

A New Range-Free Localization Algorithm in Wireless sensor networks Based on Improved Genetic Algorithm by Control on Adrenaline Levels

Zhuang Liu, Zhiyi Fang
College of Computer Science and Technology
Jilin University, Changchun
130012, P.R China
lz1227@live.cn



ABSTRACT: In this paper, a new method is proposed to improve genetic algorithm (GA) based on control on adrenaline levels and a new range-free localization algorithm is designed based on the new method to increase the accuracy of the localization in wireless sensor networks (WSN). We choose the classic localization algorithm Centroid Algorithm, to perform experiments to demonstrate the localization accuracy of the new algorithm. The result shows that the new algorithm we proposed provides better localization accuracy and it can optimize the localization algorithm in WSN.

Categories and Subject Descriptors:

C.2.1 [Network Architecture and Design] Wireless: I.1.2 [Algorithms]

General Terms: Steganography, Information Hiding, Image Tamper Detection, Dual Watermarking

Keywords: Wireless Sensor Networks, Localization Accuracy, Control on Adrenaline Levels, Genetic Algorithm

Received: 11 August 2011, Revised 21 October 2011, Accepted 27 October 2011

1. Introduction

The application of WSN has become wider and wider, most of which are location dependent. Localization algorithm can be divided into two basic types according to its positioning methods, one is range-based and the other is range-free. Since the range-based localization algorithm is not accurate enough and has high requirements on energy consumption and hardware, it does not meet the requirements of the development of WSN [1].

Since range-free assumes that a small part of the sensor nodes, which are called beacon nodes, are aware of their real coordinates, and other unknown nodes can estimate their position coordinates based on the beacon nodes as well as certain localization algorithm. [2]

Up till now, most of the mainstream algorithms rest on range-free localization algorithm, including the following algorithms:

Centroid localization algorithm [3] uses the centroid coordinate of the selected beacon nodes as the estimated coordinates of the unknown nodes. This algorithm requires a small amount of computation cost; however, it is bad in localization accuracy;

DV-hop localization algorithm [4] firstly broadcasts among beacon nodes, and then estimates the average distance of each hop by the total number of hops as well as the distance from beacon nodes coordinates. Each beacon node broadcasts the estimated distance of each hop to unknown nodes, which further estimate their distances from the elected beacon nodes according to hops and average distance, and thereby calculates their coordinates using basic trilateration survey. This algorithm grants more localization accuracy, but demands higher communication cost and computation cost;

Approximate point in triangulation algorithm (APIT) [5] applies point-in-triangulation test theory, and calculates the intersection of all the triangles which contains certain unknown nodes, and thereby obtains the coordinate of the unknown nodes by calculating the center of the intersection; This algorithm provides better localization accuracy than the DV-hop algorithm, however, the existence of the InToOut and the OutToIn errors cannot be eliminated;

Improved APIT algorithm [6], which bases on triangle-gravity-center calculation and grid scanning, considers the triangle-gravity-center as the coordinates of unknown nodes. It can reduce the InToOut and the OutToIn errors by setting fiducial values to select legal nodes, which could be used to simulate the movement of unknown nodes. This algorithm can somehow reduce the InToOut and the OutToIn errors and it provide better localization accuracy than the APIT algorithm, but the improvement is not obvious;

Multilevel Coordinate Localization (MCL) method [7], which assumes coordinates of the beacon nodes are unknown as well, calculates coordinates of the beacon nodes using the same algorithm via the same coordinate system, and then compares the solutions with the real coordinate to find out the error, and thereby uses the error to correct the estimated unknown nodes' coordinates. This algorithm provides better localization accuracy than the DV-hop algorithm, nevertheless, it requires more computation and communication cost than the DV-hop algorithm;

And the improved DV-hop algorithm [8] of Li's as well is applied based on regularly moving beacon nodes and received signal-strength indicator. This algorithm selects one movable beacon node while assumes all other nodes as unknown nodes, and then locates all other nodes via their communication with the regularly moving movable beacon nodes. The advantage of this algorithm is that it provides better localization accuracy than the DV-hop algorithm, nevertheless, the disadvantages are that it considerably changes the hardware, and demands rather high communication cost and extra energy cost to move, and hence unsuitable for popularization.

In this paper, a new range-free location algorithm is presented based on the optimizing genetic theory of biology. This new method is an amendatory algorithm to improve the performance of the current localization algorithms, so that we choose the most classic and most brief localization algorithm, Centroid Algorithm, to perform experiments to demonstrate the localization accuracy of the new amendatory algorithm.

2. Description of the new idea

It is amazing that evolutionary diversification of dogs develops in biosphere. In the past 10000 years, the diversity of dogs all arose from one homogeneous population of wolves and is related to humans' interpose and choice.

Humans always need to establish a cooperative and concomitant relationship with dogs, so dogs must not be ferocious as wolves, and a necessary factor in dog's characteristic is submissiveness. However, dogs' characteristic depends on their genes.

Dmitri Belyaev, Russian geneticist, researched the reproductive process of wild fox in Siberia [9]. He chose the tamest ones from each offspring generation of fox cub and bred them. But by the tenth generation, there are different features appearing from wild fox, such as down ears, up tails, bark voice, white hair, and tender characteristics and so on. Belyaev checked the adrenaline levels of the latest generation, and found that the hormone that controls the response to external stimulation was far lower than wild fox. The adrenaline is on a biochemical pathway which has something to do with upper characteristics, especially for white hair. Therefore, it can be concluded that the control on the adrenaline levels

can influence the choice of genes and trigger a diversity of genetic variation. Thus, humans can change the direction of genes' choices simply by selecting better-looking individuals. This is the method of choosing genes' characteristics based on controlling adrenaline levels.

Therefore, in localization of WSN, we can also set the locations of nodes as the population of a new species (WSN nodes), and find out the parameters standing for "good-looking" feature, and then we can change the thinking of the localization algorithm.

We propose the improved genetic algorithm based on the control on adrenaline levels, and the parameter of "adrenaline-level-control" and the parameter of "appearance-control", and then we use them in the calculation of offspring generation of genetic algorithm to accelerate the convergence of the algorithm. In the next sections of algorithm description and algorithm implementation, we will explain what "adrenalin-level-control" and "appearance-control" stand for.

3. Algorithm description and analysis

We propose a new algorithm- the Improved Genetic Algorithm Based on Control of Adrenaline Levels: IGACAL.

Briefly, we regard the problem which we confront as a population in creature. According to the algorithm of specific problem, we find out the collection of population and run the algorithm to get the first generation of the population. According to the Genetic Algorithm (GA) [10], we set the population size, string length, probability of performing crossover, probability of mutation, convergence function and termination criteria. A better generation corresponding convergence function will be found out and we will survey the new generation based on the old generation to find out the parameter of control on adrenaline levels- "adrenaline-level-control", and parameter of appearance control- "appearance-control". Just like that people select more suitable ones from one generation of the population by choosing the individuals with beautiful hair, we run the original algorithm of the problem based on the new parameters we find out and raise one new generation of this population about this problem. Repeat the above process until the satisfying generation arises.

In accordance with Centroid algorithm, each unknown node will set the central of its neighbor beacon nodes' coordinates as its own estimated coordinate. Figure 1 shows that the Centroid algorithm can give rise to some mistakes in localization.

In figure 1, B1, B2, B3, B4, B5, and B6 are beacon nodes; A is the central coordinate of the six beacon nodes; but U is the reality coordinate of the right unknown node, therefore, the distance between U and A is the mistake of the localization.

Given a net-laying demand, with the control on adrenaline

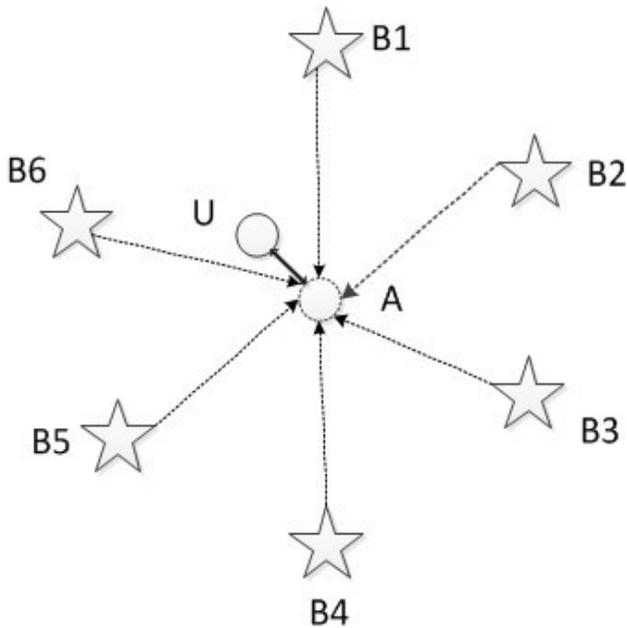


Figure 1. Mistake in Localization of Centroid Algorithm

levels theory, we do not know what the parameter of adrenaline-level-control and the parameter of appearance-control should be like at the beginning. Therefore, based on simulation experiments, we gain all of the estimated coordinates of each unknown node by Centroid algorithm, and devise the first generation of population through random variation on these estimated coordinates, and the population size is half of the amount of unknown nodes. In GA, the probability of performing crossover is standing for the probability of the genes exchange when two individuals reproducing and its value should be about 50%. The probability of mutation is standing for the probability of the genes mutation of the offspring individuals and its value should be about 0.1%. Through a mass of experiments and the experience of GA, we assume that the probability of performing crossover is 0.7, the probability of mutation is 0.005, the convergent function is the mean-square deviation between the estimated coordinate (formula 1) and the reality coordinate of the unknown nodes, and the termination criteria is that the mean-square deviation declines.

Formula 1:

$$f = \sqrt{\frac{\sum_{i=1}^n (x_i - x'_i)^2 + (y_i - y'_i)^2}{n}} \quad (1)$$

In formula 1, (x_i, y_i) is the estimated coordinate of unknown node, $i = 1 \sim n$; (x'_i, y'_i) is the reality coordinate of the unknown node, $i = 1 \sim n$; n is the amount of the unknown nodes; f is the mean-square deviation mentioned above. After a number of generations of GA process, one individual's mean-square deviation is less than that obtained from Centroid algorithm, and we find out that the sum of all the distances between each unknown node's estimated coordinate and its neighbor beacon nodes' coordinates is increasing.

Formula 2:

$$\left(\sum_{i=1}^n \left| \sum_{j=1}^m \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right| \right)_k > \left(\sum_{i=1}^n \left| \sum_{j=1}^m \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right| \right)_{k+1} \quad (2)$$

In formula 2, (x_i, y_i) is the estimated coordinate of unknown node, $i = 1 \sim n$; (x_j, y_j) is the coordinate of the neighbor beacon node of unknown node (x_i, y_i) , $j = 1 \sim m$; n is the amount of the unknown nodes, and m is the amount of the neighbor beacon nodes of one unknown node; k is the number of the generation of the population.

The statistics of the sum of the distances between unknown nodes' estimated coordinate and their neighbor beacon nodes' coordinates is shown in figure 2

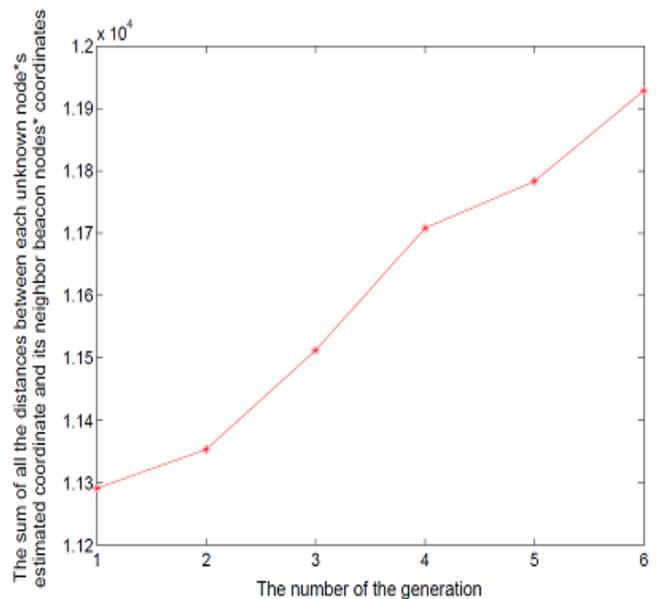


Figure 2. Distances Between Unknown nodes' Estimated Coordinate and their Neighbor Beacon Nodes' Coordinates

In figure 2, the horizontal abscissa represents the number of the generation of the population; vertical coordinate represents the sum of all the distances between each unknown node's estimated coordinate and its neighbor beacon nodes' coordinates.

From figure 2, the sum of the distances between each unknown node's estimated coordinates and its neighbor beacon nodes' coordinates of one qualified generation is greater than the previous generation. Therefore, we assume the parameter of adrenaline-level-control as the sum of the distance between each unknown node's estimated coordinates and its neighbor beacon nodes' coordinates in each generation of the population.

Based on the above, we need to choose more suitable and reasonable neighbor beacon nodes assembly and we can delete the existing neighbor beacon nodes or insert

virtual neighbor beacon nodes. We choose to insert some virtual neighbor beacon nodes. The sum of the distance between the unknown node's estimated coordinates itself and its neighbor beacon nodes' coordinates will be increased quickly when the estimated coordinate unknown node move to the furthest neighbor beacon node and far away from the nearer neighbor beacon nodes. Therefore, we insert a virtual neighbor beacon node nearby the furthest neighbor beacon node of the estimated coordinate unknown node. Assume that the unknown node is i , and the total number of neighbor beacon nodes is n_i , we assume n_i+1 as the parameter appearance-control.

Therefore, we gain the way to control the population genes by appearance-control. Focusing on the population of Centroid algorithm in WSN, we can use this appearance-control parameter- " n_i+1 ", to optimize the localization accuracy.

4. Algorithm implementation

Step 1: run a Centroid algorithm on WSN, and calculate the sum of all the distances between each estimated coordinate unknown node and its neighbor beacon nodes' coordinates and add them all, defined as d_1 . Operate random variation on these estimated coordinates obtained from Centroid algorithm to establish GA's original generation of the population and set half of the amount of the unknown nodes as the number of the population size. Calculate the mean-square deviation of the distances between unknown nodes and their neighbor beacon nodes of the best individual in the first generation, set the mean-square deviation as the convergence function, and set the probability of performing crossover, probability of mutation and termination criteria. Run GA.

Step 2: if the generation with smaller mean-square deviation is found, we calculate the sum of all the distances between each estimated coordinate unknown node and its neighbor beacon nodes' coordinates and add them all which is defined as d_2 . It is undoubtedly that $d_2 > d_1$, thereupon, we calculate the distances between each estimated coordinate unknown node and its neighbor beacon nodes' coordinates and rank them, and define the assembly of these distances as $\{D\}$.

Step 3: assume the appearance-control parameter as the number of each unknown node's neighbor beacon nodes, defined as n , set $n = n + k$, k is the adjusting value. In Centroid algorithm, $k = 1 \sim 3$. k is defined as 1 in this essay. Select neighbor nodes whose distance with the neighbor unknown node mentioned above is d , $d \geq \max\{D\} - m$, m is the distance parameter which controls the number of the neighbor beacon nodes to find virtual beacon node and establish a new assembly, in which p elements exist. Calculate the central coordinate of these p beacon nodes as the virtual beacon node and then insert it into the assembly of the unknown node's neighbor beacon nodes.

Step 4: insert all of the virtual neighbor beacon nodes into its neighbor unknown node and turn to step 1 until we find the individual which satisfies the termination criteria.

5. Algorithm simulation and comparison

The new algorithm proposed in this article performs simulation experiments via matlab7, and the CPU is Intel(R) Xeon® X3430@2.40GHz 2.53Hz, the memory is 4.00G, and system is Windows7.

First of all, experiments are performed in a 100m*100m area, 100 nodes and 30 beacon nodes are positioned randomly, with communication radius of 30m. 100 times of simulation experiment are performed.

Figure 3 shows that the new algorithm amended by IGACA has better accuracy on localization

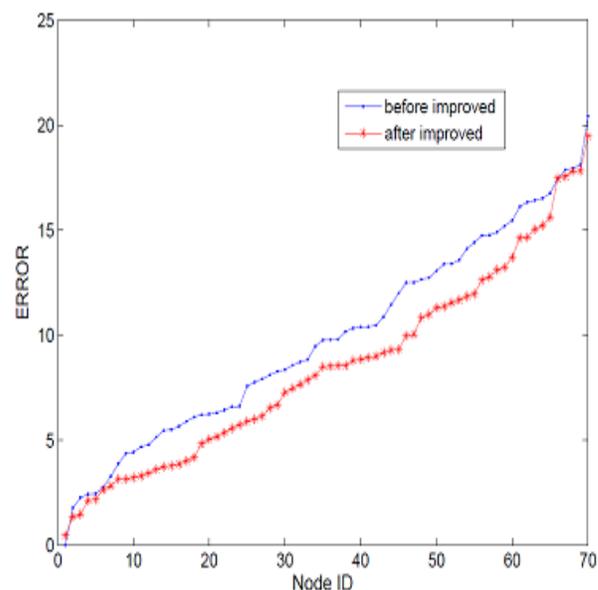


Figure 3. Comparison of Localization Accuracy between Improved and before Improved

In figure 3, the horizontal abscissa represents the node ID; vertical coordinate represents the statistics of distance error of the localization; red line represents the localization accuracy after improved; blue line represents the localization accuracy before improved.

We can conclude that the localization accuracy after improved is better than that before improved and the average improvement is $(9.8513-8.4841)/9.8513 = 13.88\%$. The number "9.8513" and the number "8.4841" stand for the "average error" before and after improving respectively.

"Average error" stands for formula 3:

Formula 3:

$$\left(\sum_{j=1}^m \sqrt{(x_{i1} - x_j)^2 + (y_{i1} - y_j)^2} \right) / m \quad (3)$$

In formula 3, (x_{i1}, y_{i1}) is the estimated coordinate of unknown node, (x_i, y_i) is the real coordinate of unknown node, and m is the number of the unknown nodes.

6. Conclusion and Future Application

We propose an algorithm- the Improved Genetic Algorithm Based on Control of Adrenaline Levels: IGACA. We take a localization algorithm, Centroid algorithm, in WSN for example, and design the amendment algorithm to apply. Simulation experiments verify that the amended Centroid algorithm enhances the accuracy of localization. IGACA can also be used to amend other algorithms in order to accelerate the convergence of the GA.

For example, in DV-Hop, the estimated distance in one hop can cause hundreds of times of error. In addition, the estimated distance of each hop of beacon node has error so that the more hops the more errors. We can use IGACA to choose suitable number of hops.

In APIT, the number of the triangles is the crucial factor which affects the accuracy of localization. The area acreage is too big when the triangles are few, and the more triangles, the more IntoOut or OuttoIn errors. We can use IGACA to choose suitable number of beacon nodes as the point of the triangles.

Reference

- [1] Akyildiz I. F, Su W., Sankarasubramaniam Y., et al. (2002). Wireless sensor networks: a survey [J]. Computer Networks, 38(4) 393-422
- [2] Limin, Sun., Jianzhong, Li., Yu, Chen., Hongsong, Zhu (2005). Wireless Sensor Networks [M]. Beijing: Tsinghua University Press.
- [3] Bulusu, N., Heidemann, J., Estrin, D. (2001). Density adaptive algorithms for beacon placement in wireless sensor networks [EB/OL]. (2001-05-15) [2005-12-11].
- [4] Niculescu, D., Nath, B. (2003). DV based positioning in Ad hoc networks [J]. Telecommunication Systems, 22: 1-4, 267-280
- [5] He Tian, Huang Chengdu, Brian M. (2003). Range-free localization schemes for large scale sensor networks[C] // MobiCom'03. San Diego: IEEE Press, 81-95
- [6] Zhou Yong, Xia Shixiong, Ding Shifei, Zhang Lei, Ao Xin, (2009). An Improved APIT Node Self-Localization Algorithm in WSN Based on Triangle-Center Scan [J]. *Journal of Computer Research and Development*. 46(4) : 566-574.
- [7] Zhiyi Fang, Zhuang Liu, Yongbo Ma, Hongyu Sun, Naiji Ren. (2011). A New Coordinate Correction Localization Theory and Its Implementation Mechanism in Wireless Sensor Networks. *JICS(Journal of Information and Computational Science)* Vol. 8 (1) 23-28
- [8] Ruixue, Li., Zhiyi, Fang., Tingting, Yi (2011). Improved DV-hop Localization algorithm based on regularly moving beacon (RMAN) and received signal strength indicator (RSSI) and its performance analysis. *Journal of Jilin University* (Engineering and Technology Edition). 41(2) 435-441.
- [9] Dogs, More Dogs [TV Programme], (2004). NOVA Education, Feb. 03, available from: http://www.pbs.org/wgbh/nova/teachers/programs/3103_dogs.html (Accessed Feb. 03. 2012).
- [10] Zhou Ming, Sun Shudong, (1999). Genetic Algorithms: Theory and Applications [M]. Beijing: National Defence Industry Press.