ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

Enhancing Parallelism for K-Nearest Neighbor Query Processing using Customized Greedy Perimeter Stateless Routing in Wireless Sensor Networks

[1]Vijayaraja. V, [2]Prasanna Kumar. R, [3]Mukunthan. MA, [4]Bharathi Mohan. G
[1]Department of Computer Science and Engineering, Jaya Engineering College, Chennai, India
[2]Research Scholar, Anna University of Technology, Coimbatore, India

Abstract
In the Parallel itinerary based KNN query processing we have routing phase, KNN boundary estimation and spatial irregularity which uses the routing protocol called Greedy Perimeter Stateless Routing (GPSR) .This GPSR offers routing support for Wireless Sensor Network (WSN).However GPSR was designed for the symmetric links (bidirectional reachable), but sensor networks are often asymmetric in nature. GPSR suffers by energy inefficiency as it has to trace through all the nodes in the boundary for reaching the destination. In the prior work spatial queries are propagated in the sensor nodes and it is energy efficiency and query accuracy is determined by using the parallel itinerary based KNN query processing techniques. But the problem it is applicable for the symmetric link thereby to optimize the energy efficiency and query accuracy we propose the Customized Greedy Perimeter Stateless Routing (C-GPSR), modified version of GPSR is proposed which identifies optimal route based on energy utilization and overcome problems in GPSR so that the feasibility of using GPSR in asymmetric WSN can be increased thereby improve the overall energy efficiency in parallel itineraries'. The simulation using Network Simulator results proves that the energy and delay is minimized and hence the proposed protocol outperforms the existing routing protocol for WSN.

Keywords
K-Nearest neighbor query, Modified Greedy Perimeter Stateless Protocol (M-GPSR), Wireless Sensor Networks (WSN).

I. Introduction

A. Wireless Sensor Networks
Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications [3] and digital electronics have enabled the development of low-cost, low-power, multifunctional sensor nodes that are small in size and communicate untethered in short distances. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, leverage the idea of sensor networks based on collaborative effort of a large number of nodes. WSN nodes are battery powered, the routing protocol should consume less energy and also it should be ensured that the information transfer delay should be less. Due to the importance of geographical features in WSN applications, spatial queries [7] that aim at extracting sensed data from sensor nodes located in certain proximity of interested areas become an essential function in WSNs. Use of this WSN is implemented in the data mining concept which proved to be effective one in terms of energy consumption and query latency. Existing system deals with the use of GPSR [2] in the routing phase, KNN boundary estimation and spatial irregularity in order to route the query using this GPSR protocol which has some issues in the performance like GPSR is not applicable to asymmetric links thereby increase the energy consumption during the spatial irregularity region. In this paper we optimize the performance in terms of energy consumption and query latency we go for Customized-GPSR[2] where it can identifies optimal route based on energy utilization and overcome problems in GPSR so that the feasibility of using GPSR in asymmetric WSN can be increased.

B. Itinerary based KNN query processing in Parallel Concentric Itinerary based KNN (PCIKNN)
Itinerary-based query processing algorithm [1] based on optimized parallel concentric-circle itineraries, namely PCIKNN. We have three Phases 1) Routing phase 2) The KNN boundary estimation phase 3) The query dissemination phase. It is shown with example in Fig.1. Initially, a KNN query, issued at a source node, is routed to the sensor node nearest to the query point q (referred the home node) at the routing phase. Next, in the KNN boundary estimation phase, the home node estimates an initial KNN boundary (i.e., the solid boundary line circle in Fig. 1a), which is likely to contain K nearest sensor nodes from q. Finally, in the query dissemination phase (as shown in Fig. 1b), the home node propagates the query to each node within the estimated initial KNN boundary. While the KNN query propagates along certain well-designed itineraries, query results are collected at the same time.

C. GPSR in PCIKNN
GPSR protocol [13] is the earliest geographical routing protocols for adhoc networks which can also be used for WSN environment. The GPSR adapts a greedy forwarding strategy and perimeter forwarding strategy to route messages. It makes uses of a neighborhood beacon that sends a node's identity and its position. GPSR, PCIKNN [1] adopts two modes (i.e., the greedy mode and the perimeter mode) in the KNN query propagation. In the greedy mode, a message is forwarded by selecting the sensor node making the most progress toward the destination. When a void region is encountered, PICKNN switches to the perimeter mode. After bypassing voids, it changes back to the greedy mode and continues to move forward.

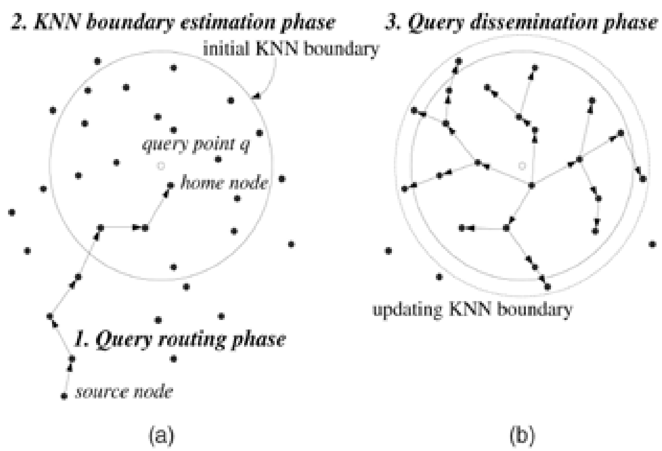

Fig.1: An overview of itinerary-based KNN query processing

D. Customized-GPSR

The proposed routing scheme is based on the fact that the energy consumed to send a message to a distant node is greater than the energy needed for a short range transmission. GPSR protocol is extended using aggregation node or head set node [2]. Aggregation node is responsible for transmitting messages to the distant home node and routing the query is decided using the respective head set members. The head set is decided on a routine basis with reference to the energy level of the signal received to the home node at the time of reception of result in the source node. This C-GPSR is used in the KNN boundary estimation when sensor nodes are dynamic and posse's asymmetric links.

II. Overview Of PCIKNN With Customized GPSR

The below diagram gives the overall idea of PCIKNN with Customized GPSR see fig 2.In the diagram mentioned below source node is the arbitrary node where we propagate the initial spatial query to it. And home node which estimates the boundary estimation after the itinerary propagation of query using C-GPSR in the routing phase.

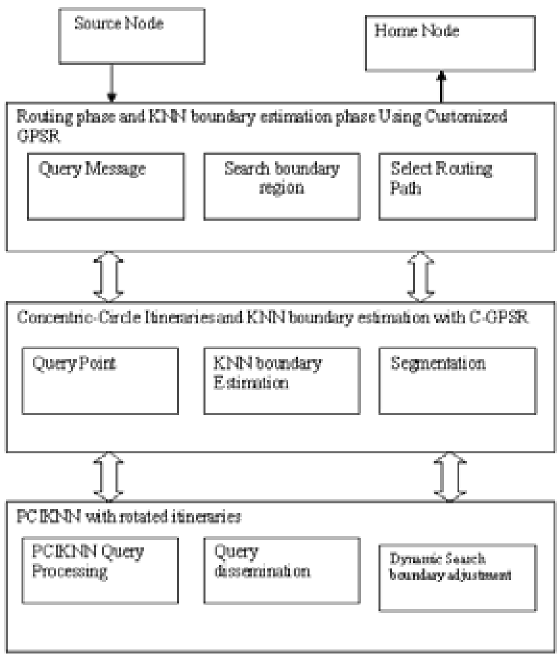

Fig.2: Architectural diagram of PCIKNN with C-GPSR

A. Routing phase and KNN boundary estimation phase using Customized GPSR

Routing is decided using the respective head set members in the sensor nodes where source node is called the arbitrary node [1,2]. The head set is decided on a routine basis with reference to the energy level of the signal received to the source node at the time of reception of spatial query. A query message with its Q is geographically routed from the sink node s to the nearest neighbor (i.e., the home node np, where p denotes the number of hops along the routing path) around the query point q. From the Fig.1 the information of the sensor network is gathered along with the routing procedure without the aid of any infrastructure. Upon receiving Q and the collected information from the previous phase, the home node estimates a searching boundary, named KNN boundary, with radius R by using an efficient (specifically, linear time) KNNB algorithm. The estimated boundary is not fixed and will be dynamically adjusted (by the other nodes) as long as additional information is available in the next phase.

B. Design of Concentric-Circle Itineraries

Given a query point q and an estimated KNN boundary [1], the area within the boundary can be divided into multiple concentric-circle itineraries. KNN query is first propagated along branch-segments in each sector. Along the branch-segment, a Q-node broadcasts a probe message and collects partial results from D-nodes within the region width of w

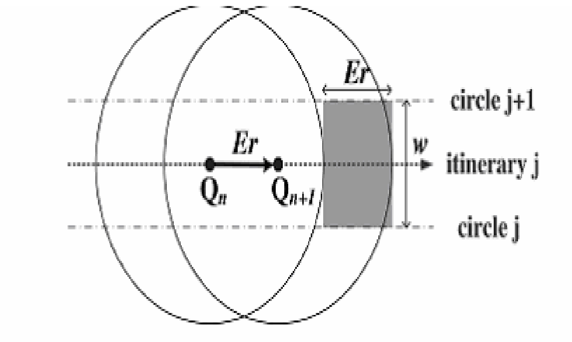

Fig.3: Design of Concentric circle.

For each sector as in fig 3, when the KNN query reaches one of the concentric-circles, two KNN query threads are forked to propagate along the two peri- segments, while the original KNN query continues to move along the branch-segment. To propagate a KNN query in two peri-segments, the Q-node in the branch segment first finds two Q-nodes in peri-segments and evenly divides the partial query result collected to these two Q-nodes.

C. Rotated Itinerary Structures of PCIKNN

Here a rotated itinerary structure in PCIKNN is dealt in order to known the boundary estimation. Without loss of generality, we consider an example, where the number of concentric-circles is 4 (i.e., C = 4). Fig.4, shows an example of a rotated itinerary structure in PCIKNN, where the bold line shows an itinerary structure within a sector. As shown in Fig. 3, for each concentric-circle, the branch-segment is rotated by θ degree. Assume that S is the number of sectors and C is the number of concentric-circles.

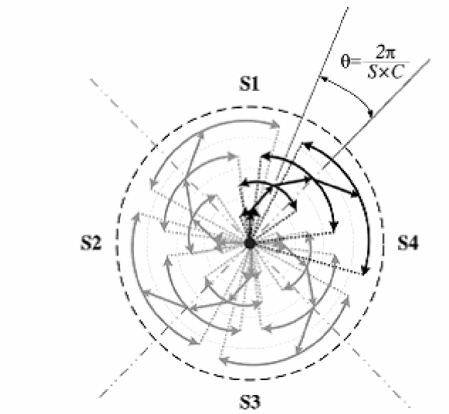

Fig.4: Parallel concentric itineraries in PCIKNN

The rotation angle θ is set to 2π/S × C such that the return-segments of all concentric-circles are evenly distributed within a KNN boundary. The dotted straight lines in Fig. are the return-segments. For each concentric-circle, sensor nodes along the return-segments are different, and thus, the energy consumption of sensor nodes along the return-segments is balanced. While the rotated itinerary of PCIKNN deals with the energy exhaustion problem of sensor nodes along the return-segments, the query latency is slightly increased because the total length of branch-segments in one sector is increased.

D. Dynamic KNN boundary estimation

Network information collected during the routing path of KNN query is used to derive the network density, which, in turn, is used to estimate a KNN boundary. In the fig 5,we describe how PCIKNN updates these two values during the routing phase. Message transmitting from node Ni to node N+1 The gray area is the newly explored area, denoted by EA i. The number of sensor nodes in EA i is denoted by inc i+1. By adding inc i+1 to Num, we have the updated number of nodes encountered so far.

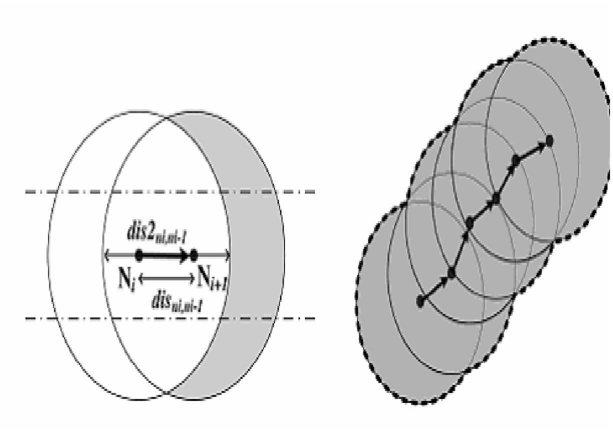

Fig. 5: Coverage areas estimated in the routing phase.

III. Performance Evaluation

A. Simulation Model

Our simulation is implemented in Network Simulator 2 [4,5]. There are 1,000 sensor nodes randomly distributed in a 500 _ 500 m2 region and the transmission range of a node is 40 m. For each sensor node, the average number of transmitting or receiving messages is 30 ms. In our default settings, sensor nodes are static. For each query, the location of a query point q is randomly selected. The default value of K for each KNN query is 100. A sensed datum is 4 bytes long and the query result is not aggregated. The broadcasting period of beacon messages is 3 s. A KNN query is considered as answered when the query result is returned to the source node. In each round of experiment, five queries are issued from randomly selected source nodes. Each experimental result is derived by obtaining average results from 10 rounds of experiments. For a fair comparison, we obtain the result of PCIKNN and compare with PCIKNN using Customized GPSR with the minimum latency/ energy consumption.

B. Experimental Results on KNN Boundary Estimation with C-GPSR

We evaluated the proposed KNN boundary estimation With C-GPSR, we set the value of K to be 350. To avoid the effect of network boundary, query points in the middle region (i.e., 100 m _ 100 m) are selected. The linear regression function of PCIKNN is set to H(dist(NiNi+1)) = (-76.9166* dist(NiNi+1) + 4999.0903) by linear regression technique in [10]. The optimal KNN boundary is the average distance of the Kth distant nodes of all queries derived by the experiments. As shown in Fig. 6, PCIKNN with customized GPSR is very close to the optimal KNN boundary under various network density. However, the boundary estimated in PCIKNN has little decrement in the optimal KNN boundary. This performance graph shows that query latency is decreased by using the PCIKNN with C-GPSR when compared to PCIKNN and thereby improve the efficiency in query latency.

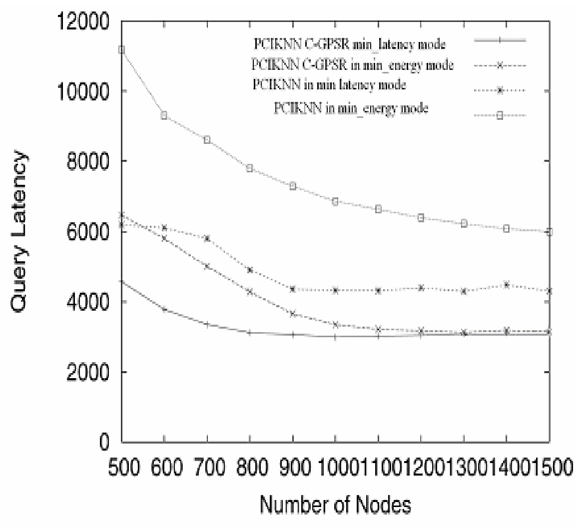

Fig. 6 : Query Latency performance graph.

C. Performance with respect to Energy Consumption

In this experiment, we evaluate the proposed PCIKNN with C-GPSR by varying the value of K from 50 to 400. As shown in Fig. 6, PCIKNN with rotated itineraries has a high query accuracy. Fig. 7 depicts the query latency under PCIKNN and PCIKNN with rotated itineraries. As shown, the latency of PCIKNN with rotated itineraries slightly increases because the length of branch segments is increased. It can be seen in Fig.6b that the total energy consumption of PCIKNN and PCIKNN with C-GPSR is almost the same. Thus, both PCIKNN and PCIKNN with C-GPSR consume a similar amount of energy since all sensor nodes within the same KNN boundary are involved. However, from the standard deviation of energy consumption PCIKNN with the C-GPSR has a more balance energy consumption distribution in both the minimum latency and the minimum energy mode.

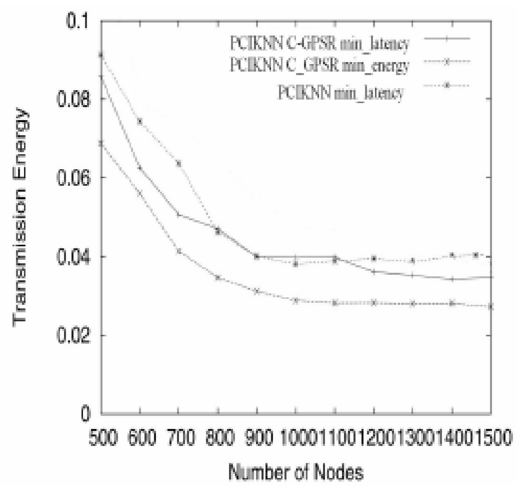

Fig. 7 : Energy consumption performance graph.

IV. Conclusion

In this paper, we proposed Customized GPSR for PCIKNN in a wireless sensor network. The Customized Greedy Perimeter Stateless Routing (C-GPSR), modified version of GPSR is proposed which identifies optimal route based on energy utilization and overcome problems in GPSR so that the feasibility of using GPSR in asymmetric WSN can be increased thereby improve the overall energy efficiency in parallel itineraries'. The customized GPSR routing protocol strives to address the unique requirements for sensor network applications. It provides an energy-efficient routing protocol with the ability to route messages from node to node and guarantees the routing of query under situations where non-uniform transmission ranges exist in the spatial irregularity regions. The simulation using Network Simulator results proves that the energy and delay is minimized and hence the proposed protocol outperforms the existing routing protocol for WSN.

References

[1] Tao-Yang Fu, Wen-Chih Peng, Wang-Chien Lee," Parallelizing Itinerary-Based KNN Query Processing in Wireless Sensor Networks." IEEE Transactions on knowledge and Data engineering, vol. 22, pg 711-729, may 2010.

[2] L. Nithyanandan, G. Sivarajesh, P. Dananjayan, "Modified GPSR Protocol for Wireless Sensor Networks." International Journal of Computer and Electrical Engineering, Vol. 2, No. 2, April, 2010.

[3] F.L.Lewis,"Wireless Sensor Networks."Technologies Protocol and Applications ed.D.J.Cook and S.K.Das , John Wiley,New York ,2004.

[4] Study of Network Simulator-2, S-38.3148 Simulation of data networks / fall-07.

[5] K. Waclena, "Lists and keyed lists." [Online]. Available: http://www2.lib.uchicago.edu/keith/tcl-course/topics/lists.html

[6] "Ns-2 trace formats." [Online]. Available: http://nsnam.isi.edu/nsnam/index.php/NS-2 Trace Format

[7] M. Demirbas, H. Ferhatosmanoglu, "Peer-to-Peer Spatial Queries in Sensor Networks," Proc. Third Int'l Conf. Peer-to-Peer Computing, pp. 32-39, 2003.

[8] D. Estrin, R. Govindan, J. Heidemann, "Next Century Challenges: Scalable Coordination in Sensor Networks," Proc.ACM/IEEE MobiCom, pp. 263-270, 1999.

[9] H. Ferhatosmanoglu, E. Tuncel, D. Agrawal, A.E. Abbadi,"Approximate Nearest Neighbor Searching in Multimedia Databases,"Proc. 17th IEEE Int'l Conf. Data Eng. (ICDE), pp. 503-511,2001.

[10] D. Goldin, M. Song, A. Kutlu, H. Gao, H. Dave, "Georouting and Delta-Gathering: Efficient Data Propagation Techniques for Geosensor Networks," Proc. NSF Worshop GeoSensor Networks,2003.

[11] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan,"Energy-Efficient Communication Protocol for Wireless Microsensor Networks," Proc. 33rd Hawaii Int'l Conf. System Sciences,pp. 8020-8029, 2000.

[12] G.R. Hjaltason, H. Samet, "Distance Browsing in Spatial Databases," ACM Trans. Database Systems, vol. 24, no. 2, pp. 265-318, 1999.

[13] B. Karp, T.H. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," Proc. ACM/IEEE MobiCom, pp. 243-254, 2000.

[14] Y.B. Ko, N.H. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks," Wireless Networks, vol. 6, no. 4,pp. 307-321, 2000.

[15] S.J. Leon, Linear Algebra with Applications. Prentice Hall, 2002.

V.Vijayaraja is working as an Assistant Professor, Department of Computer Science, Jaya Engineering College, Tamilnadu, India. He received his B.E. Degree in Electronics and Instrumentation Engineering from Annamalai University and M. Tech Degree in Computer science and Engineering from Dr. M. G. R University. He has about 14 years of teaching experience and 2 years of research experience in the field of Wireless sensor networks. He is the life member of I.S.T.E.

R.Prasanna Kumar is a Research Scholar, Department of Computer Science, Anna University of Technology, Coimbatore, Tamilnadu, India. He received his B.E. Degree in Computer science and Engineering from Madras University and M. Tech Degree in Computer science and Engineering from Dr. M. G. R University. He has about 8 years of teaching experience and 2 years of research experience in the field of Data Mining.

M.A Mukunthan is working as an Assistant Professor, Department of Computer Science, Jaya Engineering College, Tamilnadu, India. He received his B.E. Degree in Computer science and Engineering from Madras University and M.E Degree in Computer science and Engineering from Anna University. He has about 8 years of teaching experience and 2 years Industrial experience.

`

G.Bharathi Mohan is a Research Scholar, Department of Computer Science, Anna University of Technology, Chennai, Tamilnadu, India. He received his B.E. Degree in Electrical and Electronics Engineering from Madras University and M. Tech Degree in Information Technology from Anna University. He has about 5 years of teaching experience.