves in direction $v_t \in \mathbb{R}^2$ with speed $\|v_t\|$ its position after Δt is given by

$$p_{t+\Delta t} = p_t + \Delta t \cdot v_t + a_t \qquad (3)$$

where a_t is a zero mean random noise. The covariance of a_t is positively correlated to the distance $\Delta t \cdot \|v_t\|$ that p_t is going to travel and to the recent prediction result $\|z_t - p_t\|$. The latter correlation is a mechanism to modify the prediction noise depending on its prediction performance so far, i.e. the more agreement with the prediction and the observed data, the less is the prediction noise.

---

[1]The rationale behind equation (1) is that *physical entities seek to move as efficiently as possible*, which is also known in physics as the principle of least action [Feynman *et al.*, 1965].
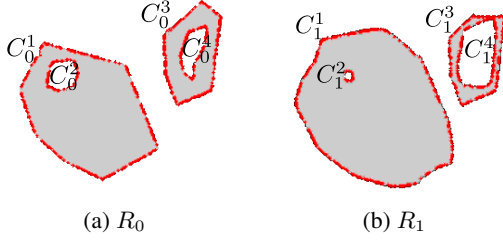
(a) R_0 (b) R_1

Figure 2: Depicted is region R_t for $t = 0$ and $t = 1$.

(a) An overlay of R_0 and R_1. Assignments are without adjustment.

(b) An overlay of R_0 and R_1. Assignments are after adjustment.
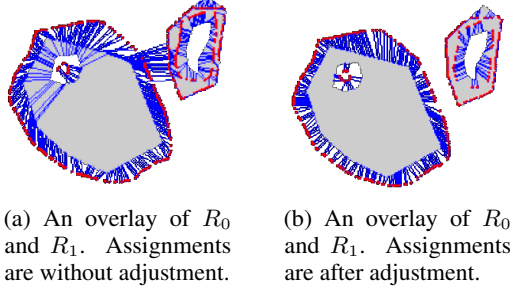
Figure 3: An example of the trend identification between evolving regions. The boundary points are indicated by the red dots. Each individual assignment between two points z_t and z_{t+1} is highlighted by a blue line.

Output The output of our approach is an approximation of region boundary $\partial R_{T+\Delta t}$. Our approximation ensures that with a high probability the real region will be inside the approximated region. The method also determines the topological change of regions (e.g. merge, split, disappearance).

4 Trend-Based Prediction Algorithm

In this section we present the prediction algorithm in detail. We first describe how the trend of a change is identified from two consecutive snapshots of a region boundary. Then we describe how this information is used to predict changes.

4.1 Qualitative Adjustment of Trend

Equation (1) in Section 3 is an instance of the *assignment problem* that can be solved by the Hungarian method [Kuhn, 1955] in $O(n^3)$ time. However, as mentioned in Section 3, equation (1) does not always lead to meaningful assignments, particularly if the the topology of the boundaries is complex. An example is illustrated in Fig. 2. In this example region boundary ∂R_t consists of four closed curves $C_t^1, C_t^2, C_t^3, C_t^4$, where C_t^1 contains C_t^2 and C_t^3 contains C_t^4 at time $t = 0$ (Fig. 2a). At time $t = 1$ the topological structure of region boundary ∂R_t remains the same, but each of the curves $C_t^1, C_t^2, C_t^3, C_t^4$ has either shrunken or expanded (Fig. 2b). As a result a one-to-one assignment of points in $C_0^1, C_0^2, C_0^3, C_0^4$ to their counterparts in $C_1^1, C_1^2, C_1^3, C_1^4$ is not possible, leading to false assignments of boundary points (Fig. 3a).

Our method solves this problem qualitatively by analysing the topological relations between closed curves of the boundaries, and by ensuring that all assignments will only be between corresponding curves.
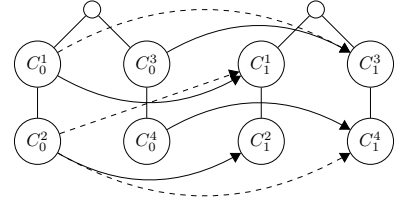

Figure 4: A graphical representation of the assignment between ∂R_0 and ∂R_1. Each line with an arrow is an assignments before adjustment. Each dashed line stands for an assignment that violates a certain rule. We resolve these violations in breadth-first order.

To this end we first construct a *containment tree* G_{R_t} that represents the topological relations between closed curves of ∂R_t. In this tree, each curve C is represented by a vertex v_C and the background is represented by the root. A vertex v_C is a parent of $v_{C'}$, if C contains C' and no other curve C'' lies in between, i.e. there is no C'' that is contained by C and at the same time contains C'.

Then, we compute the initial assignment f_t between ∂R_t and ∂R_{t+1} by solving the assignment problem in equation (1), and represent the assignment graphically using the containment trees G_{R_t} and $G_{R_{t+1}}$ (Fig. 4). We say that there is an assignment between a vertex v_{C_t} and $v_{C_{t+1}}$ when there is an assignment between the underlying curves C_t and C_{t+1}, i.e. $f_t(C_t) \cap C_{t+1} \neq \emptyset$. We then check whether there are any *conflicts* or *ambiguities* in the assignment by traversing the containment tree G_{R_t} in breadth-first order and examining its vertices. Given an assignment f_t between v_{C_t} and $v_{C_{t+1}}$

- a *conflict* arises when v_{C_t} and $v_{C_{t+1}}$ are not at the same level in the containment tree, or when there is no assignment between their parents. The conflict is resolved by preventing such an assignment;

- an *ambiguity* arises when there are assignments between v_{C_t} and $v_{C_{t+1}}$ and between v_{C_t} and $v_{C'_{t+1}}$ with $C'_{t+1} \neq C_{t+1}$. Such an ambiguity is natural, if a topological change (e.g. merge, split) occurs between the transition from R_t to R_{t+1} and is sometimes hard to resolve. Ambiguities allow us to classify and to explain topological changes (cf. Table 1).

After identifying all conflicts and ambiguities, we recalculate the assignment f_t between R_t and R_{t+1} while preventing assignments between conflicting curves. This is obtained by setting the distances between boundary points of conflicting curves so high that the Hungarian algorithm will avoid these assignments. Note that we specifically designed our method in this way as the correspondence of nodes between the two containment trees is not necessarily known.

4.2 Boundary Point Prediction

Once we obtained the trend analysis $z_0, \ldots, z_T$ of each boundary point p_t, we can then estimate its current location as well as predict its future location. To this end, we apply the Kalman filter [Kalman, 1960], which is a data fusion method that has

Algorithm 1: Boundary point prediction

Input : A sequence $z_0, z_2, \ldots, z_T$ of observed boundary points and $\Delta t \in \mathbb{N}$.

Output : The location of $p_{T+\Delta t}$ and the covariance $\Sigma_{T+\Delta t}$ of its noise.

1   $p_0 \leftarrow z_0, \quad p_1 \leftarrow z_1, \quad \Sigma_1 \leftarrow I$ `// initialise`
2   **for** $t \leftarrow 1$ **to** T **do**   `// Update current state`
3     $(p_{t+1}, \Sigma_{t+1}) \leftarrow \text{PREDICT}(p_{t-1}, p_t, 1, \Sigma_t)$
4     $(p_{t+1}, \Sigma_{t+1}) \leftarrow \text{UPDATE}(p_{t+1}, \Sigma_{t+1}, z_{t+1})$
5   **return** $\text{PREDICT}(p_{T-1}, p_T, \Delta t, \Sigma_T)$ `// pT+Δt,ΣT+Δt`

6   **Function** $\text{PREDICT}(p_{t-1}, p_t, \Delta t, \Sigma_t)$
7     $v_t \leftarrow p_t - p_{t-1}$
8     $Q_t \leftarrow \Delta t \cdot \|v_t\| \cdot \|z_t - p_t\| \cdot I$
9     $p_{t+1} \leftarrow p_t + \Delta t \cdot v_t$
10    $\Sigma_{t+1} \leftarrow \Sigma_t + Q_t$
11   **return** p_{t+1}, Σ_{t+1}

12   **Function** $\text{UPDATE}(p_t, \Sigma_t, z_t)$
13    $D \leftarrow r/\sqrt{|B_r(z_t) \cap \partial R_t|}$
14    $S_t \leftarrow D \cdot I$
15    $K_t \leftarrow \Sigma_t(\Sigma_t + S_t)^{-1}$ `// Kalman gain`
16    $p_t \leftarrow p_t + K_t(z_t - p_t)$
17    $\Sigma_t \leftarrow \Sigma_t - K_t\Sigma_t$
18   **return** p_t, Σ_t

been widely used for guidance, navigation, and control of vehicles and for object tracking in computer vision. The Kalman filter is an ideal choice as it is the optimal estimator that satisfies previous assumptions. For more details about Kalman Filter see [Kalman, 1960] or [Russell and Norvig, 2009, Section 15.4].

The Kalman filter requires a *state transition model* as well as an *observation model*. The state transition model describes how the state of a system at time t evolves linearly to a state at time $t + 1$. Our state transition model is given in equation (3) and the observation model, which describes the observation (or measurement) z_t of the true location p_t, is given in equation (2). The detailed integration and prediction procedure is implemented in Algorithm 1. In line 8 of the algorithm, Q_t is the covariance of the noise a_t in the state transition model, where I is the 2×2 identity matrix. The term $\|z_t - p_t\|$ measures the agreement between the prediction model and the observed boundary points as addressed in Section 3. In line 14 S_t is the covariance of the noise b_t in the observation model, where D reflects the density of boundary points in a neighbourhood of z_t and is defined as $D = r/\sqrt{|B_r(z_t) \cap \partial R_t|}$, where $B_r(z_t)$ is an open disk with center z_t and radius $r > 0$.

4.3 Outer Approximation and Qualitative Prediction

Given the predicted boundary points $\partial R_{T+\Delta t}$ whose locations have Gaussian noises, we can now approximate $\partial R_{T+\Delta t}$ by building an outer boundary that most likely contains the region at time $T + \Delta t$. We call this the *outer approximation* of a

Table 1: Six basic topological changes. The changes are given both pictorially and with containment trees. The left containment tree is before a change, and the right containment tree is after the change. The region boundaries where the change takes place are coloured in red.
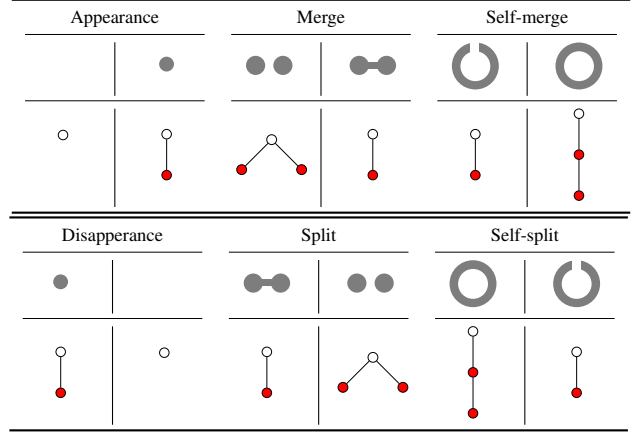
$\partial R_{T+\Delta t}$. Using our method we are also able to predict certain topological changes that might occur, as is described below.

Outer approximation To outer-approximate $R_{T+\Delta t}$ we need to adapt the outer-approximation to the noise of each predicted boundary point of $\partial R_{T+\Delta t}$, i.e. if the noise has a large covariance then its outer-approximation has to be large and vice versa. For this purpose we use the Mahalanobis distance

$$D_M(p) = \sqrt{(p - p_{T+\Delta t})^T \Sigma_{T+\Delta t}^{-1}(p - p_{T+\Delta t})}$$

which measures how far a point p is away from a boundary point $p_{T+\Delta t}$ relative to its covariance $\Sigma_{T+\Delta t}$, and build an α-concave hull of $\partial R_{T+\Delta t}$ in a way that the hull does not intersect points with $D_M(p) \leq d$ for a radius d. Thus the threshold value d controls the size of the outer approximated area and can be varied to suit different prediction requirements.

An admissible concave hull should enclose the entire region while including as few points that are outside the region as possible. We use [Edelsbrunner *et al.*, 1983] to obtain an admissible α-concave hull. The outer approximation of a region R_t is denoted as $\alpha(R_t)$

Qualitative Prediction The method predicts topological changes of the region during the transition from R_T to R_{T+1} by detecting changes of their containment trees. We detect 6 types of topological changes according to [Jiang *et al.*, 2009], namely, appear, disappear, merge, split, self-merge and self-split (Table 1).

Additional Layer Information Integration Our method is able to integrate additional information. Specifically, due to the nature of Kalman filter, we can trace the path along each predicted point and its previous states. This allows us to adjust the resulting points based on domain-specific models

and their previous states. Integrating different models may require different methods. For example, during a bushfire, isolation areas such as a highway or water body can prevent fire from burning across. Here we show the integration of isolation areas in a bushfire as an example. We assume fire cannot burn on or jump over isolation areas but has to extend along the edge. Then the predicted points will be filtered out under the following two circumstances:

Given the predicted point p_t and its previous two states p_{t-1} and p_{t-2}

1. if p_t is on the isolation area

2. if either of the path between p_t and p_{t-1} and between p_{t-1} and p_{t-2} intersects with the isolation areas.

While this is intuitive, it demonstrates the ability of our method to integrate with domain specific models.

5 Evaluation

We now evaluate our proposed method using real-world data from two different domains as well as generated data. We built our own simulator for generating test data as that allows us to generate data that is more complicated and has more erratic changes than what we saw in the real data. For each evaluation, our method takes a sequence $R_{T-w+1}, \ldots, R_T$ of w recent snapshots to predict the outer approximation of regions at $T+1$. For our analysis we varied both w and the threshold d as defined in Section 4.3. We measure the accuracy of the predicted outer approximation $\alpha(R_t)$ using the metrics of Recall (R) and Precision (P):

$$R = \frac{|\alpha(R_t) \cap R_t|}{|R_t|}, P = \frac{|\alpha(R_t) \cap R_t|}{|\alpha(R_t)|}$$

where $|R_t|$ refers to the size of the area enclosed by R_t.

Given our outer approximation task, recall is more important than precision because we care more about how our method can cover the region rather than if the prediction points are true positive. However, precision should stay in a reasonable range, as a prediction that covers the full map is useless.

5.1 Experimental Results on Real-world Data

We applied the method to two different spatial domains, namely wild fire progression and also cloud movement, where it is more likely that sudden changes or topological changes of a region occur. We use a real-world data set for each domain and the average precision and recall under different settings are summarized in Table. 2.

Wild Fire Progression: From the data base of USDA forest service[2], we obtained a sequence of progression maps of a real wild fire. Each map depicts the boundary of the wild fire on a certain date. The time gap between consecutive maps varied between 1 and 15 days. As our method is able to take only two snapshots as input for prediction, data with different time gap between adjacent snapshots can be dealt with. There have been several sudden changes in the boundary of the fire and the method can successfully deal with these
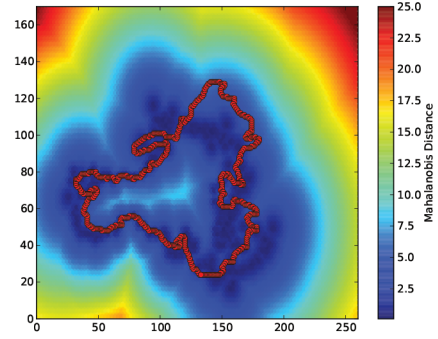
---

[2]http://www.fs.fed.us/nwacfire/ball/

Figure 5: Prediction and outer approximation of R_3 based on R_0, R_1, R_2 from Fig. 1. The red dots are the ground truth in Fig. 1d. It shows that the approximated boundary (shown in dark blue) based on $d = 2.5$ covers most of the actual boundary points. The one (shown in light blue) based on $d = 5$ covers all the actual boundary points.

changes. Fig. 5 shows an example where the method predicts the fire boundary based on the snapshots displayed in Fig. 1. Although there is a significant non-linear spatial change from R_2 to R_3, the method is still able to overcome it with a reliable outer approximation.

There are several reasons why we could not compare our approach with other wild fire prediction methods. One problem was that fire maps used to test these approaches were not publicly available, relevant papers only contain a few images of the data that was used. Another problem is that these existing tools require a detailed specification of all kinds of parameters about fuel and weather, which was not available to us. While this made it impossible for us to compare our method, it clearly demonstrates one major advantage of our method, namely that it can be applied to any types of changing regions without having to know any domain parameters.

Cloud Movement   We also obtained video data of moving clouds that is recorded using a fisheye lens (see [Wood-Bradley et al., 2012] for detail). Predicting cloud movement is particularly interesting for estimating power output of solar power plants. Before using the data, we had to flatten all images to reflect the actual shapes and movement of clouds. Unlike the wild fire case, the clouds movement is almost linear. However, the problem is still challenging because the topological structure is very complex in such a data set. Therefore adjusting the assignment of matching points with qualitative information is useful here. We compared the results with and without qualitative adjustment. The result shows that after adjusting the point assignment, our algorithm performs better on detecting the inner structure of the region (Fig. 6). Moreover both the recall and precision are generally better than the result without adjustment.

In the end, we compared our method with a state-of-the-art method developed in [Junghans and Gertz, 2010] described in Section 2. The results show that our method outperforms their method on all the data sets (see Table 2 for a summary).
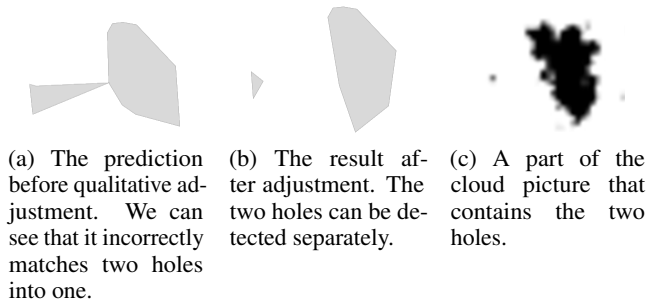
(a) The prediction before qualitative adjustment. We can see that it incorrectly matches two holes into one.

(b) The result after adjustment. The two holes can be detected separately.

(c) A part of the cloud picture that contains the two holes.

Figure 6: Comparasion between prediction results with and without qualitative adjustment.

5.2 Experimental Results on Generated Data

We applied the proposed method to a simulated environment of forest fire where fire regions undergo extensive change. The results confirm that the method can correctly identify the spatial changes of multiple evolving regions in a environment with unknown dynamics.

The underlying structure of the simulator is a two-dimensional cellular automaton (CA). It has been shown that CA are expressive enough to simulate highly complex phenomena from simple rules. Each cell has two properties: *fuel* and *heat consumption*. The amount of fuel determines how long the cell will remain on fire and the amount of heat consumption determines how long it takes for a cell of forest to become a cell of fire. The CA model simulates a global wind that has the eight cardinal directions. The wind speeds up the progression of the fire along its direction and slows down the progression in the other directions.

We evaluate the method in scenarios where the global wind randomly changes its direction and magnitude every Δt time steps. Therefore, there can be sudden changes in the wind's direction (e.g. a north wind becomes a south wind) or magnitude (e.g. a strong wind suddenly disappears). We use this setting to test the capability of the method in handling uncertainty within a highly dynamic system.

5.3 Discussion

The results in Table 2 show that our method (M_1) performed better than the state-of-the-art method (M_2) on all the data sets. Table 2 shows the average recall and precision obtained using M_1 with different pairs of w and d on each data set.

M_1 achieved an average recall of more than 90% on all the data sets while M_2 obtained an average of 83%. M_2 approximates a region using the oriented minimum bounding rectangle. Compared with the proposed method, this outer-approximation tends to include more false positive points. Besides, M_2 is not able to handle holes and sudden spatial changes such as disappearances of regions. Therefore, the precision rate of M_2 is much lower than M_1.

The results clearly reflect the relationship between d and the recall as well as the precision rate. Because a larger d results in a larger outer approximation, the recall increases and the precision decreases when d increases. The result also indicates that running M_1 with a larger window size (w) may lower down the precision. Since the regions in those data sets

Table 2: M_1: the proposed method. M_2: the method in [Junghans and Gertz, 2010]. w : window size. d : Mahalanobis threshold. R : recall. P : precision.

	d	Fire		Cloud		Simulation	
		R	P	R	P	R	P
$w = 2$							
M_2	–	0.80	0.62	0.84	0.4	0.79	0.54
M_1	0.5	0.86	0.83	0.98	0.55	0.88	0.85
	1.5	0.93	0.73	0.98	0.47	0.93	0.75
	2.5	0.97	0.56	0.98	0.44	0.96	0.71
$w = 3$							
M_2	–	0.87	0.61	0.85	0.35	0.82	0.51
M_1	0.5	0.88	0.76	0.97	0.54	0.94	0.83
	1.5	0.94	0.66	0.97	0.45	0.95	0.78
	2.5	0.98	0.52	0.98	0.41	0.98	0.68
$w = 4$							
M_2	–	0.89	0.61	0.87	0.31	0.84	0.48
M_1	0.5	0.85	0.76	0.98	0.43	0.98	0.80
	1.5	0.94	0.62	0.97	0.39	0.99	0.75
	2.5	0.98	0.47	0.97	0.37	0.99	0.65

underwent strong spatial changes over time, a larger window size introduced more uncertainties in the point prediction. As a result, the covariance of the prediction has been made larger by the method to accommodate these uncertainties, which eventually relaxed the outer approximation.

The precision rate obtained on the cloud movement data set is significantly lower than the one obtained on the wild fire data set. This is mainly because the cloud data contains many small connected components and holes. The concave hull of the outer approximation may merge a hole with a connected component if they are too close. However, the recall rate is generally higher than the wild fire data set. This is because the cloud movement can be considered as linear in most cases if the wind is not strong, which fits our assumption well.

6 Conclusions

We developed a general method that is able to predict future changes of evolving regions purely based on analysing past changes. The method is a hybrid method that integrates both a probabilistic approach based on the Kalman filter and a qualitative approach based on the topological structure of a region. It first identifies the trend of changes using qualitative analysis, and then predicts the changes with the Kalman filter. It does not need detailed a parameter setting as is required by existing domain specific prediction tools. We evaluated our method using both real world data from two different application domains and also by using generated data that allowed us to test our method under more difficult scenarios.

The main motivation of developing a general prediction method is the ability to use it in situations where predictions are required without having detailed knowledge about a domain or about the particular parameter setting of a specific situation at hand. This is particularly useful in disaster sce-

narios where time is the most critical factor and where fast and reasonably accurate predictions can save lives and prevent further destruction. As our experimental evaluation demonstrates, our general method is very good at providing accurate predictions at different application domains and thus can be used as the method of choice in time critical situations and in situations where detailed knowledge about a given scenario is not available. However, any available domain parameters can be integrated into our method as an additional layer.

Acknowledgements

References

[Alexandridis *et al.*, 2008] Alex Alexandridis, D Vakalis, Constantinos I Siettos, and George V Bafas. A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through Spetses Island in 1990. *Applied Mathematics and Computation*, 204(1):191–201, October 2008.

[Bhatt, 2014] Mehul Bhatt. Reasoning about space, actions, and change. In *Robotics: Concepts, Methodologies, Tools, and Applications*, pages 315–349. IGI Global, Hershey, 2014.

[Blake and Isard, 1998] Andrew Blake and Michael Isard. *Active Contours*. Springer London, London, 1998.

[Duckham, 2013] Matt Duckham. *Decentralized Spatial Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[Edelsbrunner *et al.*, 1983] Herbert Edelsbrunner, David G Kirkpatrick, and Raimund Seidel. On the shape of a set of points in the plane. *Information Theory, IEEE Transactions on*, 29(4):551–559, 1983.

[Feynman *et al.*, 1965] Richard Phillips Feynman, Robert B. Leighton, and Matthew Linzee Sands. *The Feynman Lectures on Physics: Mainly mechanics, radiation, and heat*. Addison-Wesley, 1965.

[Galton, 2000] Antony Galton. *Qualitative Spatial Change*. Oxford University Press on Demand, December 2000.

[Galton, 2009] Antony Galton. Spatial and temporal knowledge representation. *Earth Science Informatics*, 2(3):169–187, May 2009.

[Ghisu *et al.*, 2015] Tiziano Ghisu, Bachisio Arca, Grazia Pellizzaro, and Pierpaolo Duce. An optimal cellular automata algorithm for simulating wildfire spread. *Environmental Modelling & Software*, 71:1–14, September 2015.

[Jiang *et al.*, 2009] Jixiang Jiang, Michael Worboys, and Silvia Nittel. Qualitative change detection using sensor networks based on connectivity information. *GeoInformatica*, 15(2):305–328, October 2009.

[Junghans and Gertz, 2010] Conny Junghans and Michael Gertz. Modeling and prediction of moving region trajectories. In *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming*, IWGS '10, pages 23–30, New York, NY, USA, 2010. ACM.

[Kalman, 1960] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, March 1960.

[Kuhn, 1955] Harold W Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, March 1955.

[Liu and Schneider, 2011] Hechen Liu and Markus Schneider. Tracking continuous topological changes of complex moving regions. In *Proceedings of the 2011 ACM Symposium on Applied Computing*, SAC '11, pages 833–838, New York, NY, USA, 2011. ACM.

[Peng *et al.*, 2014] Zhenzhou Peng, Shinjae Yoo, Dantong Yu, Dong Huang, Paul Kalb, and John Heiser. 3d cloud detection and tracking for solar forecast using multiple sky imagers. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing*, pages 512–517. ACM, 2014.

[Rochoux *et al.*, 2014] Mélanie C. Rochoux, Sophie Ricci, Didier Lucor, Bénédicte Cuenot, and Arnaud Trouvé. Towards predictive data-driven simulations of wildfire spread – Part I: Reduced-cost Ensemble Kalman Filter based on a Polynomial Chaos surrogate model for parameter estimation. *Nat. Hazards Earth Syst. Sci.*, 14(11):2951–2973, November 2014.

[Russell and Norvig, 2009] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, 3. edition, December 2009.

[Santé *et al.*, 2010] Inés Santé, Andrés M. García, David Miranda, and Rafael Crecente. Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*, 96(2):108–122, May 2010.

[Singh *et al.*, 2015] Sudhir Kumar Singh, Sk Mustak, Prashant K. Srivastava, Szilárd Szabó, and Tanvir Islam. Predicting spatial and decadal LULC changes through cellular automata markov chain models using earth observation datasets and geo-information. *Environmental Processes*, 2(1):61–78, February 2015.

[Wood-Bradley *et al.*, 2012] Philip Wood-Bradley, José Zapata, and John Pye. Cloud tracking with optical flow for short-term solar forecasting . *Australia and New Zealand Solar Energy Society Conference (Solar 2012)*, 2012.

[Worboys and Duckham, 2006] Mike Worboys and Matt Duckham. Monitoring qualitative spatiotemporal change for geosensor networks. *International Journal of Geographical Information Science*, 20(10):1087–1108, November 2006.