Distributed Coverage Hole Detection Algorithm Based on Čech Complex

Wang Yuchen[1], Lu Jialiang[1(✉)], and Philippe Martins[2]

[1] SJTU-ParisTech, Elite Institute of Technology, Shanghai, China
lucassjtu@gmail.com, jialiang.lu@sjtu.edu.cn
[2] Institut Telecom, TELECOM ParisTech, LTCI, Paris, France
philippe.martins@telecom-paristech.fr

Abstract. Coverage problem is essential to Wireless Sensor Networks on energy efficient deployment and monitoring. In this paper, we propose a distributed Čech complex algorithm for coverage hole detection in WSNs. Based on our algorithm, each node takes only local information to build Čech sub-complex. Simulations on randomly deployed nodes show that the algorithm achieves a comparable accuracy and a much lower communication cost than a centralized Čech complex construction. Furthermore, it can be combined with distributed Rips complex algorithm to gain an even better performance.

Keywords: Coverage problem · Wireless Sensor Network
Čech complex · Distributed algorithm

1 Introduction

As the theoretical foundation of Internet of Things, Wireless Sensor Network (WSN) is a collection of nodes provided with wireless communication capability, limited computing ability, and sensors to detect physical signals in the environment in which they are deployed. However, the deployment of nodes is always either randomly or highly effected by other restrictions. Thus, it is important to study the coverage problem as fundamental issues in a WSN at the very beginning. In general, coverage problem of WSN reflects how well an area is monitored or tracked by sensors [1]. It plays an important role in many superior applications like energy saving, disaster recovering, load balancing and solving deployment problem. For example, researchers have developed an energy saving algorithm for wireless network based on detecting whether a coverage hole appears if certain nodes are shut down or weakened [2].

In the paper [3], Martins uses Čech complex to represent the coverage states of a node system and propose an algorithm to construct the Čech complex based on the coverage information of each node. In order to deal with its high complexity, a parallel version of this algorithm is already given as well [3]. However, the sensor nodes of WSN have also limited communication, computation powers and even unreliable physical links. Therefore, there is a need to process the

B. Li et al. (Eds.): ChinaCom 2017, LNICST 237, pp. 165–175, 2018.
https://doi.org/10.1007/978-3-319-78139-6_17

information gathered by the sensors locally and then send them back to a central processing unit [4], which means a distributed Čech complex algorithm that we are going to propose in the remainder of this paper is needed in practice.

2 Mathematic Preliminaries

2.1 Simplicial Homology

The homology of the simplicial complex gives us topology information, including connectivity and coverage, about the deployment of WSN nodes. Simplicial homology is a type of homology which results when the spaces being studied are restricted to simplicial complexes and sub-complexes.

Definition 2.1 ***(simplex):*** *Given a set of vertices V and an integer k, a k-simplex is an unordered subset of $k+1$ vertices $[v_0, v_1, v_k]$ where $v_i \in V$ and $v_i \neq v_j$ for all $i \neq j$ [6].*

As represented in Fig. 1, a 0-simplex is a vertex, a 1-simplex is a segment of line, a 2-simplex is a filled triangle, a 3-simplex is a filled tetrahedron, etc.

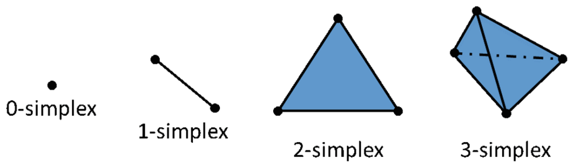


Fig. 1. Examples of simplices

Any subset of vertices included in the set of the $k+1$ vertices of a k-simplex is a face of this k-simplex. Thus, a k-simplex has exactly $k+1(k-1)$-faces, which are $(k-1)$-simplices [7].

Let X be a simplicial complex. For each $k \geq 0$, we define a vector space $C_k(X)$ whose basis is a set of oriented k-simplices of X. If k is greater than the highest dimension of X, let $C_k(X) = 0$. We define the boundary operator to be a linear map $\partial : C_k \rightarrow C_{k-1}$ as follows:

$$\partial[v_0, v_1, \ldots, v_k] = \sum_{i=0}^{k} (-1)^i [v_0, v_1, \ldots, v_{i-1}, v_{i+1} \ldots, v_k] \tag{1}$$

This formula suggests that the boundary of a simplex is the collection of its faces [6]. For example, the boundary of a segment is its two endpoints and a filled triangle is bounded by its three segments.

Consider two subspaces of $C_k(X)$: cycle-subspace and boundary-subspace, denoted as $Z_k(X)$ and $B_k(X)$ respectively. Let ker be the kernel space and im be the image space. By definition, we have:

$$Z_k(X) = ker(\partial : C_k \rightarrow C_{k-1}) \tag{2}$$

$$B_k(X) = im(\partial : C_{k+1} \rightarrow C_k) \tag{3}$$

$Z_k(X)$ includes cycles which are not boundaries while $B_k(X)$ includes only boundaries. A k-cycle u is said homologous with a k-cycle v if their difference is a k-boundary:
$[u] \equiv [v] \Leftrightarrow u - v \in B_k(X)$. A simple computation shows that $\partial \circ \partial = 0$. This result means that a boundary has no boundary. Thus, the k-homology of X is the quotient vector space:

$$H_k(X) = Z_k(X) \setminus B_k(X) \tag{4}$$

Definition 2.2 ***(Betti number):*** *The k-th Betti number is the dimension of* $H_k(X)$

$$\beta_k = dimH_k = dimZ_k - dimB_k \tag{5}$$

This number has an important meaning for coverage problems. The k-th Betti number counts the number of k-dimensional holes in a simplicial complex. For example, the β_0 counts the number of connected components while β_1 counts the number of coverage holes, etc. [6].

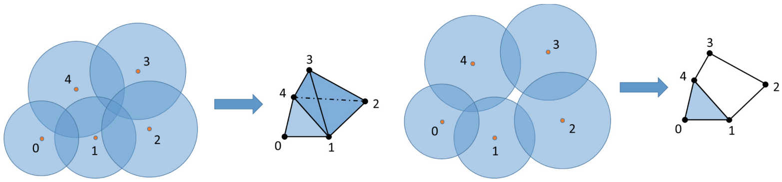
2.2 Čech Complex

The definition of Čech complex is given by Martins in paper [3]:

Definition 2.3 ***(Čech complex):*** *Given a collection of cover sets U, the Čech complex of U, denoted as C(U), is the abstract simplicial complex whose k-simplices correspond to nonempty intersection of* $k+1$ *distinct elements of U.*

We choose $\epsilon(\omega)$ to be the cells coverage range R_{cov}, thus, the Čech complex will represent exactly the coverage states of our system. Graphically, each cell, which indicates a node and its coverage zone, is symbolized by a vertex. A covered space between cells corresponds to a triangle, tetrahedron, etc. filled with colors. A coverage hole between cells is represented by a non-filled triangle, rectangle, etc.

Figure 2 shows how we use Čech complex to symbolize all the coverage information of a network, while forsake all the other insignificant information like the size and position of each coverage zone. The construction of Čech complex needs the exact position and coverage range information of all nodes in WSN, which results in huge communication cost in practice.



(a) A filled tetrahedral and a triangle (b) An empty hole and a triangle

Fig. 2. Čech complex representation

2.3 Rips Complex

The definition of Rips complex is given by Martins in paper [3]:

Definition 2.4 ***(Rips complex):*** *Given a metric space* (M, d)*, a finite set of points* V *on* M *and a fixed radius* ε*, the Rips complex of* V*,* $R_\epsilon(V)$*, is an abstract simplicial complex whose* k*-simplices correspond to unordered* $(k+1)$*-tuples of point in* V *which are pairwise within distance* ε *of each other.*

As an approximation of Čech complex, Rips uses only connectivity information, while coverage range is unknown to Rips complex [8]. Consequently, its accuracy and performance are decided by the ratio γ between communication range and coverage range. For $\gamma \le \sqrt{3}$, Rips complex will not miss any coverage hole, while may detect fake one. For $\gamma \ge 2$, Rips complex will not detect any fake coverage holes, while may miss an existing one. For $\sqrt{3} < \gamma < 2$, Rips complex will not only miss an existing coverage hole, but also detect fake ones. The proof can be found in paper [9].

3 Distributed Čech Complex Algorithm

3.1 Basic Idea

The existing parallel Čech complex algorithm proposed in paper [3], divides assemble of all nodes in a WSN into several sub-domains according to their coordinates on one axe. Then, by constructing Čech sub-complexes and integral all the connection information of those sub-complex into a complete Čech complex, a global Čech complex is obtained. This algorithm is aiming at reducing the computation time to construct the full complex for a large network. Therefore, it is our task to bring forward a distributed version of the algorithm aiming at reducing communication cost and being applicable on real WSN nodes. Compared with geometry based distributed hole detection algorithms like the one in paper [10], our algorithm will not require the boundary information of the network. To start up, our test scenario will be judging whether there are coverage holes or not in a network.

The most basic distributed idea showing as Algorithm 1 is that we let every node collect the position information of all nodes within the communication range R_{com} as their sub-domains. Then every node constructs their Čech sub-complexes on their own sub-domains. Finally, if any sub-domains finds coverage holes, there are coverage holes. Otherwise, there are not.

Unfortunately, in contrary with the centralized Čech complex constructed with acknowledge of all nodes, Čech sub-complexes have only parts of nodes, and thus may not only miss existing coverage hole, but also detect fake non-existing coverage hole. We will demonstrate it in the following section. Consequently, additional process is needed in order to obtain a better approximation to centralized one.

Algorithm 1. Basic distributed idea

$S_0 = \emptyset$; {sub-domain of the node}
S_0 = all nodes within communication range;
C_0 = Construct Čech complex (S_0);
{$C_0.Betti_1$ is the 1-th Betti number of C_0}
if $C_0.Betti_1 > 0$ **then**
  There are $C_0.Betti_1$ coverage holes
else
  There are no coverage holes
end if

3.2 Algorithm Design

Preliminaries

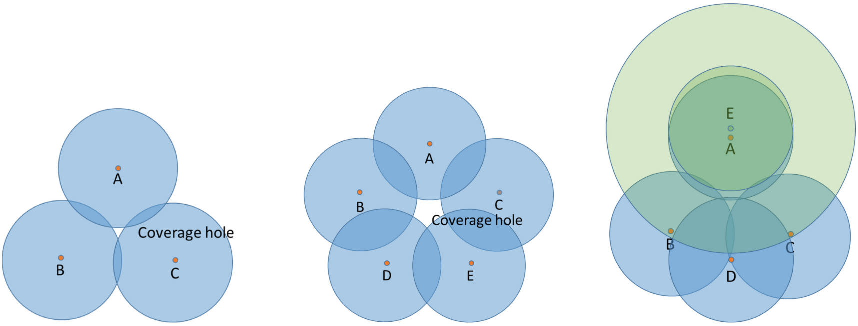
Proposition 1: *When range ratio $\gamma \geq 2$, which means communication range is at least 2 times greater than coverage range, all the nodes connecting with node v_0 can directly communicate with v_0.*

The demonstration of Proposition 1 is obvious and it is the foundation of the following discussion and demonstration. We will thus consider $\gamma \geq 2$ as default from now on.

Triangular Hole Detection

Proposition 2: *When range ratio $\gamma \geq 2$, a triangular coverage hole will definitely be detected by all the nodes on its boundary who construct the Čech sub-complex of all nodes within their communication range.*

Demonstration: As a consequence of Proposition 1, for a triangular coverage hole shown in Fig. 3a, any of its boundary coverage zones must be directly connected with the other two. For Algorithm 1, this indicates that the other two boundary nodes are in its sub-domain for sure. So, the triangular coverage hole will definitely be detected by distributed Čech algorithm run by node A, B or C because they construct local Čech complexes which provide exactly the topology of their sub-domains.



(a) Triangle (b) Pentagon (c) Fake hole

Fig. 3. Coverage holes

Non-Triangular Hole Detection

For the same reason, if we want to guarantee the detection of holes with more edges, we can let nodes collect multi-hop neighbor nodes information in order to ensure the awareness of vertexes which are not directly connected with it.

Proposition 3: *With n-hop neighbor node information, for all the:*

$$k \leq 2n + 1, k \in N \tag{6}$$

coverage holes with k edges can be guaranteed to be detected by at least one node.

Demonstration: Figure 3b shows an pentagonal example, the node A is guaranteed to see node B and C since they are within its communication range ($\gamma \geq 2$), while it cannot see node D or E. For the same reason, the node B is guaranteed to see node D. Therefore, if every node has 2-hop neighbor nodes information, node A is also guaranteed to see node D through node B and see node E through node C. Symmetrically, all the nodes on the boundary of a pentagonal coverage hole can now see each other. Čech complex of sub-domains with n-hop neighbor node information will thus never miss coverage holes with less than $2n+1$ edges.

Fake Hole Exclusion

However, a triangular coverage hole could be covered by a fourth node D and remains no longer a hole like the one shown in Fig. 3c. Furthermore, this fourth node D may not be visible to a fifth node E which can see node A, B and C. Consequently, node E may detect fake coverage holes. In the case above, we can call node D a missing one to node E. This kind of fake coverage holes is actually not only limited in triangular case, but also in any non-triangular cases.

Proposition 4: *A fake coverage hole found by a node in its Čech sub-complex of all nodes in their sub-domain can be excluded by additional 1-hop neighbor nodes information and re-construct the Čech sub-complex.*

Demonstration: It is obvious that the missing one must be certainly connected with all the nodes on the boundary of the fake coverage hole in order to cover the whole fake area. For example, in Fig. 3c, node D must be connected to node A, B and C to cover the whole fake hole.

According to the Propositions 3 and 4, we can ameliorate Algorithm 1 to Algorithm 2 mainly in 2 step: the construction of sub-domain with n-hop neighbor nodes information and the verification of coverage holes. However, this algorithm cannot guarantee to exclude all fake cases. There is an exceedingly infrequent case where the suspected missing nodes we enlarge into the sub-domain form new fake coverage holes. We can identify the coverage hole to fix that, which takes a lot of extra cost. Nevertheless, Algorithm 2 already performs marvelously on detecting no fake hole in practice.

Algorithm 2. Distributed Čech complex algorithm

```
S_0 = ∅; {sub-domain of the node}
S_0+ = n-hop neighbor nodes;
{guarantee to find holes with at most 2n+1 edges}
C_0 = construct Čech complex (S_0);
{C_0.Betti_1 is the 1-th Betti number of C_0}
if C_0.Betti_1 > 0 then
   S_1=(1-hop neighbor nodes of S_0)-S_0
   for all n_i ∈ S_1 do
      if n_i connect to at least 3 n_j ∈ S_0 then
         S_0+ = n_i
      end if
   end for
   C_1= construct Čech complex (S_0);
   if C_1.Betti_1 > 0 then
      There are C_1.Betti_1 coverage holes
   else
      There are no coverage holes
   end if
else
   There are no coverage holes
end if
```

The Algorithm 2 is a distributed Čech complex algorithm that has the same complexity for constructing a Čech complex, but has significantly decreased the execution time and communication cost by reducing the number of nodes in it. Now that we obtain a guaranteed detection on triangular holes, with negligible chance to detect fake holes.

3.3 Combination with Rips Complex Algorithm

In paper [9], an important corollary about the range ratio γ and the Rips complex is proven:

Corollary 4.1: *When $\gamma \geq 2$, if there is a hole in Rips complex $R_{R_{com}}(V)$, there must be a hole actually.*

On the one hand, the distributed Rips complex algorithm proposed in paper [9] reaches a good approximation of centralized one with a high accuracy (over 99%) with no fake detection and can detect almost all the non-triangular coverage holes (>99%) [9]. On the other hand, with only direct neighbor nodes information for all nodes and 2-hop neighbor node information for nodes who detected coverage holes, our distributed Čech complex algorithm is guaranteed to detect all triangular coverage holes. In additional, distributed Rips complex acquires only connectivity information based on communication range, which is already included in the demand of our distributed Čech complex algorithm. Thus, we believe that combining the two distributed algorithms is valuable. For range ratio $\gamma \geq 2$, we can obtain exactly the accuracy of distributed Rips complex algorithm for detecting only non-triangular coverage holes.

4 Simulation and Performance

4.1 Simulation Settings

In order to prove the feasibility to combine our algorithm with the distributed Rips algorithm, we use almost the same simulation setting with what is presented in paper [5]. WSN nodes are randomly deployed in a 100 m×100 m square flat zone as target field according to a Poisson Point Process (PPP) with intensity λ selected from 0.002 to 0.0095 with interval of 0.0005. There are not fence sensors located along the edges of the square. The coverage range of all nodes is fixed to 10 m and the communication range is fixed to 20 m. Figures 4, 5 and 6 give different networks distribution examples under low, middle and high intensity. All the simulations are repeated 1000 times. The distributed Čech complex algorithm runs on different thread in simulation.

Two crucial values are recorded during each simulation. Firstly, the miss rate represents the percentage for distributed Čech complex algorithm to make wrong decision, the lower the better. Secondly, the communication cost is estimated under simple flooding protocol with message cache, which means nodes will not

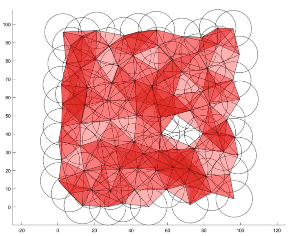


Fig. 4. $\lambda = 0.009$

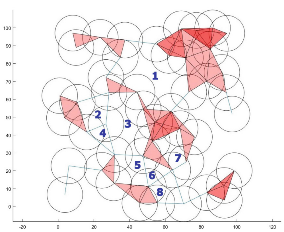


Fig. 5. $\lambda = 0.005$

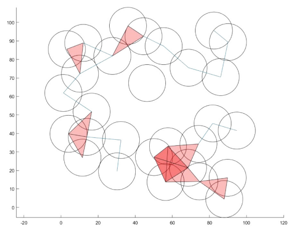


Fig. 6. $\lambda = 0.0025$

retransmit the same message twice. The unit of communication cost is defined as sending one location message to a direct neighbour node. Finally, we compare the communication cost of distributed and centralized algorithm as a ratio. This relative indice shows better the improvement in communication cost.

4.2 Random Deployment Results

Figure 7 shows the results of miss rate and Fig. 8 shows communication cost in accordance with intensity for random PPP deployment. We can have a clear view that for middle and high intensity (>0.005), 2-hop is practically enough to reach a high accuracy (over 99%). This is because that coverage holes with more edges have lower probability to appear as the intensity grows. For low intensity (<0.005), 3-hop has perceptible improvement over 2-hop and remains a high accuracy (over 98%).

On the other hand, Fig. 9 shows the cost ratios between distributed and centralized Čech algorithm. We can see that the communication cost ratio decreases as the intensity increases and converge at around 10% for 2-hop and 30% for 3-hop even under a number of node relatively low (<100). We can thus conclude that an acceptable trade off between accuracy and communication cost is realized. For all the intensity, we manually process a sample survey checking all the

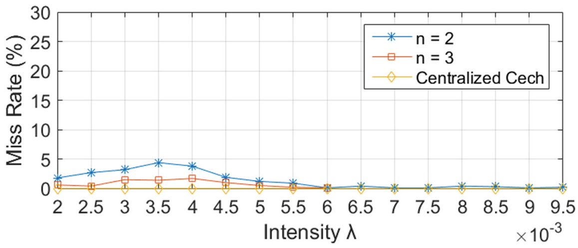


Fig. 7. Miss rate for random deployment

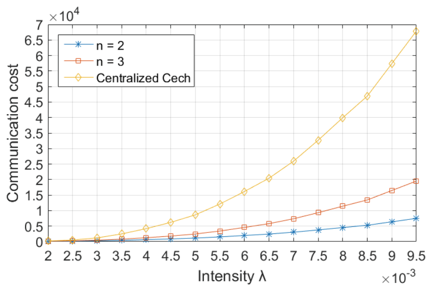


Fig. 8. Communication cost

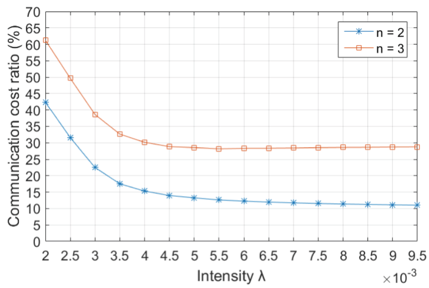


Fig. 9. Communication cost ratio

missing coverage holes and no missing triangular coverage hole is found. Besides, there are three fake coverage holes detected. They all share the same reason that we explained in Sect. 3.

5 Conclusion and Future Work

In this paper, we bring forward a distributed Čech complex algorithm for low cost WSN nodes with communication range at least two times larger than sensing range, which has for now two following usages. On the one hand, it can be used simultaneously with distributed Rips complex algorithm to obtain a detection on all coverage hole with an accuracy over 99%. On the other hand, it can be used independently to obtain an accuracy over 98% or even more with much lower communication cost comparing with centralized Čech complex algorithm and acquires no boundary information comparing with other geometry based algorithms. With adjustable number of hop, larger the network is, easier it becomes for us to find the compromise between accuracy and cost of both communication and computing.

Based on what we have already achieved, there are some perspectives we can proceed in the future. First and foremost, we can identify coverage holes found by separated nodes according to their boundary nodes. Current version of distributed Čech algorithm only takes into account the information on Betti numbers. A complete Čech complex construction from sub-Čech complex can be achieved as well, but a more light-weight representation should be considered. Then, besides the information on Betti numbers, distributed algorithm which can construct a complete Čech complex from sub-Čech complex is also possible. In the end, the hardware experiment of the distributed Čech complex algorithm on real WSN testbeds like IoT-Lab is already on our schedule.

Acknowledgements. This research was supported by 863 under grant No. 2015AA015802, NSF of China under grant No. U1401253 and No. 61373155.

References

1. Huang, C., Tseng, Y., Lo, L.: The coverage problem in three-dimensional wireless sensor networks. J. Interconnection Netw. **08**(03), 3182–3186 (2015). vol. 5
2. Le, N.K., Martins, P., Decreusefond, L., et al.: Simplicial homology based energy saving algorithms for wireless networks. In: IEEE International Conference on Communication Workshop. IEEE (2015)
3. Martins, P.: Homologie simpliciale applique aux rseaux sans fil. Doctorat ParisTech, 2013-ENST-00xx (2013)
4. Roosta, T., Mishra, S.M., Ghazizadeh, A.: Robust estimation and detection in ad hoc and sensor networks. Ecology **87**(2), 372–381 (2006)
5. Yan, F., Vergne, A., Martins, P., Decreusefond, L.: Homology-based distributed coverage hole detection in wireless sensor networks. IEEE/ACM Trans. Netw. **23**(6), 1705–1718 (2015)

6. Le, N.K., Martins, P., Decreusefond, L., et al.: Construction of the generalized Čech complex Complex. In: IEEE Vehicular Technology Conference, pp. 501–510 (2015)
7. Vergne, A., Decreusefond, L., Martins, P.: Simplicial homology for future cellular networks. IEEE Trans. Mob. Comput. **14**(8), 1712–1725 (2013)
8. Yan, F., Martins, P., Decreusefond, L.: Accuracy of homology based approaches for coverage hole detection in wireless sensor networks. In: IEEE International Conference on Communications, pp. 497–502 (2012)
9. Yan, F., Martins, P., Decreusefond, L.: Connectivity-based distributed coverage hole detection in wireless sensor networks. In: IEEE Global Telecommunications Conference, pp. 1–6 (2011)
10. Ma, H.C., Sahoo, P.K., Chen, Y.W.: Computational geometry based distributed coverage hole detection protocol for the wireless sensor networks. J. Netw. Comput. Appl. **34**(5), 1743–1756 (2011)