Cut Graph Based Information Storage and Retrieval in 3D Sensor Networks with General Topology

Yang Yang, Miao Jin, Yao Zhao, and Hongyi Wu

Presenter: Hongyi Wu

University of Louisiana at Lafayette

Distributed Data Storage and Retrieval

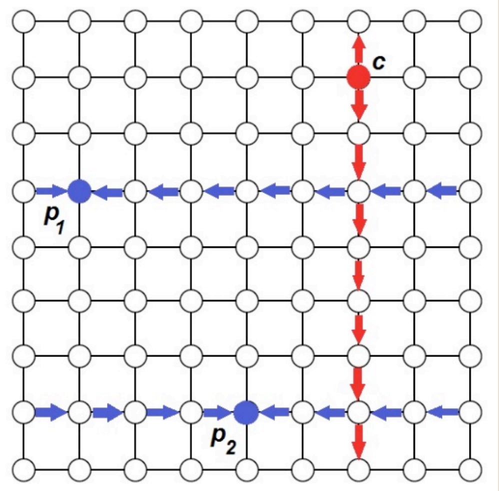
- Distributed in-network data-centric processing aims to reduce energy consumed for communication and establish a self-contained data storage, retrieval, aggregation, and query sensor system which focuses more on the data itself rather than the identities of the individual network nodes.

Outline

- Introduction
  - An overview of distributed data storage and retrieval algorithms in 2D networks
  - Challenges in 3D networks
- Our approach in 3D networks
- Simulations
- Conclusion and future work

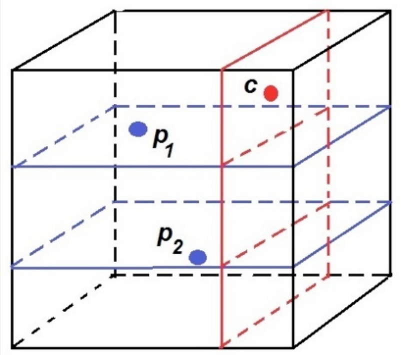
An Overview of Distributed Data Storage and Retrieval Algorithms in 2D Networks

- Existing in-network data storage and retrieval algorithms in 2D networks
  - Geographical hash table based schemes
  - Double-ruling based schemes
    - Efficient aggregated data retrieval
    - Well balanced load across the network
    - Better fault tolerance

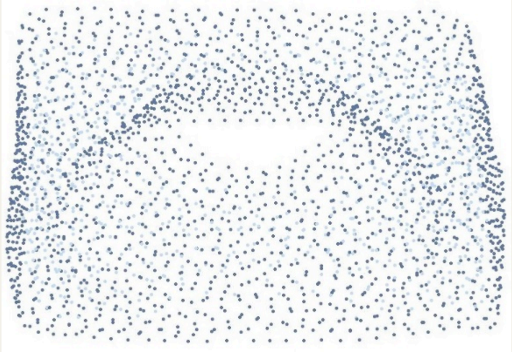


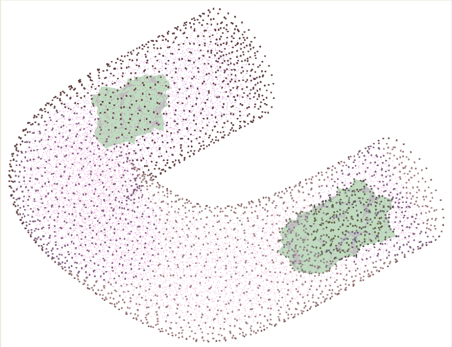
Challenges in 3D Networks

- A naive double-ruling based scheme requires a network with a regular cube shape and uniform node distribution.



- While Many practical 3D sensor networks are topologically different from a cube.



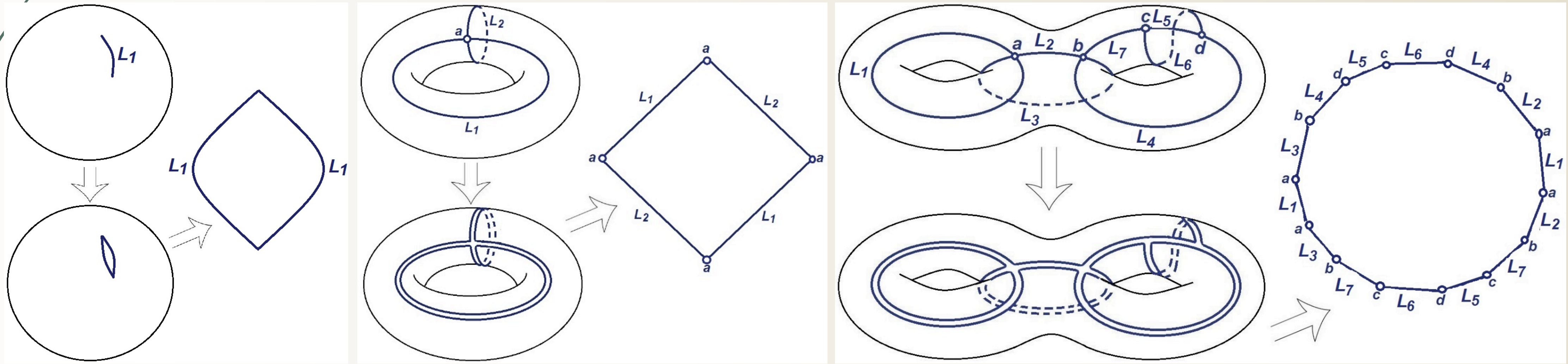


Challenges in 3D Networks

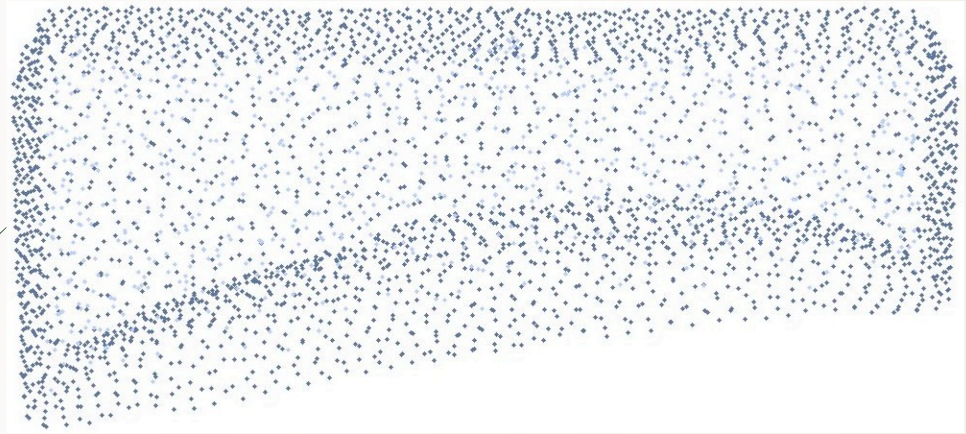
- The delivery of data and query to the mapped geo-localizations for in-network data storage and retrieval requires location information.
- Except the cost to equip GPS, some application scenarios even prohibit the reception of satellite signals by part or all of the sensors, rendering it impossible to solely rely on global navigation systems.
- Moreover, guaranteed geographic routing with local information only and constant storage at each node is non-trivial and even impossible in general 3D sensor networks.

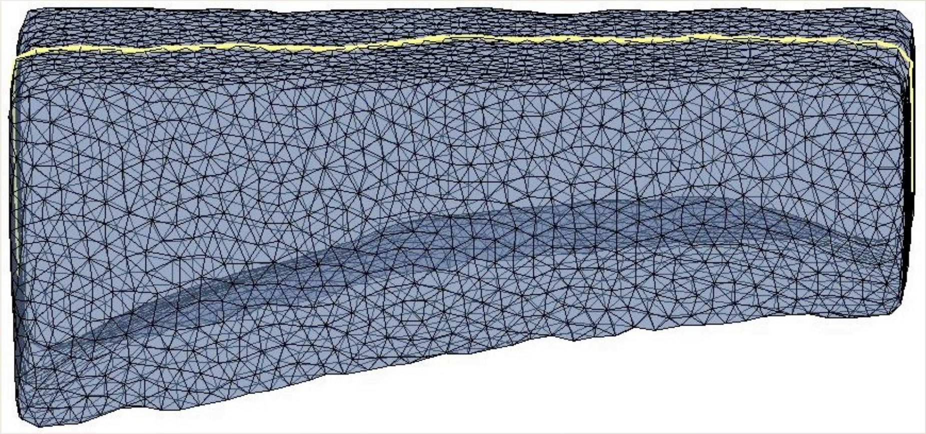
Cut Graph

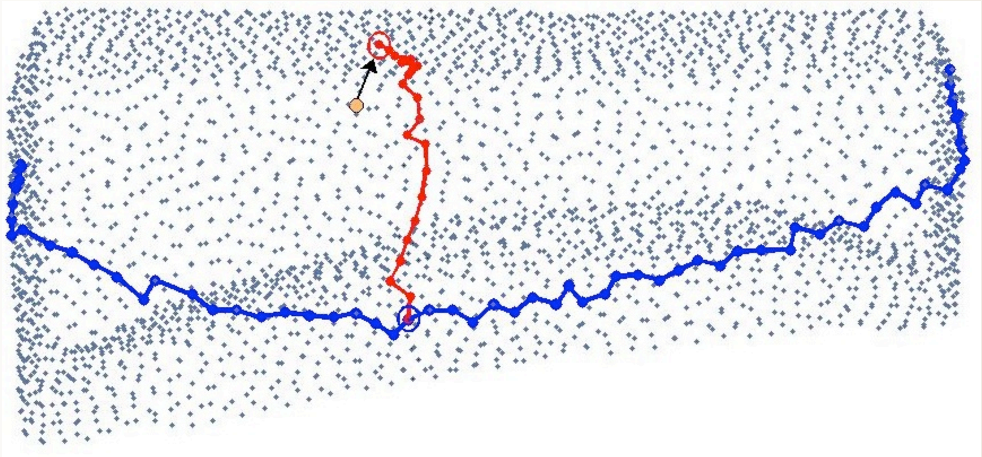
- Any closed surface (e.g., a surface without a boundary) can be opened into a topological disk (e.g., a surface with one boundary) by cutting along an appropriate set of edges called cut graph.

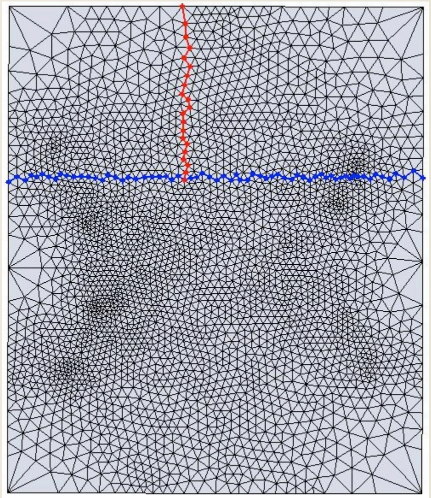


Overview of Our Approach

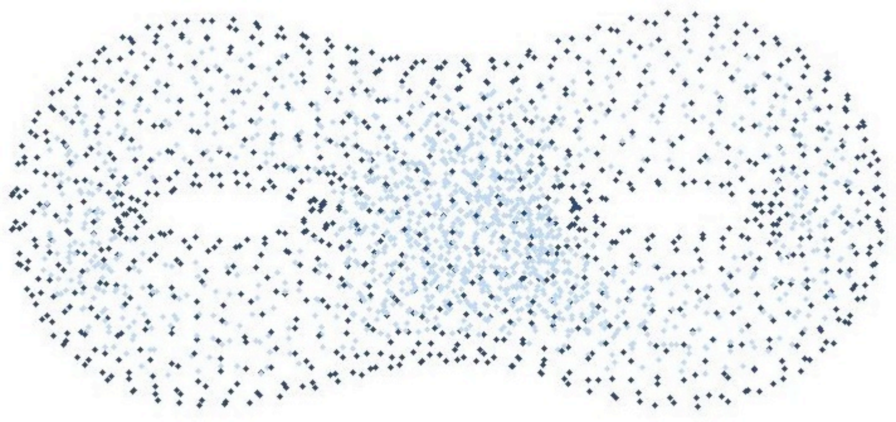


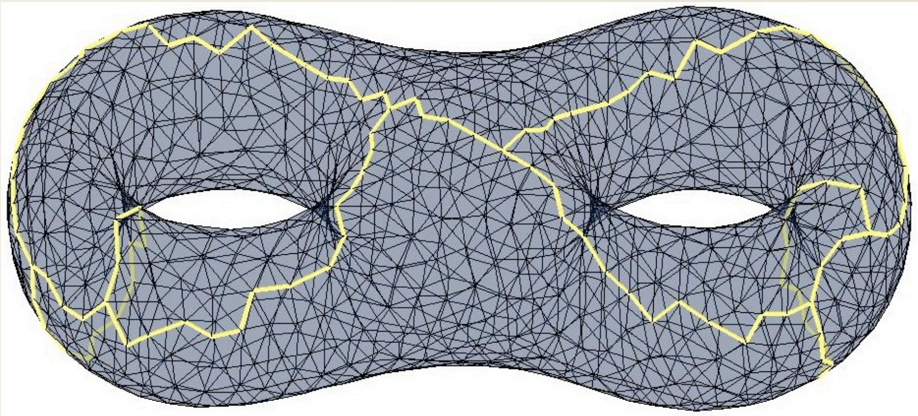


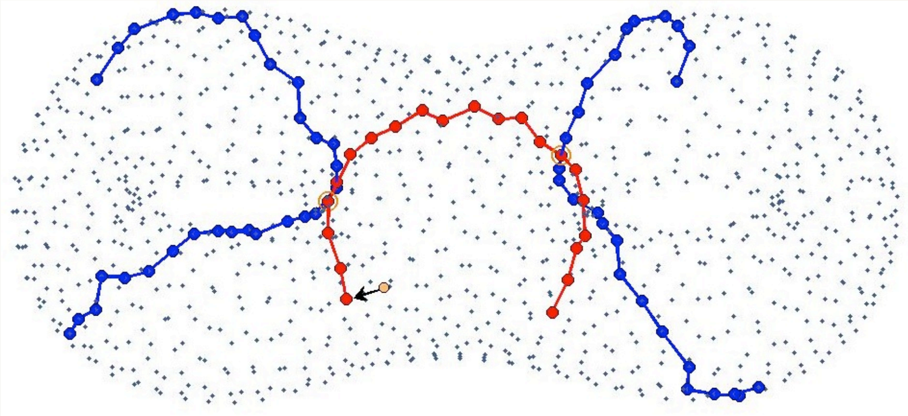


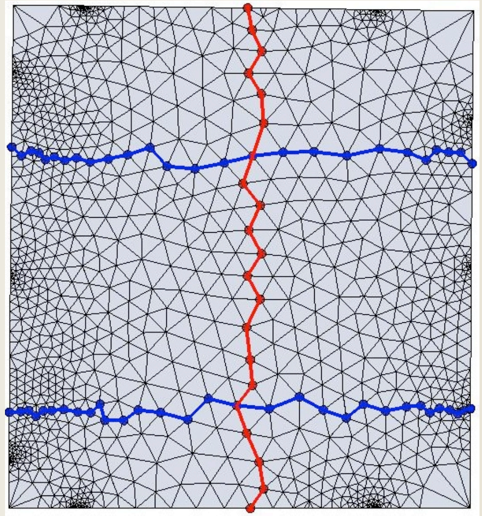


Overview of Our Approach









Computing Cut Graph of Boundary Surface

- Given a sensor network deployed in 3D, we first identify its boundary nodes and then extract a triangular mesh structure of the boundary of the network, denoted as M.
- The algorithm randomly chooses one triangle face and mark its three edges. Triangles grow with the width first way. At each step of the growing, all the marked triangles always form a topological disk, and the marked edges form the boundary of the disk.
- When all the triangles have been marked, the marked edges form a cut graph of M. We can cut M into a topological disk D along the computed cut graph.

Generating Planar Rectangle Virtual Coordinates

- Tool: Discrete Surface Ricci Flow
- Input: a topological disk ***D***
- Output: ***D*** is mapped to an aligned planar rectangle
- Computing Planar Rectangle Virtual Coordinates
  - Uniformly pick four vertices along the boundary of the computed cut graph surface and assign their target Gaussian curvatures $\pi/2$. Assign the target Gaussian curvature of all other vertices 0.
  - Discrete surface Ricci flow continuously deform the edge length of the triangular mesh and the final edge length induces a planar rectangle embedding of ***D***.

IMPLEMENTATION

- Data Replication
- Data Retrieval
- Delivery of Data and Query
- Storage

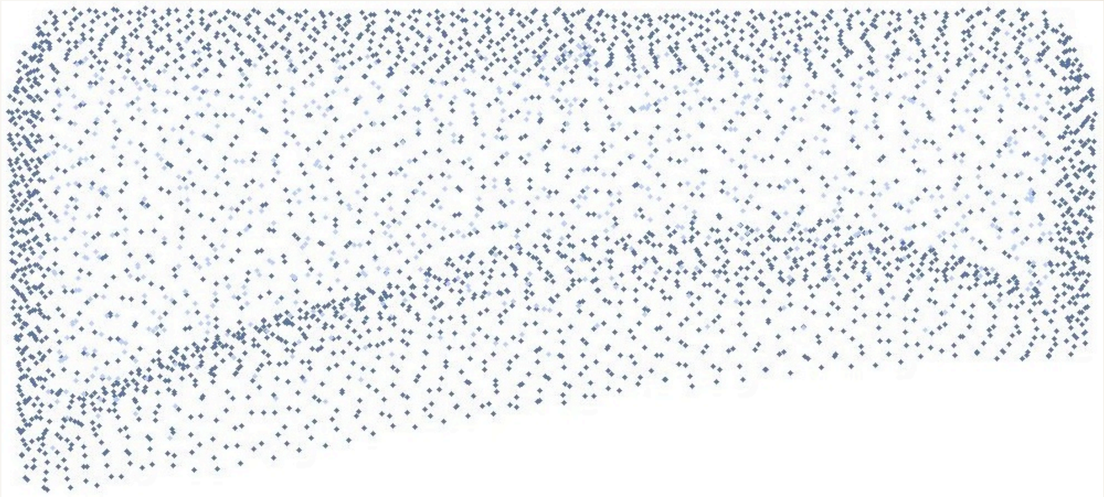
Time Complexity and Communication Cost

- Denote the size of nodes of a network as n, the size of its boundary nodes as m, n>>m.

	Cut Graph	Ricci Flow	Routing to Boundary
Time Complexity	O(m)	$-C\log\varepsilon/\lambda$	O(n)
Communication Cost	O(m)	$-C\log\varepsilon/\lambda$ dm	O(dn)

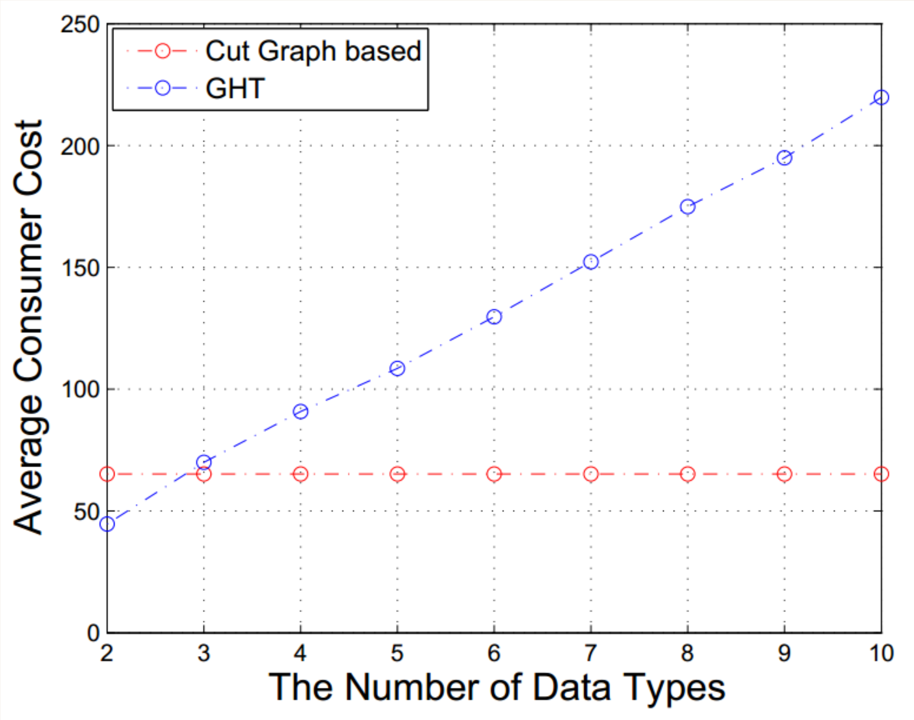
SIMULATIONS

- Seabed model
  - The network has 4369 number of nodes and the average number of neighbors of each node is 13.79.
  - Data storage and retrieval costs are measured by the number of hop counts needed to store or retrieve data.
  - Traffic load on each node is measured by the number of messages passing through it.



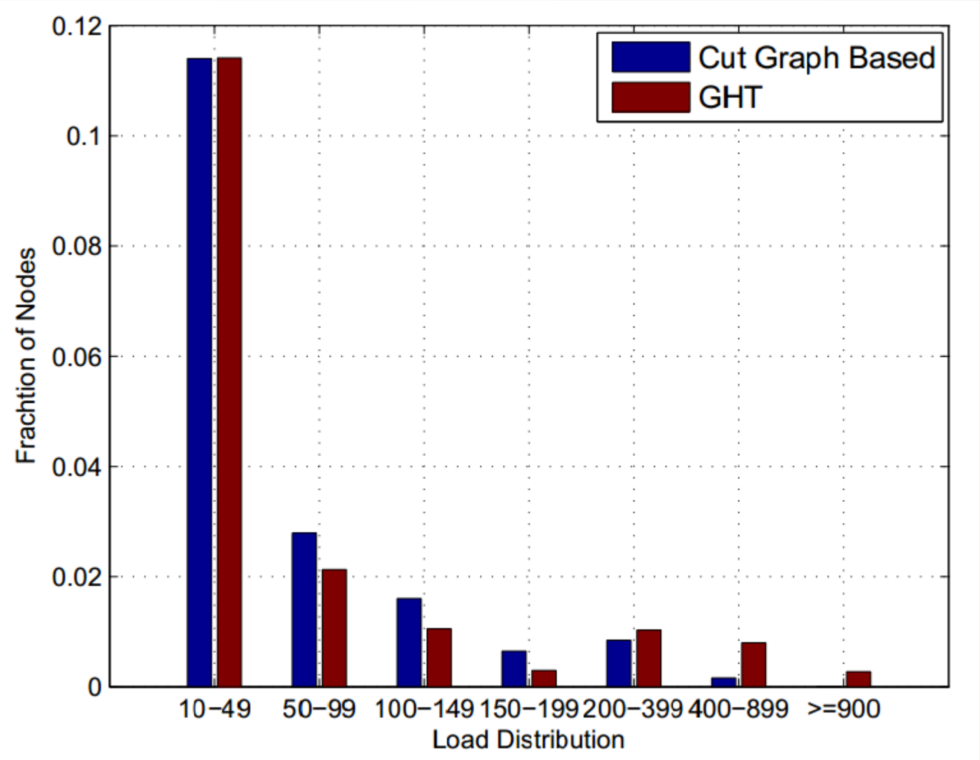
SIMULATIONS

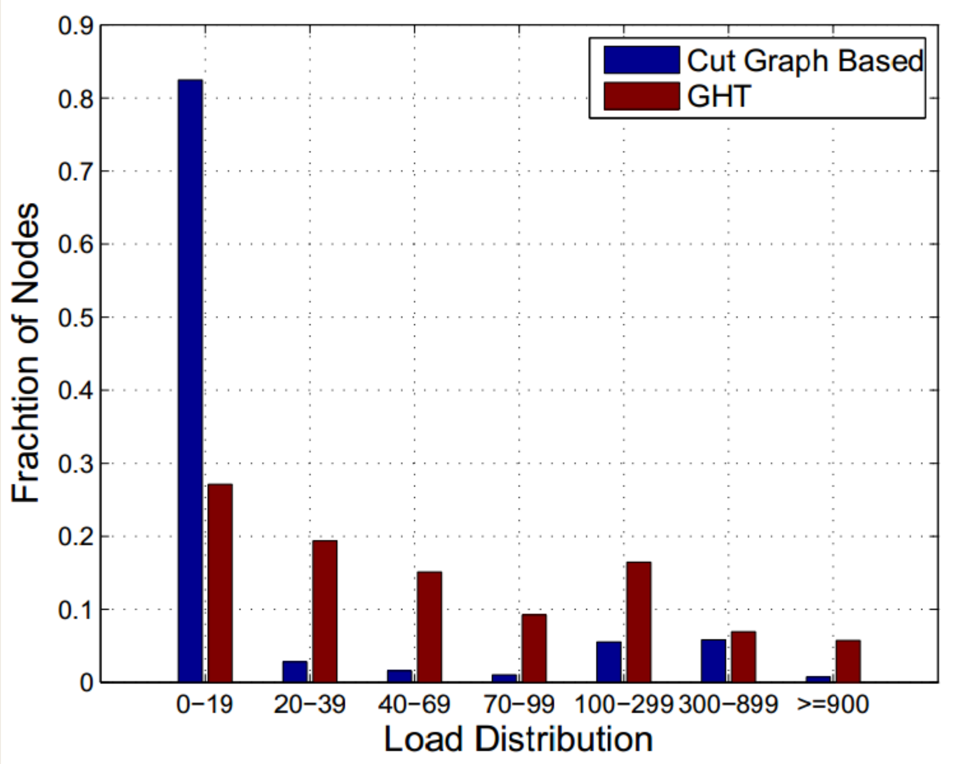
- Aggregated Data
  - The retrieval cost of cut graph based scheme is fixed
  - The retrieval cost of GHT based scheme increases proportional to the number of data types



SIMULATIONS

- Load Distribution
  - A network stored data with one data type
  - A network stored data with ten data types





CONCLUSION AND FUTURE WORKS

- A location-free cut graph based double-ruling scheme for large-scale 3D sensor networks.
- A data query travels along a simple curve with the guaranteed success to retrieve aggregated data through time and space with different types across the network.
- Future work: a dynamic network with a local recovery scheme for the presence of nodes' failures and replacement.

Questions?