OPTIMAL SIMPLEX DISTRIBUTION IN HOMOLOGICAL SENSOR NETWORKS*

ABUBAKR MUHAMMAD† AND MHEQUB HAYAT ‡

Key words. Computational algebraic topology, sensor networks, graph mappings

AMS subject classifications. 55U99, 94A99, 94C15

Abstract. In this work, we investigate optimal mappings of simplices in sensor network that solve problems like coverage and routing using topological (homological) methods. Such mappings are important to enable a distributed implementation on practical real-world networks with bandwidth, energy and computational limitations. We discuss aspects of this problem and outline solutions.

1. Introduction. In recent years, several properties in networked sensing and distributed systems have been modeled by using topological spaces and their topological invariants. The unifying theme in these approaches has been that the local properties of a network, as dictated by local interactions among agents, can be captured by certain topological spaces. The solution to a global problem related to the network is then given by a topological invariant of that topological space. These spaces are mostly combinatorial in nature and are generalizations of graphs. Successful applications of such modeling include coverage problems for sensor networks [1, 2]; concurrency modeling in asynchronous distributed systems [4]; and routing in networks without geographical information [3].

For a practical utilization of these studies on distributed systems, it is vitally important that the global topological invariants can be computed in a decentralized manner with manageable complexity and scalability. In [5, 6] the author focussed on one class of such problems, i.e. the computation of homological groups in networks, and outlined decentralized methods of computing the homology groups that are implementable on real networks. The methods have been outlined in the form of tokenized protocols that take into account network characteristics such as medium access and bandwidth limitations.

In this work, we investigate one particular aspect of this problem in more detail, i.e. the problem of mapping higher order simplicial structures onto the physical resources of the network, namely the sensor nodes. In graphical models, such mappings are usually straightforward. Since only vertices and edges need to be mapped, the vertices are assigned injectively to the nodes while the edges represent node interactions by communication or perception, resulting in a totally decentralized scheme. However, this natural mapping is not available when higher order network interactions are modeled by simplices. In [6], one candidate scheme was given with the observation that there are many solutions of the problem and the selection of the most optimal solution is critical for a real world implementation.

2. Simplicial Mapping Problem. Let X be a d-dimensional simplicial complex modeling some important characteristic of a sensor network. For a precise defini-

*This work was funded by a LUMS faculty startup grant.

†Department of Electrical Engineering, LUMS School of Science & Engineering, Lahore, Pakistan 54792. abubakr@lums.edu.pk

‡Department of Electrical Engineering, LUMS School of Science & Engineering, Lahore, Pakistan 54792. mhayat86@gmail.com

tion of simplices and simplicial complex, please see [5]. Let the set of k-simplices of X be denoted by X^k. E.g., the set of vertices is denoted by X^0 and the set of edges by X^1. For $k > 1$, X^k denotes the higher order simplices of the complex. Let a network be made up of n sensor nodes and let the set of nodes be denoted by $N = \{1, \ldots, n\}$. Then, we define the simplicial mapping as follows.

Find a set of maps, $f^k : X^k \to N$, for $0 \leq k \leq d$ such that $f_k^{-1}(i) \subseteq X$ is a subcomplex of X denoting the k-simplices *owned* by a sensor node i and

$$X = \bigcup_{k=1}^{d} \bigcup_{i=1}^{n} f_k^{-1}(i).$$

From the properties of a simplicial complex, it can be seen that a particular choice of f^k induces all lower dimensional maps f^j for $j \leq k$. Therefore, we only investigate $f^d : X^d \to N$ with the property that

$$X = \bigcup_{i=1}^{n} F^i,$$

where we use the notation $F^i = f_d^{-1}(i)$ for convenience. The only requirement on the map is that it needs to be *deadlock free*. In other words, the map should be defined in a way such that it can only be interpreted in a unique assignment of each simplex to only one sensor node.

One extreme example of such an assignment is a map that assigns all the simplices to only one *super node*. This will result in a totally centralized computation, in which all network information is collected at one computational resource (a data center) and the homology group computations are carried out using standard algorithms that do not incorporate issues of parallelism or distribution. Another example is the protocol `SimplexOwn` given in [6] by the author. There, the map is defined in the following way. Each k-simplex is a $k+1$-tuple of vertices $\{v_0, v_1, \ldots, v_k\}$. If the vertices are mapped injectively to the sensor node IDs $\{1, \ldots, n\}$, then a unique deadlock free assignment of simplices is to assign each simplex to the lowest ID numbers of its member vertices (in other words sensor nodes). This assignment has been described in the form of a tokenized protocol.

One motivation for designing more complicated assignment maps is to force the sub-complexes owned by each node be as large as possible, so that the total number of complexes are small in number. One potential candidate could be a protocol that favors simplex ownership by nodes that have a higher degree than its neighbors. This setup is likely to favor the concentration of more global information into a small number of nodes so that a distributed algorithm may not be too far away from a corresponding global algorithm in terms of error and convergence rate. On the other hand, such a scheme may put an unacceptably large computational load on a small number of nodes, thus under-utilizing the computational resources present on the idle nodes. Therefore, there is a clear need to find assignments that optimize the costs by striking a balance between these conflicting needs. In this work, we give some concrete answers to finding those optimal simplex assignment.

3. Results. We briefly outline our results below. Details will be given in a full version of the paper.

3.1. Feasible Simplex Mappings. As mentioned above, a feasible map only needs to be *deadlock free*. We prove that the set of all deadlock free simplex maps is a subset of the k-set agreement problems studied by [4]. We conjecture that the two problem sets are in fact equivalent.

3.2. Cost of Simplex Mapping. We have formulated a cost formula for each simplex mapping that computes the total power consumed of a distributed algorithm running at the simplex level. The cost incorporates computational and communication loads at each node by using computational energy/instruction, communication energy/broadcast, instruction cycles times and bandwidth. We further assume that that the energy of computation is some known multiple m_1, of the energy of computation. Similarly, we relate instruction cycle and bandwidth by a second multiple m_2, to get a parameterized cost formula at the node level in terms of m_1 and m_2.

We remind the reader that computational loads and communication loads capture the two conflicting requirements for which a balance is required. More computational loads in smaller number of nodes result in less communication loads and translate to a more centralized architecture and more hierarchical network structure with possibly better algorithmic convergence rates. A less or more even computational load shifts the balance to a more decentralized scheme with more requirements for communication. However, communication is more expensive in terms of energy and ultimately hits bandwidth limitations.

3.3. Statistics of Simplex Mappings. Moving towards cost minimization, we first observe that usually the network topology is not fixed or very structured in a real network. Therefore, we carried out a statistical analysis of the simplex distribution for arbitrary simplex assignments. This gives us *probability distributions* for the computational and communication loads at the nodes. This approach allows us to attack the optimization problem without assuming a certain topology.

3.4. Optimal Simplex Mappings. We write down the costs as the *entropy* of the load distribution and find a distribution that maximizes this cost entropy. For specific cases, we have successfully constructed explicit maps that carry out such an optimization. We have also analyzed the distributions that are induced by the optimal simplex mappings.

3.5. Further Results. We are further investigating the following questions in relation to the above findings.

1. For what choices of m_1, m_2 is it possible to extremize the optimal map to a completely centralized solution? This question is important to justify the need for a distributed algorithm against real energy budgets in a network.
2. In all optimal simplex mappings, is it necessary to map simplices to a member node only? We have some evidence on why this is true but a formal proof remains to be done.
3. How well does a random assignment of simplices to nodes fare viz-a-viz an optimal assignment using the above machinery? The main obstacle in such analysis is how to define explicitly a random assignment as a feasible deadlock free mapping.

3.6. Simulation Results. We give below some simulation results to confirm the theory outlined in this project. In all simulations, a unit-square with identified boundaries (i.e. a torus) is sampled in a pseudo-random fashion to generate N sensor node locations $x_i \in [0,1]^2$, $i = 1\dots, N$. Pseudo-random sampling is achieved by a lattice scheme given by $x_i = (i/N, \{i\alpha\})$ where $\{u\} = u - \lfloor u \rfloor$ is the fractional part of $u \in \mathbb{R}$ and $\alpha = \frac{\sqrt{5}+1}{2}$, the golden ratio. From these node locations, simplicial com-

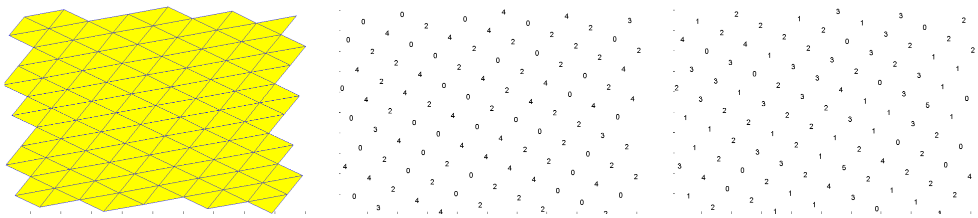

FIG. 3.1. *[Left] Simplicial complex of a* 100 *node network sampled on the torus (simplices on periphery omitted to remove clutter). [Center] Number of simplices mapped to each node by a SimplexOwn protocol using a median map. [Right] Results of SimplexOwn protocol using a random assignment map.*

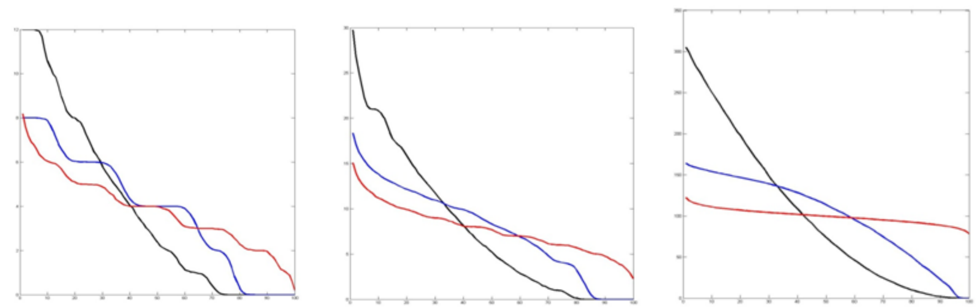

FIG. 3.2. *Simplex distribution histograms for three sensor diameters: 0.16 units (left), 0.20 units (center) and 0.32 units (right). Each network is tested for three SimplexOwn protocols: Min (black), Median (Blue) and Random (Red). In all simulations, random assignment produces a better load balancing than other schemes (measured via distribution entropy).*

plexes are generated using the Rips-complex construction for different sensor range distances, as given in [2]. We sample pseudo-randomly on a torus to remove edge effects in network creation and to create networks with statistically uniform simplex density.

Different simplex own protocols are simulated on these networks. The resulting mappings are recorded in the form of histograms to visualize the distribution of simplices onto the nodes. The results of these simulations have been summarized in the figures above.

REFERENCES

[1] R. Ghrist and A. Muhammad, "Coverage and Hole-Detection in Sensor Networks via Homology," *The Fourth International Conference on Information Processing in Sensor Networks* (IPSN'05), UCLA, Los Angeles, CA, April 25-27, 2005.

[2] V. de Silva, R. Ghrist and A. Muhammad, "Blind Swarms for Coverage in 2-D," *Robotics: Science and Systems*, MIT, Cambridge, MA, 2005.

[3] Q. Fang, J. Gao, L. Guibas, V. de Silva, and L. Zhang, "Glider: gradient landmark-based distributed routing for sensor networks," *Proc. IEEE Infocom*, 2005.

[4] M. Herlihy and N. Shavit, "The Topological Structure of Asynchronous Computability," *Journal of the ACM*, vol. 46, no. 6, pp. 858–923, November 1999.

[5] A. Muhammad and M. Egerstedt, "Control Using Higher Order Laplacians in Network Topologies," in *Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan , 2006.

[6] A. Muhammad and A. Jadbabaie, "Decentralized Computation of Homology Groups in Networks by Gossip", American Control Conference, 2007.