Research Article
Optimal Sink Position Selection Algorithm for Wireless Sensor Networks

[1]Wenbo Zhang, [1]Lingjiao Yuan, [1,2]Xiaobo Tan and [2]Hai Zhao
[1]Department of Information Science and Engineering, ShenYang Ligong University,
Shenyang, 110159, China
[2]Department of Information Science and Engineering, Northeastern University,
Shenyang, 110006, China

Abstract: The Optimal Sink Position Selection Algorithm is proposed based on the maximum demands (OSPSA). In this algorithm the communication demands of the nodes and the communication failure probability between the nodes and the sinks are considered. The sinks multiply cover the key nodes to satisfy the maximum demands to improve the quality of service. Furthermore, the characteristics are analyzed in theory. Simulation experiments are conducted to analyze and compare the relationships between the failure probability, the coverage radius and the maximum coverage demands. Moreover the effects between the number and the maximum coverage demands and the effects between the coverage and the maximum coverage demands are also compared.

Keywords: Multiple coverage, optimal selection algorithm, sink location, wireless sensor network

INTRODUCTION

A wireless sensor network (Priyadip *et al*., 2008) is composed of a large number of tiny sensor nodes which can do computations and communicate wirelessly. The sensor nodes are easily deployed but hard to be re-collected and recharged. So the energy consumption is an important metric to sensor network, since it is directly related to the network operation lifetime. The network has been used in so many prospects such as military surveillance, environmental monitoring, industrial production, healthcare and many other fields (Xiao *et al*., 2008). A wireless sensor network is application oriented, so its architecture is very complex. The network nodes deployed in the monitored region form a self-organization network (Niu *et al*., 2006). Through multi-hop relay, the sensing data could be transmitted to the base station (sink node). Eventually, with long distance or temporary sink link, the data from the entire monitoring region is transmitted to the remote management center (Ganeriwal *et al*., 2008).

The existing studies typically assume that only one base station existed in the network. The sink node is located in the center or randomly distributed in the network. For a large-scale wireless sensor network (Yuan *et al*., 2010), the sensor nodes are difficult to transmit sensing data to the sink node through one hop. Therefore, the routing algorithm is a hot topic for wireless sensor network (Li *et al*., 2008). In a wireless sensor network the nodes could be grouped into clusters and there is a special node called cluster-head which has more resources and, thus, is more powerful than the common-nodes (Younis *et al*., 2004). Furthermore, cluster-heads are responsible for sending data to the sink node. The sink node communicates with the observer, which is a network entity or a final user that wants to have information about data collected from the sensor nodes. However, the algorithm of routing or clustering needs to transfer a large amount of control information which will consumes the sensor energy. In addition, for the impact of monitoring environment, the communication links between the sensor node and the sink node are subject to interfered which will lead to the link failures (Anastasi *et al*., 2009). Especially for hot monitoring region, how to guarantee the reliable and timely data transmission must be considered Jia *et al*. (2008).

Based on the above analysis, a large-scale wireless sensor network Optimal Sink Position Selection Algorithm (OSPSA) was put forward in this study. By optimally deploy multiple sink nodes, more data transmission demands from sensor nodes can be satisfied. Especially for some important sensor nodes, even in the case of the wireless communication link failures, multiple sink nodes can effectively ensure the data transmission demands (Tang *et al*., 2010). In this study it assumed that the location of the sensor nodes is known and the sink nodes can be arbitrarily deployed in a certain region. Through the optimal sink position selection algorithm, the sink nodes can multiply cover the key sensor nodes to improve the reliability of data transmission.

In this study, we propose the Optimal Sink Position Selection Algorithm based on the maximum demands (OSPSA). In this algorithm the communication demands of the nodes and the communication failure probability between the nodes and the sinks are considered. The sinks multiply cover the key nodes to satisfy the maximum demands to improve the quality of service. Furthermore, the characteristics are analyzed in theory. Simulation experiments are conducted to analyze and compare the relationships between the failure probability, the coverage radius and the maximum coverage demands. Moreover the effects between the number and the maximum coverage demands and the effects between the coverage and the maximum coverage demands are also compared.

THE OPTIMAL SINK POSITION SELECTION ALGORITHM

System model: In a wireless sensor network, supposed that N was the number of node , M was the number of sink node, h_k was communication demands between any sensor node K and sink node. In order to take advantage of limited sink node to cover as much as possible node demands and to play the maximum service capacity of the sink node, the model can be expressed as follows:

$$Maximize \sum h_k y_k \qquad (1)$$

That was, by setting the location of the sink node, to covers the maximum demands of the sensor nodes:

$$y_k - \sum_{i=1}^{N} a_{ki} X_i \leq 0 \ \ K = 1,2 \dots, N \qquad (2)$$

As the formula (2) shown above, if and only if the sink node was set on node i and the dwastance between the sink node and sensor node k was not greater than the coverage radius, the sensor node k can be covered. For coverage these conditions should be met: the sink node was deployed at i as the center; the maximum communication distance between the sink node and sensor node was the radius (here, supposed that the sink node and sensor nodes have the same maximum communication distance). When the sensor node k was included in the coverage circle, the node k was multiply coverage. The coverage radius was expressed as D. In addition, the actual setting of sink nodes should be in the available ranges:

$$\sum_i X_i \leq M \qquad (3)$$

$$X_i = 0, 1 \ I = 1, \dots N$$
$$Y_k = 0, 1 \ k = 1, \dots N \qquad (4)$$

$$y_k = \begin{cases} 1 & \text{Node k is covered by sink node} \\ 0 & \text{Node k isn' tcovered by sink node} \end{cases}$$

$$X_i = \begin{cases} 1 & \text{Node k is covered by sink node} \\ 0 & \text{Node k isn' t covered by sink node} \end{cases}$$

Obviously, if the objective function (1) can be obtained under the constraints of (2), (3) and (4), we can get the maximum coverage demand in the restricted condition.

In formula (1), all sink nodes are in the normal communication with the sensor node on default. However, wireless sensor network is using wireless communication mode; then that was to say, communication link failure frequently appears between the sink node and the node within range of coverage. Set the failure probability was p ($0 \leq p \leq 1$) and in wireless sensor network the failure probability of communication between the sink node and sensor nodes was equal and unrelated, so when the node K was covered by m sink nodes, the success probability of communication was:

P {the success probability of communication}
= 1-P {the failure probability of communication}
= 1-P^m \qquad (5)

Further assumed that $H_{k,m}$ was the coverage demand of the node k when the node k was covered by m sink node nodes:

$$H_{k,m} = \begin{cases} h_k & \text{communicat ion success} \\ 0 & \text{communicat ion failure} \end{cases} \qquad (6)$$

The expectation of $H_{k,m}$ was:

$$E(H_{k,m}) = h_k(1 - P^m) \ \ \forall k, m \qquad (7)$$

The number of sink node which covers node k increases from m-1 to m and the increasing amount of Node K cover expectation was:

$$\Delta E\ (H_{K,m}) = E\ (H_{K,m}) - E\ (H_{K,m-1})$$
$$= h_k P^{m-1}(1 - P) \ \ m = 1,2, \dots M$$

The sink node which can cover node k was determined by $\sum_{i=1}^{N} a_{ki} X_i$ in formula (2).

Using the above definition, the model (1) - (4) was improved as follow:

Max.
$$\sum_{k=1}^{N}\sum_{j=1}^{M}(1 - P)P^{j-1}h_k\ y_{jk} = \sum_{k=1}^{N}\sum_{j=1}^{M}W_j\ h_k \ (8)$$

St. $\sum_{j=1}^{M} y_{jk} - \sum_{i=1}^{N} a_{ki} X_i \leq 0 \ \forall$ \qquad (9)

$$\sum_{i=1}^{N} X_i \leq M \qquad (10)$$

$$X_i = 0,1, \dots, M, \quad \forall i \tag{11}$$

$$y_{jk} = 0,1, \quad \forall i, k \tag{12}$$

In which,

$$y_{jk} = \begin{cases} 0 & \text{If node k is at least covered by j sink node.} \\ 1 & \text{Sink node which covers node k is less than k.} \end{cases}$$

X_i was equal to the number of the sink node located at the node i.

$$W_j = (1 - P)P^{j-1} \; j = 1, \dots M$$

The objective function for each value of K about j was concave, it shows that:

If $y_{jk} = 1$, then $y_{1k} = y_{2k} = \dots y_{jk} = 1$

If $y_{jk} = 0$, then $y_{1k} = y_{2k} = \dots y_{jk} = 0$

Therefore, the objective function can be expressed as:

$$\sum_{k=1}^{N}(1 - P) P^{j-1} \sum_{j=1}^{M} h_k y_{jk}$$

In which, the inner sum represents the number of demands which was at least covered by j sink nodes (for each value of j). For each such item, font coefficient can be understood as weights of the number of maximum demand, which are at least covered j times.

Model properties:
Theorem 1: In OSPSA model, the number of sink node increases from M_1 to M_2, the maximum coverage demands meets the condition $Demands(M_1) \leq Demands(M_2)$.

Supposed that wireless sensor network contains N nodes, when the sink node number was M_1, the coverage demand set was C = {c_1, c_2,... c_N}, the optimal position set of the sink node was L = {$l(1)$, $l(2)$, …$l(M_1//0$}. When the sink node number was M_2, the coverage demand set was $c' = \{c', c', \dots c'_N\}$ the optimal position set of the sink node was $L' = \{l'(1), l'(2), l'(M_2)\}$

According to the above two cases analysis, due to $M_1 < M_2$, there must be $M_1 + 1 < M_2$ or $M_2 = M_1 + 1$.

1. $M_2 = M_1 + 1$

    Assumed that $l'(i) = l(i), i = 1,2, \cdots M_1$, $l'(M_2)$ was any point in the candidate set. Due to the emergence of the M_2 sink node, coverage

relationships between sink nodes and sensor nodes are likely to change, namely $C \neq C'$.

i.   For any node i in the network, there was no $C \neq C'$. The node coverage times doesn't change because of M2 sink node join, that was $C \neq C'$. So the formula $\sum_{k=1}^{N} h_k (1 - P^{c_k}) = \sum_{k=1}^{N} h_k (1 - P^{c_k})$ was founded and then $Demands(M_1) \leq Demands(M_2)$.

ii. The coverage times of the sensor node I was $c'_j > c_j$, according to the model definition, the M_2 sink node covers the sensor node i, At this time the whole network coverage times set changes, $C \neq C'$. If $c'_j > c_j$, then:

$$1 - P^{c'_i} > 1 - P^{c_i} \Rightarrow h_i\left(1 - P^{c'_i}\right) > h_i(1 - P^{c_i}) \Rightarrow$$
$$\sum_{i=1}^{N} h_i (1 - P^{c'_i}) > \sum_{i=1}^{N} h_i (1 - P^{c_i}) \tag{13}$$

Through the formula (13), the inequality $Demands(M_1) < Demands(M_2)$ was established. In the assumption (1), $Demands(M_1) \leq Demands(M_2)$ was met combining (i) with (ii):

2. $M_2 > M_1 + 1$

    Because of $M_2 > M_1 + 1$, there must be M' which makes $M' = M_1 + 1$ and $M' < M_2$. We can get $Demands(M_1) \leq Demands(M_2)$ from the proof of formula(1). From the analysis of basic model above we can get that if $M_2 > M_1 + 1$ then $Demands(M_1) \leq Demands(M_2)$.

Combining the proof between (1) and (2), if $M_1 > M_2$ then:

$$Demands(M_1) \leq Demands(M_2)$$

Theorem 1: It shows that when the number of sink node increases, the maximum coverage demands was in growth trend, which can make service performance become more superior.

Theorem 2: In OSPSA model, if the failure rate in the network increases from p_1 to p_2, then the maximum coverage demands meet $Demands(Sp_2) \leq Demands(Sp_1)$. Assumed that wireless sensor network contains N nodes and M sink nodes; when the failure rate was p_1, the optimal position set of the sink node was L= {$l(1)$, $l(2)$,…$l(M)$}, the maximum coverage demands was $Demands(SP_1)$; When the failure rate was p_2, the optimal position set of the sink node was $L' = \{l'(1), l'(2), \cdots l'(M)\}$, the maximum coverage demands was $Demands(SP_2)$.

When the failure rate was p_1, $Demands(SP_1)$ was the maximum coverage demands. The corresponding optimal layout scheme was L = {/(1), /(2),... /(M)}, So under other scheme the coverage demands must be less than or equal to $Demands(SP_1)$. In other words, the optimal position set $L' = \{l'(1), l'(2), \cdots l'(M)\}$ was used when the failure rate was p_2. The coverage demands met $Demands(P_1) \leq Demands(SP_2)$ when the network failure rate was p_1.

Because $Demands(P_1)$ and $Demands(SP_2)$ corresponding to the position set was consistent and the failure rate meets $p_1 < p_2$, in the network H_M represents demands covered by M times. We can get the maximum coverage demands Demands(p) by substituting $H_1, H_2,... H_M$ into the model:

$$Demands(P) = H_1(1 - P) + H_2(1 - P_1^2) + \cdots + H_M(1 - P^M)$$

From $p_1 < p_2$, we can get that:

$$Demands(P_1) - Demands(P_2) = [H_1(1 - P1 + H21 - P12 + ... + HM1 - P1M - [H2H11 - P2 + H21 - P22 + ... + HM1 - P2M = H1(P2 - P_1) + H_2(P_2^2 - P_1^2) + ... + H_M(P_2^M - P_1^M) > 0$$

So $Demands(P_1) \leq Demands(P_2)$ was founded.

From Theorem 2, the service capability of wireless sensor networks within the cover radius was subject to the constraints of the failure rate, the failure rate increased and the services declined.

Theorem 3: In OSPSA model, if the coverage radius increases from D_{c1} to D_{c2}, then the maximum coverage demands met $Demands(D_{c1}) \leq Demands(D_{c2})$.

Proof: Assumed that wireless sensor network contains N nodes and M sink nodes; when the coverage radius was D_{c1}, the coverage set was C = {c_1, c_2,... c_N}, the optimal position set was L = {/(1), /(2),... /(M)}. When the coverage radius increases to D_{c2}, the coverage set was $c' = \{c', c', ... c'_N\}$, the optimal position set was $L = \{l'(1), l'(2), \cdots l'(M)\}$.

Because the coverage radius increases from D_{c1} to D_{c2}, it can lead the coverage status of the sensor node to change. We divide into 2 categories for discussion:

1. If the network does not exist any one sensor node k, whose coverage status changes due to the increase of the coverage radius, then the entire network coverage set will not change, that was $C = \acute{C}$. Therefore, the success probability of service within the coverage radius of each node

will not change, that was $1 - P^{C'k} = 1 - P^{C_k}$, $K = 1, 2, ..., N$. Thus, if the coverage demands of each sensor node in the network are summed up, we can get:

$$\sum_{K=1}^{N} h_K (1 - P^{C'k}) > \sum_{k=1}^{N} h_k (1 - P^{c_k})$$

That was,

$$Demands(D_{c1}) = Demands(D_{c2})$$

2. If the network exists the sensor node K, when its coverage times changes, the coverage set will change too, that was $C \neq C'$. According to the model, not only the coverage set may vary with the increase of the coverage radius, but also the optimal location set may also be affected.

i. If the optimal position set does not change along with the change of the coverage radius then $L \doteq L$. Because of $C \neq C'$, there was at least one sensor node K which makes $c'_k > ck$. Thus, the success probability of service within the coverage radius meets $1 - P^{c'k} > 1 - P^{c_k}$. Both sides of the inequality multiply node demand number h_k and then sum, $\sum_{K=1}^{N} h_K (1 - P^{C'k}) > \sum_{k=1}^{N} h_k (1 - P^{c_k})$, still found. That was the maximum coverage demands exists the relationship $Demands(D_{c1}) < Demands(D_{c2})$ when the coverage radius was D_{c1} and D_{c2}.

ii. If the optimal position set changes along with change of the coverage radius then $L' \neq L$. L' was the optimal layout scheme when the coverage radius was D_{c2}, so under other scheme the coverage demands must be less or equal to $Demands(D_{c2})$. At the same time, according to the evidence (1), because $Demands(D_{c1})$ and $Demands(D_{c2})$ corresponding to the location schemes are L and the coverage radius meets $D_{c1} < D_{c2}$, $Demands(D_{c1}) < Demands(D_{c2})$ was founded. Combining these 2 points in the current assumption, the inequality $Demands(D_{c1}) \leq Demands(D_{c2})$ was founded.

From (i) and (ii), we can get that when the coverage set changes, $Demands(D_{c1}) \leq Demands(D_{c2})$ was founded.

When the coverage radius increases from D_{c1} to D_{c2}, there was $Demands(D_{c1}) \leq Demands(D_{c2})$ between the corresponding maximum coverage demands.

In Theorem (3), the coverage radius influences the service ability of the wireless sensor network. Namely the coverage radius was bigger; network service ability was possibly more formidable.

Table 1: Nodes location and the demand number

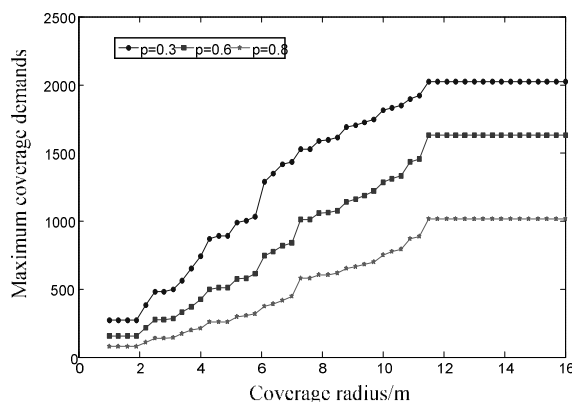| No.	(x_i, y_i)	h_i	No.	(x_i, y_i)	h_i	No	(x_i, y_i)	h_i	No	(x_i, y_i)	h_i
1	(1, 7)	51	11	(4, 10)	84	21	(7, 19)	71	31	(8, 16)	87
2	(2, 6)	2	12	(10, 3)	30	22	(13, 19)	81	32	(12, 1)	64
3	(1, 8)	33	13	(4, 0)	48	23	(6, 14)	53	33	(19, 8)	27
4	(17, 1)	75	14	(12, 7)	6	24	(9, 13)	51	34	(17, 1)	39
5	(8, 6)	80	15	(20, 7)	10	25	(14, 17)	76	35	(3, 14)	35
6	(6, 4)	41	16	(9, 20)	52	26	(10, 1)	85	36	(17, 7)	32
7	(15, 11)	41	17	(16, 18)	73	27	(19, 7)	95	37	(16, 11)	96
8	(8, 17)	26	18	(9, 9)	5	28	(4, 2)	10	38	(7, 8)	55
9	(17, 5)	10	19	(5, 19)	99	29	(11, 7)	55	39	(15, 3)	71
10	(7, 12)	61	20	(13, 10)	51	30	(10, 0)	91	40	(1, 1)	89

Fig. 1: Relationship between critical distance and maximal demands coverage
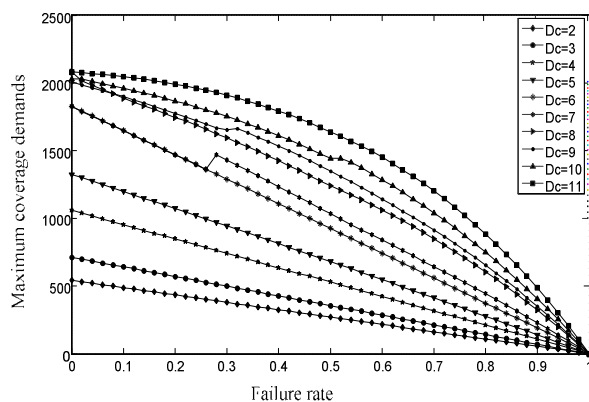

Fig. 2: Relationship between failure rate and maximal demands coverage

SIMULATION ANALYSES

In order to study the relationships between the failure rate p, the coverage radius D_C and the maximum coverage demands, we can assume that in a wireless sensor network monitoring object was the region of $20 \times 20 \, m$; The network consists of 30 sensor nodes, their location and demand number distribution was shown in Table 1; Preset sink node was 3; the coverage radius D_C and the failure rate p are variables.

When the failure rate of sink node was respectively $p = 0.3$, $p = 0.6$ and $p = 0.8$, the relationship between the coverage radius and the maximum coverage demands was shown in Fig. 1.

In Figure 2, the curves change trend was similar between the coverage radius and the maximum coverage demands, the maximum coverage demands increases along with the increase of the coverage radius. This study in the sink node location optimization model increases the coverage number of sensor nodes as far as possible. Multiple coverage makes the possibility of service increase within the specified distance, so as to protect the network global service capability. Therefore, when sink node was 3, the sensor nodes in the network
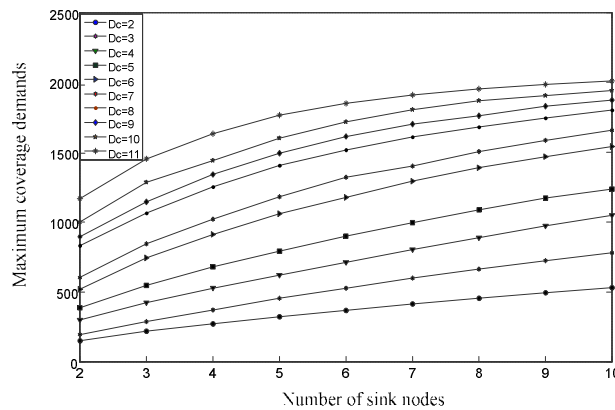
Fig. 3: Effect of varying sink nodes' number on maximal demands coverage
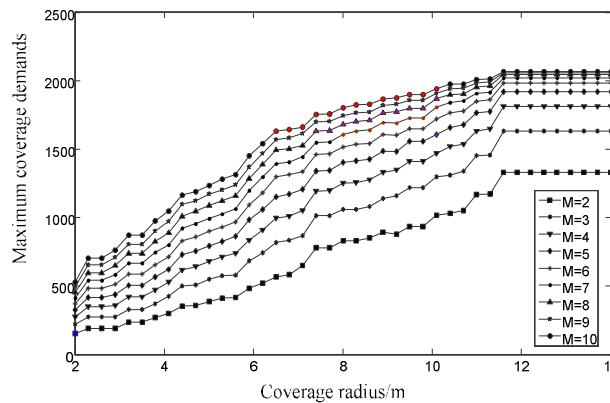

Fig. 4: Effect of varying critical distance on maximal demands coverage

may be in four states: uncovered, 1 times covered, 2 times covered and 3 times covered. Meanwhile, the variety of state of the node demands are different, the state which demands was the most make the curve show the corresponding feature. When *p* was 0.6 and D_C was 7.3, double coverage begins to occupy the leading position in the network performance influence factor. That the coverage radius was less than 7.3 mainly shows one coverage feature. Similarly, when D_C was 11.5, almost all nodes are overwritten 3 times. Increase the coverage radius will not improve the network service performance D_C = 7.3 was a sensitive point, network service performance was very sensitive near the point; small changes can lead to substantially improve the service performance.

The changes of the maximum coverage demands can be a reflection of changes in the performance of wireless sensor networks. That the failure rate effects on the maximum coverage demands reflects the restrictive relationship between the failure rate and network service performance. Setting different coverage radius, the

relationship curve of failure rates and the maximum coverage demands was shown in Fig. 2. 10 curves in the figure respectively represent the relationship between the failure rate and the maximum coverage demands.

From Fig. 2 that the maximum coverage demands decreases along with the increase of the failure rate, which we can find if the failure rate was high, sink node position optimization cannot effectively enhance network service performance; when the failure rate was very low, the network service performance will not improve due to the sink node position optimization. This shows that there was a favorable failure rate interval for guiding the sink node position optimization. In this interval, the sink node position optimization algorithm can be very good to play multiple coverage advantages.

As for Fig. 2, the failure rate interval was [0.3, 0.7]. If $p \in [0, 0.3]$, the current state of wireless sensor network was relatively stable and reliable and the demands sent by the sensor nodes can be almost received by the sink node. Since global demands can be basically serviced, adding multiple coverage will not

have a great impact on network performance, so the sink node position optimization was not dominant; if $p \in [0.7, 1]$, network communication was in poor condition, the sink node position optimization was no advantage.

Assume that the failure rate of sink node was p = 0.6, preset sink node and the coverage radius D_C are variables and the coverage radius was from 2 to 10. We can get 10 curves which respectively represent the relationship between the sink node and the maximum coverage demands as show in Fig. 3. In which, the top curve represents the coverage radius was 11, the bottom curve describes the coverage radius was 2. Intuitive judgment, the 10 curves show a consistent trend, which was that the increase of sink node will bring the increase of the maximum coverage demands. It was different that the size of the coverage radius influences on the overall level of the maximum coverage demands, the larger the coverage radius, the more the maximum coverage demands.

Assume that the number of sink node was known, the influence of the coverage radius on the maximum coverage demands was shown in Fig. 3. The figure gives the relationship between the coverage radius and the maximum coverage demands in the different sink node number. In which, the number of sink node from bottom-up was 2-10. Figure 4 shows the effect of varying critical distance on maximal demands coverage.

It was not difficult to find that the number of sink nodes and the coverage radius will lead to the change of the maximum coverage demands and change laws are consistent. Along with the increase of the coverage radius, the phenomenon of multiple coverage was more and more obvious; the maximum coverage demands also increase. The maximum coverage demands are an important indicator to measure the ability of network service. Its growth also means the improvement of service performance in the wireless sensor network.

CONCLUSION

In wireless sensor network, the selection of the base station location has a significant impact on the quality of network service. So the optimal sink position selection algorithm was proposed in this study. Anglicizing communication failure probability in a certain case, node communication demand was maximum satisfied by multiple coverage of the key node. Namely base station location optimization can cover up a maximum number of the node demands. On this basis, the characteristics are analyzed in theory and simulation experiments verify the validity of the model OSPSA.

REFERENCES

Anastasi, G., M. Conti, M.D. Francesco and A. Passarella, 2009. Energy conservation in wireless sensor networks: A survey. Ad hoc Network, 7(3): 537-568.

Ganeriwal, S., L. Balzano and M. Srivastava, 2008. Reputation-based framework for high integrity sensor networks. ACM. T. Sensor Network., 4(3): 1-37.

Jia, J., J. Chen and G.R. Chang, 2008. Efficient cover set selection in wireless sensor networks [J]. Acta Aut. Sinica, 34(9): 1157-1162.

Li, F., W. Xu, X. Liu *et al*., 2008. The energy efficient real-time cluster routing protocol in wireless sensor-actuators networt [J].Computer research and development, 45(1): 26-33.

Niu, R., P.K. Varshney and Q. Cheng, 2006. Distribured detection in a large wireless sensor network [J]. Int. J. [3] Inform. Fusion, 7(12): 380-394.

Priyadip, R. and P.K. Varshney, 2008. Distributed detection in wireless sensor networks using dynamic sensor thresholds [J]. Int. J. Distrib. Sensor Networks, 4(1): 5-12.

Tang, W. and G. Wei. 2010. Energy efficient base station location optimization algorithm in wireless sensor networks [J]. J. Commun., 31(9): 1-7.

Xiao, B., C. Hekang and S. Zhou, 2008. Distribute localization using a moving beacon in wireless sensor networks [J]. IEEE T. Parall. Distr., 19(5): 587-600.

Younis, O. and S. Fahmy, 2004. Heed: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks [J]. IEEE T. Mobile Comput., 3(4): 660-669.

Yuan, X. and W. Yang, 2010. Decision fusion of dynamic large-scale wireless sensor networks [J]. J. Eectron. Inform., 32(1): 2976-2980.