Landmark-Based Information Storage and Retrieval in Sensor Networks

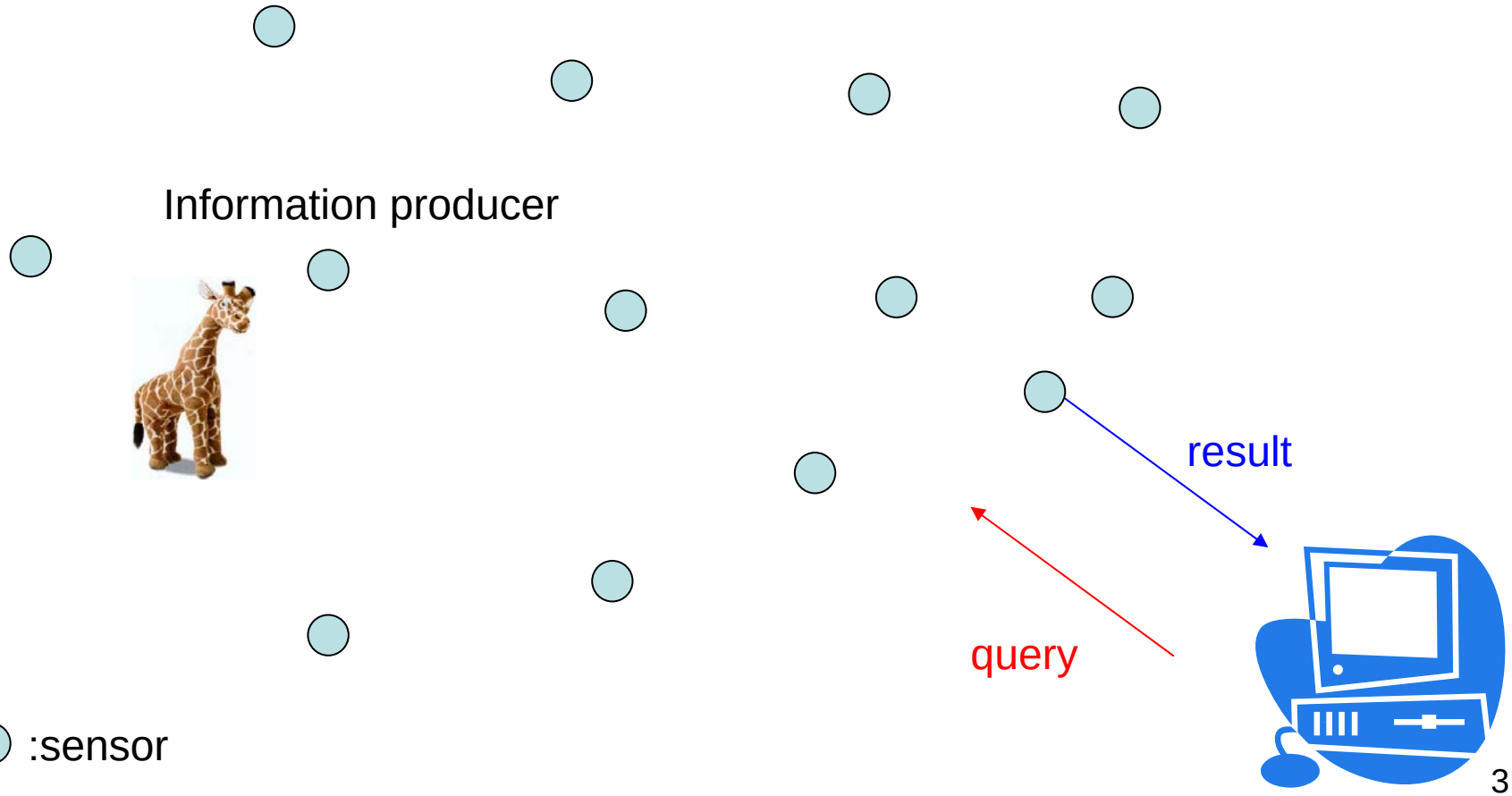
INFOCOM 2006

Speaker: Cheng-Han Wu

Outline

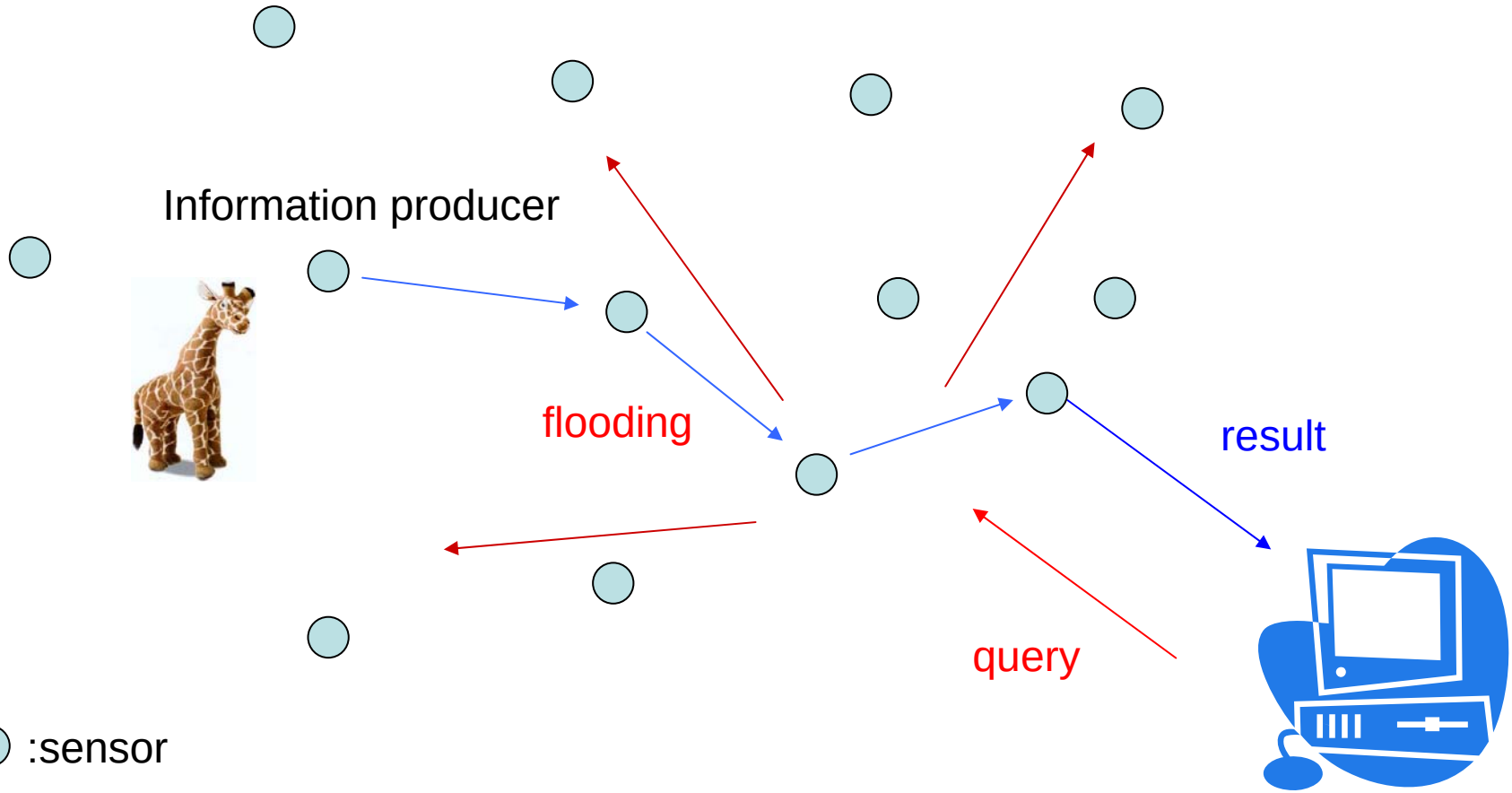
- Introduction
  - Geographic Hash Table (GHT)
  - Double-ruling
- Landmark-Based Information Storage and Retrieval
- Simulation
- Conclusion

What problem we want to solve?



An Information Brokerage

- An *information brokerage* scheme is a mechanism that carries out data publication, data replication for the information producers (a.k.a. *producers*) and data retrieval for the information consumers (a.k.a. *consumers*).


Information producer
flooding
result
query
:sensor
Information consumers

DHT(Distributed Hash Table)

- Data is named and stored via its content.
- A piece of data is hashed to a node based on its content, using a common hash function known to the producers and consumers.
- The node stores the data of the producers and acts as a a reservoir from which the consumers can retrieve its desired data.

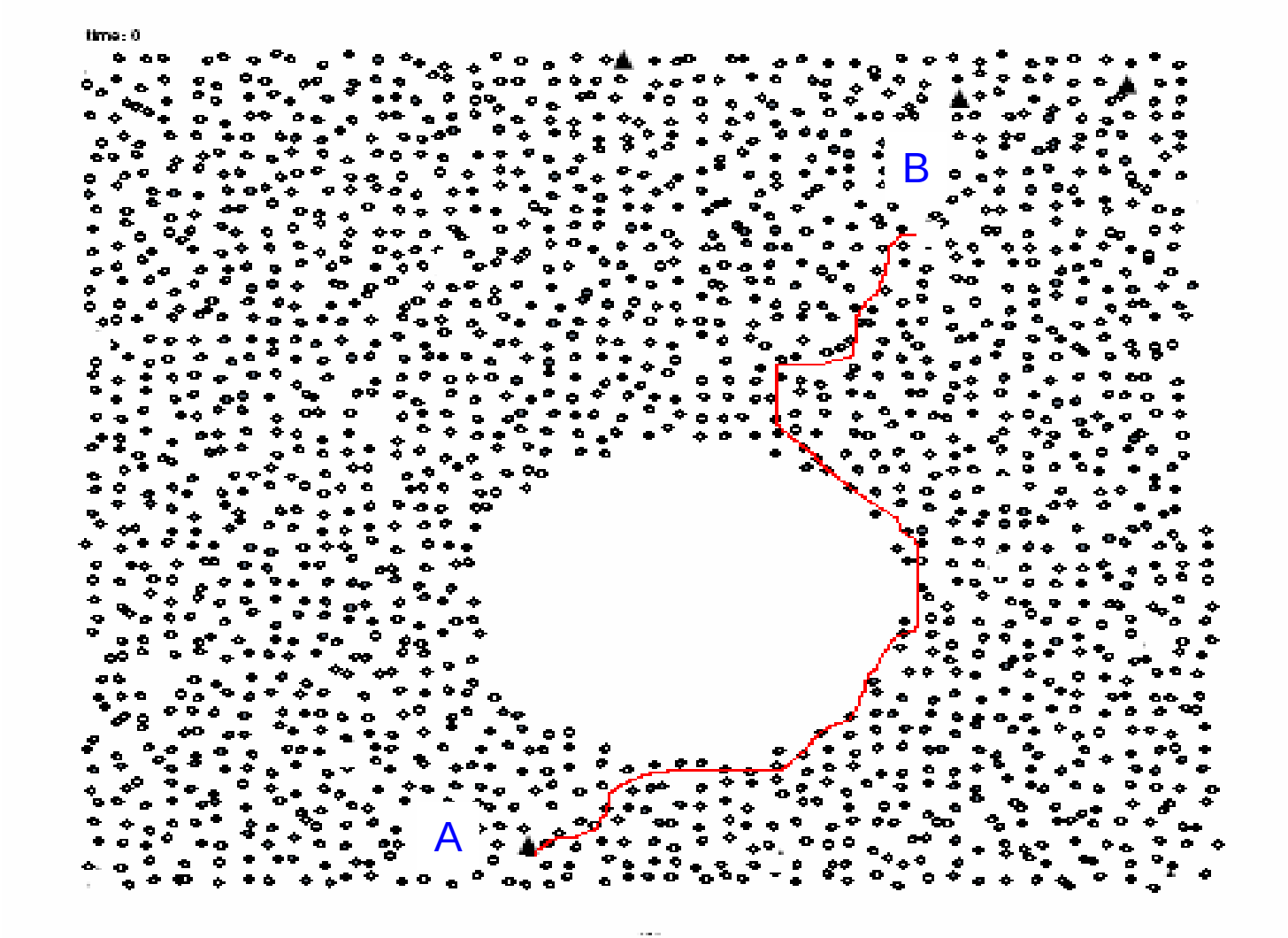
GHT (Geographic Hash Table)(1/3)

- combines the DHT idea with geographic naming and routing.

- uses a geographic routing approach ,GPSR, as the underlying routing scheme.

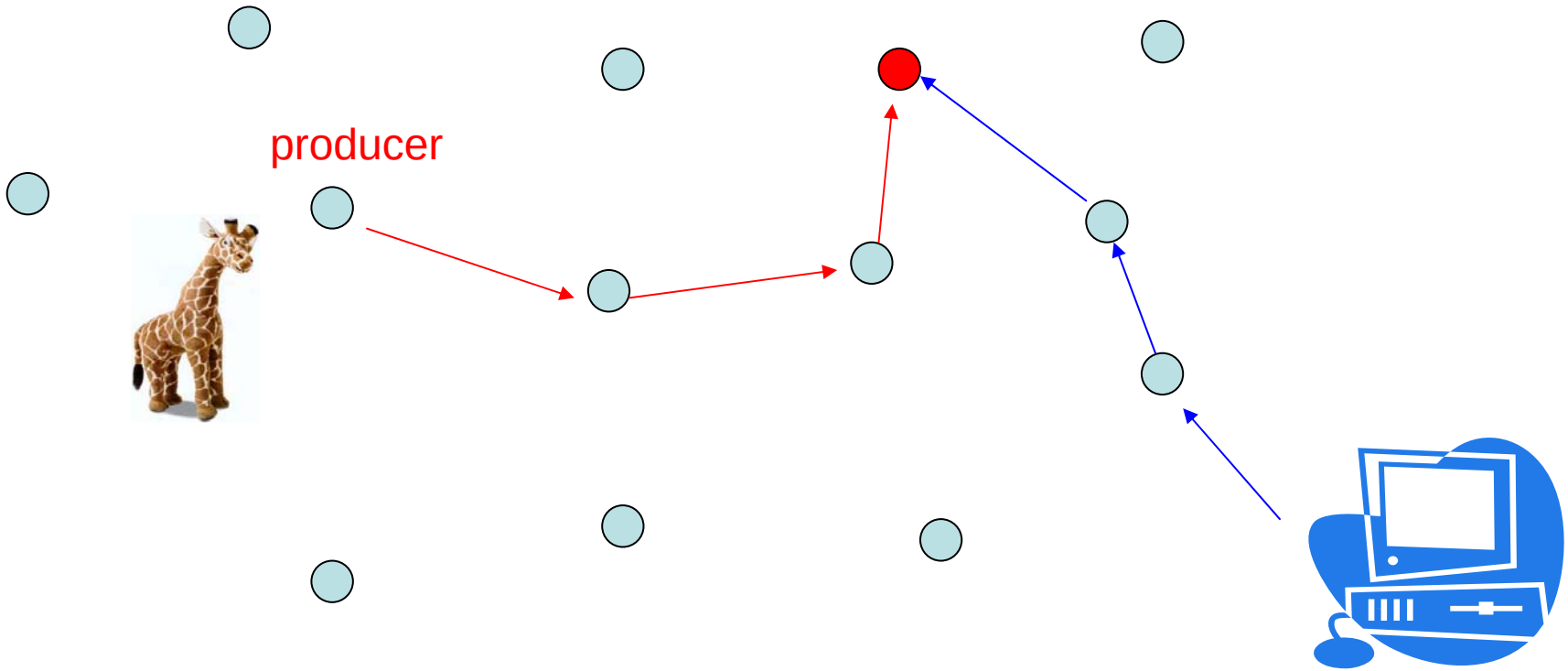
GHT (Geographic Hash Table)(2/3)

GPSR



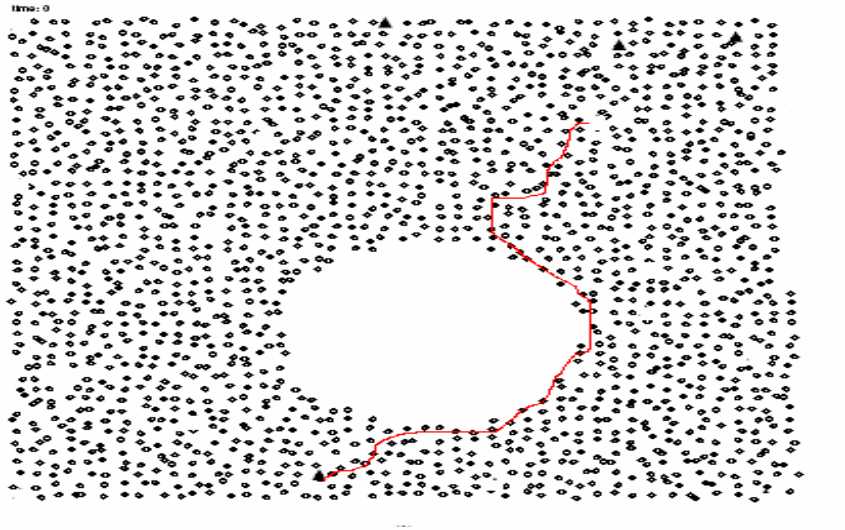
GHT (Geographic Hash Table)(3/3)

The name of data —hash→ the coordinate of a node

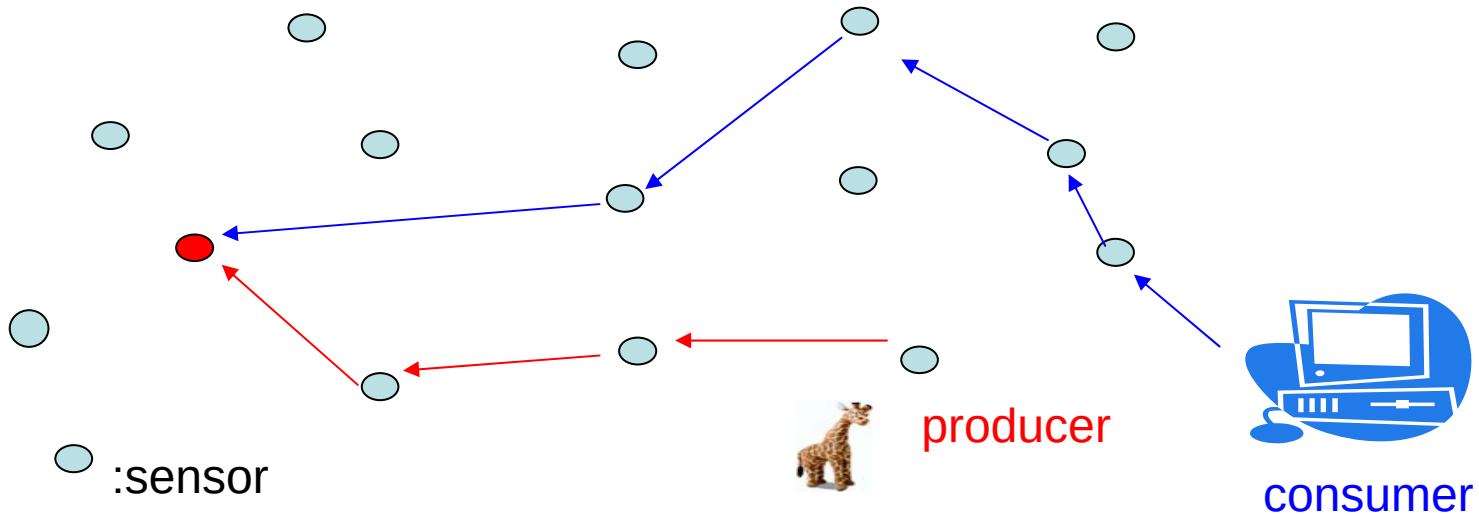


The Disadvantage of The GHT

- Nodes on the boundaries of holes are used more often.



- GHT is not locality aware.



The Advantage of LANDMARK-BASED DATA-CENTRIC STORAGE AND RETRIVAL

- Nodes on the boundaries of holes are not used more often.
- It is locality aware.
- It allows the possibility of data aggregation
- It achieves the above goals without requiring node's geographic information as in GHT.

LANDMARK-BASED DATA-CENTRIC STORAGE AND RETRIVAL

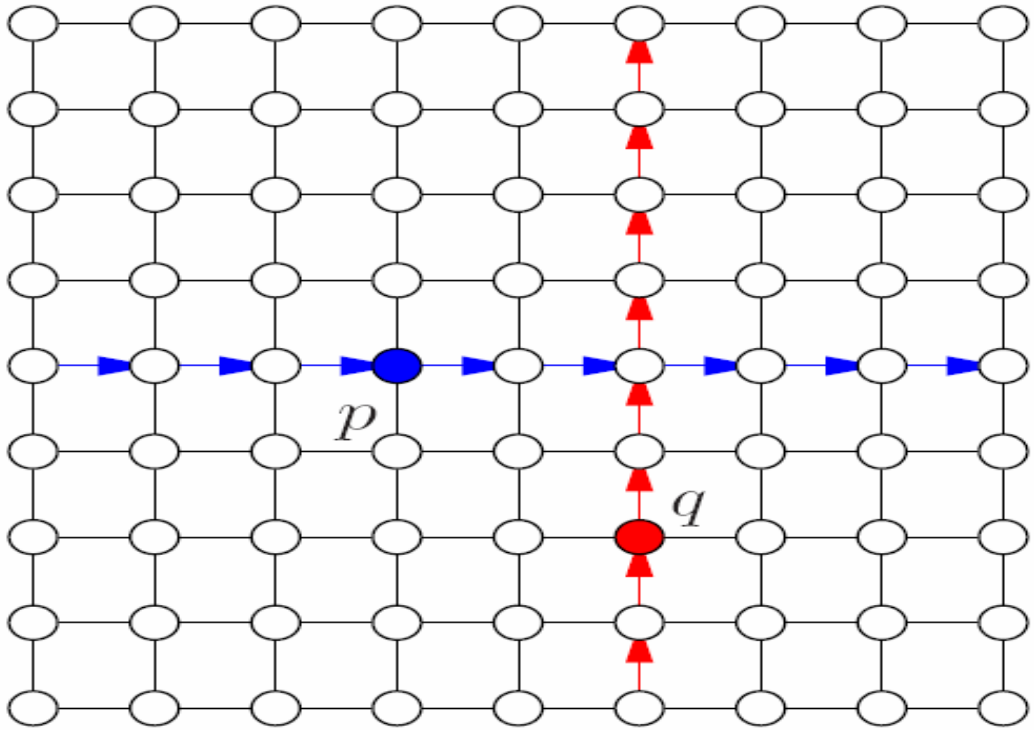
- integrates the Distributed Hash Table (DHT) idea and the double-ruling idea.

- uses the Gradient Landmark-based Distributed Routing protocol (GLIDER) as the underlying routing scheme.

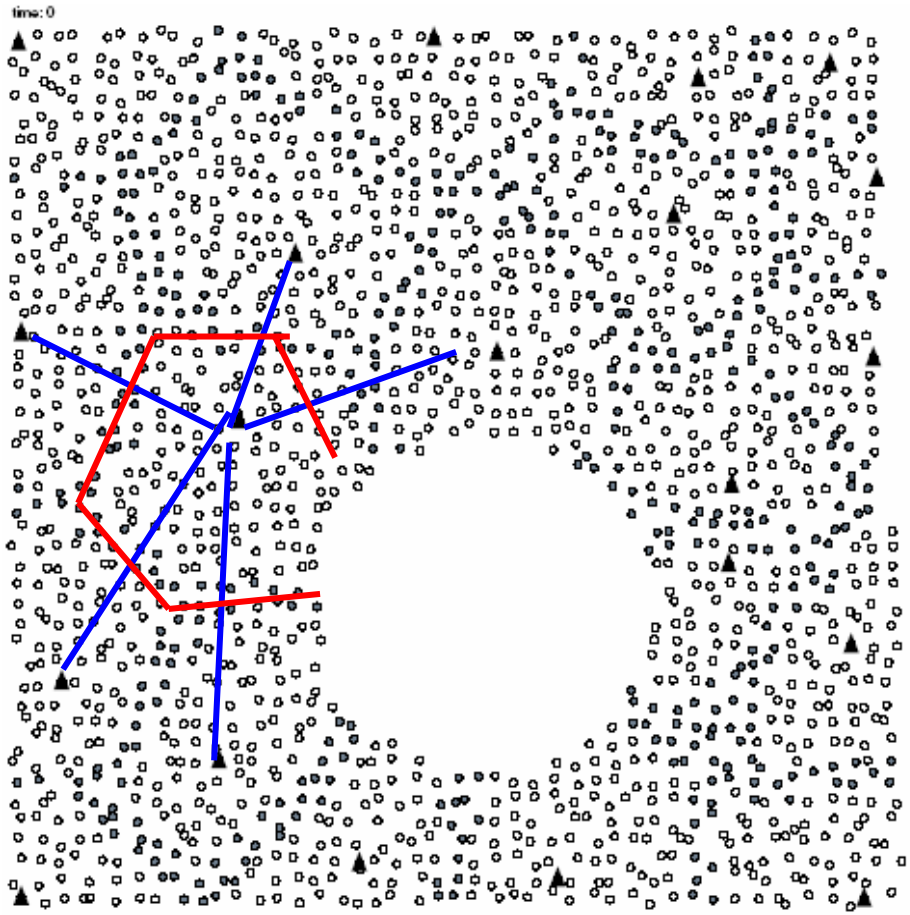
Double-ruling

- Data is not stored at a single node or its nearby neighbors, but at nodes that follow a one-dimensional curve while a data request travels along a set of nodes that follow another one-dimensional curve. Therefore, successful retrieval is guaranteed if every retrieval curve intersects every data storage curve.

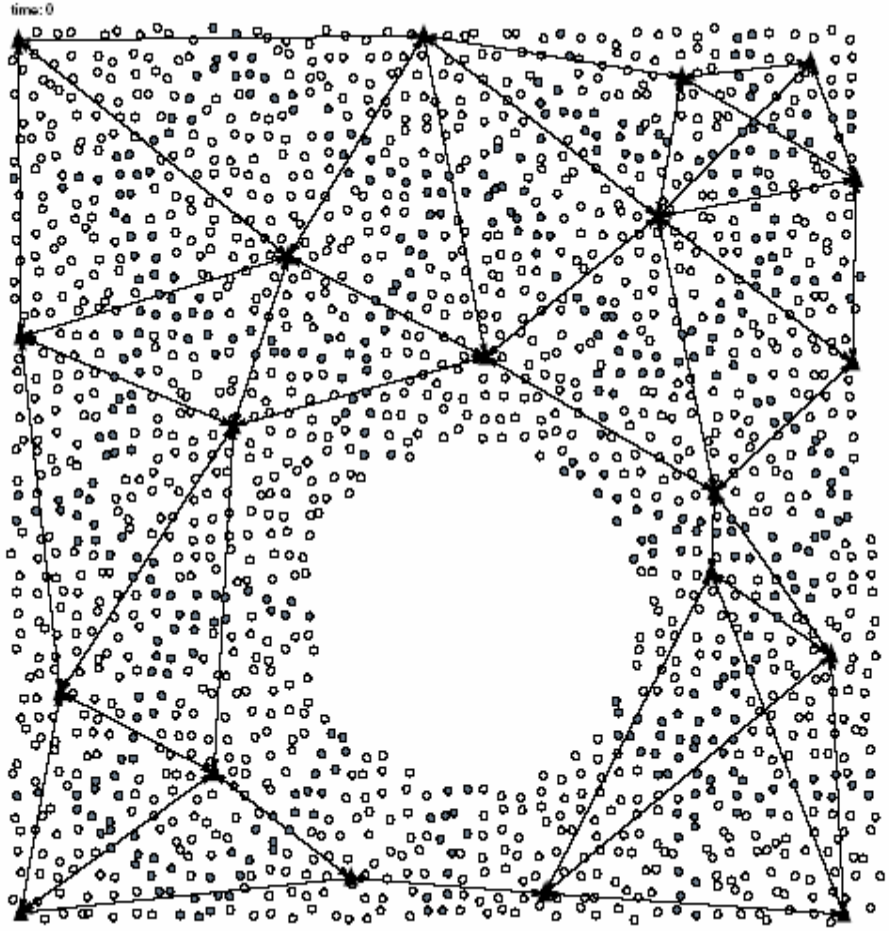
- The curves are functions of only the locations of producer and the consumer respectively.



Gradient Landmark-based Distributed Routing protocol (GLIDER)



The landmark Voronoi complex (LVC)

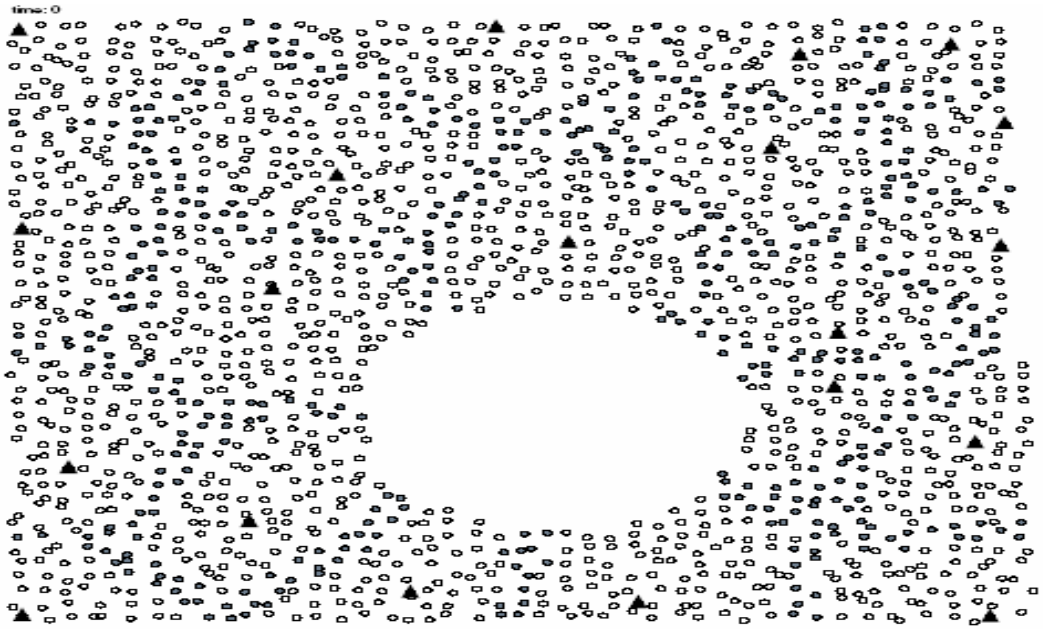


The combinatorial Delaunay Graph(CDG)

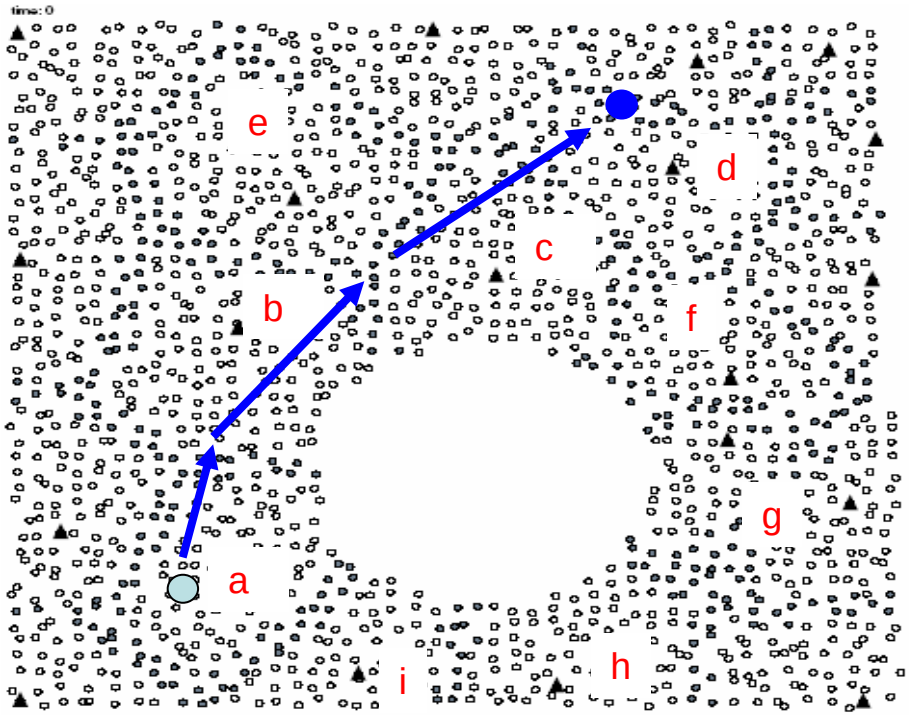
▲ :landmark

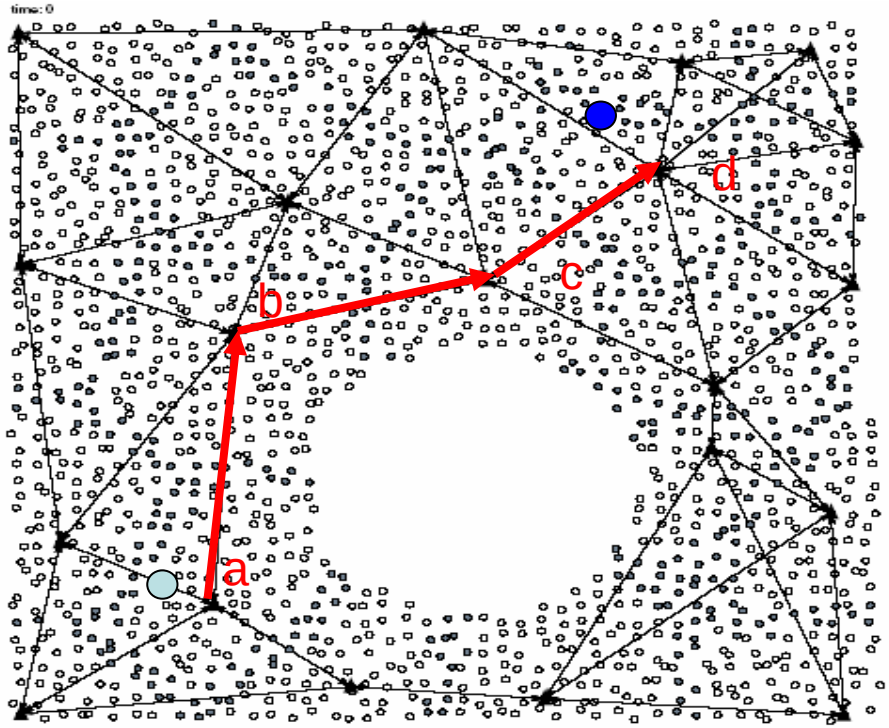
Gradient Landmark-based Distributed Routing protocol (GLIDER)

- Each nodes in sensor network knows the Combinatorial Delaunay Graph (CDG)
- Each node in sensor network knows how to transmit a packet to a node in the same Voronoi cell.
- The nodes in the same Voronoi cell know how to transmit a packet to the the landmark of the neighbor Voronoi cell.



Example



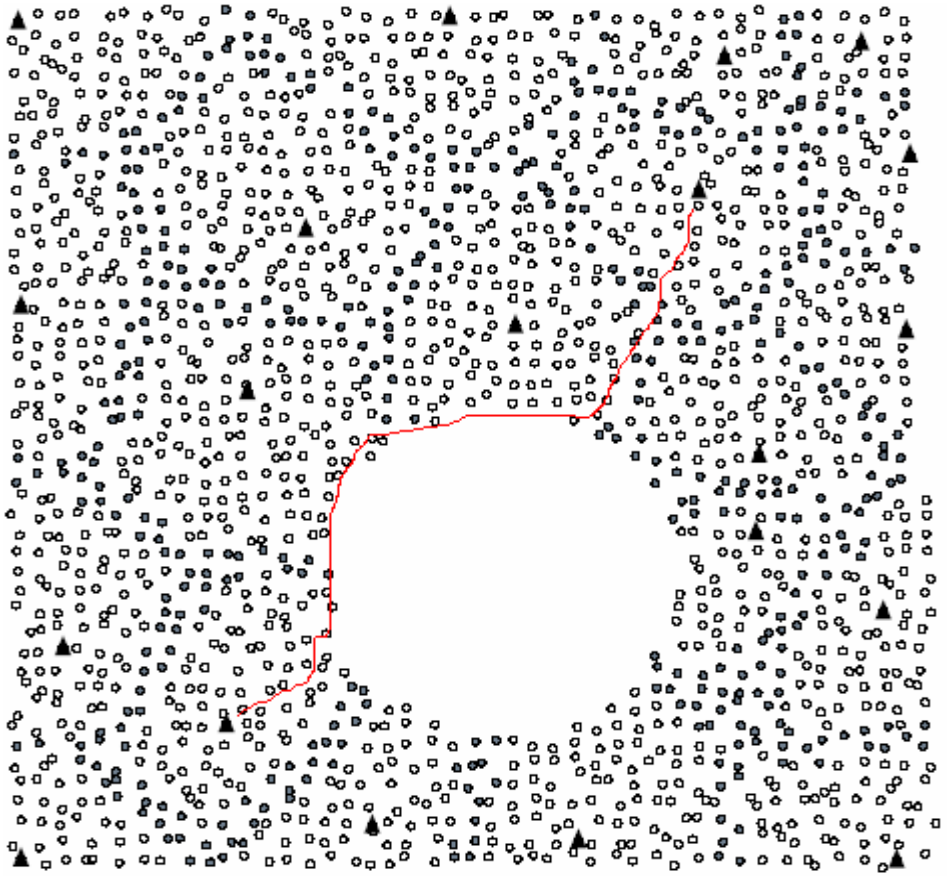


CDG

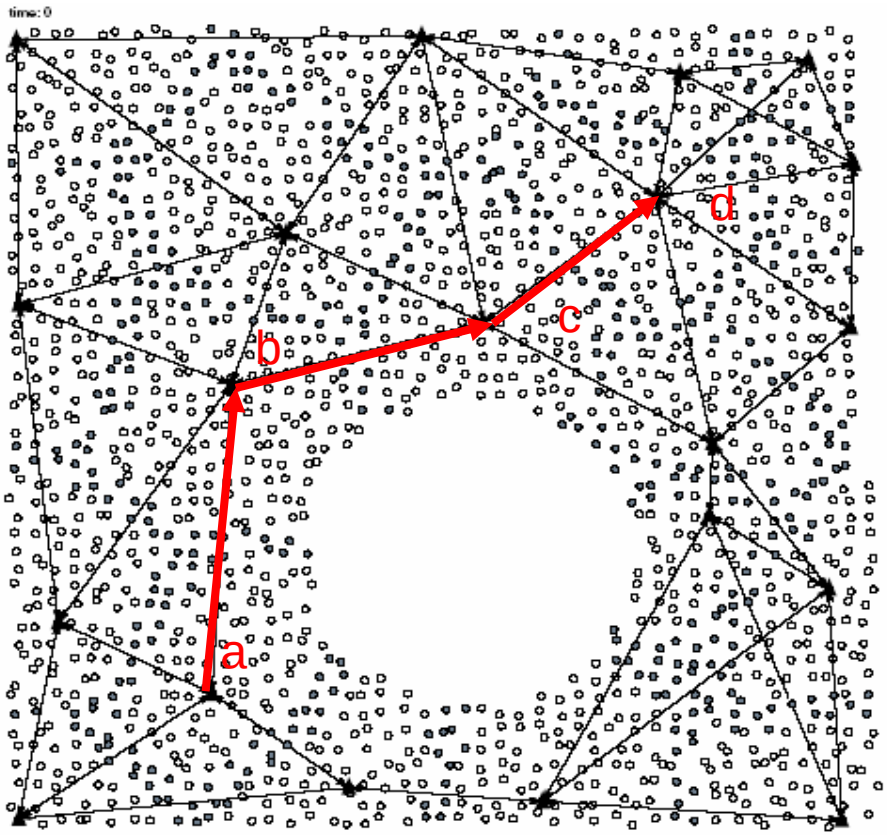
a → b → c → d

The Advantage of The GLIDER

- Load balancing is better
- without location information



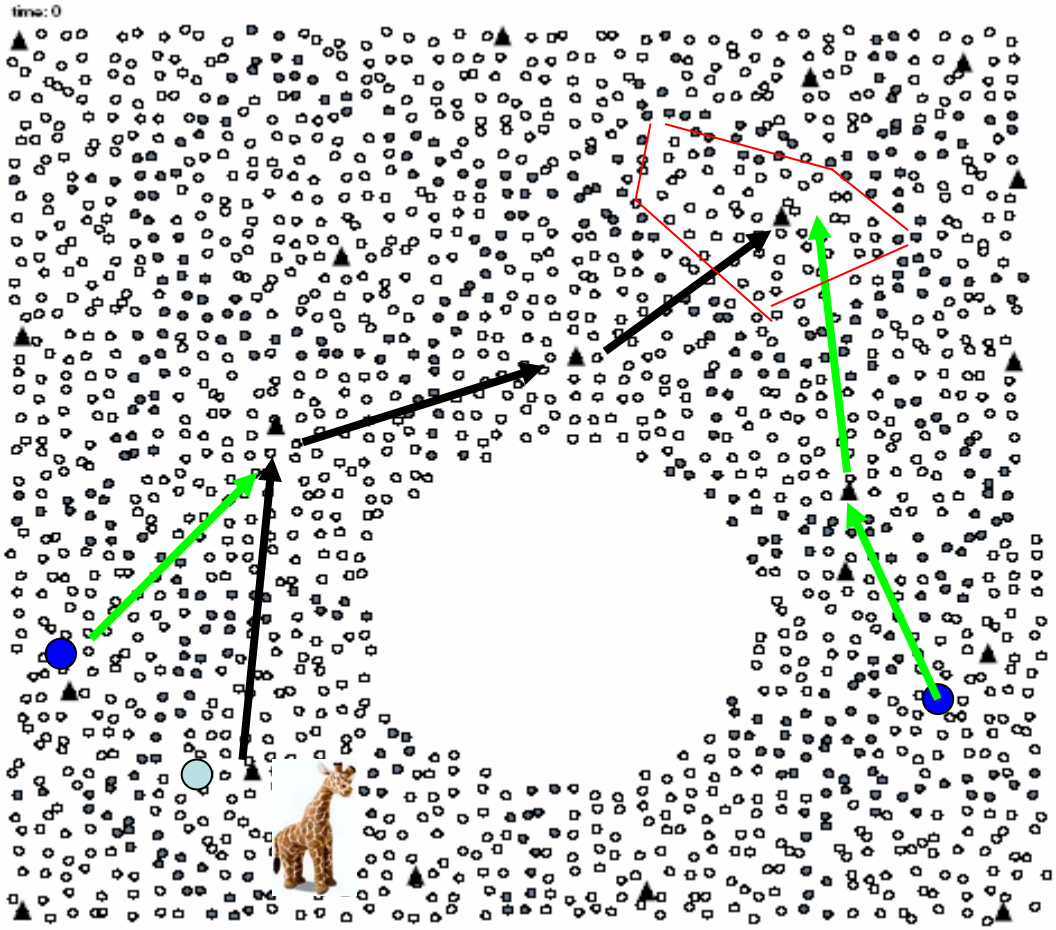
GPSR routing

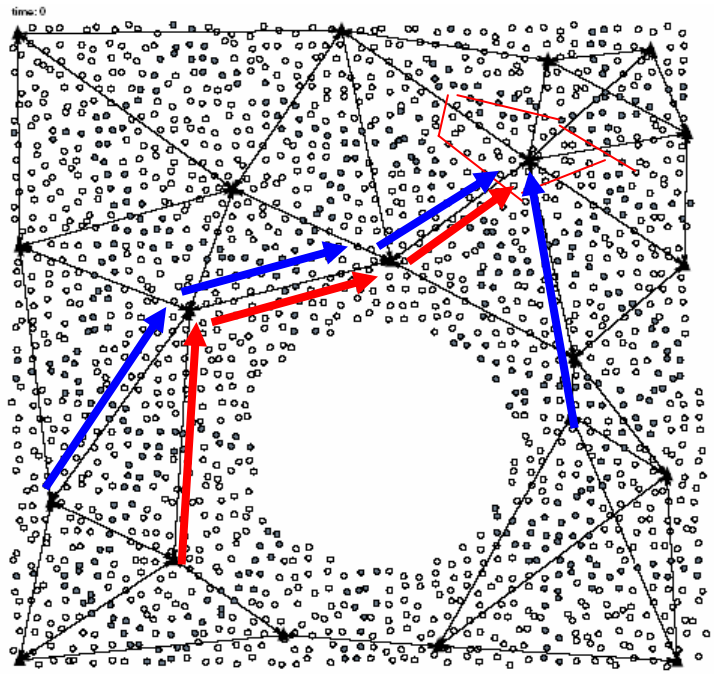


GLIDER routing

LANDMARK-BASED DATA-CENTRIC STORAGE AND RETRIVAL

The name of data —hash→ the landmark ID





CDG

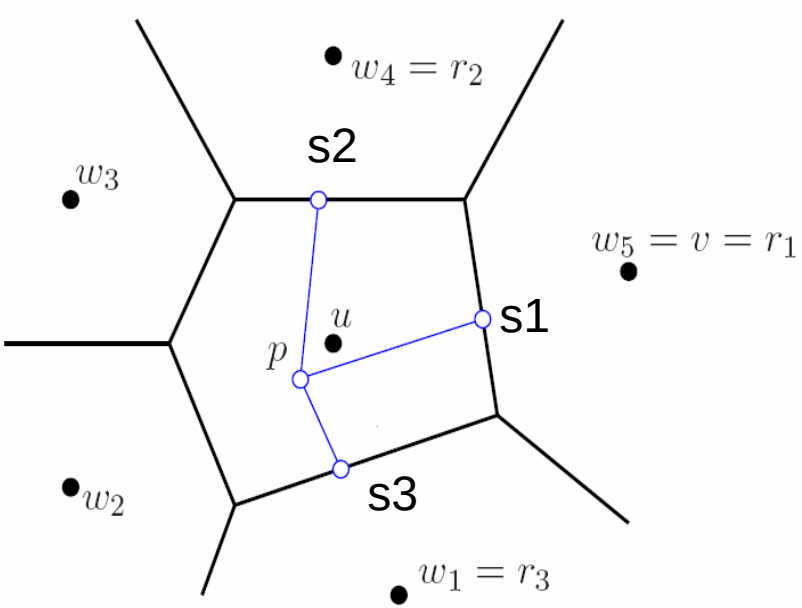
:producer

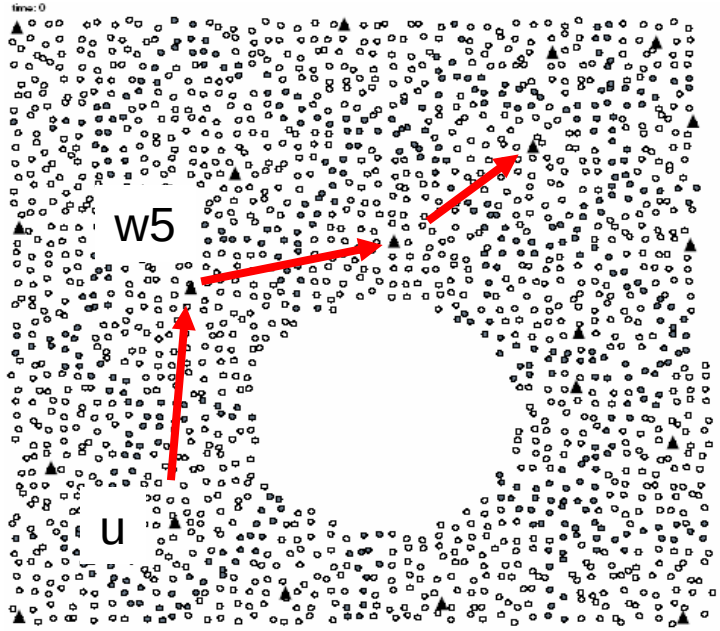
:consumer

Data Replication within a Voronoi Cell

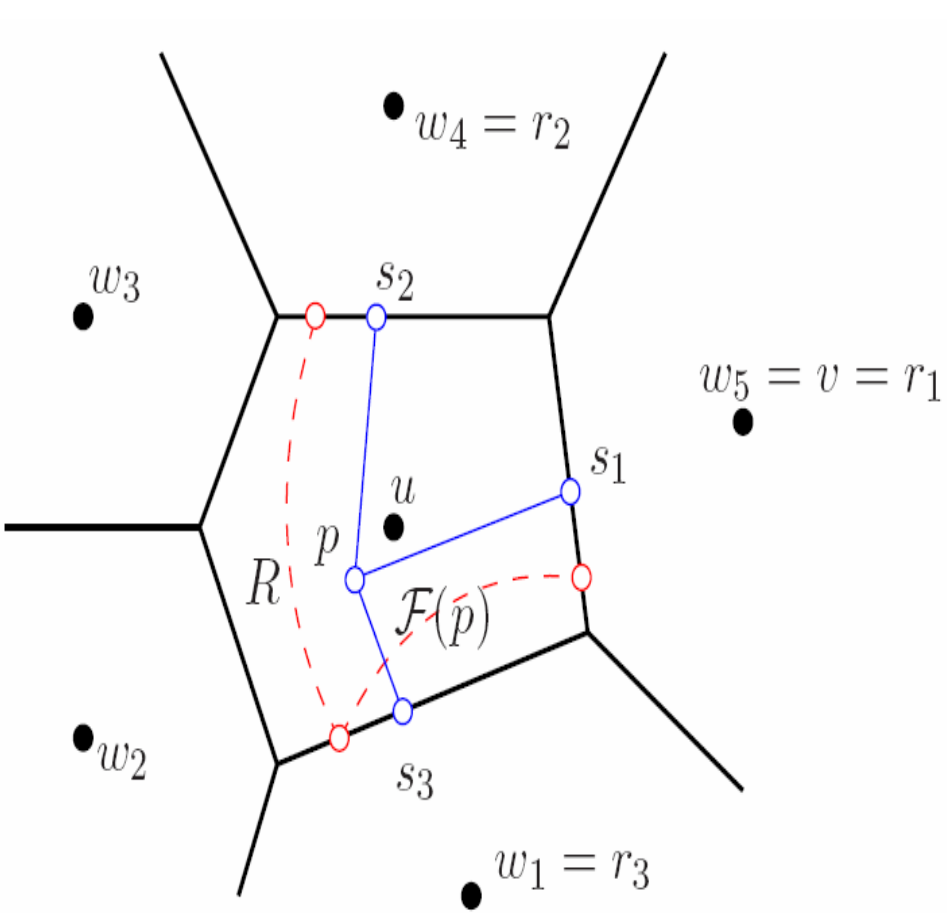
- Step1: select p, r1, r2, r3

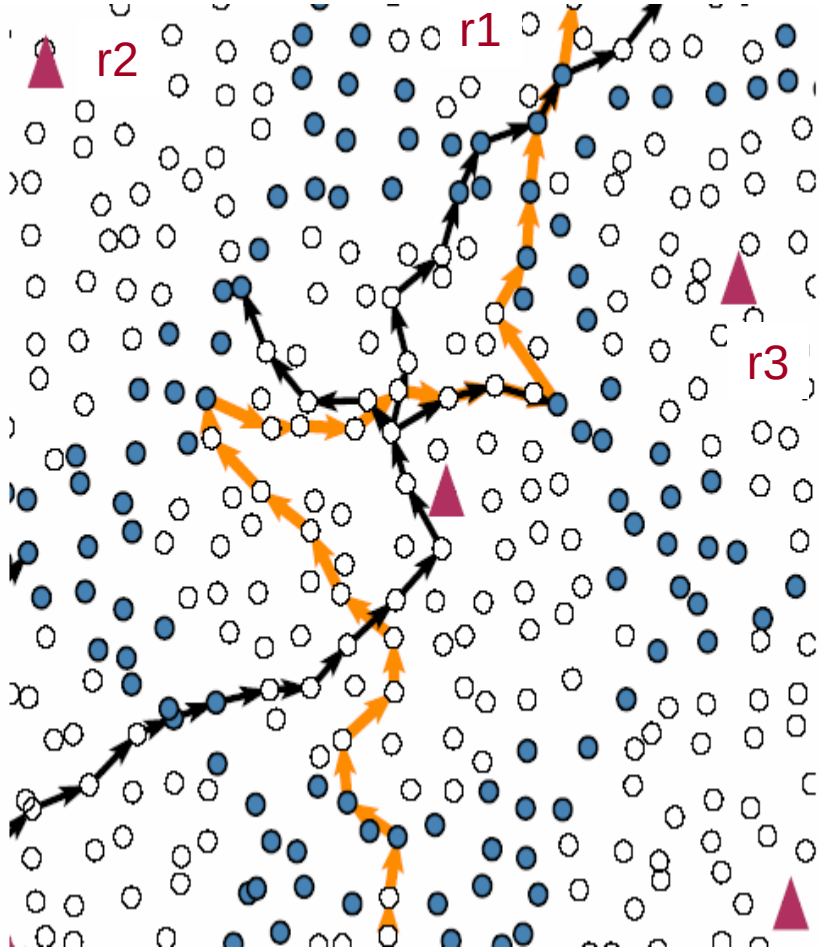
- Step2: replicate data along a path from p to each of the three guides r1,r2,r3





Data Retrieval within a Voronoi Cell





The Advantage of LANDMARK-BASED DATA-CENTRIC STORAGE AND RETRIVAL

- Nodes on the boundaries of holes are not used more often. (due to the GLIDER routing protocol)
- It allows the possibility of data aggregation.
- It is locality aware.
- It achieves the above goals without requiring node's geographic information as in GHT.

  (due to the GLIDER routing protocol)

Simulations

- The simulator is written in C++ and runs the distributed algorithm.

  - The simulated network has 2000 nodes distributed on a perturbed grid in a 316m by 316m field.
  - The radio range is 11 meters.
  - Each node can communicate with a set of neighbors nearby.
  - The average number of neighbors for each sensor is 6.2 .
  - Among 2000 nodes.23 are chosen as landmarks.

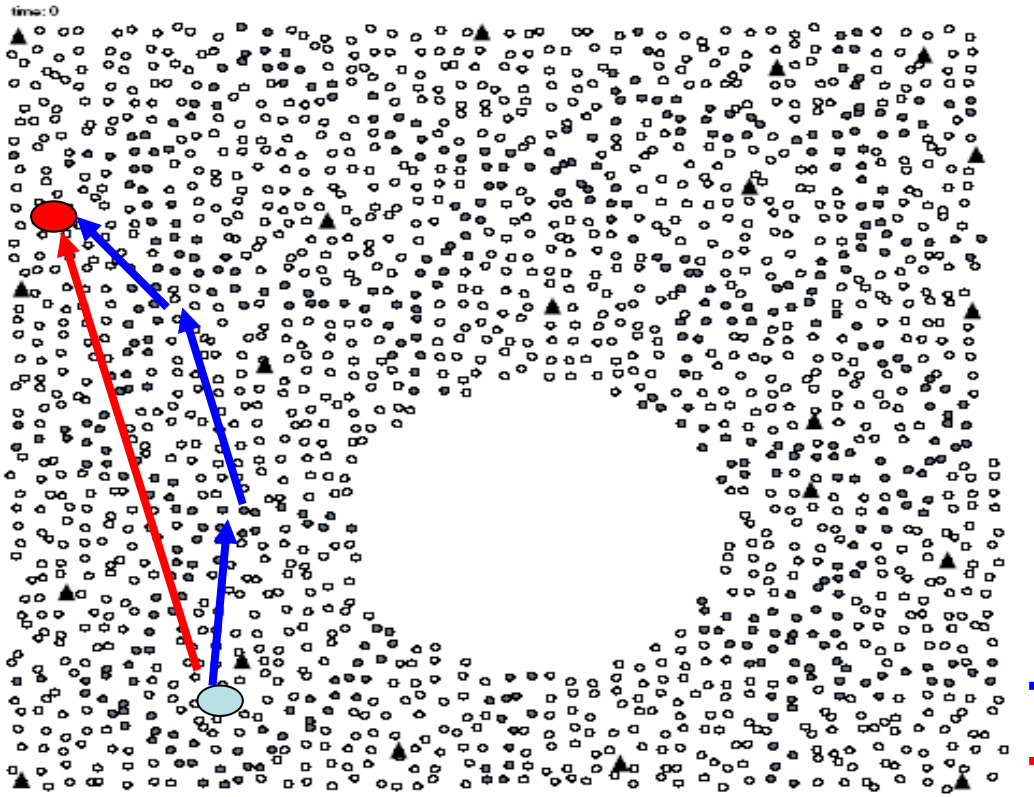
Simulations

- The cost for producer
- Locality awareness and the cost for consumers
- Load balancing

The Cost for Producer

number of producers per tile	1	2	3	4
cost ratio to GHT	1.63	1.24	1.04	0.98

tile: Voronoi cell



:producer

:GLIDER routing path

:GPSR routing path

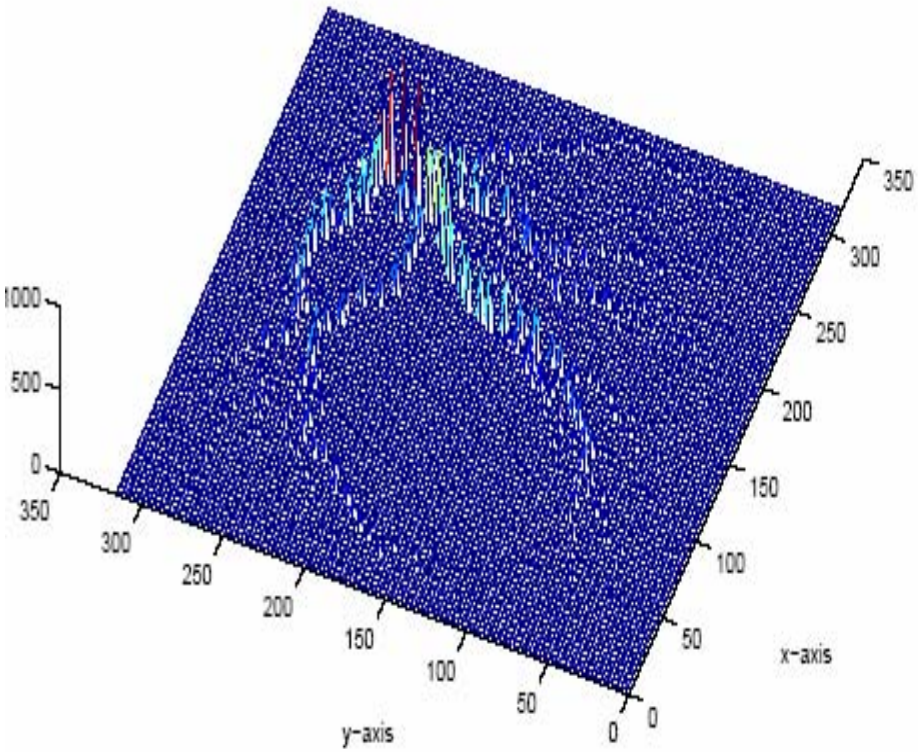
Locality awareness and the cost for consumers

disposition (hops)	≤ 4	8	12	16	20	24
cost ratio to GHT	0.17	0.16	0.44	0.47	0.91	0.95

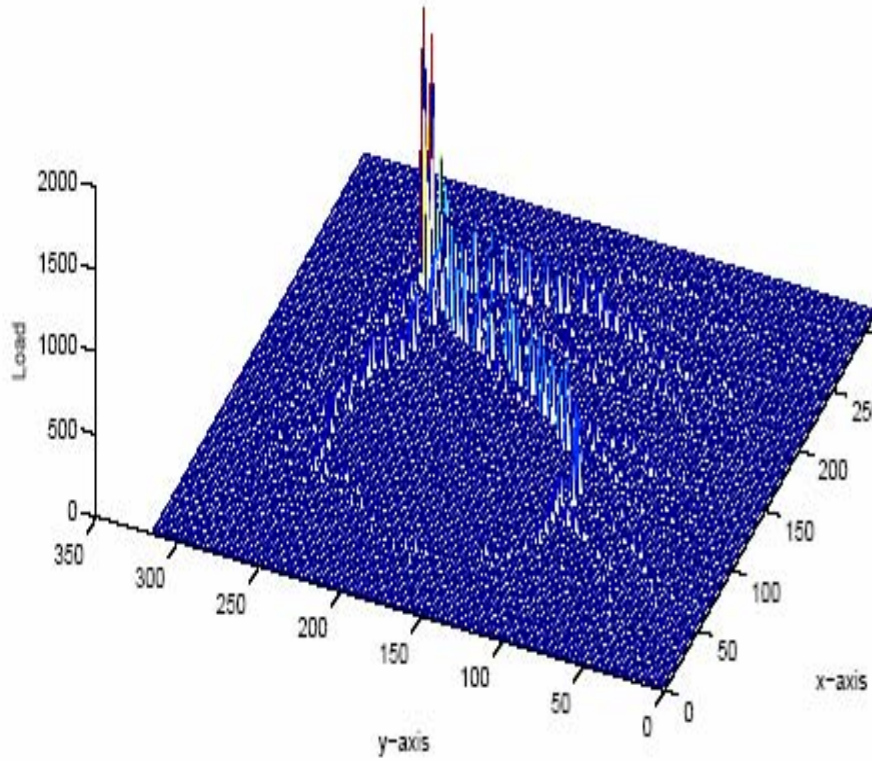
Disposition: the distance between the consumer and the producer

Load Balancing

Max load value for this graph is 793



Max load value for this graph is 1615

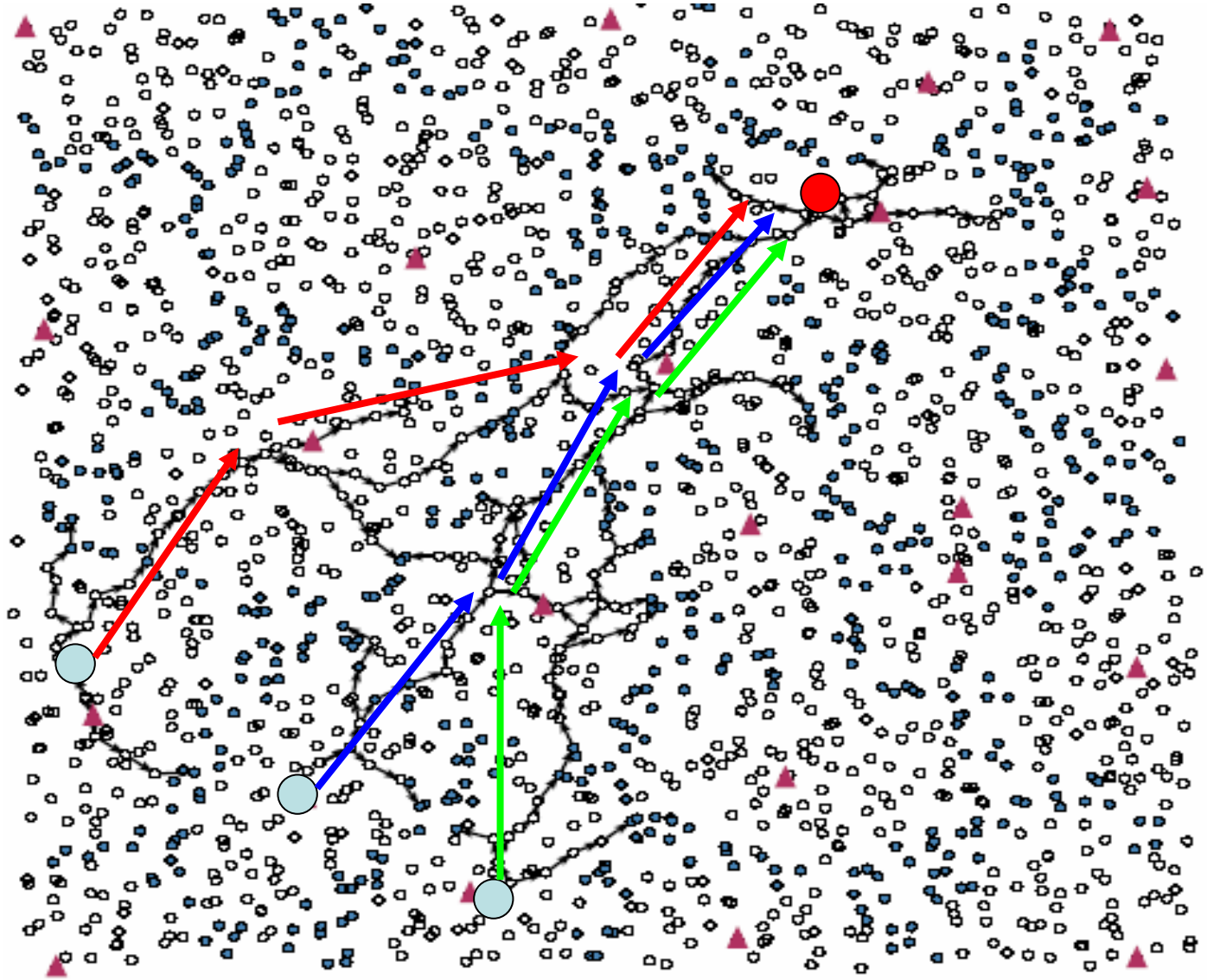


Landmark-Based information storage and retrieval

GHT

Conclusions

- The author proposed a data-centric, location-free, landmark-based information brokerage scheme for sensor networks.
- The efficiency of this scheme is fully manifested when there are multiple consumers and/or producers appearing throughout the network.
- Such advantages are attributed to the fact that we aggregate producer data before they reach the hashed tile when multiple producers of the same data are in close proximity of one another, as well as to the fact that consumers can potentially retrieve data on their way to the hashed tile.



:producer