

无线传感网中改进的节能分级数据汇聚机制

杨光松¹,石江宏²,陈辉煌²

YANG Guang-song¹, SHI Jiang-hong², CHEN Hui-huang²

1.集美大学 信息工程学院,福建 厦门 361021

2.厦门大学 通信工程系,福建 厦门 361005

1.School of Information Engineering, Jimei University, Xiamen, Fujian 361021, China

2. Department of Communication Engineering, Xiamen University, Xiamen, Fujian 361005, China

E-mail: Gsyang@jmu.edu.cn

YANG Guang-song, SHI Jiang-hong, CHEN Hui-huang. Modified energy-saving hierarchical data aggregation mechanism for wireless sensor network. *Computer Engineering and Applications*, 2009, 45(15):122–125.

Abstract: Considering the distributions of node's energy consumptions in Wireless Sensor Networks(WSN), a modified hierarchical data aggregation protocol is proposed. Firstly, this paper sets up a simple model to analytically evaluate the energy consumption in WSN by using aggregation mechanism. This paper also shows that energy consumption of an aggregation node increases so rapidly in the neighborhood of the sink node that these nodes will die and shorten the lifetime of the whole network. Secondly, a modified hierarchical data aggregation strategy in order to achieve balanced energy depletion by reduce the cluster head selection probability near the sink is proposed. Simulation results show that, this method can reduce the energy consumption efficiently for sensor networks. So, it can prolong the lifetime of the whole networks.

Key words: Wireless Sensor Network(WSN); aggregation; energy-efficient; hierarchy

摘要: 考虑 WSN 中节点的能耗分布情况,提出了一种改进的多级汇聚协议。首先建立分析模型评估了在 WSN 中采用多级汇聚机制时的能耗情况,证明在 SINK 附近的汇聚节点的能量增加很快,从而会耗尽能量,缩短整个网络的生存期。然后提出了一个改进的多级汇聚协议,通过改变通过减少离 SINK 附近的汇聚节点选择概率,可以均衡网络性能。仿真结果证明,该方法可以有效地降低能量,从而延长网络生存期。

关键词: 无线传感网; 汇聚; 能量有效; 分级

DOI:10.3778/j.issn.1002-8331.2009.15.035

文章编号:1002-8331(2009)15-0122-04

文献标识码:A

中图分类号:TP393.03

1 引言

在无线传感器网络(Wireless Sensor Network, WSN)中,电池容量、通信带宽等资源十分有限,因此在网络协议设计时,既要考虑减少干扰,满足链路的 QoS 条件,又要考虑减少能量的损耗。

在大范围环境监测的 WSN 应用中,需要长时间的数据收集、传送,由于传感节点靠电池供电,节能显得十分必要。无线传感网中节点的能耗主要是来自数据传输和处理的耗能,例如,1 bit 的数据传到 100 m 外的目的节点的耗消的能量,相当于 CPU 执行 300 条指令的耗能^[1]。在无线信道上,能量消耗与传送的信息数量成正比,与接收距离的指数成正比。为了达到节能目的,目前有许多方法被提出。第一种方法是改变传输模式,用短距多跳传送代替长距离传送,从而节能^[2];第二种是基于数据处理(压缩、运算、存储等)所耗损的能量,对于用户所需

要的数据,尽可能去除冗余,减少数据传输量,以延长生存期^[3],此方法称为数据聚集(Data Aggregation)或者数据融合(Data Fusion)^[4-6],它将不同传感节点监测到的资料,先行经过整合、滤除重复信息后,再传往目的节点,从而能减少数据重传次数,缩小数据量。

数据汇聚技术是减少能耗、消除数据冗余、增加从源节点到基站的有用信息流、延长网络寿命的重要方法。有许多文献讨论了 WSN 中的汇聚问题,文献[7]提出了 LEACH 分簇算法,每个节点都按固定的概率成为簇头,建立簇头来汇聚各节点的数据,文献[8]使用不同的簇大小来均衡网络能量,文献[9]提出多级数据聚集算法,介绍了自动选择簇头、成簇、汇聚的过程,虽然证明了通过多级分簇使区域内全体电量降低,但没有讨论各个节点的能量分布,也没有讨论其生存期。针对此算法存在的不足,提出一种新的,根据离 SINK 近的距离来决定汇集点

基金项目:福建省自然科学基金(the Natural Science Foundation of Fujian Province of China under Grant No.A0710021);集美大学优秀青年骨干教师基金资助(No.2008B002)。

作者简介:杨光松(1968-),男,博士,副教授,主要研究领域无线网络;石江宏(1968-),博士,副教授,主要研究领域为无线通信;陈辉煌(1944-),男,博导,教授,主要研究领域为无线通信。

收稿日期:2008-03-26 **修回日期:**2008-05-29

的概率,从而可以又能减轻能量消耗,延长生存期,最后通过仿真证明了其有效性。

2 无线传感网中分级汇聚方法

2.1 分级汇聚方法简述

文献[10]提出了多级数据汇聚方法(Hierarchical Energy-Efficient Protocol for Aggregator Selection,HEPAS),如图1,协议假设监测区域随机分布 n 个传感节点,各节点掌握了其他节点的位置信息,每个节点有理想的同步,中央有一个SINK作为数据处理中心。传感节点分成簇,每个簇有一个簇头作为汇聚节点AN(Aggregation Node)。传感节点SN(Sensor Node)将感测到的数据送往离自己最近的GN,在此将其他传感节点的数据进行压缩汇聚后,以分级的方式向SINK传送数据:首先将所有的SN分成0级,从这些SN中选出一个子集作为第一级的 AN_1 ,又从第一级的 AN_1 中选出第二级的 AN_2 ,如果有 h 级,则SINK作为最高的 $h+1$ 级。按此方法, i 从1到 h 的顺序选项出第 i 层的汇聚节点 AN_i ,最后形成多级的分簇。传感节点将数据发往最近的一级AN,一级AN收集数据后,送往最近的二集AN,逐级上传,直至送达SINK。

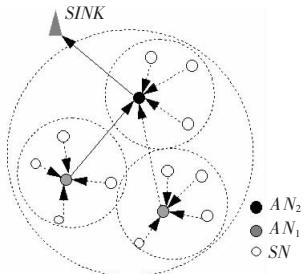


图1 多级数据汇聚模型

HEPAS是一个随机的分布式算法,其工作过程分为以下几个阶段,

(1)选出第 i 层的汇聚节点 AN_i ($i=0,1,\dots,h$)

在 $i-1$ 层的节点 SN_{i-1} 中,以概率 $p_{1i} \in [0, k_i/(k_{i-1}-1)]$ 自动地成为 AN_i , k_i 是第 i 层的汇聚节点数,显然, $k_0=n$, $k_{h+1}=1$ 。一个第 i 层的分簇,由一个 AN_i 和多个 SN_{i-1} 组成。

(2)形成分簇

按第(1)步形成 AN_i ,便通知其他周围的 $i-1$ 层节点 SN_{i-1} ,它们收到信息后,将数据发往最近的 AN_i ,于是形成了以 AN_i 为中心的分簇。为了覆盖所有的观察区域,AN的覆盖范围 b_i 按下式设置^[10]:

$$b_i = \frac{2a}{\sqrt{k_i}} \quad (1)$$

(3)在第 $i-1$ 层的节点中,如果没有收到第 i 层的 AN_i 的通知,知道自己不属于任何一个簇时,按满足下式的概率成为簇头:

$$p_{1i} + p_{2i} (1-p_{1i}) (1-p_{1i} b_i^2)^{k_{i-1}-1} = \frac{k_i}{k_{i-1}} \quad (2)$$

其中,右边是所有节点中,成为AN的比例。左边第一项是所有节点中第一阶段选出的AN的比例,第二项是在不属于任何簇的节点中选出AN的比例。如果 p_2 满足上式,则能够保证选出的选出最优能量的AN。

(4)按第(3)步选出的 AN_i ,以自己为中心, b_i 为半径通知附近的按(1)~(3)步未找到归宿的节点。形成一个新的分簇。

(5)经第(1)~(4)步处理后,在 $i-1$ 级中,不属于任何AN

的 SN_{i-1} 直接把信号送往SINK。

2.2 分级汇聚法中的能量分析

2.2.1 分析模型

下面用图2的简单模型对分级汇聚方法的耗能进行分析,首先对单级汇聚的耗能进行分析。以SINK为中心,半径为 a 的区域内分布有 n 个节点,其中有 k 个节点被选为AN,则SN的密度为 $n/(\pi R^2)$,AN的密度为 $k/(\pi R^2)$,AN的簇半径按式(1)给出,所以,SN数量为 $n_c = n \cdot (b/a)^2$ 。

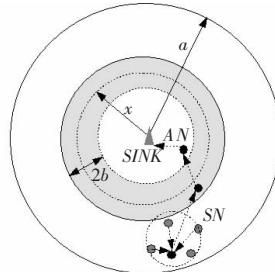


图2 能量分析示意图

由于发送消耗的能量与发送距离 d 和发送的数据长度 r 有关,一个节点中继一个单位的数据到距离 d 的另一个节点,则其耗能为^[11]:

$$E = r \times \alpha_1 + r \times \alpha_2 \times d^l \quad (3)$$

其中, α_1 和 α_2 分别是电路能量消耗指数和天线能量消耗指数,取决于硬件情况,而 l 是衰减指数。

一个节点到另一个节点使用能量最小的距离称为特征距离,此值取决于使用的硬件及环境。

$$d_{char} = \sqrt{\alpha_1 / \alpha_2} (l-1) \quad (4)$$

2.2.2 能量分析

每个节点能量消耗包括三个部分:一为压缩数据的耗能;二为AN接收来自自己通信范围内SN的能量;三为中继其他AN到SINK的能量。考虑图2中半径为 x ,宽度为 $2b$ 的园环内的节点的耗能情况。

假设所有的SN数据都送往AN,每个AN接收的数据为每个分组 r bit,共有 k 个AN,一个AN接收数据的速率为每个周期 rn/k 。

(1)环内压缩数据的耗能

$$E_{a0} = k_r \times f_a(nr/k) \quad (5)$$

其中, k_r 是环内的AN数, $f_a(nr/k)$ 是汇聚能量消耗函数,代表每个AN压缩 nr/k 数据的耗能, $f_a(x) = \beta x$ 是压缩函数。

(2)每个AN接收簇内SN的耗能

在一个以 (x_c, y_c) 为中心的簇内,数据传送到簇头AN的总的距

$$\frac{n}{\pi a^2} \times \iint_{(x, y \in c)} \sqrt{(x-x_c)^2 + (y-y_c)^2} \, dx \, dy \quad (6)$$

如果每个SN选择距离其最近的AN作为簇头,便在网络内构成了一个voronoi图,每个簇对应于一个voronoi区,于是上式可以表示成为 $2/3ak^{-3/2}$,考虑耗能参数 α 和节点数据速率 r ,传感节点到AN的总的耗能消耗:

$$E_{a0} = k_r \times \frac{n}{\pi a^2} \times \iint_{(x, y \in c)} \sqrt{(x-x_c)^2 + (y-y_c)^2} \, dx \, dy = \\ 2/3k_r \alpha n r k^{-3/2} \quad (7)$$

(3)AN转发环外数据的耗能

区域内环外围节点数为 $n_{outside} = \pi(a^2 - (x+b)^2)$,所以外围AN

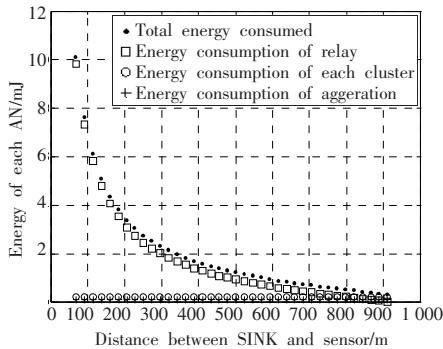


图3 单级汇聚中AN能耗与节点距离关系

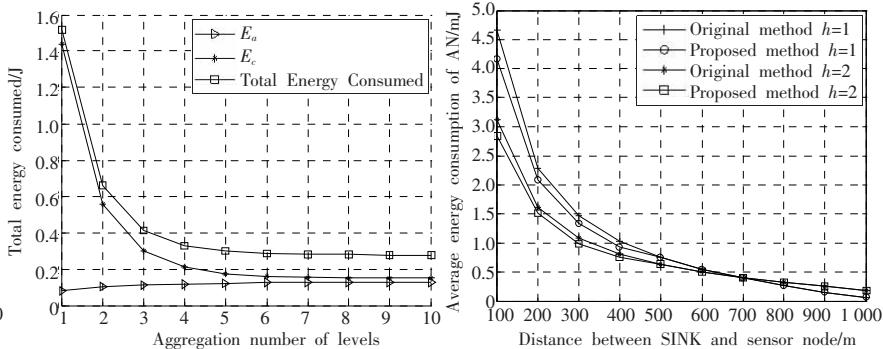


图4 多级汇聚能量分布

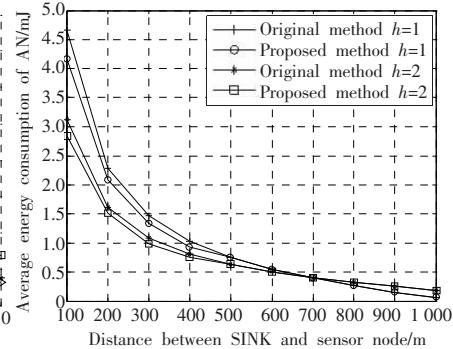


图5 能量消耗的比较结果

所有产生的数据都通过环内的AN来转发,为简化分析,假设每个AN中继环外的数据量是相等的。它们合并自己的数据和中继数据,并将其传递给SINK或者离SINK更近的节点。因此,AN转发环外节点的耗能为:

$$E_{ci} = n_{outside} \times n \times r / k \times d_{char} \times \alpha \quad (8)$$

其中, d_{char} 为式(4)表示的最优能量距离, α 为传送单位距离的耗能。

采用表1的数据,可以得到一级的各汇聚节点的平均耗能,如图3所示。图中可见,离SINK越近,耗能越大。这是由于使用多对一模式,为AN既要进行数据汇聚,又要转发其他节点的耗能。离SINK近的节点,负荷加重。因此消耗能量比较快,整个网络的生存期缩短了。因此会导致较早耗尽能量。从图3中还可观察到,能量消耗主要来自AN转发数据的耗能,压缩耗能和簇内处理的平均耗能基本不随距离改变。

表1 仿真参数

参数	意义	取值
n	区域中SN数量	10 000
a	网络半径	1 000 m
α_1	通信电路耗能	5×10^{-8} J/bit
α_2	天线耗能	1×10^{-8} J/bit/m ²
l	衰落指数	2
r	数据速率	160 bit/s
β	汇聚能量消耗率	5×10^{-8} J/bit
γ	数据压缩率	25%
c	数据压缩负荷	32 bit

在多级汇聚的情况下,再析多级的耗能,传感节点分成簇,每个簇有一个簇头,其他节点周期性地将自己的信号传送到簇头上去,簇头将数据收集并压缩,最后送到SINK。

根据文献[10],可以计算出各级进行数据汇聚时的耗能为:

$$\sum_{i=1}^h E_a + \sum_{i=0}^h E_{ci} \quad (9)$$

E_c 代表将数据从 i 级传送到 $i+1$ 级消耗的能量。 E_a 表示在 i 级数据汇聚的耗能。根据式(9)可以绘出其分析曲线如图4,图中可见,随着级数的增加,网络的总耗能呈现下降的趋势,说明采用多级汇聚确实可以节省能量。 E_a 随级数的变化甚微,而 E_c 随级数的增加而减少。所以总耗能主要是来自数据传输的耗能。

3 改进的多级数据汇聚方法

3.1 协议存在的问题

根据2.2节的能量分析,如采用以上的方法,在整个监测

区域内,各层汇聚点的选出概率是一样的,因此存在以下的问题:

(1)如果位于区域边缘的节点成为AN的话,要将汇集的数据向相反方向的SINK传送,造成数据传送的浪费。

(2)由于数据是依靠节点相互中继转发进行多跳传送,因此,SINK周围的节点耗电较大。

提出的多级数据汇聚方法的主要思想是:根据传感节点离SINK的距离来确定汇聚节点选择概率,避免中心节点的电量尽早消耗,从而均衡整个网络的能量,以提高整个网络的性能。

3.2 协议工作过程

在2.2.2节中,第(1)、(2)耗能必须由AN完成,而第(3)的处理非AN也可以达到,SINK附近和区域边缘相比,数据中继次数比较多,在SINK附近的高层AN,既要充当AN,也要充当中继节点,因此,电量消耗很快。所以,为了让高层的AN不在SINK附近出现,让中继处理在非AN节点中进行,以分散负载。

为了控制SINK附近节点选择为AN的概率,对2.2节所述的算法进行改进,其AN选择、成簇、汇聚等过程与之类似,其第(2)、(4)、(5)步与其一样,仅在第(1)、(3)步进行改动,主要改进有以下两方面,

(1)2.2节步骤(1)的AN选出概率如下所示:

$$\begin{cases} p_{li}(x)=0, & \text{if } x \in (0, \frac{b_i}{2}] \text{ and } x \in (\frac{b_i}{2}, a - \frac{b_i}{2}) \\ p_{li}(x) \in [0, \frac{k_i}{k_{i-1}}] & \text{if } x \in (a - \frac{b_i}{2}, a_i] \end{cases} \quad (10)$$

k'_{i-1} 是在半径为 $[\frac{b_i}{2}, a - \frac{b_i}{2}]$ 的圆环区域内的AN的数量,这样

选取的目的是为了降低在SINK附近的SN被选为簇头的概率。

(2)2.2节步骤(3)选出 i 级的AN($i=2, \dots, h$)的概率满足如下条件:

$$\begin{cases} p_{li}(x) + p_{2i}(x)(1-p_{li}(x)) \left(1 - \frac{p_{li}(x)b_i^2}{a^2}\right)^{k'_{i-1}-1} = \frac{k_i}{k'_{i-1}} \\ \text{if } x \in (0, \frac{b_i}{2}] \text{ and } x \in (\frac{b_i}{2}, a - \frac{b_i}{2}) \\ p_{li}(x) = 0 \text{ if } x \in (a - \frac{b_i}{2}, a_i] \end{cases} \quad (11)$$

按以上步骤执行后,减少了靠近SINK的节点成为AN的概率,外围数据通过其他非AN的节点进行转发,因而避免了一些AN节点过早地耗尽能量,从而可以延长网络的生存期。

4 仿真结果及分析

本章通过仿真对提出的算法进行评估,采用图2所示的场景,区域内随机分布的每个节点按3.2节所述的方式以确定的

概率选出簇头,进行数据汇聚。假设数据收集是同步的,每周期包括了数据收集、汇聚、通信的过程,AN 将收到的数据和自己监测到的数据进行整合,按压缩率为 γ 进行压缩,如果簇内的节点数为 m ,压缩后的数据长度为:

$$r = m \times r \times \gamma + c \quad (12)$$

这里, c 为压缩时附加的负荷。

各级 SN 将数据发送到 AN,在 AN 进行整理压缩后,送到上一层 AN,进行 h 级汇聚,直到送达 SINK。由于要经过其他节点进行多跳传送,为保证最佳能量,路由协议采用文献[11]所述的 CDPR (Characteristic Distance Progressive Routing) 协议,并按式(4)定义的最优能耗距离进行多跳传送。

仿真参数见表 1,测量一个周期内的能量消耗,分别统计在二级汇聚中,分别位于第一级的 AN 和第二级 AN 的耗能。结果如图 5 所示,显示了在不同的级数以及离 SINK 不同距离的情况下,各节点的耗能情况。图中可见,SINK 越附近,消耗电量越大。而且一级节点耗能大于二级节点,这是由于上级分簇中的 AN 将下级节点中的数据进行收集、压缩,送往一级的数据量显减少了。图中还可看到,无论是一级节点,还是二级节点,采用提出的汇聚方法,其耗能特性都有一定程度的改善。体现在 SINK 附近的能耗减少,这样,避免了这些节点过早耗尽能量,从而可以延长网络的生存期。

5 结束语

在进行广范围地域监测的无线传感网络中,最终的数据必须交付 SINK 进行处理,SINK 附近的节点耗能相对较大。提出了一种改进的多级数据汇聚方法,通过减少 SINK 附近节点成为 AN 的概率,避免了这些节点过早地耗尽能量,从而延长网络的生存期。通过仿真实验,将现有方法与提出的方法进行了比较,仿真结果证明,提出的方法能改善能量分布,从而节省能量。

下一步的研究内容将根据 SINK 距离的不同,考虑采用不同的分簇大小、节点密度,或者采用不同的 AN 选择概率,从分级的观点更深入地进一步考虑能量平衡问题。

(上接 94 页)

平均跳数较大(在 5~10 跳之间),说明没有找到最优路由,随着网络运行时间的增加,每次路由的平均跳数趋向稳定(在 4~10 之间)。50 节点情况下也存在相同的规律,而 25 节点情况下由于节点数量少,总体上每次路由的平均跳数都稳定于 2 跳到 3 跳之间。

网络刚开始工作时,无论 25、50 还是 100 节点,平均路由发现时间都比较大,因为还没有缓存到目的节点的路由;随着网络工作时间的增加,平均路由发现时间逐渐减少而趋于稳定,其原因是由于 DSR 采用了 Cache 缓存机制,它可以将节点所知的路由信息存储起来。随着节点数量的增加,平均路由发现时间也变长,这是因为网络规模增大,每次路由的平均跳数增加,中间节点处理的延迟和路径的延迟增大。

5 小结

固定节点模型和移动节点模型的比较中看到,DSR 的路由发现机制并不受节点移动性的影响,在网络拓扑结构不断变化的网络中,通过基于最少区域模型的仿真框架,每次路由发现的时间即每次源节点获得一条到达相对位置已改变目的节点

参考文献:

- [1] Tilak S, Abu-Ghazaleh N B, Heinzelman W. A taxonomy of wireless micro-sensor network models[J]. Mobile Computing and Communications Review, 2002, 1(2): 1~8.
- [2] Ramanathan R, Rosales-Hain R. Topology control of multihop wireless networks using transmit power adjustment[C]// Proc IEEE INFOCOM 2000 Conference, Tel Aviv, Israel, 2000: 354~366.
- [3] von Rickenbach, Wattenhofer R. Gathering Correlated Data in Sensor Networks[C]// Proc 2004 Joint Workshop on Foundations of Mobile Computing, 2004: 64~65.
- [4] Krishnamachari B, Estrin D, Wicker S. The impact of data aggregation in wireless sensor networks[C]// Proc International Workshop of Distributed Event Based Systems(DEBS), Vienna, Austria, 2002: 1~4.
- [5] Yee C, Kumar P S. Sensor networks: Evolution, opportunities, and challenges[C]// Proceedings of the IEEE, 2003, 91(8): 1248~1250.
- [6] Choi J Y, Lee J W, Lee K, et al. Aggregation time control algorithm for time constrained data delivery in wireless sensor networks[C]// VTC 2006, Melbourne, Australia, 2006: 563~567.
- [7] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks[J]. IEEE Trans on Wireless Comm, 2002, 1(4): 660~670.
- [8] Soro S, Heinzelman W B. Prolonging the lifetime of wireless sensor networks via unequal clustering[C]// Proc of the 19th IEEE Int'l on Parallel and Distributed Processing Symposium. San Francisco: IEEE Computer Society Press, 2005: 236~240.
- [9] Mas-Colell A, Whinston M D, Green J R. Microeconomic theory[M]. UK: Oxford University Press, 1995.
- [10] Chen Y P, Liestman A L, Liu J. A hierarchical energy-efficient framework for data aggregation in wireless sensor networks[C]// The 2nd International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks 2005(QSHINE'05), 2005.
- [11] Bhardwaj M, Garnett T, Chandrakasan A. Upper bounds on the lifetime of sensor networks[J]. Proc IEEE Int Conf Comm, 2001, 3: 785~790.

新的路由的时间并不受节点的移动性影响而趋于稳定,可明显减少路由跳数,增加路由计算效率。仿真结果表明,对于 DSR 路由协议而言,随着网络运行时间的推移,路由量,每路由平均跳数和平均路由发现时间趋于稳定,从而说明在网络规模不是很大时 DSR 路由协议的动态路由适应性较好。

参考文献:

- [1] 陈敏. OPNET 网路仿真[M]. 北京: 清华大学出版社, 2004.
- [2] 王文博. OPNET Modeler 与网络仿真[M]. 北京: 人民邮电出版社, 2003.
- [3] OPNET Technologies. OPNET Modeler Documentation, Version 10.0. A[S]. 2003.
- [4] 石怀伟, 王华, 张念军, 等. OPNET 网络仿真技术及其应用分析[J]. 计算机工程与设计, 2006, 27(17): 3309~3311.
- [5] Johnson D, Maltz D, Hu Y. The Dynamic Source Routing protocol for mobile Ad hoc networks(DSR)[C]// Internet Draft, IETF, 2004.
- [6] 闻英友, 赵建立, 赵林亮, 等. 基于拓扑的 MANET 路由协议性能与网络覆盖密度相关性的研究[J]. 计算机研究与发展, 2005, 42(4): 684~689.
- [7] 符琦, 蒋之霞, 蒋瑞林. 基于 NS2 的 Ad hoc 网络路由协议的模拟实现[J]. 计算机工程与应用, 2006, 42(5): 149~151.