The analysis of the high reliability of the routing protocol based on wireless ad hoc networks in automation of production processes

Zhang Yuxue[*,1] **Shan Chen**[2]

[1]*Suzhou Vocational University. Suzhou,JiangSu, 215104 ,P.R.China*
[2]*Changchun University of Science and Technology, Changchun, JiLin,130022, P.R.China*

Abstract

In automation of production processes, we often establish wireless ad hoc networks, in which, the process of establishing high reliability has important significance for routing protocols. In this paper, some protocols which are more at less lack of reliability based on the current research caused by various factors are analyzed and discussed . On the basis of this, the paper proposes an improved reliability of wireless ad hoc network routing protocol, the key point of the agreement protocol design completed construction. With the simulation method to analyze the agreement, get the agreement with respect to the current can improve the reliability of results of various routing protocol 16.5%. Demonstrate the validity of the wireless ad hoc network routing protocol, to help to locate persons and objects better, and enhance the efficiency of testing, analysis and maintenance in automation of production processes

Keywords: WIRELESS NETWORK, AUTOMATION OF PRODUCTION PROCESSES, ROUTING PROTOCOLS, NODE-POSITIONING ALGORITHM

1. Introduction

Routing protocol is the key technology of wireless network applications. Because of the special nature of wireless ad hoc network application environment and the application conditions, the routing protocol must have robustness and high adaptability to adapt to the complex and changeable environment and constantly changing network structure. Location service protocol in mobile ad hoc network is the basis of the routing protocol, location allocation in ad hoc network routing protocol, the universal current routing protocols are analysed, the advantages and disadvantages of them, then according to the characteristics of current routing protocols were analysed. The final design of the virtual layer routing protocol based on geographical location, characteristics of the agreement is nodes do not understand their own position, for each node as cluster after cluster head node and gateway node is determined by the modified VHLS, and based on the VHLS algorithm to form their own virtual hierarchy. A node leave a cluster, cluster head sends the request to the corresponding cluster head, this information to other cluster for the first notice, so that all the nodes to the cluster node has left. the Node joins a cluster, cluster head sends the request to the corresponding cluster head, verify the legitimacy, authenticity, and then the same process. After the demise of the cluster head nodes, not to its Beacon in a certain time after retraction, confirmed the death, then the above processing. This paper proposed a routing protocol based on virtual geographic position with high security for the characteristics of the design, using the improved distance localization algorithm of the agreement. Through the simulation using NS2 software analysis, that the protocol can improve the ratio of security routing protocol based on virtual geographic location of the 20% traditional.

While the calculation speeds of the algorithm protocol does not decrease.

2. The Analysis of Routing Protocol Reliability

For example, in military applications, the wireless ad hoc network technology to combat network battle-

field communication system access to more advanced. In some cases, wireless ad hoc network and the end of terminal network access internet. Because the network using different routing protocols and transport layer service wireless ad, the need for the corresponding conversion in the internet. Routing protocol for wireless ad hoc network can be divided into table driven routing protocol (Table-driven), on-demand (On-demand) routing protocol of two kinds[1-3]. Table driven routing protocol (Table-driven Routing Protocol), we are also known as proactive routing protocol, is the routing protocol routing table[4]. Proactive routing protocol routing found a similar strategy with traditional routing protocols, routing information by broadcasting packets periodically exchange routing information, active node, route discovery. When the network topology changes, node sends a routing update message, receive routing updates to nodes will update their routing table, so as to ensure the routing information consistency, timeliness and accuracy. Typical proactive routing protocols for wireless mobile: OLSR (Optimized Link State Routing), DSDV (Destination Sequenced Distance Vector), WRP (Wireless Routing Protocol), STAR (Source Tree Adaptive Routing). On demand routing protocol (On-demand Routing Protocol), also known as reactive (Reactive) routing protocol, is when needed to find the routing selection. Node does not save the routing information timely and accurately. When the source node sends a message to the destination node , the source node by initiating route discovery process in the network, find the corresponding routing, began to send the message, in order to improve the efficiency, a node can be found for subsequent transmission routing saved in the cache usage. On demand routing is mainly composed of a route discovery (Route Discovery) and route maintenance (Route Maintenance) two parts. It has the advantages of routing information broadcast does not require periodic, save some cyber source. Drawback is that a data packet is sent, if not to a destination node routing, packets need to wait for the route discovery caused by delay. Wireless mobile reactive routing protocol typical: AODV (wireless ad hoc network On-demand Distance Vector), DSR (Dynamic Source Routing), TORA (Temporally Ordered Routing Algorithm). Routing protocol for wireless ad hoc network itself does not provide security mechanism, so the security problem has become one of the hot researches. So far, domestic and foreign researchers put forward many improved security routing protocol, to sum up, there are the following: SEAD, Arianna, DSDV, DSR, GLS, and GHLS etc. But these routing protocols are

characteristic of certain, is not applicable to the entire mobile ad hoc network. This study mainly used in mobile ad hoc network in large and complex, at this point, we particularly analysed VHLS related improvements for our application of the existing routing protocols.

The network location can be divided into two categories of wireless ad hoc, node self-localization and external orientation. The former is based on nodes with known locations, in accordance with a positioning mechanism (distance or non-distance) to determine the node position. The latter is calculated relative position information by external target information. External positioning method is currently the most widely used. the most mature of the positioning system is a global positioning system GPS, it has high precision, strong anti-interference ability, good real-time[5-8]. But it is restricted by the cost, power consumption, scalability and the use environment conditions, GPS cannot adapt to the wireless ad hoc network node localization,such as wireless nodes in the network have the characteristics of groups with limited resources, random deployment, communication, vulnerable to environmental interference or node failure, so positioning mechanism must satisfy the self-organization, robustness sex, low power consumption, distributed computing requirements; in the battlefield environment, GPS will be disturbed and unable to use. According to the characteristics of Mantes, the mechanism and algorithm realization self-positioning of nodes in wireless ad hoc networks is very necessary. Node self-localization method is currently mainly two ways in accordance with the distance or not to locate in accordance with the distance. Because there are many cannot solve the problem of external positioning mode, so the self-localization algorithm has been paid more and more attention. Although many localization precision is not high is not practical, so consider the composite application of the node localization technology in the US, set by adopting the composite positioning technology to design our routing protocol. A location algorithm based on high requirements for hardware, but the positioning precision is high; the range free localization algorithm has low requirement to hardware, but the positioning accuracy is not high. In order to pursue higher positioning accuracy while the hardware requirements are not too high, the corresponding appeared positioning algorithm distance independent and distance combination[9-10]. The node localization technology of this complex in the US, in order to obtain a better safeguard is the design of composite high security of our routing protocol.

3. The Design of Routing Protocol With High Reliability

The current representative algorithms are: Euclidean algorithm, a universal positioning algorithm, robust positioning algorithm, N composite positioning algorithm. The Euclidean algorithm: the algorithm requires nodes with RSSI ranging capability, based on the determined by unknown nodes and anchor nodes in 2 hops. First broadcast in the network anchor node beacon signal with its ID and position, and the TTL domain settings of the signal is 2, namely, the beacon can transmit only 2 hops. When a node is an unknown node can be from 2 known distance of each other and directly adjacent to the anchor node receives the signal, it can calculate the self and the anchor node distance. When unknown nodes obtain and distance of 3 or more anchor nodes after positioning itself. The universal positioning algorithm: the algorithm to specify the unknown node calculation and finally accepted the position estimation, and upgraded as anchor nodes conditions, in order to reduce the influence of errors. This algorithm includes the start and circulation in 2 stages. In the initial stage, all nodes exchange messages, if a node's neighbours are less than 3, is marked as solitary node, the anchor node labelled terminal node. Phase of the cycle, if unknown nodes in the neighbour nodes of non-isolated node number less than 3, is marked as solitary node; if the terminal node is greater than 3, is positioning itself, and according to the number of terminal nodes have positioning by consistency of position and neighbour node in the computation. The objective function va- lue is lower than the node all neighbour nodes to accept the position estimation, a new terminal node. If all of a node's neighbours have become terminal or isolated node, then the algorithm ends, otherwise the next cycle. Robust positioning algorithm: the algorithm is divided into three stages ranging, localization and refinement, in the distance, first use the distance vector protocol, so that all the nodes obtain network hops from anchor nodes. The second stage, the other position apart from anchor nodes and jump, computing network average hop distance of anchor nodes, and then used as a correction value is broadcast to the network. When receiving the corrected value, node calculation and anchor nodes according to the hop distance. N jump composite location algorithm: the algorithm using ultrasonic ranging technology, is divided into 3 stages, the first stage is the formation of cooperation: according to the criteria, sub trees to generate multiple constraints consisting of unknown node and anchor node integrity or super constraints configuration in the network, called collaborative sub tree. Each configuration including n variables (coordinates) and at least n are nonlinear equations. Is not in the sub tree contains node cooperative algorithm processing stage positioning. The second stage is a rough estimate of unknown node positions, then through the rough position of unknown nodes and anchor nodes distance to estimate the unknown node.

At present, the centric algorithm for convex programming algorithm, distance vector algorithm and APIT algorithm are the main representative algorithms. Centric algorithm: in this algorithm, anchor nodes periodically to the adjacent node broadcast the anchor node ID and location information. When the unknown node receives a packet from the different number of anchor node exceeds a certain threshold K or receives after a certain period of time, will determine the centric position for itself the anchor node polygon. Convex programming algorithm : the inter node communication point-to-point connection as the geometric constraint of node position, the whole network is modelled as a convex set, the node localization problem is transformed into a convex constrained optimization problem, and then use the semi definite programming and linear programming method to obtain a global optimal solution, to determine the node position. Distance vector algorithm: the algorithm between the distance between unknown nodes and anchor nodes in network, between the average distance of each hop and two hop product representations. The algorithm first uses distance vector protocol, so that all the nodes obtain network hops from anchor nodes; then, anchor nodes to calculate the average distance of each hop in networks, and its use flooding broadcast to network. So each node received only first values obtained, and discard all later, this strategy ensures that the vast majority of node can receive the value from the anchor node nearest. The APIT algorithm: the basic idea of the algorithm for unknown nodes listen all audible anchor nodes, and from these the anchor node any 3 non collinear anchor nodes constitute a triangle, the approximation test method in determination of unknown node is in a triangle, triangle test all combinations, we can determine the multiple contains unknown node triangular region. The intersection of these triangles is a polygon, it identifies contains unknown node region smaller. This region can be a grid scanning algorithm, grid scanning algorithm and triangle intersection compared with geometric method, can reduce the calculation, and reduce the amount of calculation. Then, calculate the centric of the polygon, and the centric as the estimated position of the unknown node.

In location based routing protocol, nodes by GPS

method to obtain their own position, get obtaining-information node location by location service. The location of the neighbour nodes can jump through the radio to obtain, this belongs to the neighbour information to maintain the contents of the agreement. According to participate in the nodes of location-based services, location-based services can be divided into «part of the node» in the location of the service and the «all nodes involved» in the location of the service. According to the position of each location server maintenance information can be divided into «part of the node position» and «all node position». Both combined, location services can be divided into the part of the location of the nodes involved in part of the service, the position of the nodes involved in all of the parts service, and all the position of the nodes involved in part of the service, all nodes involved in all the four kinds of the location of the service. In know themselves and the location of the other nodes; node packet forwarding could be carried out. Pa- cket forwarding strategy can be divided into, greed message forwarding. Node forwards the packet to a neighbour near obtaining-information direction; the neighbour node choice depends on the optimization criterion of the algorithm. Node forwards the packet to multiple jump neighbours near obtaining-information direction and decentralized packet forwarding. Location services is the foundation of the location based routing protocols. At present there are mainly centralized location services, distributed services, grid location and grading service four location services. Below is the introduction respectively. Centralized location service: in a centralized location service network, all nodes with the location server to interact, to update their location information, or get other node location information. This type of location service network has a hot; hot spots will bring robustness of location service. Distributed location service: the location of each node to maintain a position library, library contains all the nodes in the network location information. Location library table of contents is with node identifier, relative to the direction and distance of the nodes and the table of time. Each node by flooding position «message» to other nodes is updated the location of the library. The message of location and each includes a TTL field survival time; the transmission distance is used to control the message. This position by flooding message to update the location of each node repository is the basic idea of whole distributed location service. Different protocols have its specific strategy Location services in some routing protocols that node distance between effect, namely, the farther the distance of two nodes, the greater the relative mo-

bility of each other; It is thought that the mobile node have autonomy, namely nodes location update trigger depends only on the mobile movement rate. So, in their location in the message, the location of the TTL value low message sending high frequency; In addition, according to the mobile node of autonomy, the routing protocol based on mobile rate to dynamically adjust the frequency. In the grid location services, the network is divided into a grid. Grid is defined as the smallest grid order is 1, the n order grid is a grid of $(n + 1)$ order a quarter. Node location information is distributed in the networks of multiple servers. Location server is not specified, but according to certain algorithm is selected, this avoids the server's election, the design and implementation of the agreement. The realization of the grid location service mainly includes the following aspects: the choice of the location server, GLS through a hash function, the nodes of uniquely named mapping for a globally unique ID. And define a node :closest node is greater than the smallest ID of the nodes in the node ID. A node A to have a place in each grid server, and the location server is the order grid nodes of the closest A; Limited position update, each node in 2 HELLO message broadcast. HELLO message includes all node ID, location, the average speed of mobile, obtaining neighbour ID and position, etc. In this way, the node can be learned within two jumps within the neighbour node ID and the change of location information. In practical process, the UPDATE message is made up of the nodes in each grid closest one by one order to forwarding; Location query: GLS server selection algorithm ensures the location query of the selected server and location update of the selected server is the same. When a node S need the location information of query node D, it would gradually nodes D closest order query, and eventually get the position information of nodes D. Hierarchical network location service will divide the area into alternating each other. Area at the design stage to determine, not dynamic, the size of the area depends on many factors, such as node mobility, node density, the transmission power and wave propagation characteristics, etc. Of course this also can be divided into simple geographical division. Some routing protocols in the location service network topology is divided into two levels, one is the node level, regional level. To achieve classification of link-state routing protocols, each node supports two types of control packets, namely node link status messages and regional link status message. These routing protocols include nodes all the neighbours, and the first broadcast in the area. The ID of the list also includes node is connected area, and using the method

of flooding the entire network broadcasting. Through these two packets, nodes can maintain two routing table, routing table is an area, through short circuit by the algorithm; Two is a regional routing table.

4. The Modified Routing Protocol Based on Virtual Geographic Locatioin

Due to the geographical position is through the information packet forwarding, location based routing protocols do not need to use the route table lookup. In the case of know purpose of node position, the current node can forward position, so the forward delay compared with the passive on-demand routing protocols is smaller; Don't need to maintain the routing table stateless. Because there is no routing table maintenance, so the location based routing protocol is a stateless routing protocol, it has good expansibility, can better adapt to a large-scale network, also more suitable for network topology changes frequently. The downside is that the location service need extra overhead. Location based routing protocols rely on location services, and the realization of the network location service requires extra overhead. This is the routing protocol has to focus on the problem of the consideration. But under the condition of the network size is not too big, if adopt the centralized location services, is a lot of implementation costs to relatively small; the application of location information relatively drab. Now, most of the location based routing protocols are distance as an application of geography information and as a basis for the forwarding of a message, the greedy forwarding unstable link, etc. Based on regional classification of link-state routing (ZHLS) protocol to divide network into mutually overlapping area (zone), and for each area are identified using the unique ID number, this is a bit like GLS agreement. Area at the design stage to determine, not dynamic, the size of the area depends on many factors, such as node mobility, node density, transmission power, and radio waves propagation characteristics, etc. This can be divided into simple geographical division, also can be divided based on wave propagation. Node by GPS method so as to obtain their own location, and on the basis of calculating its regional identity.

Based on the combination of blind spots to locate and blind area routing algorithm using the distribution of local neighbours Angle information, through local search algorithm to locate the current node by greedy algorithm is unable to be local blind area. Then through several times of neighbour information exchange and blind area of information fusion, grasp the current node global blind area information. When we check out after all blind spots in the network and

the corresponding blind area, can set up the network system of blind area. In the blind area system based on the greedy routing algorithm, take advantage of the blind area information to guide the routing node hands, bypass the objective point in its corresponding blind area in the blind spot, to be able to in a timely manner to avoid routing hole, avoid cavitation's routing. This algorithm can efficiently guarantee the reliable data transmission, the number of hops, load balance, communication overhead and so on, this algorithm is superior to the previously proposed some geographical routing protocols. In choosing the reference node, a node from its neighbouring reference node reference information, when it obtained a series of reference nodes information, it will use a method to sort these reference node classification. Show their residual values and reference node selecting the lowest surplus value of reference node under the algorithm a stage in the application. For a coordinate is X, Y, Z node of surplus value is defined. Its surplus value is equal to that point to the distance of each reference node and the reference node of the square of the difference standard coverage. In other words, the surplus value is the reference node to the node and the reference node is calculated distance of different area of the sum of the radius. From the initial reference node, the value of the surplus value should be close to zero. Composite distance measurement, after a series of reference nodes when the node selection, it will be collected from each of the reference node distance. For some measurement technology, when a node receives the news from the reference node, it needs some location information should be included in the reference node. For a precise measurement technology, the node can collect information of a series of samples, until the samples information changes tend to be 0. Node can use the information at this time. the sample, that is, the distance between the node and the reference node. Nodes can also refuse to apply some reference node, the reference node information provided sample is unstable. Position estimation, it is up when a node, the distance to the reference node position can be estimated out of its own. In this method, it is using linear equation and Taylor series. This method is through Taylor expansion formula, to approximate nonlinear equations into linear equations. Parameter is obtained by linear iteration needs. But, unfortunately, in some cases, this approach will produce great error; the result of vibration amplitude is very large. Many kinds of technology are applied to reduce the happening of this kind of situation. Carefully choose the start of system parameters. Use can accept parameters, such can make the system reach

the balance of a static. Cycle phase, if the position of a node to obtain a reasonable budget, it can be in the agreement as a new reference node, node need most now is through the limit standard of reference nodes. As a result, a large part of the nodes are in position to get a good budget, budget only more accurately node position, in this way, can enlarge the coverage of network, and can avoid the generation of large error. New reference node is added to the known exact location nodes, at the same time, improve the position precision of the budget. By the above methods, and constantly budget precisely the locations of some nodes, these already know the precise location of the nodes and can be used as a reference node in the network application. In this way, through continuous cycle, can get accurate location, most of the nodes on the network in the application of wireless sensor network can be a very important role. That is we are going to use the basic principles of compound node location technology. Neighbourhood, the blind area information fusion process is shown in figure 1.
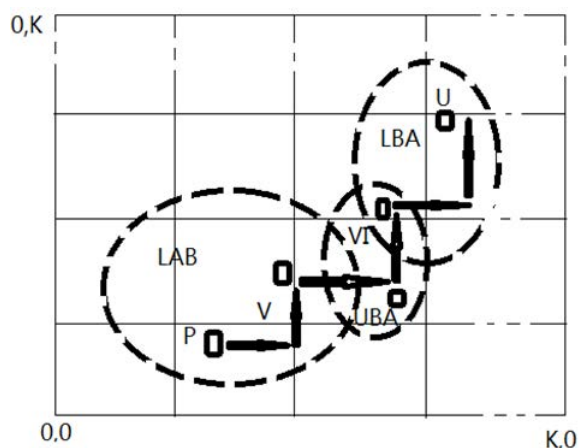

Figure 1. The blind area information fusion process

Assuming that the current node p in the picture, after localization blind area search algorithms local blind area of p is LBA (p). Through the blind spot of information exchange between neighbours, mastered the jump after a neighbour's blind area information, the itch a neighbour's local blind area for the LBA (I). And V connection UI and perpendicular bisector of VI, and p between two perpendicular bisector of communication beyond the range of the regional flag is Fi. Blind area in the first time in the process of information exchange and fusion, when the objective point in Fi area, a neighbour of the current node p that I must jump neighbour distance objective point nearest node. Point p will inevitably by point I forwarding packets. , therefore, the case of a neighbour's local I blind area LBA (I), if there are areas in the Fi, the fall in the centre portion of the Fi will also become blind area of p, namely the point p on the direction of the I is

blind spots for BA (I) the door Fi. Calculate all jump neighbours' blind area of p, p of the blind area should be (BA(I) the knife) U (BA (2) the door F2)... U (BA (I) knife)... U (BA) (n) nine UBA don't). At this point, the point p of the range is extended to the blind spot information of two jump range. At the same time, point p a jump neighbour's blind area of records may also be changed, so the blind area information exchange and fusion between neighbours again. In this way, each node of local blind spot information will spread out gradually, and tend to be convergent, ultimate stability after the blind spot of information is recorded in the global blind area. Every node in the nodes of the blind area is actually a network of local convergence blind area and fusion. Each node to master himself, and a jump after the neighbour node blind spot information, you can clearly understand what area in the network is unreachable by greedy algorithm, which area are through which unreachable neighbour nodes using the greedy algorithm. Methods is to divide network into k x k grid, each grid computing is the blind spot of each node is located in scope. Excluded from the blind spot information recorded in a blind area of each node, the size of the blind area table for K2hit, each bit marked the corresponding grid is blind area, and if the blind area, the position 1, if it is not blind area, then the position 0. Each node holds a blind area table and table jumped all neighbours. Such a record method is simpler, and pay the set operation is very convenient, also can be a very good save the storage space. However, the size of the grid can bring certain error to blind area of judgement. In the ideal situation, i.e., k is large enough, so that there is only one in each grid node, so that you can accurately know if each node is located in the blind spot of the current node.

5. The Simulation Analysis

On the one hand, in view of the simulation protocol details, need an efficient processing bytes, group head and run on a large number of data sets algorithm of a language system, in this case, the speed is very important, and c + + language run very fast. on the other hand, most of the network simulation research is need to change the network parameters or configuration, or quickly research a series of scenarios. In this case, repeatedly change the parameters or configuration of the time is very important, tall language although run slower, but provide the flexibility to change parameters and configuration, greatly saves network simulation parameters change or read the time of different configuration. NS2 to use a combination of two languages, in the network movement scene change (especially in Ad Hoc network) and speed of data processing has achieved a certain optimization.

Network researchers by writing Tall script to invoke the simulation element simulation in Octal class library.

In the composite simulation test for the algorithm of blind spots to locate and blind area, we adopt is the same with compound node positioning simulation software, namely the NS2. The main parameters of the simulation environment for: 800 m by 800 m square area randomly distribute the 500-2500 nodes, average node degree of change in the range of 5 to 25. Network UDG conditions, MAC layer USES the 802.11 protocol, wireless transmission radius of 5 m. In the 100 s of the simulation time, randomly selected node to send 100 packets. Each set of data is measured by the five randomly generated network data and take the average. Comparing the routing protocol for the current application of more GPSR routing protocol is proposed. By comparing the performance of their security is to verify the effectiveness of the combined location routing. Face under complex blind area to locate and blind algorithm simulation test,The set to the packet to discard the packet routing is empty. We in the random distribution of network nodes, artificially set an empty, empty size is 200 x 200 m^2. In GBIR algorithm, we choose the network is divided into 8 *8 * 16, 32 * 32 * 64, comparing the small pieces of the grid side length respectively communication radius 2 times, 1 times, and half a quarter. As the web hole increases, the number of hops increases GPSR is very obvious. However, the algorithm on the number of hops is hardly affected by the size of the hole, only related to the network density. It is enough to prove the security of this compound algorithm in routing security performance is better.

6. Conclusions

In view of the wireless ad-hoc network routing technology used in automation of production processes and its location based routing service technology development present situation, key technologies and development trend of question for study. Then in wireless ad-hoc network routing protocol based on virtual layer location service based on the safety of the compound design, this part includes the location of the current problems of service technology, the design method of composite and the process, the innovation embodied in the combined location routing technology, the algorithm design and simulation verified the effectiveness of the compound agreement. Another key of the article is for the combined location routing security policy research, in this section is mainly for composite location based routing protocol of the basic principle and design idea were studied.

Through the subsequent simulation test, further evidence show that the combined location routing protocol is of availability and security.This can enhance the efficiency of testing, location, analisis and maintenance heavily in automation of production processes

References

1. Hofmann T , "ACM Transactions" , Information Systems,vol. 22(1),pp.89-115, 2004
2. Golub G, Kahan K , "Calculating the singular values and pseudo-inverse of a matrix", the Society for Industrial and Applied Mathematics, vol. 2(2) pp.24-25, 1965
3. Lee D D, Seung H S , Nature 401(6755) pp.91-788,1999
4. Lee D D, Seung H S , "Algorithms for non-negative matrix factorization Advances" , Neural Information Processing Systems, pp.56-62, 2000
5. Salakhutdinov R, Mnih A, "Probabilistic matrix factorization Advances" ,Neural Information Processing Systems ,vol.20(3) ,pp.43-51, 2008
6. Mislove A, Marcon M, Gummadi Krishna P, Druschel P, "Measurement and analysis of online social networks Proceedings " in Proc. 7th ACM SIGCOMM Conference . Internet Measurement. San Diego, CA, USA, 2007, pp.29-42,
7. Blei D M, Ng A Y, Jordan M I , "Latent Dirichlet allocation" , Machine Learning Research ,vol.3(3) pp.993-1022, 2003
8. Shah H H, Banerjee , "A Bayesian co-clustering In Altman R. ed".in Proc. the ICDM 2008. Washington: IEEE Computer Society Press, 2007,pp. 53-90
9. Sajda P, Du S, Parra L, et al , "Recovery of constituent spectra using non-negative matrix factorization"in Proc SPIE, 5207, pp.21 -31, 2003
10. Liu N N, Yang Q , "Eigenrank: A ranking-oriented approach to collaborative filtering" in Proc. the SIGIR 2008 83-90. Available at: http://www.cs.ust.hk/~qyang/Docs/2008/SIGIR297-liu.pdf