Traffic Load Distribution in Large-Scale and Dense Wireless Sensor Networks

Qinghua Wang, Tingting Zhang
Department of Information Technology and Media
Mid Sweden University, SE-85170 Sundsvall, Sweden
Email: {qinghua.wang, tingting.zhang}@miun.se

Abstract—Because of the dominating many-to-one communication pattern, traffic load is not evenly distributed over the nodes in a wireless sensor network (WSN). In a large-scale WSN, the multi-hop routing even deteriorates this traffic load imbalance by accumulating traffic flows along their paths to the sink. Understanding the traffic load distribution can guide the network-wide energy allocation, direct the design of routing algorithms, and optimize the node deployment in WSNs. In this paper, we consider a large-scale and dense WSN with nodes evenly deployed in a disk area, and find the traffic load distribution over the nodes as a function of their distances from the sink. Further, the effects of network scale and routing hop length on traffic load distribution are also investigated. The traffic loads of individual nodes are found to be in direct proportion to the radius of the network and in inverse proportion to the mean routing hop length, while independent of network density. The results presented in this paper are verified through extensive simulation experiments.

Index Terms—Traffic load distribution, wireless sensor network.

I. INTRODUCTION

In large-scale wireless sensor networks (WSNs), the dominating communication pattern is a large number of sensor nodes deliver their sensed information to one or a few data sinks through multi-hop transmission [1], [2]. This kind of communication pattern drastically imbalance the traffic load distribution across the network where the nodes close to a sink experience heavy traffic loads. Since communication is believed to dominate the energy consumption of a sensor node [3] and sensor nodes are usually provided with limited energy resources, the imbalanced traffic load distribution is very harmful and the nodes close to a sink could die early thus renders the rest of the network to be useless.

To counter or alleviate the harm resulting from the imbalanced traffic load distribution, new methodologies of data dissemination and resource allocation for WSNs need to be developed. It is clear that we still need more understanding about the traffic load distribution in WSNs for the future development of an improved WSN solution. In this paper, we consider those large-scale and dense WSNs. The need for a large-scale WSN is clear. For example, wide-range monitoring basically requires large-scale WSNs. A large-scale WSN considered in this paper should have a network radius which is at least longer than several routing hops. As to the need for a spatially dense WSN, ref. [4] has stated that it is driven by at least two requirements: i) reliable decision-making in the face of unreliable individual components; and ii) superior spatial localization of transient events of interest. The large-scale and dense WSNs have a few nice features which include: i) the routing hop length is negligible compared to the radius of the network; ii) the path of packet forwarding can be viewed as a straight line. Benefiting from these special features, we formulate the expression of the traffic load over the sensor nodes as a function of their distance from the sink for a large-scale and dense WSN deployed in a disk area. Further, the influence of the network scale and the routing hop length are analyzed, and the results are verified through simulations. What is a bit surprising is that the analysis results for the considered large-scale and dense WSNs also coincide with the simulation results for normal-scale and low density networks. This makes the results presented in this paper have more practical significance.

In the following of this paper, Section II introduces the related work in the area of sensor network traffic load analysis. Section III then presents the large-scale and dense sensor network scenario considered in this paper, together with the assumptions. Section IV presents our analytic results step by step. The analytic results are verified by extensive simulation experiments and the experiment part is presented in Section V. Finally, Section VI concludes this paper.

II. RELATED WORK

The special traffic load distribution in multi-hop networks including WSNs has been paid attention to by many researchers. The authors of [5]–[7] realize that the imbalanced traffic load distribution can render one part of nodes die earlier than others, thus degrades the network performance. To counter the negative effect of the imbalanced traffic distribution on network performance, new routing algorithms which refer the remaining energy reserves and other kinds of path capacity measurements are proposed. The authors of [8] consider the load balancing problem of uniformly distributed traffic demands in a unit disk. By deliberately routing traffic along slightly longer paths instead of shortest paths, the highly congested links are avoided and a particularly flat traffic load distribution is achieved. The authors of [9] address the problem of balancing the traffic load in multi-hop wireless networks with uniformly distributed point-to-point communication. They develop a routing algorithm called

Curveball Routing which can avoid the crowed center and have a performance not much worse than the optimum.

Obviously, formalizing the traffic distribution in WSNs can give researchers more insights in improving future WSNs, including developing future routing algorithms. Noori and Ardakani [10] analyze the traffic distribution in a 1-dimensional linear WSN and explicitly give the traffic load expression for sensor nodes at different locations in the linear WSN. They observe that the traffic load does not increase monotonically as the node distance from the sink reduces, if the shortest path routing is adopted and the network is not dense. Using their analysis, one can find the position of the nodes which carry the highest level of traffic and hence form the network bottleneck.

This paper studies the traffic load distribution in large-scale and dense planar WSNs. Unlike [10] where the shortest path routing is adopted, we however use a hypothetical routing algorithm based on a weak routing assumption, i.e. the number of relay nodes traversed by a packet is linearly proportional to the distance between the source and the destination. Further, we find that the traffic loads at individual nodes are mainly dependent on their relative distances to the sink, while independent of the underlying routing strategy (differentiated by the mean routing hop length) and the network scale. This result is particularly useful when designing general-applicable efficient WSNs.

III. WSN SCENARIO AND ASSUMPTIONS

We consider a WSN of n sensor nodes evenly deployed in a disk area of radius R. There is a sink node located in the center of the deployed disk area. All the deployed sensor nodes regularly sense the physical phenomenon around and forward that information to the sink directly or by multi-hop transmission.

We assume the considered network is large-scale and dense [1]. The large-scale assumption means that the routing hop length is negligible compared to the diameter of the network. The denseness assumption means the forwarding node can always find the next relay node along the line segment connecting the forwarding node and the sink, which is also the final destination for all data messages.

A routing assumption is also imposed: the number of routing hops during the transmission of a packet is linearly proportional to the distance between the source and the destination. This routing assumption is realistic in the sense that it is consistent with many well-accepted routing algorithms and protocols. For example, the shortest path routing algorithm looks for a routing path where all hops have a length close to the maximum radio range. A routing algorithm with geographic knowledge can easily build a routing path where all hops equal to or are close to a certain optimal length [11]. Actually, many research works [11], [12] have pointed out that to minimize the path cost, the optimal routing policy should

choose relay nodes on an equidistant basis along the line that connects the source and the destination, and there exists an optimal hop length which is decided by the link cost model. With regard to the popular reactive routing protocols such as AODV [13], it has been found in [14] that except for the first hop, the average number of hops grows linearly as the distance between the source and the destination increases in one demensional MANETs.

In this paper, we suppose the proposed routing assumption is a good model for at least the above mentioned routing strategies and consider a generalized routing algorithm. This generalized routing algorithm selects the next relay for each transmitting node using two steps: 1. The source node identifies its ideal relay point as the point which lies on the line segment connecting the source node and the sink, and which is also closer to the sink with one hop distance h; 2. The source node selects the deployed node which is within its maximum radio range and which is also the closest one to its ideal relay point as its next hop relay. In step 1, the source node will directly forward its packet to the sink if its distance from the sink is less than one hop distance h. In step 2, there will be no relay if the sink is the closest deployed node to the identified ideal relay point. Obviously, we have $h \ll R$ given the large-scale assumption.

Given the randomness of event occurrence in the sensing area, we can reasonably assume that all sensor nodes have the same average sensing rate (i.e. the amount of packets originated during a unit time period), and this average sensing rate is represented by $\bar{o}$ in the following.

In this paper, we are only going to investigate the "stable-state" traffic load distribution when no node fails because of the energy constraint. Thus, the energy constraint will not be considered in the following.

IV. TRAFFIC LOAD ANALYSIS

In this section, we find the traffic load distribution over the nodes as a function of their distance from the sink.

We define the traffic load at a given node as the total amount of packets handled on that node during a unit time period. Obviously, the total amount of packets handled at a node includes the amount of packets originated at that node and the amount of packets relayed by that node. For any node, $\bar{o}$ should be the expected amount of packets originated at that node during a unit time period, and this conclusion is straightforward. However, to find out the amount of packets relayed by a given node, one has to firstly find out how many child nodes the given node has. As for the definition of child node, we say node b is a child node of node a or node a is a parent node of node b if node b needs node a to act as a relay when forwarding its packets to the sink.

Consider node S shown in Fig. 1. Node S has a distance of r_S from the sink. Obviously, node S has to forward all the traffic directed to it as well as those directed to its occupying proximity region. We say a node is occupying its proximity region if all traffic directed to this region is handled thus forwarded by the node in question. According to the
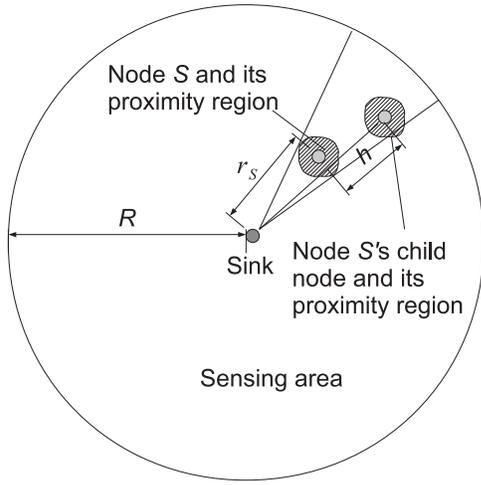
Fig. 1. Sketch map for a sensor network deployed in a disk area. The definition of a node's child node is shown.
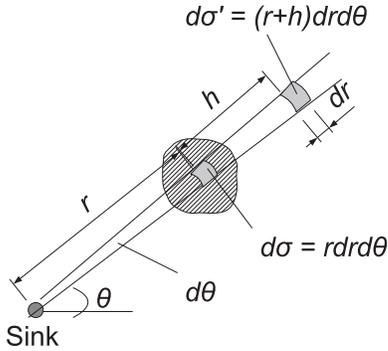

Fig. 2. An infinitesimal region in a node's proximity region.

routing algorithm prsented in Section III, a node is the closest deployed node to any point in its proximity region. In this article, we consider that there is no obscure region, i.e. all the points in the sensing area must belong to some node's proximity region, but they cannot be simultaneously within multiple nodes' proximity regions.

With the concept of proximity region, we can further define the occupied region by a node's i-th hop child nodes. Since the considered network is symmetric and dense, the most efficient routing strategy is forwarding packets along the straight line determined by the source node and the sink. Thus the line segment connecting a source node and the sink must pass through the proximity regions of this source node's parent nodes (e.g. node S is the parent node of another node shown in Fig. 1). With specialty, a midpoint on this line segment and is i hops away from the source node must be within the proximity region of the source node's i-th hop parent node.

Let S_S be the proximity region of node S in Fig. 1. For each infinitesimal $d\sigma$ in S_S which is defined by $d\sigma = r\, dr\, d\theta$ in the polar coordinates (see Fig. 2), all the points which lie on the straight line connecting a point in $d\sigma$ and the sink and also i hops farther away together form a region $d\sigma'$ which is defined by $d\sigma' = (r + i\,h)\, dr\, d\theta$. Similar to that

the summation of $d\sigma$ forms S_S, the summation of $d\sigma'$ forms the occupied region by node S's i-th hop child nodes. Thus, the area of the occupied region by node S's i-th hop child nodes is $\sum d\sigma' = \iint_{S_S}(r + i\,h)\, dr\, d\theta$. Recall that the network is dense, we have the approximation $\iint_{S_S}(r + i\,h)\, dr\, d\theta \approx (r_S + i\,h)\iint_{S_S} dr\, d\theta$, where r_S is node S's distance from the sink. In an even-distributed network, each node occupies an equal proximity region, i.e. $r_S \iint_{S_S} dr\, d\theta$ is the area of any node's proximity region. Together with the area expression of the occupied region by node S's i-th hop child nodes, we can estimate the number of node S's i-th hop child nodes and that is $1 + i\,h/r_S$.

Suppose the m-th hop child nodes are the farthest child nodes of node S, the total number of child nodes for node S would be $\sum_{i=1}^{m}(1 + i\,h/r_S)$. Since each child node and node S itself periodically originate packets at the average rate $\bar{o}$ packets per time unit, the expected traffic load on node S during a unit time period can be expressed as:

$$Traffic(r_S) = \sum_{i=0}^{m}(1 + i\,\frac{h}{r_S})\bar{o} \qquad (1)$$

Note that node S is only a representation of those nodes deployed in the sensing area. We can generalize the traffic load expression for any node which has a distance r from the sink, and that is:

$$Traffic(r) = (m + 1)(1 + \frac{m\,h}{2r})\bar{o} \qquad (2)$$

Given that the deployed area has a finite radius R, we have $m = \lfloor \frac{R-r}{h} \rfloor$ in the above formula.

(2) applies to all deployed sensor nodes except when a node is too close to the sink. A realistic application scenario does not have an infinite high node density. Thus we could not simply use a node's distance from the sink to represent the distance from the sink for other points in this node's proximity region, when the node is too close to the sink. Further, not all traffic directed to a node's proximity region is relayed by the node of interest if that node is too close to the sink. In the Appendix A, we explain the possible reduction in a node's observed traffic load when the node gets fairly close to the sink. Normally, we can set the lower limit of (2)'s supporting region to be $\frac{2R}{\sqrt{n}}$ based on the observation in the Appendix A and the simulation results in Section V. Finally, the average number of nodes whose traffic load doesn't follow (2) is about 4.

We can see that the number of deployed nodes n is not present in (2). This contradicts to the intuition that increasing the number of deployed nodes may simultaneously aggravate the traffic load situation at each single node. Actually, the only factor that influences the traffic load is the node's distance from the sink, when the routing and the network radius are fixed.

But the network radius R and the routing hop length h do have influences on the traffic load, even though they can be considered as constants at most of the time. These influences

are not clearly expressed by (2). In the following, we derive an approximated traffic load expression, from which the relations between the network radius R, the routing hop length h and the traffic load can be clearly observed.

Let $m = \frac{R-r}{h} - \epsilon$, $\epsilon \in [0, 1)$ in (2). We have

$$
\begin{aligned}
Traffic(r) &= (\frac{R-r}{h} - \epsilon + 1)(1 + \frac{R-r-\epsilon h}{2r})\bar{o} \\
&= \frac{R-r+(1-\epsilon)h}{h} \frac{R+r-\epsilon h}{2r}\bar{o} \\
&= \frac{(R^2 - r^2) + (R + r - 2R\epsilon)h - \epsilon(1-\epsilon)h^2}{2r\,h}\bar{o}
\end{aligned}
$$

Let $r' = r/R$ as a node's unified distance from the sink. We further have

$$
Traffic(r') = \frac{(1 - r'^2) + (1 + r' - 2\epsilon)\frac{h}{R} - \epsilon(1-\epsilon)(\frac{h}{R})^2}{2r'\frac{h}{R}}\bar{o}
$$

Use the knowledge that $h \ll R$ in a large scale network, and we arrive at

$$
Traffic(r') \approx \frac{1 - r'^2}{2r'}\frac{R}{h}\bar{o} \qquad (3)
$$

(3) clearly states that the traffic loads on individual sensor nodes are in direct proportion to the radius R of the network, and in inverse proportion to the routing hop length h. For networks where the network radius and the routing hop length are fixed, the traffic load only varies with the node's distance from the sink. We can even see from (3) that the traffic load is a monotonically decreasing convex function for $r' \in (0, 1]$. This may seem to contradict to the observation in [10] where a traffic load reduction can be observed when a sensor gets closer to the sink. Remembering that we consider a dense network scenario in this paper and (3) also doesn't apply to the scenario when a node gets too close to the sink (see the appendix), the result presented here doesn't violate those acquired in [10].

One more observation that can be directly seen from (3) is that the ratio of the traffic loads at any two deployed nodes is independent of the routing hop length, and also independent of the network scale in a unified symmetric WSN. This observation can be used to guide the energy allocation and node deployment for general-applicable WSNs.

V. SIMULATION RESULTS

We firstly simulate a large-scale and dense sensor network where 10,000 sensor nodes are gridly deployed in a disk area of radius 100m. Thus, the average distance between sensor nodes is around 1.8m. We set the routing hop length to be 10m and all sensor nodes have the same packet generation rate which is 1 packet per time unit. The traffic load distribution over the whole deployment area is shown in Fig. 3. We find that the nodes can experience quite different traffic loads even if they have similar distances from the sink, and this is exhibited by the ridges observed in Fig. 3. The reason for the presence of ridges is explained by the grid topology used, which provids an organized even-distribution but the node density is not the same in the radial direction. To counter
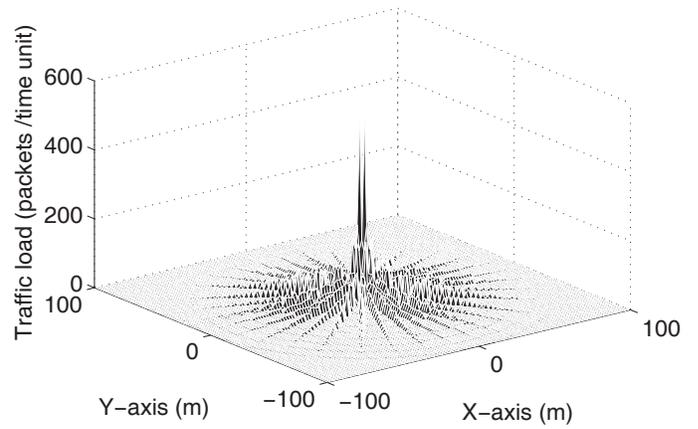

Fig. 3. Traffic load distribution over the disk deployment area ($n = 10,000$, $R = 100m$, $h = 10m$).
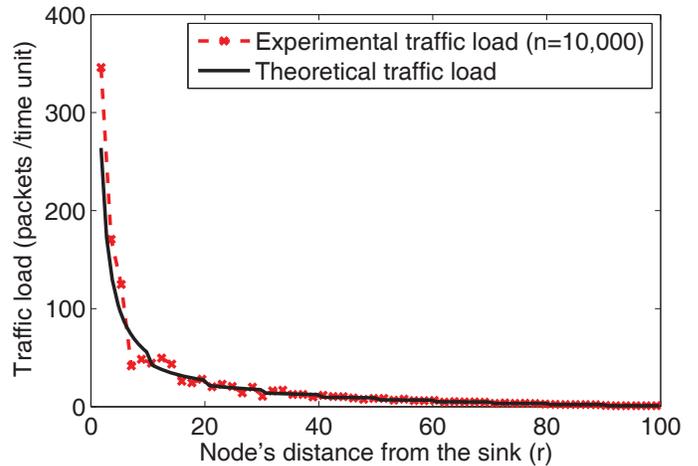

Fig. 4. Traffic load distribution in a network where $n = 10,000$ nodes are *gridly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.

the effect of the traffic asymmetry in the radial direction, we take the average of the traffic loads experienced by all nodes located within the annulus region of inner radius $r - \triangle r/2$ and outer radius $r + \triangle r/2$, as the estimate of the traffic load for nodes which are r away from the sink. Here, $\triangle r$ is the edge length of the grid and $\triangle r = \sqrt{\frac{\pi R^2}{n}}$. Fig. 4 shows the result of the estimated traffic load as a function of r with the theoretical result (calculated from (2)) shown as a comparison. For both the experimental result and the theoretical result, the smallest r considered is $\triangle r$. From Fig. 4, we can see the estimated traffic load from the experiment oscillate around the theoretical traffic load, and the experimental traffic load is converged to the theoretical value when $r \to R$. The oscillation observed for small r is due to the accumulation of the asymmetric traffic flow along its path to the sink.

To investigate the influence of node density, we also simulate networks where 20,000 and 100,000 sensor nodes are gridly deployed in the disk area of the first experiment. We also use the same network parameters as that used in the first
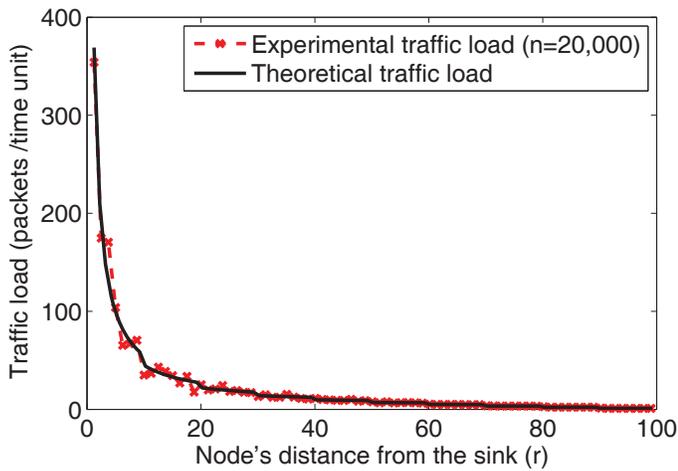
Fig. 5. Traffic load distribution in a network where $n = 20,000$ nodes are *gridly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.
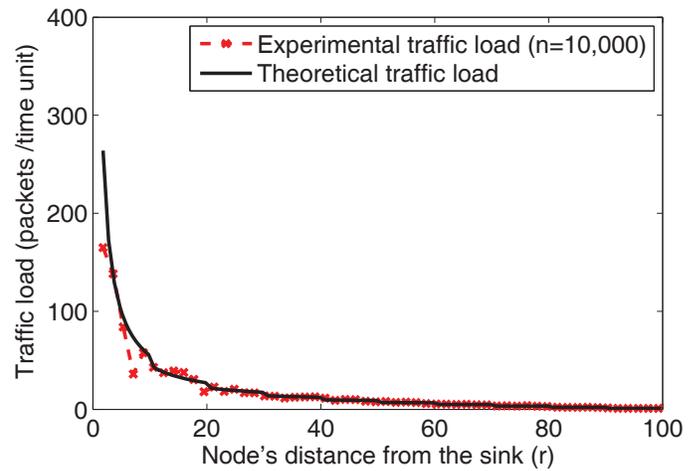

Fig. 7. Traffic load distribution in a network where $n = 10,000$ nodes are *randomly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.
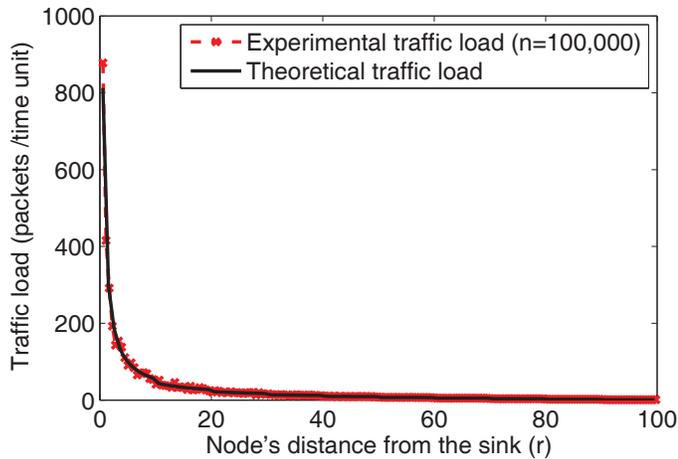

Fig. 6. Traffic load distribution in a network where $n = 100,000$ nodes are *gridly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.
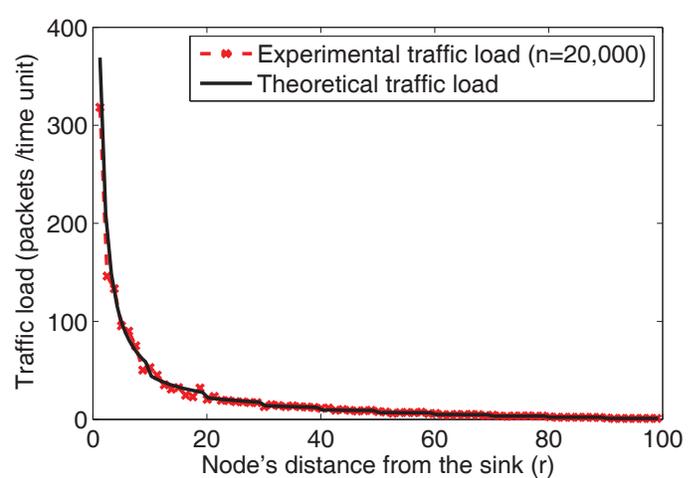

Fig. 8. Traffic load distribution in a network where $n = 20,000$ nodes are *randomly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.

experiment. The results acquired in the same way are shown in Fig. 5 and Fig. 6. It can be seen that the experimental traffic load gradually converges to the theoretical one in the whole support region of r when the number of deployed nodes is increased. This means the topology influence on the traffic distribution is smaller when the network becomes dense.

The ridges are not observed on the traffic load distribution over the nodes in a randomly deployed network. However, neighboring nodes also experience traffic load fluctuation, and the traffic load distribution becomes smooth when the node density is increased. Similar results for randomly deployed networks are shown in Fig. 7, 8, 9.

Although the theoretical result about the traffic load distribution in Section IV is achieved for large-scale and dense sensor networks, it may be interesting to check how different it is when the network is of normal size and low density. Thus, we make a new experiment of simulating the traffic load distribution in normal size and low density networks.

Concretely, there are only 250 sensor nodes and they are evenly deployed in a disk area of radius 50m. The routing hop length used this time is 15m, and the packet generation rate is still 1 packet per time unit. Fig. 10 and Fig. 11 separately show the case when the sensor nodes are deployed gridly and the case when the sensor nodes are deployed randomly. We see the theoretical traffic load distribution even coincides with the experienced traffic load distribution when the network is of normal size and low density. This is because every sensor node is supposed to be the geometrical center of its proximity region in an evenly distributed sensor network. Only a small error will be introduced when the geometrical center's distance to the sink is used as the estimate of its whole proximity region's distance to the sink in the integral approximation $\iint_{S_S} r \, dr \, d\theta \approx r_S \iint_{S_S} dr \, d\theta$ used in Section IV (A detailed analysis about the error introduced in this approximation is presented in Appendix B). This result enhances the practical
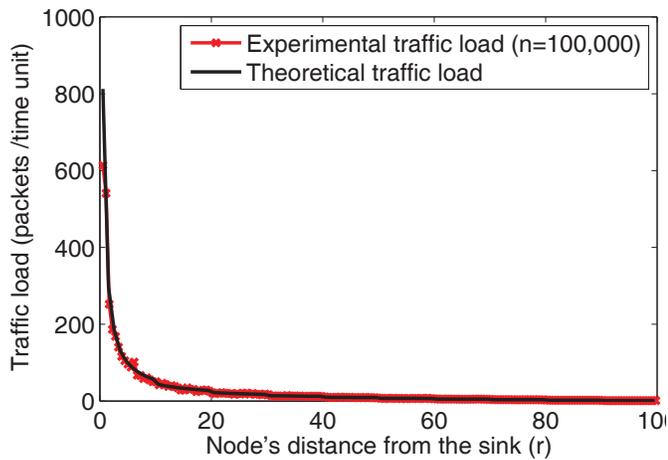
Fig. 9. Traffic load distribution in a network where $n = 100,000$ nodes are *randomly* deployed in a disk area of radius $R = 100m$ and the routing hop length is $h = 10m$.

Fig. 11. Traffic load distribution in a network where $n = 250$ nodes are *randomly* deployed in a disk area of radius $R = 50m$ and the routing hop length is $h = 15m$.
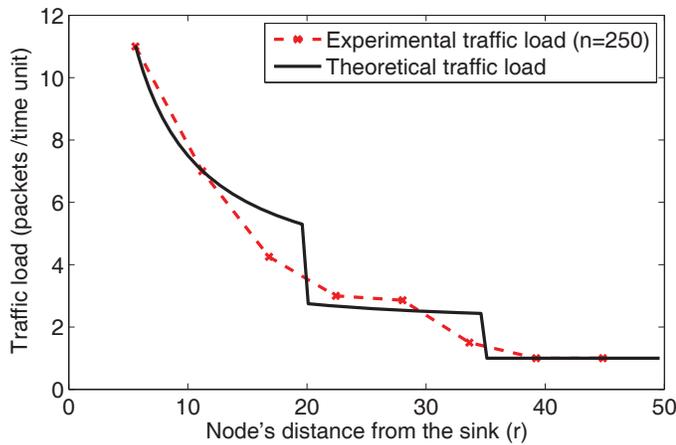

Fig. 10. Traffic load distribution in a network where $n = 250$ nodes are *gridly* deployed in a disk area of radius $R = 50m$ and the routing hop length is $h = 15m$.
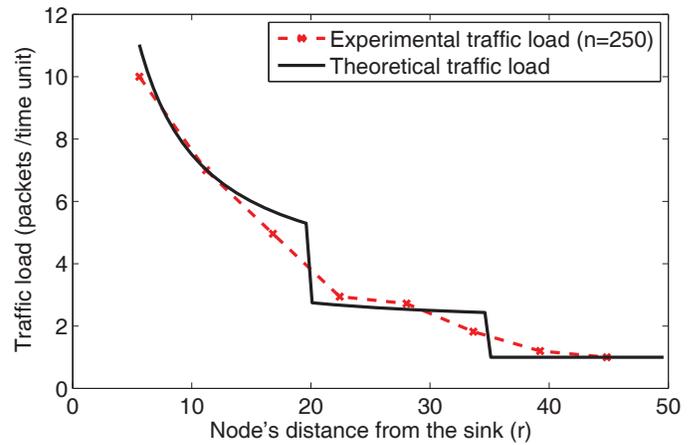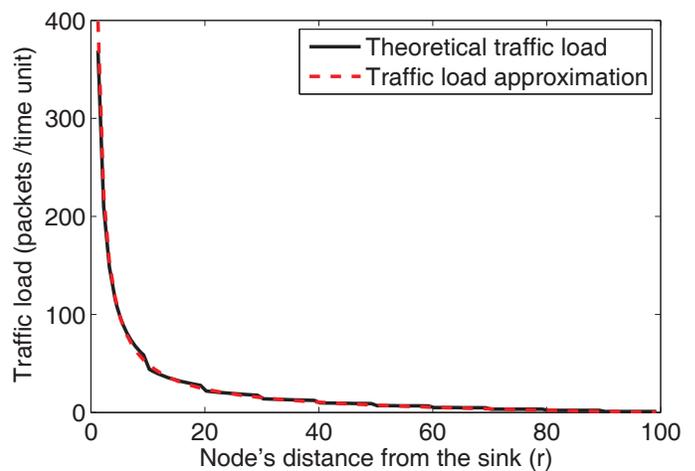

Fig. 12. Comparison between the precise theoretical traffic load and the approximated traffic load ($R = 100m$, $h = 10m$).

significance of the presented theoretical traffic load distribution.

Another experiment is also made to show that the traffic load expressed by (3) is a good approximation to that expressed by (2), thus all the corollaries based on (3) are reasonable. This experiment is based on calculation. We use parameters $R = 100m$ and $h = 10m$ in (2) and (3). The result is shown in Fig. 12.

VI. CONCLUSIONS

The special communication pattern in WSNs leads to that the traffic accumulates on its routing path to the sink. In this paper, we found the traffic load distribution over the deployed nodes as a function of their distance from the sink in the even-distributed large-scale and dense sensor networks. It was shown that the traffic load over a node increases when the node gets closer to the sink, and a traffic load blowout is expected in the proximity of the sink. However, a node may experience traffic load reduction when it gets too close to the sink. In large-scale and dense sensor networks, the theoretical traffic load distribution presented had a near-optimum fit to the experimental traffic load for almost all deployed nodes except those which are too close to the sink. Further, the theoretical traffic load distribution presented even had a good fit to sensor networks of normal size and low node density. This enhances the practical significance of the presented theoretical traffic load distribution.

In the future, the results presented can be used to find the proper node distribution that can efficiently balance the traffic load distribution over the nodes or for efficient energy allocation thus the network lifetime is maximized. While the network scenario used in this work is a symmetric network deployed on a disk area, the results presented are ready to be used for asymmetrically deployed networks providing load-balanced trees [15] are firstly constructed in such networks.
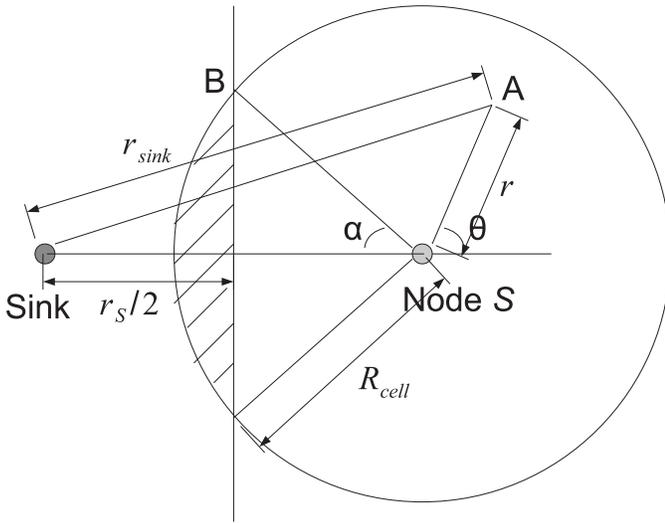
Fig. 13. Traffic load reduction experienced by a node getting too close to the sink

Fig. 14. A node and its proximity region shown in polar coordinates

APPENDIX A

In this appendix, we use an example to illustrate the possible reduction on the observed traffic load when a node gets too close to the sink.

Assume the proximity region occupied by a node is a round surface centered at the node of interest. Consider node S in Fig. 13. It has a distance of r_S from the sink, and the radius of its proximity region is R_{cell}. According to the routing algorithm explained in Section III, there will be no relay node if the sink is the closest node to the ideal relay point. The ideal relay point can only be closer to the sink than to its occupying node if its occupying node's distance from the sink is smaller than a threshold. In this special case, node S can only experience traffic load reduction when $r_S \leq 2R_{cell}$. And the part of reduced traffic load is those directed to the shadow area in Fig. 13.

Let r be the distance of node S to a point A in its proximity region. According to the law of cosines, point A's distance to the sink r_{sink} can be expressed as $r_{sink} = \sqrt{r_S^2 + r^2 - 2r_S r \cos(\pi - \theta)}$, where θ is point A's angle. The traffic load experienced by node S is the integral of the traffic load density at all points inside node S's proximity region, but outside of the shadow area.

Let α be the angle from point B to the sink as that shown in Fig. 13. The traffic load at node S is given by:

$$Traffic_S = \int_{-\pi+\alpha}^{\pi-\alpha} \int_0^{R_{cell}} Traffic(r_{sink})\rho r dr d\theta$$
$$+ \int_{\pi-\alpha}^{\pi+\alpha} \int_0^{\frac{r_S}{2\cos(\pi-\theta)}} Traffic(r_{sink})\rho r dr d\theta$$

where $\alpha = \arccos\frac{r_S}{2R_{cell}}$, $\rho = \frac{n}{\pi R^2}$, and $R_{cell} = \frac{R}{\sqrt{n}}$ in an even-distributed network. $Traffic(r_{sink})$ is defined by (2) and (3) in Section IV. Obviously, the limit of the traffic load reduction can be experienced when the traffic directed to half
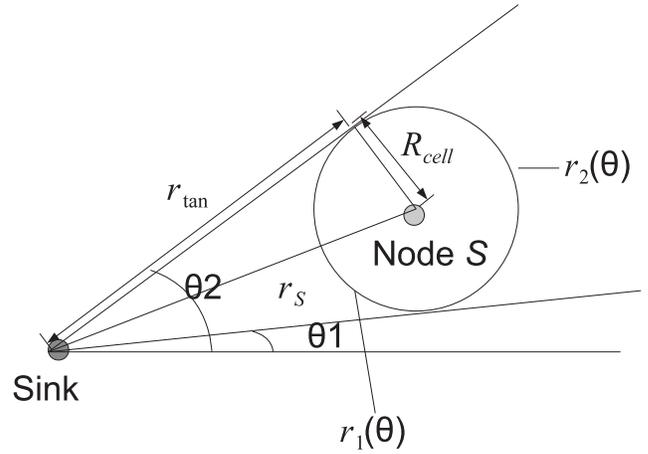
of the proximity region does not need a relay. The traffic load has a singularity at the point $r_S = 0$, where node S only experiences sensing traffic and the traffic load is $\bar{o}$ packets per time unit.

APPENDIX B

In this appendix, we analyze the possible errors introduced in the integral approximation $\iint_{S_S} r\, dr\, d\theta \approx r_S \iint_{S_S} dr\, d\theta$ appeared in Section IV. We use the same assumptions and symbols as those used in the first appendix.

Consider node S in Fig. 14. It has a distance of r_S from the sink, and the radius of its proximity region is R_{cell}. Two tangents exisit and their angles are θ_1 and θ_2. The distance between the origin and a tangential point is r_{tan}. With these definitions, we can do some analysis on the equation on the left side of the integral approximation:

$$\iint_{S_S} r\, dr\, d\theta = \int_{\theta_1}^{\theta_2} \int_{r_1(\theta)}^{r_2(\theta)} r\, dr\, d\theta$$
$$= \int_{\theta_1}^{\theta_2} d\theta \int_{r_1(\theta)}^{r_2(\theta)} r\, dr$$
$$= \int_{\theta_1}^{\theta_2} \frac{1}{2}[r_2^2(\theta) - r_1^2(\theta)]d\theta$$
$$= \int_{\theta_1}^{\theta_2} \frac{1}{2}(r_2(\theta) + r_1(\theta))(r_2(\theta) - r_1(\theta))d\theta$$

(4)

We can also do some analysis on the equation on the right side of the integral approximation:

$$r_S \iint_{S_S} dr\, d\theta = r_S \int_{\theta_1}^{\theta_2} d\theta \int_{r_1(\theta)}^{r_2(\theta)} dr$$
$$= \int_{\theta_1}^{\theta_2} r_S(r_2(\theta) - r_1(\theta))d\theta$$

(5)

To have the equality between the equations on both sides, we must have $r_S = \frac{1}{2}(r_2(\theta) + r_1(\theta))$. From Fig. 14, it can be

easily seen that $\frac{1}{2}(r_2(\theta) + r_1(\theta))$ is always a value between r_{tan} and r_S. Thus, the biggest error is less than $r_S - r_{tan}$ when r_S is used to approximate $\frac{1}{2}(r_2(\theta) + r_1(\theta))$.

Because $r_{tan} = \sqrt{r_S^2 - R_{cell}^2}$, the error is less than 0.5% of r_S when $R_{cell} = \frac{1}{10}r_S$ and 13.4% of r_S when $R_{cell} = \frac{1}{2}r_S$. This means the introduced error becomes bigger when the node density is decreased, and the introduced error also becomes bigger when a node gets closer to the sink. However, the introduced error is acceptable even in a bad case.

The angle difference between the two tagents also have influences on the value of the integral. When a node gets closer to the sink, this angle difference increases which will further amplify the error introduced in the estimation of $\frac{1}{2}(r_2(\theta) + r_1(\theta))$. However, the angle difference can not be more than π. Thus the error won't be badly amplified by the angle difference.

REFERENCES

[1] Q. Wang, T. Zhang, and S. Pettersson, "Bounding the information collection performance of wireless sensor network routing," in *Proc. of the 5th Annual Conf. on Communication Networks and Services Research (CNSR'07)*, May 2007, pp. 55–62.

[2] Q. Wang, T. Zhang, and S. Pettersson, "An effort to understand the optimal routing performance in wireless sensor network," in *Proc. of the IEEE 22nd Int. Conf. on Advanced Information Networking and Applications (AINA'08)*, March 2008, pp. 279–286.

[3] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. of the 33rd Hawaii International Conference on System Sciences*, January 2000.

[4] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 51–60, March 2002.

[5] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'00)*, March 2000, pp. 22–31.

[6] S.-C. Huang and R.-H. Jan, "Energy-aware, load balanced routing schemes for sensor networks," in *Proc. of the 10th International Conference on Parallel and Distributed Systems (ICPADS'04)*, July 2004, pp. 419–425.

[7] I. Teixeira, J. F. de Rezende, and A. de Castro P. Pedroza, "Wireless sensor network: Improving the network energy consumption," in *Proc. of the XXI Simposio Brasileiro de Telecommunicações (SBT'04)*, 2004.

[8] E. Hyytiä and J. Virtamo, "On traffic load distribution and load balancing in dense wireless multihop networks," *EURASIP Journal on Wireless Communications and Networking*, 2007, article ID 16932.

[9] L. Popa, A. Rostamizadeh, R. M. Karp, C. Papadimitriou, and I. Stoica, "Balancing traffic load in wireless networks with curveball routing," in *Proc. of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc'07)*, September 2007, pp. 170–179.

[10] M. Noori and M. Ardakani, "Characterizing the traffic distribution in linear wireless sensor networks," *IEEE Communications Letters*, vol. 12, no. 8, pp. 554–556, August 2008.

[11] S. Lee, B. Bhattacharjee, and S. Banerjee, "Efficient geographic routing in multihop wireless networks," in *Proc. of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'05)*, May 2005, pp. 230–241.

[12] M. Bhardwaj, T. Garnett, and A. Chandrakasan, "Upper bounds on the lifetime of sensor networks," in *Proc. of the 2001 IEEE Int. Conf. on Communications (ICC'01)*, 2001, pp. 785–790.

[13] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proc. of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'99)*, Feb. 1999, pp. 90–100.

[14] C. H. Foh, J. W. Tantra, J. Cai, C. T. Lau, and C. P. Fu, "Modeling hop length distributions for reactive routing protocols in one dimensional manets," in *Proc. of IEEE International Conference on Communications (ICC'07)*, June 2007, pp. 3882–3886.

[15] H. Dai and R. Han, "A node-centric load balancing algorithm for wireless sensor networks," in *Proc. of IEEE Global Telecommunications Conference (Globecom'03)*, December 2003, pp. 548–552.