Noname manuscript No.
(will be inserted by the editor)

Qualitative change detection using sensor networks based on connectivity information

Jixiang Jiang · Mike Worboys · Silvia Nittel

Received: date / Accepted: date

Abstract The research reported in this paper uses wireless sensor networks to provide salient information about spatially distributed dynamic fields, such as regional variations in temperature or concentration of a toxic gas. The focus is on deriving qualitative descriptions of salient changes to areas of high-activity that occur during the temporal evolution of the field. The changes reported include region merging or splitting, and hole formation or elimination. Such changes are formally characterized, and a distributed qualitative change reporting (QCR) approach is developed that detects the qualitative changes simply based on the connectivity between the sensor nodes without location information. The efficiency of the QCR approach is investigated using simulation experiments. The results show that the communication cost of the QCR approach in monitoring large-scale phenomena is an order of magnitude lower than that using the standard boundary-based data collection approach, where each node is assumed to have its location information.

Keywords Sensor network · topology · qualitative changes · spatio-temporal data

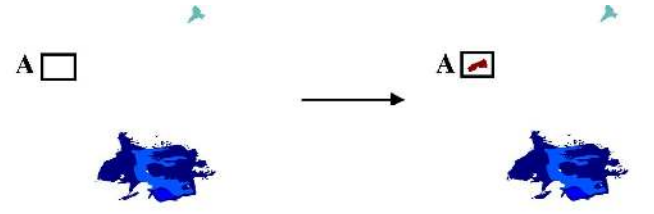
1 Introduction

The research reported in this paper uses wireless sensor networks to provide qualitative information about dynamic happenings in large-scale geographical phenomena, especially the qualitative changes. There are many examples where wireless sensor networks can be utilized for monitoring environmental change, modeled as fields. Examples include flooding, pollution plumes, temperature, and ocean salinity. For the purpose of

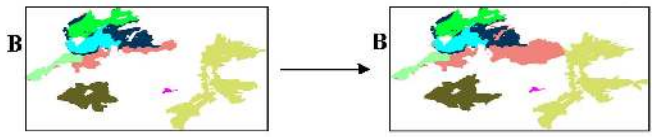
J. Jiang
NCGIA, University of Maine, ME, 04468, USA
E-mail: jixiang.jiang@maine.edu

M. Worboys
NCGIA, University of Maine, ME, 04468, USA
E-mail: worboys@spatial.maine.edu

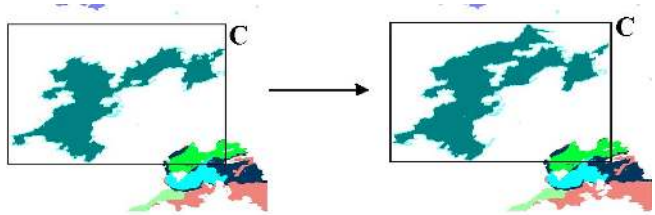
S. Nittel
NCGIA, University of Maine, ME, 04468, USA
E-mail: nittel@spatial.maine.edu



(a) Appearance



(b) Merge



(c) Self-Merge

Fig. 1 Examples of qualitative changes during the spreading of a Yellowstone Park wild fire

an example, we consider a video showing the progression of the 1988 Yellowstone Park wild fire [25]. In the first pair of snapshots shown in Fig. 1(a), a qualitative change of type '(fire zone) region appearance' can be observed within the area enclosed by the bounding box A. In the second pair shown in Fig. 1(b), a qualitative change of type 'regions merge' can be observed within the area enclosed by the bounding box B, in which two fire zone regions become connected. In the third pair shown in Fig. 1(c), a qualitative change of type 'region self-merge' can be observed within the area enclosed by the bounding box G, in which a fire zone region merges with itself and forms a hole. These are examples of different types of qualitative changes.

Geographical information of dynamic phenomena is important, and most sensing applications focused on capturing, processing and reporting the geographical information in the form of spatial-temporal data. Besides that, qualitative changes of dynamic phenomena often describe significant moments, and in many applications it is useful to have information about these kinds of qualitative changes. For example, in the case of wildfire, fire fighters might be interested if the fire zone regions split and become

disconnected, so that they can reorganize the team accordingly. They might also be interested in merging fires, as it sometimes slows down the burn when the fires are burning over each other. In the case of a flood, the emergency services may be interested in the appearance of an island in the flood, because this indicates the locations of possible safe areas. Instead of emerging from the flood directly, it is also possible that an island is formed by the flood engulfing a piece of land. In this scenario, people on the island become separated and may have difficulty escaping. Therefore, rescue from such a newly formed island might have higher priority in the overall hazard management strategy. This work focuses on detecting and reporting such salient qualitative changes using sensor networks.

By focusing on the qualitative information, our approach requires less energy in communication due to the small and discrete domain of qualitative properties [6]. In addition, the communication cost can be further reduced as the qualitative descriptions we focus on can be generated without location information. Existing research in environmental data collection applications often assumes that location information is available at each sensor node, either at the node itself [18, 12] or at the base station [26]. With the location information available, it is possible to generate the boundary shape or other quantitative properties, such as area, of the observed phenomena. Current technology allows sensor nodes to obtain their location information either from GPS directly or using localization algorithms. However, GPS is expensive, and localization approaches are energy-consuming especially for large sensor networks [2]. Therefore, in these applications based on location information, additional cost is needed. Another advantage of our qualitative approach is that the location information is not necessary. In this work, we do not require the sensor nodes to be aware of their geographic locations. The qualitative descriptions of salient changes are generated simply based on network connectivity, and therefore the cost in localization is reduced.

To enable the qualitative reporting approach, in this paper we first provide a formal classification of changes based on the readings and connectivity of the sensor nodes. Each class of changes is qualitatively described using a phrase in natural language. After that, a distributed algorithm is proposed for sensor networks to detect these changes and report them using the specified phrases in real-time monitoring. The major contributions of the paper include the following:

1. Identification of key features that allow us to distinguish different types of changes and to generate qualitative descriptions for the observations.
2. We show that at each particular sampling round, these key features can be completely captured based on the readings and connectivity information of a small portion of sensor nodes.
3. An energy-efficient approach is proposed for the detection of different types of changes in sensor networks. Our experiments show that this approach is able to generate qualitative descriptions for changes by sensor networks without location information, and the communication cost in monitoring large-scale phenomena is much lower than the standard boundary-based data collection approaches with location information.

2 Background

Applications in geographic information systems often require identification and manipulation of qualitative representations [3, 11]. Topology provides an important way to

abstract and generate qualitative descriptions for spatial data. Active research in this direction is the specification of static topological relations between spatial regions [8], and the qualitative reasoning of these relations [4]. Topological features also allow us to classify spatial changes. Egenhofer and Al-Taha [7] analyze and classify the spatial changes involving two regions based on their topological relations before and after the change, and the result is recorded using the conceptual neighborhood graph. Wilmsen [28] analyzes and identities topological states of regions in snapshots and derives different types of changes, including continuous changes (such as growing, shrinking, and moving), as well as discrete changes (such as splitting and merging). In the previous work [15], we propose a model that represents the dynamic topology of an areal object (a collection of region components, possibly with holes and islands), based on which different types of topological changes are specified.

Wireless sensor networks provide real-time information about the environment, and thus have the potential to play an important role in the monitoring of geographic phenomena [20]. Previous research in environmental data collection either focuses on proposing energy-efficient approaches to transmitting entire sensed data back to base stations [26, 23, 24, 19], or focuses on providing important spatial properties of the phenomena. For example, the snake model proposed by Jin and Nittel [16] is able to derive the area and centroid of a deformable 2D object over time. Recently, there is an increasing interest in considering topological information when processing sensed data. Gandhi, Hershberger and Suri [12] emphasize the topology of the isolines in a scalar field and propose an approach that approximates a family of isolines by a collection of topology-preserving polygons. Sarkar *et al.* [22] present a distributed algorithm for the construction of a contour tree to represent the topological structure of contours in a scalar field, based on which isoline queries can be enabled. Worboys and Duckham [29] provide a computational model for sensor networks to detect global high-level topological changes based on low-level 'snapshot' of spatiotemporal data. Zhu *et al.* [30] propose a distributed algorithm for the sensor networks to maintain contours (or boundaries) of a binary object incrementally as they deform, while guaranteeing that the maintained contours capture the global topological features of the object boundary. However, none of the existing work focuses on detecting and reporting of topological changes.

Another research area of sensor networks that relates to this work is the retrieval of network topology itself, especially the connectivity among sensor nodes in their communication graphs [17]. Deb, Bhatnagar and Nath [5] in their STREAM algorithm enable the retrieval of the entire network topology with predefined resolution, which allows users to make a trade-off between topological details and resources expended. Existing work has also shown that connectivity information allows the detection of the topological features of either the sensor networks themselves or the underlying phenomena. Funke and Klein [10] propose an approach that detects holes of sensor networks based on connectivity information. Wang, Gao, and Mitchell [27] use the connectivity information to derive the boundary of sensor networks. Ghrist and Muhammad [13] implement algorithms for sensor networks to detect holes of objects by means of homology derived from connectivity information. Although the detection of some topological properties of areal objects can be achieved based on the analysis of the network connectivity among sensors that observe the areal object [9], the cost for network topology retrieval is very high. Therefore, more efficient detection methods are needed.

3 Preliminary

This section provides the basic definitions and assumptions we have made about the sensor networks in our work. We assume that a large number of sensors are deployed in the sensing area. Each sensor node is initialized with a unique identifier and records the values of measurements. A node located near the boundary of the sensing area is selected to be the *reference node*, which is assumed to be located in the exterior of any regions occupied by the phenomena. The reference node enables identification of a region exterior, which is important for distinguishing certain types of qualitative changes, such as merge and self-split. The sensor nodes in the sensing area induce a Voronoi diagram, and each node n is associated with a Voronoi cell consisting of all the locations that are closer to n than to any other node. We stipulate that the sensor node deployment satisfies the following constraints:

1. **Density constraint,** sensor nodes are deployed densely enough so that a sensor measurement accurately reflects all locations in its Voronoi cell.
2. **Communication constraint,** each node communicates with exactly the nodes in its adjacent Voronoi cells.

Both constraints ensure that sensor networks generate accurate reports of qualitative changes. In real applications, the constraints may be broken due to energy exhaustion or hardware failure. It should be pointed out that failure to confirm to both constraints does not disable the whole detection approach, but it reduces the reporting accuracy (for example, sensor network may report that splitting of a wild fire is observed, but in reality the fire does not split).

Fig. 2 shows an example deployment of nodes and their associated Voronoi cells. It also shows a possible location of the reference node r.

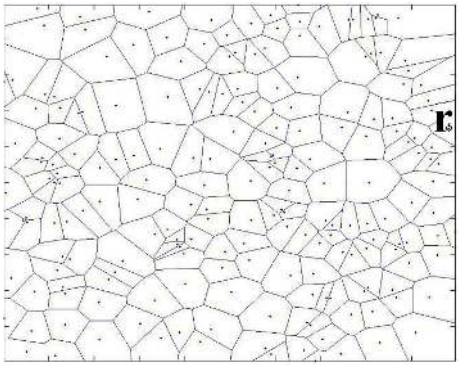


Fig. 2 Sensor network configuration

The sensors take measurements at a sequence of sampling rounds t_0, t_1, ..., t_k. We assume that the reading of a node at any of the sampling rounds is either 0 or 1. Our interpretation is that the reading is 1 if the node is in an area of high intensity (reading above a given threshold), otherwise it is 0. A change is captured by sensor readings at a pair of consecutive sampling rounds, and the qualitative descriptions of changes are generated by comparing the readings. The comparison first defines four states of nodes at sampling round t_i.

Definition 1 Let $r(n,t) \in \{0,1\}$ denote the reading of a node n at a time t. The *state* of n at a sampling round $t_i (1 \leq i \leq k)$ is defined to be a pair $h = (r(n,t_{i-1}), r(n,t_i))$, such that $h \in \{(0,1),(0,0),(1,0),(1,1)\}$

The state of a node varies with time. In the following discussion, everything is assumed to be in the same snapshot at sampling round $t_i (1 \leq i \leq k)$ unless time is explicitly specified. The states of the sensor nodes together with the node connectivity yield the following concepts that are foundations for generating qualitative descriptions.

Definition 2 Let N be a set of sensor nodes, N is said to be a *homogeneous component* if the nodes in N have the same state and induce a connected component in the communication graph. Moreover, N is defined to be a *maximal homogeneous component*, if it is impossible to find a node n in the sensing area such that (1) $n \notin N$, and (2) $N \cup \{n\}$ is a homogeneous component.

Definition 3 Let N_1 and N_2 ($N_1 \cap N_2 = \emptyset$) be a pair of homogeneous components.

1. N_1 is said to be *adjacent to* N_2 if there are nodes $n_1 \in N_1$ and $n_2 \in N_2$ such that n_1 and n_2 are direct neighbors in the communication graph. Otherwise, N_1 and N_2 are said to be *separated*.
2. N_1 is said to be *surrounded by* N_2, if any path in the communication graph that starts from the reference node and contains a node of N_1 must contain a node of N_2. N_1 is said to *surround* N_2, if N_2 is surrounded by N_1.

The generation of qualitative change descriptions does not require data from all of the nodes. Only the nodes located inside or near the locations where sensor readings change are necessary. These nodes form the transition components and the C-components defined as follows:

Definition 4 Let N be a maximal homogeneous component.

N is defined to be a *transition component*, if N consists of only nodes of state $(0,1)$ or $(1,0)$.

N is defined to be a *C-component*, if both of the following conditions are satisfied: (1) N consists of nodes of state $(1,1)$ or $(0,0)$, and (2) N is adjacent to a transition component.

The transition components consist of nodes that change their readings, whereas C-components consist of nodes that do not change. As an example, Fig. 3(A) and 3(B) show the readings of nodes at consecutive sampling rounds t_1 and t_2, respectively. Both snapshots describe a change in which a hole of a region engulfs a piece of its exterior that forms an island. The black points denote nodes with reading 1, and the white points denote nodes with reading 0. The only difference between the two snapshots is that the nodes located in the area enclosed by a polygon change their readings. These nodes form a maximal homogeneous component with state $(1,0)$, which is a transition component at t_2 by definition 4. The rest nodes are of states $(0,0)$ and $(1,1)$ at t_2, and they form other six maximal homogeneous components. Fig. 3(C) shows the geographical locations of the seven maximal homogeneous components. In the figure, region b contains the transition component. Regions a, d, and f contain the maximal homogeneous components that consist of nodes of state $(1,1)$. Regions c, e, and g contain the maximal homogeneous components that consist of nodes of state $(0,0)$. In addition, we use $\widetilde{a}$, $\widetilde{c}$, and $\widetilde{d}$ to represent the maximal homogeneous components

contained in regions a, c, and d, respectively, as shown in Fig. 3(D), 3(E), and 3(F). $\widetilde{a}$, $\widetilde{c}$, and $\widetilde{d}$ are adjacent to the transition component contained in region b. Therefore, they are C-components at t_2 by definition 4.

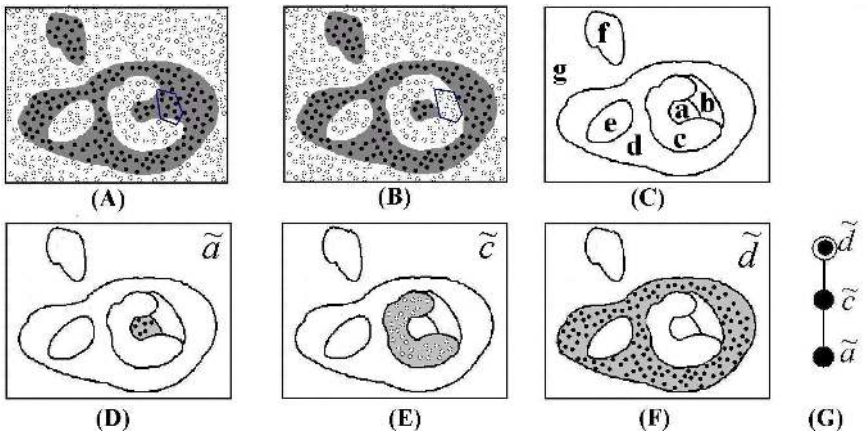


Fig. 3 An example of homogeneous components

In this work, we assume that at each sensing round, only one transition component is observed, and the region represented by the transition component is a single piece without any holes. The assumption of simple transition components simplifies our discussion. More complex situations, in which more than one transition components exist and each region represented by a transition component is allowed to have holes, will be recorded as our future work. It should be pointed out that any complex transition components can be decomposed into several simple transition components. By forcing the simple transition components to switch one after another, we could generate a sequence of descriptions for a complex change. Hence, the study of the simple changes forms the foundation for the future work.

4 Classification of changes

In this section, we classify the changes based on their topological features defined in section 3, and phrases in natural language are used to describe each type of changes in the classification result.

Similar to the discussion of binary images [21], the C-components within a snapshot have the following properties:

1. There exists exactly one C-component X which surrounds all the other C-components. X is referred to as the *background C-component* of the change.
2. The topological structure of the C-components can be represented by a rooted tree. A vertex of the tree represents a C-component, and an edge of the tree connects a pair of vertices representing adjacent C-components. The root of the tree represents the background C-component.

Returning to the example transition shown in Fig. 3, the C-component $\widetilde{d}$ is the background C-component. In addition, C-components $\widetilde{a}$ and $\widetilde{c}$ are adjacent, as well as

the C-components $\widetilde{c}$ and $\widetilde{d}$. The structure of the C-components can be represented by the rooted tree in Fig. 3(G), in which the root is indicated by a double-circled vertex.

The following three features allow us to differentiate and specify various types of changes. The changes of the same kind must have identical features, and different changes differ in at least one of the features.

1. The topological structure of its C-components
2. The state of nodes in the transition component
3. The state of nodes in the background C-component

As different rooted tree structures can be explored in a systematic way, we are able to generate the possible topological structures between the C-components. Fig. 4 lists all the rooted trees with less than 4 vertices, and example configurations of C-components represented by the rooted trees are also provided. In the figure, the geographical region that contains the transition component is indicated by a shaded area, and any vertex of the tree is placed inside the region in which the C-component represented by the vertex is located.

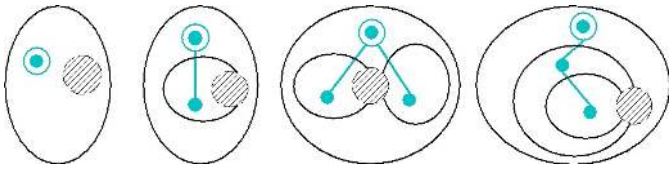


Fig. 4 Tree representations for different configurations of C-components

The classification yields different types of changes. Fig. 5 shows the classification results, in which 16 types of changes are distinguished. Based on the way in which the phenomena represented by the nodes evolve, each type of changes is described by a phrase, which is also shown in Fig. 5. These qualitative descriptions will be included in the node report. Fig. 6 gives example evolutions of regions that are described by different types of changes.

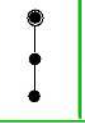
Local topological structure	State of transition component (T)		State of background C-component (B)	
	T : (0,1) B: (0,0)	T : (1,0) B : (1,1)	T : (0,1) B : (1,1)	T : (1,0) B : (0,0)
	Blob appearance (A)	Blob disappearance (B)	Hole disappearance (I)	Hole appearance (J)
	Blob enlargement (C)	Blob reduction (D)	Hole reduction (K)	Hole enlargement (L)
	Blobs merge (E)	Blob split (F)	Hole split (M)	Holes merge (N)
	Blob self-merge (G)	Blob self-split (H)	Hole self-split (O)	Hole self-merge (P)

Fig. 5 Classification of changes

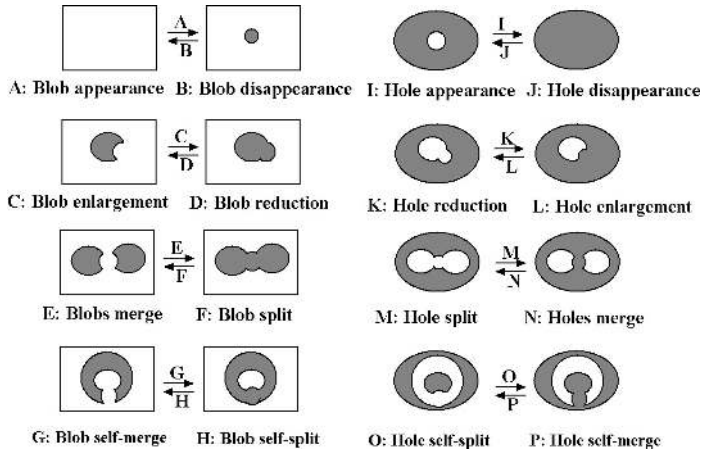


Fig. 6 Examples of specific types of qualitative changes

It should be pointed out that in addition to the specific types of changes shown in Fig. 5, there are other changes whose structures are represented by a tree with four or even more vertices. Fig. 7 gives an example of such changes and its representation tree. Such changes may not have a commonly accepted description associated with them, but they can also be detected and included in the sensing report if users are interested in them.

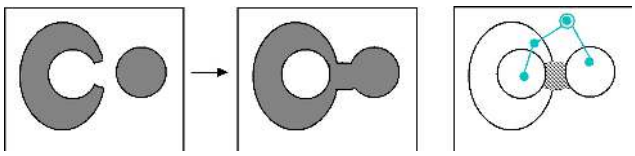


Fig. 7 An example of an uncommon type of change

5 Boundary group based representation for qualitative change detection

As discussed in section 4, in order to form the qualitative descriptions of changes, we need to identify both the transition component and the C-components together with their states, and more important, to determine the topological structure between the C-components.

Using a naive approach, we require the entire collection of nodes located in sensing area to report their readings and geographical locations back to the base station after each sensing round. With all the received data, the base station is able to determine the types of qualitative changes using centralized computation. However, it is energy-consuming to gather all of the node information to the base station.

To improve on the naive approach, only the nodes located near the boundary of the areal object are required to report after each sensing round. A boundary node is defined as a node that has a direct neighbor with a different reading. As an example, Fig. 8 (A)

and 8(B) show the boundary nodes that are required to report during the transition described in Fig. 3. Because the boundary nodes carry the necessary information for change detection, the base station is able to detect the qualitative changes based on the data received from the boundary nodes. The polygon in Fig. 8(C) encloses the boundary nodes that are identified to be part of the transition component. Fig. 8(D), 8(E), and 8(F) show the boundary nodes that represent the C-components $\widetilde{a}$, $\widetilde{c}$, and $\widetilde{d}$, respectively. By analyzing the properties of the boundary nodes, the structure of the representation tree can also be identified, as shown in Fig. 8(G).

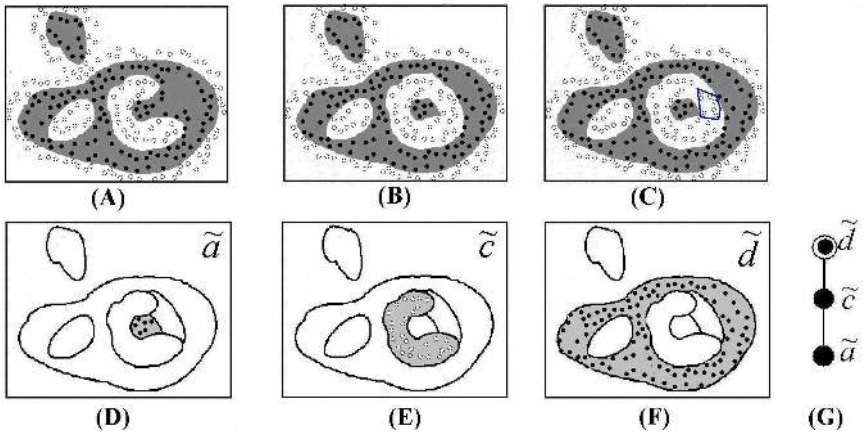


Fig. 8 Boundary nodes

The boundary-based approach reduces the number of reporting nodes. However, as we have assumed a dense deployment of sensor nodes, the number of boundary nodes may still be large. In addition, both the naive approach and the boundary-based approach assume the availability of node geographical location data. To further reduce the communication cost and to avoid the use of geographical location data, we propose the *boundary group based* approach.

In the boundary group based approach, some nodes in the sensing area form groups, and each group is assigned with a unique integer label. The structure of groups is dynamically updated after each sensing round, so that the nodes in the same group are always ensured to form a homogeneous component; that is, they all have the same state. After each sensing round, only groups located near the boundary of the areal object are required to report. A node in each of the boundary group is selected to be the group leader, which is responsible for sending data back to the base station.

The data sent from the leader of a group G represent information about the group as a whole, instead of information about a single node. The group level data include the integer label of G, the state of the nodes in G, and the neighboring label set of G, which is defined to be the set of labels of groups that are adjacent to G and whose labels are greater than G. As an example, Fig. 9(A) and 9(B) describe a basic transition. In Fig. 9(B), seventeen boundary groups labeled from 0 to 16 exist. Taking group 6 as an example, the data sent from this group to the base station include group label 6, the state of the group $(1, 1)$, and the neighboring label set $\{7, 13\}$. The data received by the base station from all the groups can be represented by the graph shown in Fig.

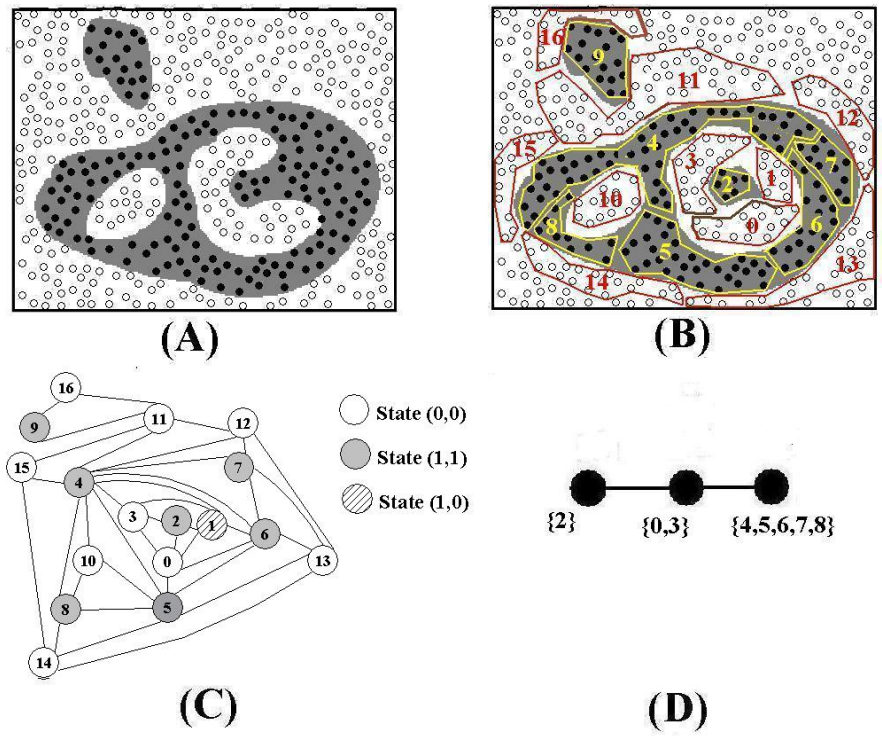


Fig. 9 Possible formation of groups

9(C). In this graph, a vertex represents a group of that label, and the color of a vertex represents the state of the group. Edges of the graph represent the adjacency relations between the groups, which are derived from the neighboring label sets of the groups.

The base station is able to identify the C-components simply based on the group level data. In this example, as the state of group 1 is $(1, 0)$, it must be contained in the transition component. By definition 4, all the other groups that are adjacent to group 1 (including the groups 0, 2, 3, 4, and 6) are contained in C-components. Consider groups 4 and 6: they are adjacent and are of the same state $(1, 1)$, so they are in the C-component. In addition, group 6 is adjacent to group 7, and both have the same state. By definition 4, group 7 must be contained in the same C-component as group 6. Similarly, groups 5 and 8 must be contained in the same C-component as group 6. But the other groups are not contained in the same C-component as groups 4 and 6. So there must be a C-component of state $(1, 1)$ that contains groups 4, 5, 6, 7 and 8. Next, consider group 2: as no other group of state $(1, 1)$ is adjacent to group 2, another C-component of state $(1, 1)$ is identified that contains group 2. Finally, consider groups 0 and 3: as both of them are adjacent and are of the same state $(0, 0)$, they are contained in the same C-component. In addition, as no other group of state $(0, 0)$ is adjacent to them, we identify the third C-component of state $(0, 0)$, which contains the groups 0 and 3.

The adjacency relations between the C-components can be identified based on the adjacency relations between the groups. A pair of C-components are adjacent if one

contains a group that is adjacent to a group in the other C-component. In this example, as group 3 is adjacent to group 4, the C-component containing groups 0 and 3 must be adjacent to the C-component containing groups 4, 5, 6, 7 and 8. Similarly, as group 2 is adjacent to group 3, the C-component containing group 2 must be adjacent to the C-component containing groups 0 and 3. Fig. 9(D) shows the C-components and adjacency relations identified based on group level data.

Finally, to identify the background C-component, we require each group to send its group hop distance to the base station, in addition to the other data. The *group hop distance* is the number of hops of the shortest path between any node in the group and the reference node. As the background C-component surrounds any of the other C-components and the reference node is located outside all the C-components, it follows that any path that connects a node in a C-component to the reference node must contain a node in the background C-component. Therefore, among all the groups identified to be contained in a C-component, the one with the minimal group hop distance must belong to the background C-component. In this example, if we place the reference node at the right-bottom of the sensing area, group 6 has the minimal group hop distance, and the C-component containing group 6 is identified to be the background C-component.

Generally, with the group level data, the following procedure can be performed in the base station to determine the type of a change.

First, the base station identifies sets of groups, each set having the following properties:

1. At least one of the groups in the set is adjacent to a group of state $(0, 1)$ or $(1, 0)$.
2. Groups in the same set have the same state.
3. Groups in the same set are connected with respect to the adjacency relation.
4. Each set is maximal with respect to properties 2 and 3.

Each set satisfying these properties is formed by groups that are located in the same C-component, and can be used to represent that C-component. By generating such sets of groups, all the C-components can be found. A pair of C-components are adjacent if a group in one of the C-components is adjacent to a group in the other C-component.

Finally, the background C-component is the one that contains the group with the minimal group hop distance.

6 Algorithms

The analysis in section 5 shows that a qualitative change description can be derived at any sampling round, as long as the boundary groups are maintained, and necessary group level data are available at the base station. This section proposes the qualitative change reporting (QCR) approach based on the foundation laid in section 5. Algorithm 1 sets out the sketch of the proposed approach, and the following subsections present each step in detail.

Algorithm 1 Qualitative change reporting

1 Initialization
1.1 Each node in the sensor network computes its hop distance to the reference node.
1.2 The base station sends out a query to the entire nodes in the whole sensing area.

2 Boundary group initialization The following steps are performed immediately after the first sensing round.
2.1 **Group formation:** After the first sensing round, groups are formed among the boundary nodes.
2.2 **Group information aggregation and reporting:** The group information is aggregated and sent back to the base station.

3 Monitoring
Sensor nodes take measurements in specified sensing rounds, and the following steps are performed in each sensing round in a distributed manner:
3.1 **Group update:** Existing groups are modified to ensure that (1) every node located near the boundary of C-components is included in a group, and (2) each group is a homogeneous component.
3.2 **Update aggregation:** The modified groups perform an update, and update messages are sent back to the base station.
3.3 **Data reporting and analysis:** Data are sent back to the base station from group leaders. The base station analyzes data it receives to determine the type of changes.

6.1 Initialization

During initialization, each sensor node computes its hop distance to the reference node, and a query request is propagated from the base station to the nodes in the sensing area.

The hop distance of a node n is the minimal number of hops that connect n to the reference node. Similar to the approach described in [10], hop distances of nodes can be computed by flooding originating from the reference node. The reference node broadcasts a $HELLO$ message maintaining a distance counter that is incremented at each hop. The minimal counter value over all messages received by a node n is the hop distance of n.

Query propagation is done by flooding originating from the base station Q, which broadcasts a query $TPQ(t_s, t_f, t_\Delta)$ to all the nodes located in the sensing area. In the query, t_s and t_f state the time to start and to finish the monitoring, and t_Δ specifies the time period between a pair of consecutive sensing rounds.

6.2 Group formation

After the first sensing round, the initial boundary groups are formed. Each node first communicates with its direct neighbors to identify the boundary nodes. A node is a boundary node if it has a direct neighbor with a different reading. Groups are formed among those boundary nodes.

During group formation, a boundary node n first waits for a random amount of time t_w in the range $[1, T_w]$, in which T_w is the maximal waiting time. After that if no existing group is found for n to join, n becomes a group leader. Each group leader propagates a group call message toward the other boundary nodes of the same reading. A boundary node joins with the same group as the sender of the group call message it first hears. During the group formation, all the boundary nodes in the same group are assigned the same label, which is the unique identifier of its group leader. Also, a routing tree is set up within each group that is rooted by the group leader and connects all the boundary nodes in the group.

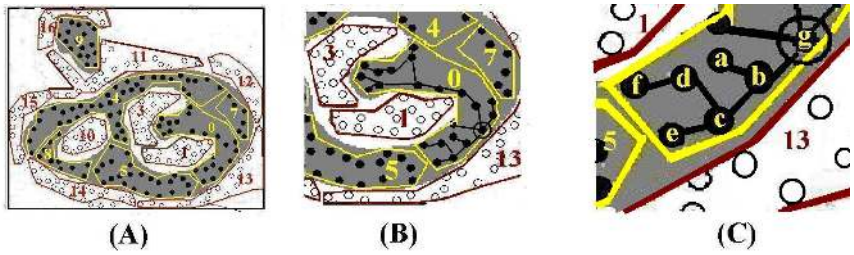


Fig. 10 An example of group formation

As an example, Fig. 10(A) shows the boundary groups formed among the boundary nodes detected immediately after the first sensing round of the basic transition in Fig. 3. Fig. 10(B) shows the details of the boundary group 0, in which the group leader is indicated by a double circle, and the routing tree built among the nodes in the group is indicated by the edges between the nodes.

6.3 Group information aggregation and reporting

After the boundary groups are formed, data necessary for qualitative change detection are aggregated along the routing trees. The aggregation result of a group includes its neighboring label set, as well as its group hop distance.

By communicating with its direct neighbors, a node n in a group observes the labels of its direct neighbors. The observed labels that are greater than the label of n form a *local label set* of n, denoted by $L(n)$. In addition, we define the *neighboring label set* of n, denoted by $N(n)$, to be the union of all the local label sets of nodes contained in the subtree rooted at n; that is, $N(n) = L(n) \cup (\bigcup_{m \in D(n)} L(m))$, in which $D(n)$ is the set of descents of n in the routing tree .

By definition, the neighboring label set of the group G is $N(r)$, where r is the root of the routing tree in G. The neighboring label sets can be computed by aggregation. During the aggregation, each node n computes $N(n)$ based on the data received from its direct children, and sends the result to its parent. If node n is a leaf node in the tree, $N(n) = L(n)$. Otherwise, let c_1, c_2,..., c_k be the direct children of n in the routing tree, $N(n) = N(c_1) \cup N(c_2) \cup N(c_k) \cup L(n)$.

As an example, Fig. 10(C) shows the details of the bottom part of the routing tree in group 0. The nodes of the part are named a to g, and the node g is the root. Each node knows its local label set, in which $L(a) = L(d) = \{1\}$, $L(b) = L(c) =$

$\{13\}$, $L(e) = \{5, 13\}$, and $L(f) = \{1, 5\}$. After the aggregation, each node knows its neighboring label set, in which $N(a) = \{1\}$, $N(f) = N(d) = \{1, 5\}$, $N(e) = \{5, 13\}$, $N(b) = N(c) = \{1, 5, 13\}$.

Similarly, with the routing tree, the group hop distance, which is the minimal hop distance of the nodes in the group, can be found at the group leader by a standard aggregation method.

After the aggregation, the group leader is able to send the necessary group information back to the base station, including the reading of the group, its integer label, its neighboring label set, and its group hop distance.

6.4 Group update

In the QCR approach, a group endures after it is formed. However, the readings of nodes in a sensing round t_i can be different from that in t_{i-1}. Therefore, the nodes in the same group formed in t_{i-1} may have different states in t_i. In order to ensure that the nodes in the same group form a homogeneous component, some groups are modified at t_i. The modifications include group partial deletions and creations.

The nodes in a group may change its reading after a sensing round. Suppose the node n in group G changes its reading. Then group G is no longer a homogeneous component, and needs to be updated. A partial deletion is performed, in which the nodes in the subtree originating at n is deleted from the group G. To perform the partial deletion, after the sensing round t_i, any labeled node of state $(0, 1)$ or $(1, 0)$ is set to be unlabeled, leaves the group, and propagates a $DESTROY$ message to its decedents in the routing tree. A labeled node that receives a $DESTROY$ message is set to be unlabeled and leaves the group.

The partial deletion of existing groups, as well as the change of the C-components, results in unlabeled boundary nodes that are not included in any groups. To report the properties of these nodes, groups are formed among the unlabeled boundary nodes. The formation procedure is the same as that described in section 6.2.

For an illustration, consider the basic transition described by Fig. 11(A) and 11(B). During the transition some nodes in group 0 change their readings from 1 to 0, which incur the partial deletions. These nodes together with their descendants leave group 0 and become unlabeled boundary nodes, as shown in Fig. 11(C). In addition, the transition also results in some new boundary nodes depicted as squares in Fig. 11(C). These boundary nodes perform a group formation procedure, and a possible result is shown in Fig. 11(D), in which new groups 2 and 6 are created, and group 0 is expanded to include two boundary nodes.

6.5 Update aggregation and reporting

After the group update, information of some groups needs to be computed. The same procedure as described in section 6.3 is performed in the new groups. In addition, the aggregation is also performed in the groups that are created in previous sensing rounds. The aggregation in these groups only takes place where nodes have different data. As an example, the nodes m, n, o, p, and q shown in Fig. 12 are the only nodes in group 0, whose neighboring label sets and group hop distances may need to be updated. As

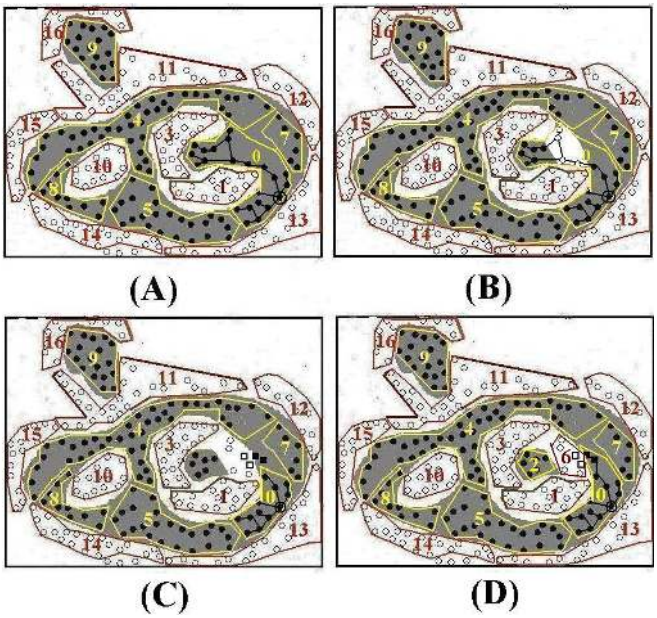


Fig. 11 An example of group update

the data in the rest of the nodes in group 0 do not change, an aggregation among them is unnecessary.

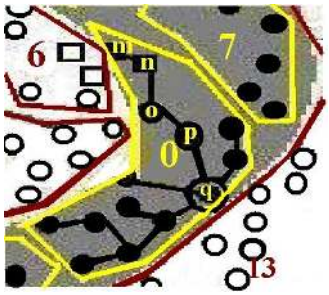


Fig. 12 An example of group update aggregation

After the update, the messages are sent from the leaders of some groups to the base station Q, which include the following.

A ***creation message*** is sent from the leader of each new group G. This message includes the label of G, the state of G, neighboring label set of G, and the group hop distance of G.

An ***update message*** is sent back from the leader of a group G whose group data changed. The update message includes the labels that are added to, or removed from, the neighboring label set of G, and the group hop distance of G, if it changes.

Table 1 Sensing reports description

Time	t_1	t_2	t_3	t_4
Size ratio	7.77%	8.45%	9.71%	11.05%
Type of change		Blob Enlarge	Blob Appear	Blob Merge
Time	t_5	t_6	t_7	t_8
Size ratio	12.85%	13.94%	14.4%	14.88%
Type of change	Blob Enlarge	Blob Merge	Blob Enlarge	Blob Enlarge
Time	t_9	t_{10}	t_{11}	t_{12}
Size ratio	13.79%	15.06%	14.03%	14.72%
Type of change	Blob Split	Blob Enlarge	Blob Split	Blob Merge
Time	t_{13}	t_{14}	t_{15}	t_{16}
Size ratio	15.49%	16.51%	17.51%	18.03%
Type of change	Blob Enlarge	Blob Enlarge	Blob self-merge	Blob Enlarge
Time	t_{17}	t_{18}	t_{19}	t_{20}
Size ratio	19.12%	19.9%	20.81%	21.8%
Type of change	Hole disappear	Blob appear	Blob Merge	Blob Enlarge

7 Experimental evaluation

An experiment was conducted by simulation in order to test the performance of the proposed approach. We used Prowler [1], a MATLAB based network simulator, as our simulation environment.

7.1 Experiment setup

In the experiment, the size of the sensing area was set to be 420 units×600 units, and nodes were deployed in a regular grid with the distance between adjacent nodes being 10 units. Therefore, 42×60=2460 nodes in total were deployed. The sensing range of a node was defined to be 15 units, so that each node was able to communicate with its 8 direct neighbors. The base station was placed in the middle of the sensing area, which collected the data and determined the type of the qualitative change in each sensing round. The sensor nodes woke up and performed the qualitative change detection in sensing rounds t_1, t_2,..., t_{20}, in which the difference between t_{i+1} and t_i was 60 seconds. The maximal waiting time T_w was set to be 10 seconds.

A sequence of 20 snapshots were generated to provide the sensing data from t_1 to t_{20}. These snapshots described a scenario, in which an areal object grew in the form of basic transitions, from a small areal object to a large one. Table 1 describes the information of the snapshots, including the ratio of the size of the areal object to the size of whole sensing area, as well as the type of qualitative changes that occurred in each sensing round. Fig. 13 shows three example snapshots at times t_1, t_{10}, and t_{20}, respectively.

We ran the qualitative change detection approach on the 20 snapshots. For comparison, the qualitative changes were also detected by the method of periodically boundary reporting (BR), as described in section 5.

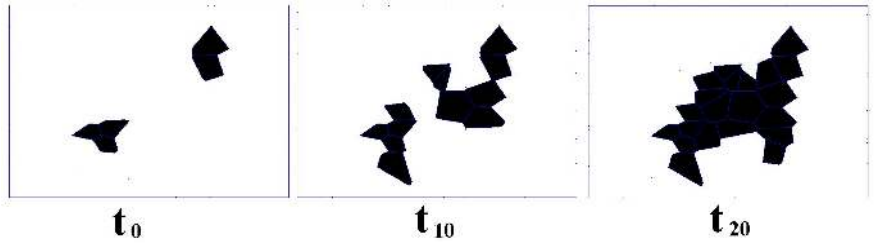


Fig. 13 Examples of experiment snapshots

7.2 Results and discussion

7.2.1 Communication cost

In the experiments, all the reported types of qualitative changes by both the QCR and the BR approach were the same as expected. Fig. 14 shows the communication cost of both approaches. The QCR approach requires higher communication cost in the first sensing round for the group formation and reporting. After that, the communication cost is reduced, as the aggregation takes place only at the locality of the change, and only the updated information is transfered to the base station. However, the BR approach requires higher communication cost than the QCR method in every sensing round, as every boundary node is required to report. The total communication cost of the QCR method is as low as 16.2% of the BR approach. In addition, as the QCR approach does not assume the availability of node locations, which also reduces the cost in handling the location information.

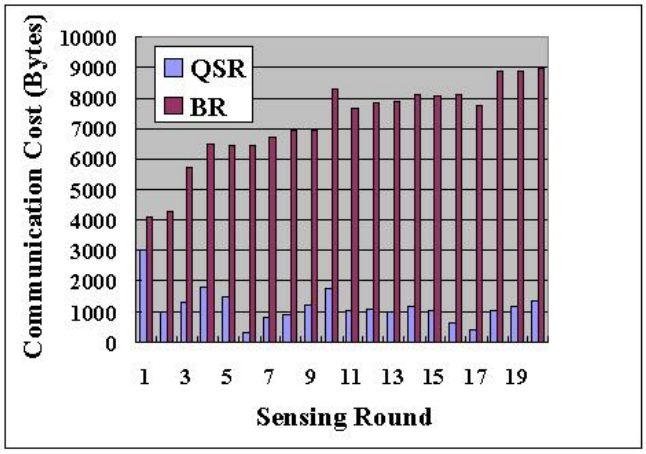


Fig. 14 Communication cost of two approaches

7.2.2 Effect of group size

The maximal waiting time T_w affects the group size in the detection. If T_w is higher, a smaller number of groups is formed, and each group contains more boundary nodes.

The total communication cost of reporting the 19 changes was affected by the average group size. In order to study this effect, we set up different T_w within the range $[5s, 30s]$ in different runs. The average size of groups formed in each run varied, so did the communication cost. The three curves in Fig. 15 show the local communication cost, reporting communication cost, and the total communication cost versus the average group size. Local communication cost refers to the number of bytes transmitted locally during the group update and data aggregation. Reporting communication cost refers to the number of bytes transmitted for the reporting. The total communication cost is the sum of both.

As shown in Fig. 15, the local communication cost increases almost linearly with the increase of average group size, whereas the reporting communication cost decreases as the average group size increases. When the average group size increases from 0 to 25, an intensive decrease of the reporting communication cost is observed, and thereafter it remains almost stable. As the combination of both the local and the reporting communication costs, the overall communication cost reaches a minimal in the range between 20 and 30 of the average group size.

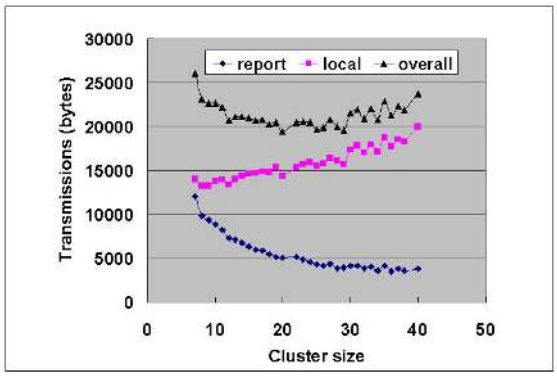


Fig. 15 Communication cost versus group size

8 Conclusions and future work

In this paper, we have performed a detailed analysis of qualitative changes that can be reported by sensor networks simply based on the network connectivity information. A complete classification of changes is provided, and each kind of change is assigned a qualitative description. An energy-efficient QCR approach is proposed to collect the important features for reporting types of observed change during real-time monitoring. The experiments show that in large-scale monitoring applications, the QCR approach is able to generate qualitative descriptions of changes simply based on the connectivity information at a low communication cost.

The work presented here assumes the transition component in each snapshot is topologically equivalent to a disk. In future work, we are going to extend the current work so that it is possible to monitor more complex changes, in which several components exist in a single snapshot and each component is allowed to have holes. One of the possible avenues for future research is to decompose the complex transition components into several simple transition components, and the changes incurred by each simple transition component is processed one after another.

Both this and our previous work [15,14] have shown that it is possible to provide qualitative descriptions of spatial-temporal data based on changing topology. This provides a foundation for further study in spatial and temporal queries in terms of dynamic topological properties. Examples of such queries include, (1) Retrieve the data describing the *appearance* of a wild fire in the forest. (2) Retrieve the data describing the typhoons that occurred last year and *split* during their evolution. Therefore, another future research direction is to incorporate the primitive qualitative changes into query languages, and finally support query manipulation in terms of the dynamic topological properties of phenomena.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant numbers IIS-0429644 and IIS-0534429. Mike Worboys' work is also supported by the National Science Foundation under NSF grant numbers DGE-0504494 and BCS-0327615.

References

1. Prowler-probabilistic wireless network simulator. `http://www.isis.vanderbilt.edu/Projects/nest/prowler/`. Last accessed June 4, 2008.
2. J. Bachrach and C. Taylor. Localization in sensor networks. In *Handbook of Sensor Networks: Algorithms and Architectures*, pages 277–310. Wiley-Interscience, 2005.
3. C. Bailey-Kellogg and F. Zhao. Qualitative spatial reasoning: Extracting and reasoning with spatial aggregates. *AI Magazine*, 24(4):47–60, 2004.
4. A. Cohn, B. Bennett, J. Gooday, and N. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1(2):275–316, 1997.
5. B. Deb, S. Bhatnagar, and B. Nath. STREAM: Sensor topology retrieval at multiple resolutions. *Telecommunication Systems*, 26(2-4):285–320, 2004.
6. M. Duckham, S. Nittel, and M. Worboys. Monitoring dynamic spatial fields using responsive geosensor networks. In *ACM-GIS 2005*, pages 51–60, Bremen, Germany, 2005.
7. M. Egenhofer and K. Al-Taha. Reasoning about gradual changes of topological relationships. In *Proceedings of the International Conference GIS - From Space to Territory: Theories and Methods of Spatio Temporal Reasoning, LNCS 639*, pages 196–219, 1992.
8. M. Egenhofer and J. Herring. Categorizing binary topological relations between regions, lines, and points in geographic databases. Technical report, Department of Surveying Engineering, The University of Maine, 1990.
9. Q. Fang, J. Gao, and L. Guibas. Locating and bypassing routing holes in sensor networks. *Mobile Networks and Applications*, 11(2):187–200, 2006.
10. S. Funke and C. Klein. Hole detection or: 'how much geometry hides in connectivity?'. In *Proceedings of the Twenty-second Annual Symposium on Computational Geometry*, pages 377 – 385, 2006.
11. A. Galton. *Qualitative spatial change*. Oxford University Press, 2000.
12. S. Gandhi, J. Hershberger, and S. Suri. Approximate isocontours and spatial summaries for sensor networks. In *Proceedings of the 6th International Symposium on Information Processing in Sensor Networks (IPSN'07)*, pages 400–409, Cambridge, MA, USA, 2007.

13. R. Ghrist and A. Muhammad. Coverage and hole-detection in sensor networks via homology. In *Proceedings of the 4th international symposium on Information processing in sensor networks*, pages 254– 260, 2005.
14. J. Jiang and M. Worboys. Detecting basic topological changes in sensor networks by local aggregation. In *16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 13–22, 2008.
15. J. Jiang and M. Worboys. Event-based topology for dynamic planar areal objects. *International Journal of Geographical Information Science*, 2009, (to appear).
16. G. Jin and S. Nittel. Tracking deformable 2D objects in wireless sensor networks. In *16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 491–494, 2008.
17. M. Li and B. Yang. A survey on topology issues in wireless sensor network. In *Proceedings of the 2006 International Conference on Wireless Networks*, page 503, Las Vegas, Nevada, USA, 2006.
18. Y. Liu and M. Li. Iso-map: Energy-efficient contour mapping in wireless sensor networks. In *Proceedings of the 27th International Conference on Distributed Computing Systems (ICDCS'07)*, page 36, 2007.
19. X. Meng, L. Li, T. Nandagopal, and S. Lu. Event contour: An efficient and robust mechanism for tasks in sensor networks. Technical report, TR-040018, Computer Science Department, UCLA, 2004.
20. S. Nittel, A. Stefanidis, I. Cruz, M. Egenhofer, D. Goldin, A. Howard, A. Labrinidis, S. Madden, A. Voisard, and M. Worboys. Report from the First Workshop on Geo Sensor Networks. *ACM SIGMOD Record*, 33(1):141–144, 2004.
21. A. Rosenfeld. Adjacency in digital pictures. *Information and Control*, 26:24–33, 1974.
22. R. Sarkar, X. Zhu, J. Gao, L. Guibas, and J. Mitchell. Iso-contour queries and gradient descent with guaranteed delivery in sensor networks. In *Proceedings of the 27th Annual IEEE Conference on Computer Communications (INFOCOM'08)*, Phoenix, AZ, USA, 2008.
23. M. Sharaf, J. Beaver, A. Labrinidis, and P. Chrysanthis. TiNA: A scheme for temporal coherency-aware in-network aggregation. In *Proceedings of the 5th International ACM Workshop on Data Engineering for Wireless and Mobile Access*, pages 69–76, 2003.
24. A. Silberstein, R. Braynard, and J. Yang. Constraint chaining: On energy-efficient continuous monitoring in sensor networks. In *Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data (SIGMOD '06)*, pages 157–168, 2006.
25. E. Sokolowsky, H. Mitchell, and S. Maher. Wildfire growth around Yellowstone National Park in 1988 (WMS). `http://svs.gsfc.nasa.gov/vis/a000000/a002900/a002909/index.html`, 2004. (Last accessed Nov. 30, 2007).
26. I. Solis and K. Obraczka. Isolines: Energy-efficient mapping in sensor networks. In *Proceedings of the 10th IEEE Symposium on Computers and Communications (ISCC'05)*, pages 379–385, 2005.
27. Y. Wang, J. Gao, and J. S. B. Mitchell. Boundary recognition in sensor networks by topological methods. In *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MOBICOM'06)*, pages 122–133, 2006.
28. D. Wilmsen. Derivation of change from sequences of snapshots. Master's thesis, The University of Maine, 2006.
29. M. Worboys and M. Duckham. Monitoring qualitative spatiotemporal change for geosensor networks. *International Journal of Geographical Information Science*, 20(10):1087–1108, 2006.
30. X. Zhu, R. Sarkar, J. Gao, and J. Mitchell. Light-weight contour tracking in wireless sensor networks. In *Proceedings of the 27th Annual IEEE Conference on Computer Communications (INFOCOM'08)*, Phoenix, AZ, USA, 2008.