Journal of Theoretical and Applied Information Technology
10th April 2013. Vol. 50 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

USING SEMI-SUPERVISED CLUSTERING ALGORITHM TO CONSTRUCT THE INTELLIGENT TRANSPORTATION SYSTEM

[1]WENJUAN TANG, [2]FENGTING ZHENG
[1]lecturer, ShangQiu Polytechnic, Henan ShangQiu , 476000, China
[2]lecturer, ShangQiu Polytechnic, Henan ShangQiu , 476000, China
E-mail: [1]edutangwenjuan@163.com

ABSTRACT

Intelligent Transportation is a modern electronic-based IT service-oriented transportation system. The prominent features of the information collection, processing, release, exchange, analysis, use the main line, to provide diversity of services for the traffic participants. Semi-supervised clustering is to use some data on the type of mark or constraints to aid the process of non-supervised clustering. The paper presents using semi-supervised clustering algorithm to construct the intelligent transportation system. The algorithm uses the tag data and constraints to meet the objective function of the clustering results. Experiment results show that the Constrained K-means algorithm can better than Seeded K-means algorithm in building ITS.

Keywords: *Clustering, Intelligent Transportation System, Semi-supervised Clustering*

1. INTRODUCTION

The transportation industry regulatory system will provide businesses, vehicles, employing basic information management capabilities, as well as vehicle traffic and operational information management functions. Including data collection, it is timely updates, and statistics and analysis functions. Based on the GIS system to achieve the positioning of commercial vehicles, it is display and monitoring functions.

System to achieve the transportation industry regulatory system is underlying the standardization of communication protocols and application layer protocol interface. Has been built right next to the appropriate technology access operations, the use of different technologies to achieve the GPS monitoring service system, communication channels and provides the control commands to the system; with the new installation of GPS vehicle drive recorder bidirectional communications, these vehicles GPS data acquisition and control instructions [1]. Right on the acceptable provincial traffic department in charge of GPS / GIS monitoring center control and management of instruction can be submitted in response to the requirements of the provincial monitoring center of the GPS / GIS data.

Vehicle control system can receive various instructions issued by the data center, including traffic plans, departure instructions, Tips, scheduling parameters, and expressed through the way of lighting or voice, to improve the accuracy of the scheduling and on synergistic effects, while the car sensor devices in real time the necessary vehicle operational information uploaded to the data center, so that the background data and centralized scheduling.

Install car photoconductivity by equipment, and vehicle wireless gateway device to connect. Passengers can not only listen to the bus out of the station site voice messages can also be a new car on the way to keep abreast of the vehicle's current location and a new car direction. Text display mode extends the voice guided by information the way to meet the information of all kinds of people receiving needs; real-time display to compensate for the limitations of the voice guide multiply information, car information services more complete and practical.

Semi-supervised clustering some data on the type of mark or constraints is in order to assist non-supervised clustering process. As many cases, the category information of the data is incomplete, semi-supervised clustering can be used to tag data of the type of information or constraints on the data gathered, extended and modified the original category tags.

Bus monitor function is based on the query status information for all bus or designated bus is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

currently running track and the historical trajectory, speed, unexpected situation, disposal and tracking functions. Bus scheduling function is automatic scheduling management, traffic schedules automatically generated, remote centralized scheduling, recovery vehicle scheduling, vehicle monitoring and adjustment of spacing, call center, driving the speed control and speed limiter settings, vehicle routing control, passenger information monitor, system monitoring, the station information feedback, fault weakening. Vehicle Information Management is the management center of the entire vehicle, the vehicle bus support, vehicle linked Counting, Lu Shan, scheduling, maintenance and other aspects, the axis of the bus business, vehicles, information management, including vehicle information, and model, engine information. The paper presents using semi-supervised clustering algorithm to construct the intelligent transportation system.

2. FUNCTIONAL REQUIREMENTS AND BUILDING CONTENTS OF INTELLIGENT TRANSPORTATION SYSTEM

Discrete electronic equipment inside the integration, standardization, integration of in-vehicle local area network interoperability and vehicle wireless gateway and data center interconnect. Vehicle wireless gateway device should provide an adequate standard hardware interface, data exchange with the current and foreseeable future automotive electronic products [2]. Vehicle wireless gateway device should have a strong computing power, ample storage space and reliable operating performance.

Through a wireless network connection, you can make the IC card machines, traffic counters and other equipment operating data reported to the center at any time; can also be a license to the head, tail, brand Yaopai, scroll inside the car, car electronics stop sign, stop and other equipment Download the guide by parameters. Multiply three-dimensional conductivity of the speech sound and light, car scrolling text changes should also be broadcast coverage multiply the terms at the same time, car electronics stop LED flashing switch; the first card, the end of brand and Yaopai voice broadcast line switch should also be a corresponding text transformation, as is shown by equation1.

$$E[\tilde{X}\tilde{X}^T] = (H^T UH)^{-1} H^T URUH (H^T UH)^{-1} \quad (1)$$

All vehicle equipment operating parameters can be remotely change the parameter information, such as the licensing of vehicles, head, tail, brand line of Yaopai, car scrolling text message, and stop the line number. Can remotely change the mailing address of the data center parameters, such as IP address, it is port number, and the name of the wireless network access line. Real-time data sampling can be changed via remote download phonebook information, the parameters of the electronic fence parameters, hysteretic station and the station of rejection, speeding alarm threshold segmentation parameters, automatic station discriminate parameters.

Guided embarkation appliances to obtain the vehicle wireless gateway processing the results of the GPS data, including the status of vehicle access points, site number, up and down the line status, and the data displayed on the screen with LED lights to achieve linkage with vehicle wireless gateway, pre-leave a substantial text information display space, to facilitate the publicity of the line operating information and the policies and regulations publicity, as is shown by equation2.

$$E\{v(k)v^T(j)\} = R(k)\delta_{kj} \quad (2)$$

Depending on the configuration, timing, or given away to upload to the data center of vehicle operation information, including: time, longitude, latitude, speed, direction. Every 5 to 10 seconds can be set to automatically store a real-time data, including time, location, speed and direction, the center real-time data can be automatically stored on the need for on-demand, to compensate for abnormal network or positioning center lost data, you can cycle store at least the last five days of real-time data. Own initiative or on demand return the logbook, including: driver information, information, interactive content, the vehicle condition, a single complete, traffic situation, and mileage and fuel consumption [3].

The above equipment to establish a connection via a wired or wireless link with the data center and interactive data, the base station in the station house equipment for reception centers departure instruction, operational statistics and frequency of notice and information to guide the Sops urge the drivers to start; vehicle docking station platform departure information bulletin screen to receive the personalized guide multiply information and convenience of the center, etc., shown by the screen text, LED text and e-site signage LED flashes way to facilitate car; mobile Zhanwu terminal equipment

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to receive a variety of grid information center, the task of query information and driving schedules, etc., to facilitate the paperless daily record, the weakening of departure in the field of people and vehicles information exchange and communication failure.

Scheduling commands (the end of the station room): scheduling order orchestration and displays the screen on the scheduling screen, vehicles from the number and departure time; arrangement of three on the grid command, scroll refresh cycle and 30 seconds, when one of departure time is coming when we will take the initiative to switch to the command interface, and outgoing vehicles and departure time is flashing; scheduling commands stored in a circular queue, when the old departure time arrives will be a new start command out of the queue, thus completing the update of the scheduling order.

Scheduling information (base station in the station house): automatically detect and download the same day scheduling information; support real-time incremental update of scheduling information; according to the current scheduling order, the corresponding driver, display the next two days to work shifts [4]. Zhan Wu mobile terminal device: receiving the start instruction from the data center; receive verification instructions from the data center personnel and vehicles; to receive from the data center staff undergo and vehicles off the assembly line and other information, as is shown by figure1.

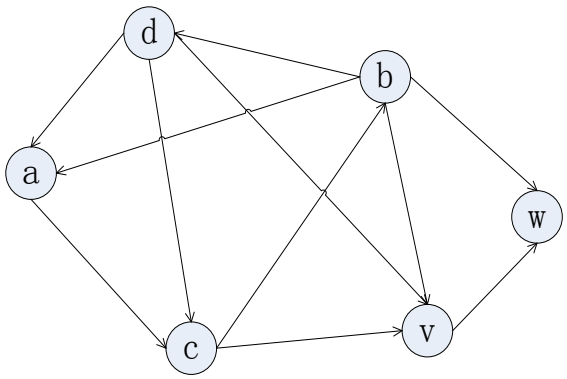

Figure1. *The Contents Of Intelligent Transportation System*

Bottom station room, can be remotely set the voice volume; can automatically apply for the information of the scheduling table; remotely change the size of the volume of the voice broadcast and automatically saved. Mobile Zhanwu terminal equipment data center can receive the latest traffic schedules and save. Line information change; mobile Zhanwu terminal equipment may apply for and download their need to manage the lines of basic personnel and vehicle configuration information passively.

Centers can provide the station house and site equipment health monitoring, abnormal immediate alarm. Room of the end station, when the paralysis of the wireless public network through wireless emergency response system to receive Center news release and on the large LED screen is displayed to guide the grid, only need to show on the big screen outgoing vehicles of self-number and departure time can be. By mobile Zhanwu terminal equipment on the vehicle operation records, driving records and the center needs a post back. Site vehicles and equipment is operating properly, can mobile Zhanwu terminal equipment sent to the command mode to data center to improve the efficiency of scheduling and retrospective, there is shown by equation3.

$$Var[X] = E[(X - \mu_X)(X - \mu_X)^T] = C_X \quad (3)$$

Automatic transformation of the signs: the lines with the number of vehicles and driver number and line capacity, automatic arrangement of working hours of the vehicle group, to meet the vehicle group (referring to the combination of vehicle and driver) the balance of workload between meet the relevant provisions of the labor Law [5]. Ban automatic transformation: the workload between the different drivers also need to be unevenly distributed within the same car group. Automatically according to the type number of the set of classes, the arrangement of the driver's daily work shift type, in order to achieve a balanced workload and labor intensity.

The basic parameter settings: Set the line length of up and down, up and down the line shift time Sanfeng quasi-point time, primary and secondary stations the last time, the name of the main and auxiliary stations, main and auxiliary stations into playing time, mileage and other parameters. The dynamic increase in the signs: based on passenger flow characteristics, and schedule type, depending on the circumstances artificially increase the part of the road signs to increase the capacity of the line. Dynamically remove the signs: based on passenger flow characteristics, and schedule type, depending on the circumstances artificially increase the part of the road signs in order to reduce the rate of no-load.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

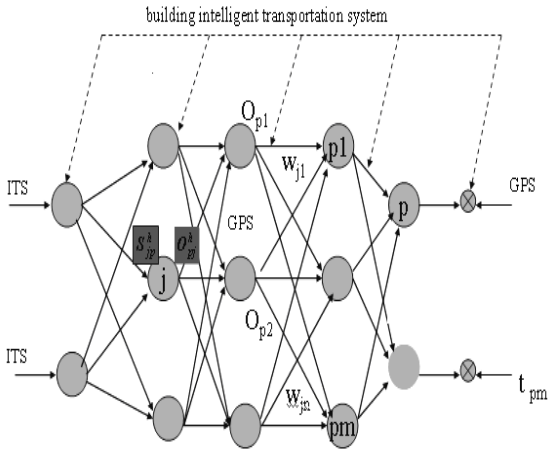

Figure2. *The Artificially Increase The Part Of ITS*

According to the scheduling table and driving schedules, their formation of the start sequence, and can automatically issued departure instructions and give the dispatcher the different state, already issued instructions about to start, the point is not issued to the state prompted [6]. Dynamic adjustment of the start sequence, it is the line of the road conditions, vehicle condition, abnormal weather, and the driver's physical condition affect the scheduling factors. Some objective causes a program can not be carried automatic scheduling system will automatically adjust the start sequence, to achieve a balanced grid.

The departure times of the First and Last it Bus Group's commitment to the people of it, no matter under what circumstances the need to safeguard the normal start of the end of the first bus [7]. When the weather or vehicles, and other abnormal system by adjusting the grade of size or reduce the operating units in the inferior means to ensure the operation of not train punctuality, as is shown by equation4.

$$\frac{\partial u}{\partial x_1} = u(x_1 + 1, x_2) - u(x_1, x_2) \tag{4}$$

Vehicle to complete a single run to reach the end station, the vehicle automatically to match with the Travel Plan, when the vehicle is yet to be completed the day of operation, continued to continue working, automatically place the pending departure area; and when the vehicle completed the day's operation, the system will automatically be placed off the assembly line, the day is no longer involved in the operation. Automatically display a variety of stops the vehicle information. Docked vehicles are mainly divided into three types: the docked vehicles in the first station and the station are not docked vehicle, abnormal vehicles. The first station is not docked vehicles: primary and secondary stations will be based on the vehicle match the station area in the two end of the vehicle in accordance with the vehicle departure time order.

Line on the vehicles contains a variety of property and status, for example: the whole car / shuttle bus is bus / shuttle / independent bus, air-conditioned car / vehicle, failure / offline / warranty / accident; you need a different the colors and images to distinguish, and configurable icon. Can different types of icons superimposed to show the various properties of the vehicle. Manual scheduling, dispatching staff by vehicle terminal to send a dispatch command to the bus, the driver should be able to quick access to scheduling command. Scheduling command history is saved in the control side, for after the tracing.

Taking into account the number of fault vehicle and repair vehicles relative to the operational vehicles is relatively small, so the main scheduling interfaces on the map interface, and repair of vehicles and all fault position of the vehicle displayed on the map [8]. This dispatcher is more intuitive, and straightforward. Smart match for the emergency rescue vehicles, timely, intelligent recovery vehicle scheduling, fault car repair information and the location of all emergency rescue vehicles, state-related, to make the quickest recovery vehicle allocation scheme, making the vehicle for repair to achieve real-time, efficient and optimal scheduling purposes, as is shown by equation 5.

$$H(x) = -\int_{-\infty}^{\infty} P(x)\log P(x)dx = \frac{1}{2}\log(2\pi e\sigma^2) \tag{5}$$

Operation of vehicles, due to the complexity of road traffic conditions, the vehicles will appear in the process of moving the vehicle Solitaire and paragraph when the phenomenon. Driver to avoid this from happening, equidistant running vehicle GPS real-time data analysis and comparison, as far as practicable, to achieve the following functions: by way of sound and light to alert the driver in front / rear of a vehicle nearby, according to this prompt appropriate deceleration / acceleration in order to adjust the spacing of vehicles.

Operation of public transport enterprises sources of income, and good operational management is the protection and support for enterprise development. The operational systems include: workforce management, management of single-point of note management, statistics query and reporting.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Workforce Management to the original scheduling plan that need to be completed manually on paper, it is electronic workflow management, saving the tedious work to need a hand-written scheduling table, to improve the efficiency of the scheduling. After the scheduling data is saved to the scheduling management of queries and the actual scheduling does not meet the dispatcher to dispatch to check and modify the class, and shifts the state of modification for the implementation. Determine the scheduling of the case; connect the printer to complete the scheduling of report printing.

3. THE RESEARCH OF SEMI-SUPERVISED CLUSTERING ALGORITHM

Clustering is a very important technology of the so-called clustering is measured according to some similarity measure, not the similarity measure or distance, and a collection of individuals according to certain criteria are divided into classes, the degree of similarity between similar individuals than different the degree of similarity between the class of individuals that do "Like attracts like" semi-supervised clustering algorithm unsupervised learning how to use a small amount of monitoring information to improve the clustering performance is now the continuous application of this article first introduces the clustering overview of the development and related technologies in the clustering process, focusing on the distance metric commonly used clustering methods and evaluation criteria for the subsequent chapters provide a theoretical and experimental basis for the previous semi-supervised fuzzy C-means clustering algorithm [9].

With the rapid development of the Internet, the vast amounts of unlabeled documents and a relatively small number of labeled documents is a general case of Web documents, and how to effectively use a small amount of label document to the cluster mass is not the label document, in order to better access to valuable information, ie semi-supervised learning problems, has become a hot topic of current research. If the monitoring data is a category labeled, and these markers data on behalf of all the relevant categories, the semi-supervised clustering and semi-supervised classification algorithm can be used for data classification. However, in many areas related to the type of tag information is incomplete, and the semi-supervised classification, semi-supervised clustering can gather data as the category with the initial labeled data, and expand, modify the existing class mark reflect the data of other law, as is shown by equation6.

$$\mu_i(k) = P\{m_i(k) \mid Z(k)\} = \frac{f_i(k)\sum_{j=1}^{n}\pi_{ji}\mu_j(k-1)}{\sum_{i=1}^{n} f_i(k)\sum_{j=1}^{n}\pi_{ji}\mu_j(k-1)} \quad (6)$$

The algorithm uses a distance measure function to meet the mark or constraints to achieve the clustering process [10]. Specific algorithms: the Mahalanobis distance based on convex optimization algorithm; Euclidean distance based on the shortest path algorithm; discrete gradient descent and get the Jensen-the Shannon; EM algorithm to improve the edit distance (string-editdistance) and other.

Clustering assumption is that the distance between the sample data with each other more recent, they have the same category. According to this hypothesis, the classification boundary must be as much as possible through the data is more sparse, in order to be able to avoid the concentration of sample data points assigned to the classification on both sides of the border. Under the premise of this hypothesis, the learning algorithm can use unlabeled sample data to analyze the sample data distribution in the sample space to guide the learning algorithm to adjust the classification boundary to make it as far as possible through the relatively sparse area of the sample data layout. For example, the Joachims transduction support vector machine algorithm, the training process, the algorithm to modify the classification hyperplane and exchange hyper plane on both sides of some of the unlabeled sample data mark, making the classification boundary to maximize the interval in all the training data allowing access to a regional data is relatively sparse, but also, as far as possible, the right to divide all sample of the data hyperplane.

Constrained K-means algorithm implementation process, the seed cluster category tag is to remain unchanged, the only non-seed data to recalculate the mean, also the mark of non-seed data may change, no noise or need to change on seeds marking [11]. Most Seeded K-means algorithm for noise on seeds, seed tag can be changed in the clustering process, in the initialization of the seed center; you can get rid of noise seed:

1) Arbitrary choice of the data as the initial center point;

2) Repeat (3) to (6) does not change until the center point;

Journal of Theoretical and Applied Information Technology
10th April 2013. Vol. 50 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3) The assignment of each of the remaining data from its recently the center of the class represented by;

4) Randomly selected a non-center point data;

5) Corresponds to the data set for each division, at least one seed;

6) Constrained K-means algorithm implementation process, the seed cluster category tag is to remain unchanged, the only non-seed data to recalculate the mean, also the mark of non-seed data may change, no noise or need to change on seeds seed tag.

The K-means is a special case of the EM (Expectation Maximization) algorithm to solve; you can use the Gaussian mixture model. The model is based on the following assumptions: all the data are independent and identically distributed Gaussian probability distribution of the same class of data, all the data constitute a Gaussian mixture distribution. Gaussian mixture model (described mixed-density distribution of the model) is the most classic, most complete statistical learning modeling algorithms.

4. THE DEVELOPMENT OF INTELLIGENT TRANSPORTATION SYSTEM BASED ON SEMI-SUPERVISED CLUSTERING

Geographic information database based on transit bus transfer GIS-based queries, provide the following information: line queries: You can query the end of the first station of bus routes, half-way through the station, the last train time, fare of their respective companies convoys, and other information, relevant information on a map marker. Site information: You can check a site where a few lines after, and displayed on the map line. Conductivity multiplied query (change of query): according to a given starting point and terminus station name or alias information, calculate the shortest distance, the price of most provinces and the site at least traveling, and can be displayed on the map dynamic simulation.

Vehicle terminal real-time acquisition to move the position of the vehicle, state, business data, business-related data including the empty passenger, mileage, prices and other taxi operators need to send data to the scheduling began carrying passengers and passengers reach their destinations information center in the empty state from time to time or from time to time send to the system. According to the need for each company vehicle operating conditions in real-time monitoring, and keep abreast of the vehicle situation and provide the basic data for the car sent a car business.

Vehicle location query: The system can query the current location of the designated vehicle and also shows the target vehicle speed, latitude and longitude, vehicle status information. The real-time vehicle tracking: the system can track the specified monitoring vehicles, real-time display on the map running routes and driving trajectory of the target vehicle [12]. Emergency alarm functions: the driver of the vehicle is in motion the process, in case of emergency, such as banditry, accident, failure, help, you can press the vehicle emergency alarm button alarm to the dispatch center, police intelligence information to be uploaded to within 2 seconds command center, the Center received the alarm, voice monitoring, power off the oil, adapter 110 and review the operation, increasing the safety of road transport.

Vehicle after inspection management: all taxi vehicle GPS location and vehicle status data are stored in the system back-end database, schedule, monitor the front desk for any one designated vehicle at a specified time period, according to a certain mode (speed, distance, etc. intuitive continuous star) in the seat map the history of track playback, to reproduce the vehicles in the corresponding period of the driving route and related information, to achieve the purpose of the vehicle after the audit management.

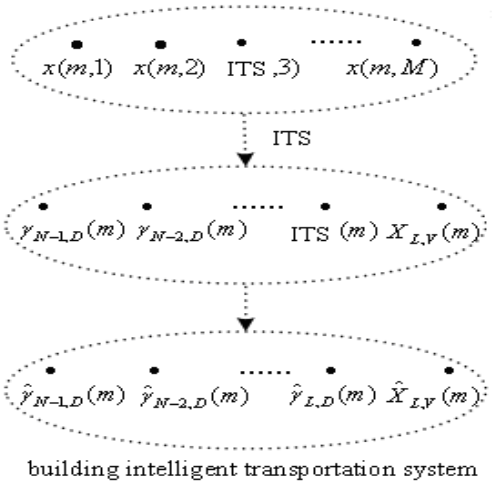

Figure3. The Development Of Vehicle After Inspection Management Diagram

Most Seeded K-means algorithm and Constrained K-means algorithm is essentially in the case of certain assumptions, the EM algorithm, Gaussian mixture model [13]. Distribution in the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

semi-supervised K-means algorithm, assuming that all the data in all categories of obedience is a "larger" Gaussian mixture distribution, the Gaussian number is the number of categories, each a Gaussian distribution, it this "small" Gaussian distribution of hybrid formation of the distribution is Gaussian mixture distribution for each category can be broken down into "smaller" Gaussian distribution, as is shown by equation7.

$$BB = \begin{bmatrix} \overline{Y} - C(output)(Cx(0) + Dd) \\ -\underline{Y} + C(output)(Cx(0) + Dd) \end{bmatrix} \quad (7)$$

The standard K-means algorithm does not provide any monitoring information, the mean of the E-step is chosen at random, and subsequent M-step is divided to the nearest mean. Each data set in the semi-supervised clustering data corresponding is in order to a mean value and the corresponding conditional distribution. E-step, data is randomly assigned to a cluster, which is equivalent to a conditional distribution of selected data from a conditional distribution.

1) For each seed $x_i \in S$, the user should belong to the category; Select the initial cluster center;

2) to (3) to (5), iterative convergence until the objective function;

3) Calculate the mean of each cluster in the data, each data re-assigned to the most similar class;

4) Re-computing the mean of the class assigned to each class;

5) Most Seeded K-means algorithm is initialized with the seed cluster centers, where center is not like K-means algorithm randomly mean, choosing instead to seed set by the center as the first cluster center;

6) Most Seeded K-means algorithm is applied to the noise of the seeds and the seed tag can be changed in the clustering process, so in the initialization of the seed center, you can get rid of noise seed.

Taxi terminal, GPS location information, driver information, work time and attendance information and other data analysis, statistical analysis of a variety of regulatory reports, including taxi peak shifting and shift statistics, monitoring of key areas, filling stations to monitor taxi Hometown monitoring, the status monitor taxi, taxi income monitoring, taxi inter-regional monitoring, limited supervision, various types of analysis reports for the monitoring of speeding within the region.

The standard K-means algorithm does not provide any monitoring information, the mean of the E-step is chosen at random, and subsequent M-step is divided to the nearest mean. Each data set in the semi-supervised clustering data corresponding to a mean value and the corresponding conditional distribution's-step, data is randomly assigned to a cluster, which is equivalent to a conditional distribution of selected data from a conditional distribution.

Successful business within a certain time does not automatically send a car, accepted the need for manual intervention, which is generally set the rules of the original computer can not find a suitable vehicle. Available through the GIS system queries to relax the conditions to obtain the corresponding vehicle, and telephone notification to the passenger car time, or unable to successfully send a car, with reasons. Special operations or passenger requirements can not meet the situation, business can be handed over to the monitor units for business processing, and sent a car to the appointment of vehicles, the system will automatically remind accept staff query whether or not to send a car successfully, or to the monitor units responsible for keeping track of business, as is shown in Figure4.

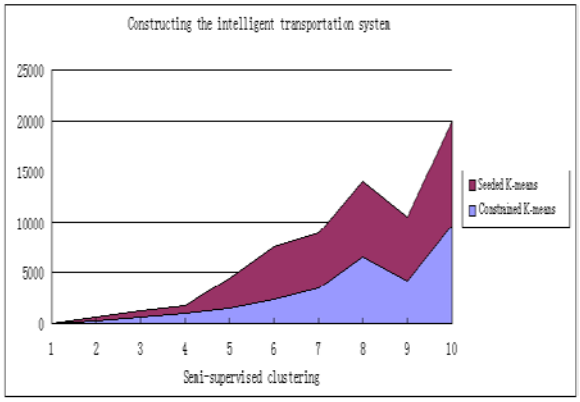

Figure4. The Building Intelligent Transportation System Based On Semi-Supervised Clustering Algorithm Diagram

Intelligent traffic control information, instruction, effective, rapid and accurate transmission between the various components of the main role of the system network communication transmission systems one for intelligent traffic management system to provide access; public security network and outside networks and intelligent traffic management system data exchange. The network consists of a network of intelligent traffic

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

management, public security network, external network and video communications network, in which external network including the Internet, wireless communications network, leased operators of wireless networks, and video communications network, see TV monitoring system construction program. In accordance with the relevant provisions of the of the Ministry of Public Security, border access platform, construction of public security network security isolation as public security networks and other network connections. Border access platform consists of the simultaneous construction of intelligent traffic management system in accordance with the construction specifications of the Ministry of Public Security. The paper presents using semi-supervised clustering algorithm to construct the intelligent transportation system.

Most Seeded K-means algorithm and Constrained K-means algorithm is essentially in the case of certain assumptions, the EM algorithm, Gaussian mixture model. Distribution in the semi-supervised K-means algorithm, assuming that all the data in all categories of obedience is a "larger" Gaussian mixture distribution, the Gaussian number is the number of categories, each a Gaussian distribution, it this "small" Gaussian distribution of hybrid formation of the distribution is Gaussian mixture distribution for each category can be broken down into "smaller" Gaussian distribution.

5. CONCLUSIONS

The paper presents using semi-supervised clustering algorithm to construct the intelligent transportation system. Intelligent traffic control system to control room covers an area of approximately 350 square meters, the height of approximately 2.4m. It as completed the decoration system, the construction of the electrical system, weak monitoring and management system, lightning protection and grounding systems, air-conditioning systems, disaster prevention system and other subsystems. Server farm deployment of an operational audit of the system, it is public security core business databases or business network operating behavior of fine-grained auditing. Parse the behavior of a network operating in line with strategy, analysis, recording, reporting, to help achieve the prevention of advance planning, things in real-time monitoring, response to violations after the compliance report, the accident track playback, to strengthen the internal network behavior regulation, to avoid detachment of the core assets (databases, networks, servers, etc.) losses, safeguard the normal operation of business systems.

REFRENCES:

[1] Hao Wang, "Wireless Sensor Networks for an Extended City Intelligent Transportation System", *IJACT*, Vol. 3, No. 5, 2011, pp. 300 ~ 307.

[2] Ma Zhenliang, Xing Jianping , Yan Shihao, Wu yong, Wu Yubing, "An Aggregation Method for Dynamic Bus Arrival Time Prediction", *AISS*, Vol. 3, No. 6, 2011, pp. 37 ~ 48.

[3] Sunghan Lim, "Advanced Road Management System using IT Technology", *IJEI*, Vol. 2, No. 1, 2011, pp. 84 ~ 90.

[4] Rahal ROMADI, Hassan Berbia, B.Bounabat, "WIRELESS SENSOR NETWORK SIMULATION OF THE ENERGY CONSUMPTION BY A MULTI AGENTS SYSTEM", *Journal of Theoretical and Applied Information Technology*, Vol 25. No. 1 ,2011,pp.189-198

[5] Xu Wei, Huang HaiPing, Wang RuChuan, Sun LiJuan, "A Novel Driver Assistance Using Kernel-Plural Discriminant Analysis Based on Fourier Transform in Vehicle Networking Environment", *IJACT*, Vol. 3, No. 11, 2011, pp. 477 ~ 485.

[6] Krishan Sabaragamu Koralalage and Noriaki Yoshiura, "OTag: Architecture to Represent Real World Objects in RF Tags to Improve Future Intelligent Transportation Systems", *JCIT*, Vol. 4, No. 2, 2009, pp. 30 ~ 48.

[7] Xia Zhanguo, Cai Shiyu, Zhang Wentao, "An Improved Semi-supervised Clustering algorithm based on Initial Center Points", *JCIT*, Vol. 7, No. 5, 2012, pp. 317 ~ 324.

[8] Shen Yan, Song Shunlin, Zhu Yuquan, "A EM Probabilistic Clustering Algorithm for Large Scale Data Sets based on Partial Constraints Information", *AISS*, Vol. 3, No. 10, 2011, pp. 20 ~ 29.

[9] WENGE CHEN , XUELIN HUANG, "RESEARCH OF DIFFERENT COLOUR MODEL OF COLOR CHARACTERISTIC IN DIGITAL IMAGE", *Journal of Theoretical and Applied Information Technology*, Vol. 45, No. 1 ,2012,pp. 049-053

[10] Yu Yuecheng, Wang Jiandong, Zheng Guansheng, Gu Bin, "Semi-supervised Distributed Clustering with Mahalanobis Distance Metric Learning", *JDCTA*, Vol. 4, No. 9, 2010, pp. 132 ~ 140

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[11]Wang Yikang, Liu Xiangguan, "Applying Fuzzy C-means Clustering and Multiple SVM to Silicon Content Prediction in Hot Metal", *AISS*, Vol. 4, No. 2, 2012, pp. 40 ~ 48

[12]JianJiao Chen, Anping Song, Wu Zhang, "Hybrid K-harmonic Clustering Approach for High Dimensional Gene Expression Data", *JCIT*, Vol. 7, No. 3, 2012, pp. 39 ~ 49.

[13] H S Guruprasad and Dr. H D Maheshappa, "EFFICIENT CHANNEL ALLOCATION BASED ON THE WEIGHTS OF VIDEOS USING AGENTS", *Journal of Theoretical and Applied Information Technology*, Vol.5, No.4, 2009, pp 101 - 109