International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 3, Special Issue 3 (September '2017)
Proceeding of - International Conference on Recent Trends in Computational Life Science and Information Technology,
Conference held at Tiruppur Kumaran College for Women,Tirupur,Tamilnadu, India.

OPTIMIZING THE CITY TRANSPORT SYSTEM USING ACM ALGORITHM

N.Revathi,
Department of Computer Science & Application,
Faculty of Park's College (Autonomous), Chinnakarai,
Tiruppur, Tamilnadu, India.

M.Madhangiri,
Department of Computer Science & Application,
Faculty of Park's College (Autonomous), Chinnakarai,
Tiruppur, Tamilnadu, India.

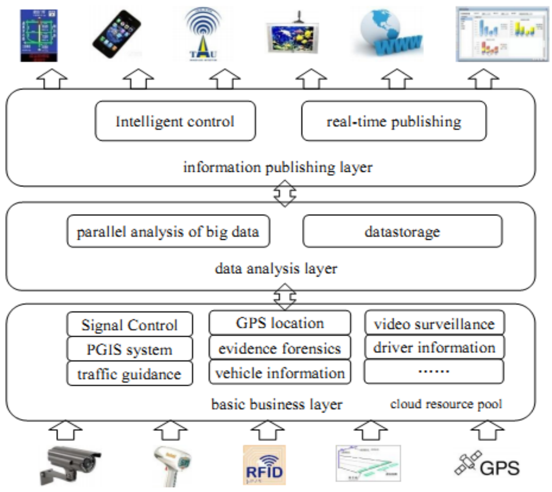
Abstract: It is a very difficult for a passenger to find out an optimal route to travel in a metropolis. A novel approach of search the optimal routes based on date mining technology is presented in this paper. In the approach, we first use the Ant Colony-based Data Miner algorithm to search and to get the candidate set of the optimal bus-routes, and then use data mining method to mine the optimal lines which is hidden in the candidate set. The goal of Ant-Miner is to extract classification rules from data. Ants often locate the shortest path among a food basis and the nest of the colony without utilizing visual information. In order to replace information about which path should be go after, ants communicate with each other by means of a chemical substance called pheromone. The algorithm is inspired by both research on the performance of real ant colonies and a few data mining concepts and principles. The mutual use of ACO and DM (the utilization of ACO algorithms for DM tasks) is a very capable direction. In this chapter, we evaluate ACO, DM, categorization and Clustering (two of the most accepted DM tasks) and focus on the use of ACO for Classification and Clustering. Furthermore, we briefly present associated applications and examples and outline possible future trends of this promising collaborative use of methods.

Keywords: Ant Colony based data mining, router, ACO, Clustering, DM.

I.INTRODUCTION

With the quick growth of society, transportation industry is also facing unprecedented challenge, if big data is used in this region, the benefit is clear, it can find out the questions of usual traffic administration system, in this article we propose an design of intelligent transportation system on big data platform, then we disuse the key knowledge in ITS: computation of bayonet traffic flow, calculation of average speed of a road, querying the travel path of a vehicle, checking and controlling the fake vehicles. Keywords: big data; Intelligent traffic system; application Transportation industry can't meet the quick development of data. The data of transportation industry have rich sources, diverse types, and new data is produced repeatedly. Customer information of rail way, road traffic, aviation industry, public transit, are recorded, and tens of billions travel records are generated every year. Operating data generated by transport companies, such as the data generated by the courier companies.

Dynamic data created by various sensors, such as induction coil at bayonet point, infrared detector, microwave detector, ultrasonic detector, laser detector, video detector, and so on, and the data are produced by GPS vehicle location tracking system and other mobile devices each year the amount of data created by the transportation industry in a city has exceeded TB level, are developing from PB level to EB.A massive data storage space and equipment are required and it must have fault tolerance and stability. Usual data processing systems are faced with the difficulty of inefficient or even failure. The information system of transportation industry after development for ten years has had a definite foundation and scale, But invention of new business, quick development of data, complexity of the data processing are not foreseen. The traffic information administration system using usual data processing method can't meet the quick increase of data, it is incompetent, collapse and failures have happened, when meting out big data recent years, new projects and renovation of the old system were carried out in some places, The rapid growth of data is still not considered, In the course of project construction and maintenance, the construction was emphasized, the maintenance was overlooked, data have not been excavated deeply, with the change of leadership, the life cycle of the system is shortened. The current administration system comes out single purpose, the lack of addition, backward technology and further problems. In the method of building the transportation information system, homogenization is serious, at the same time, development of information technology in dissimilar regions is not balanced, data acquisition is at different depths in different area, and without uniform standards,



According to the data of the road network, the demand of public travel and comprehensive examination of data, information analysis layer uses big data method, data mining method, combines with a diversity of mathematical models for real-time effective study. It can grasp the situation of the transportation system in any time, such as road congestion degree, normal speed, saturation, occupancy rate, interrupt rate. It can make extra congestion warning, traffic supervision and other intelligent transportation actions. Data analysis layer is built on Hadoop ecosystem, use profitable cheap server as hardware stage, use the open source Linux as operating system, use HDFS as file system for big information storage, use Map Reduce as a parallel calculating model, use HBase as the database for processing the data, use Hive as data warehouse, use Sqoop and Flume as tools for data integration. The information publishing layer according the result of the data examination layer, issuing traffic conditions to public, business units, industry executives, etc. by internet, mobile terminal, desktop application, report, for their travel and business resolution.

II.RELATED WORK

In recent years, smart cards and automated data collection systems (ADCS) had also been widely adopted in public transport networks around the world. These smart cards, which are portable credit card size devices that store and process data, were credited with monetary value and used for public transit fare payment [2]. Even though the preliminary purpose of smart cards is for set of fares for public transport, Bagchi and White [1] discussed the potential use of the passively recorded transaction data for travel behavior study and public transport planning. Since then various researches and case studies were done on the data collected from public transport smart card systems around the world.

Morency et al [3] explored the data mining techniques used to analyze the spatial and temporal variability of Canadian pubic transit network passengers using different card types. Asakura et al [4], constructed a origin-destination (O-D) matrix using the smart card data collected from Japan‘s public train network and applied statistical analysis to study the change in passengers‘ travel patterns when the train operator changed its train timetable. Kim and Kang [5] also attempted a similar research on Seoul public transit network by using the transaction data collected from T-Money, South Korea‘s electronic fare card system. Locally, there were also some researches done using the data collected from EZ-Link, Singapore‘s smart card used for public transit. Lee et al [6] did a case study to optimize serviceability and reliability of bus routes by running statistical analysis on the EZ-Link data collected from bus transit trips. Sun et al [7] also did a study using the EZ-Link data collected from public train transit to estimate the spatial-temporal density passenger onboard a train or waiting in the train station.

In the literature, there is a large body of work on applying statistical analysis on the smart card data collected from public transit trips. However, few researches were able to deal with the growing large amount of temporal data effectively or were there any extensive study done using the smart card data to reveal the commuters travel patterns.

Time-series data refers to data collected in a routine, continuous and sequential manner. This type of data, which typically accompanied with a timestamp, has always been collected by businesses and organizations in their daily operations. Examples of such data include sales transaction, delivery orders, stock quote prices etc. Increasingly, businesses and organizations seek to analyze these time-series data to uncover more business insights. However, the analysis of time-series data posted new challenges, as traditional data mining and statistical techniques are inappropriate when analyzing data that contain a time factor.

These challenges eventually became the motivation for the development of time series data mining techniques. According to Schubert and Lee [8], there are two main challenges in analyzing time-stamped data. Firstly, it is a tedious process to transform time-stamped data into table formats, which is suitable for the application of traditional statistical techniques. Secondly, it is challenging to apply pattern detection on time stamped data using traditional data mining as the time-stamped data might be irregularly recorded. Time-series data mining however, tried to overcome these challenges by presenting a framework technique that converts time-stamped data into time-series format suitable for analysis and pattern detection.

In one of the earliest network-based studies of urban transportation systems, Latora and Marichiori investigated the Boston subway network and highlighted the small-world properties and local and global efficiency of the network [6]. Sen et al. studied the Indian railway network and also discovered small-world properties, like small average distances, and exponential degree distribution [10]. Seaton et al. compared the urban train line networks of Vienna and Boston by extending the usage of random bipartite graph models from social networks to technological ones [11].

Angeloudis and Fisk analyzed the world's 20 largest subway systems and found that the high connectivity and low maximum vertex degree are characteristic features and highlighted the robustness of the networks against random failures [12]. Chen et al. investigated the urban transportation networks of four major cities in China and showed that the distribution of the number of routes that a stop joins follows a power-law with exponential decay, while the distribution of the number of stops in a bus route follows asymmetric, unimodal functions [13]. Xu et al. defined spaces P and L as the network of stations connected by links and the network of stops as nodes with links that only exist between two nodes if they are consecutive stops on a route [14]. They analyzed three bus transportation networks in China and also examined the small-world behavior of them in both spaces P and L. Sienkiewicz and Holyst collected and analyzed the data of the PTNs of 22 cities in Poland and found that the degree distributions in space P follow a power-law, while in space L they are exponential [15]. Ferber et al. studied the PTNs of 14 major cities of the world and found that the degree distributions follow the power-law with various exponents [16]. They also proposed an evolutionary model of growth of PTNs. Later, they performed a comprehensive complex network analysis of these networks [7]. Derrible and Kennedy investigated 33 metro systems in the world and found that most of them are indeed scale-free (the degree distributions follow a power law), but the presence of transfer hubs (stations hosting more than three lines) results in relatively large scaling factors [17].

III.BACKGROUND STUDY

PROBLEMS IN PRESENT SYSTEM

The data and it has two main goals: prediction and description. Prediction consists in using some variables in the data to find out indefinite or future values of other variables of attention, while description focuses on finding human interpretable organizations telling the data. We do not try to converse all accessible mining process; as a substitute we evaluate generic ones and recognize the interactions among them. By generic mining techniques we mean mining techniques useful on a wide range of trajectory data in a wide range of functions. We separate the techniques discussed into two groups: primary technique and secondary techniques. Given a trajectory dataset, primary methods plan at classify the trajectories based on their properties. On the other hand, secondary mining techniques are fused methods in the sense that they apply a sequence of primary mining techniques or classical statistical methods or a combination of the two.

The discussion of the methods focuses on the objective of the method, its general working principle and the data granularity level at which it works (whole trajectories versus sub-trajectories, single trajectory versus multiple trajectories, and trajectories of single moving object versus trajectories of multiple objects). The discussion ends with an attempt to identify relationships between different methods.

In traffic management Ant Colony Optimisation is increasing in popularity, with its ability to find optimised solution in situation where traditional methods fail to produce any good solution. One of the large benefits of using ant colonies to assist the planning of transport systems is their ability to adapt to changes within the environment.

In an ideal world all transport systems would run to schedule without delays, but in reality any scheduled transport is swarmed with a list of problems which could occur on any given day. These range from short delays given on a specific route for reasons such as road works or rush hour congestion, to larger problems such as a vehicle crash or an entire route closed for maintenance.

IV.PROPOSED METHODOLOGY

ACO Works

Ant Colony Optimization (ACO) is a relatively new nature inspired algorithm which was originally developed by Colorni and Dorigo in 1992. The method takes inspiration from how an ant colony is able to co-ordinate as a whole to locate and collect food without a central control organizing and communicating with individual ants in the system. By mapping some of the ant's behaviors to computerized agents, an ACO system can be used to replicate various different situations and assist in finding the optimal path across a network. Whilst the individual ants do not communication on a one to one basis, each ant is able to lay down a pheromone trail to mark a favorable path. Other ants in the system are then able to detect the level of pheromones across different trails and choose which path to follow in favour of the one with the greatest concentration.

As the levels of pheromone fades over time, only routes which are regularly traversed will remain attractive to further ants in the colony. The whole process can be broken down into a number of simple steps. Firstly, a number of foragers go out into the environment in search of food. Once they have located a food source, they will return to the nest, laying down a pheromone trail as they go, to mark the path they have followed. In the next stage, other ants from the colony will choose which path to follow in favour of the one with the highest concentration of pheromones. When each of these ants has located the food source it too will lay down a pheromone trail indicating its chosen path to the other ants in the colony. [Fig 1.2] As time passes, the pheromone values attached to each path will change. The routes most favoured by the ants will consistently have a high concentration of pheromones, while the infrequently used paths will become less and less attractive. In turn this encourages the ants to follow the more popular route often creating a single line of ants from the nest to the food source.

In Transport

Public transport is a complex multi optimization problem due to the number of factors that should be considered. The needs of the passengers need to be balanced against the needs of the operators and the whole system needs to work within the confines of the existing infrastructure. Some of these considerations include:

Passengers: Frequency of transport, Routes.

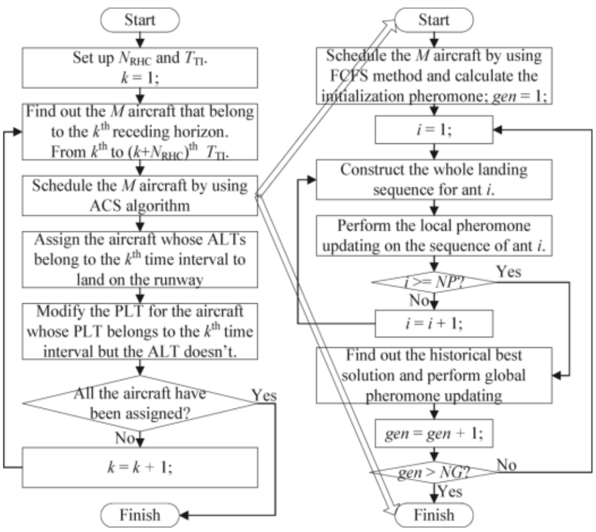
Operators: Number of vehicles, Number of drivers, Routes, Reduced idling time.

Infrastructure: Finite possible routes, Traffic or other obstructions, Limited track capacity.

Many of these may be directly contradictory such as frequency of transport and numbers of vehicles and drivers, the passengers want a frequent, regular service but running larger vehicles less regularly may be cheaper as fewer drivers and vehicles are required. These problems are traditionally solved either by experts using their knowledge of the system and trial and error to find a suitable solution. They can also be solved exactly using mathematical methods although this can be expensive for large problems. Due to these restrictions nature inspired methods have proved to be useful in solving these problems.

Initial solutions were found using the nearest neighbor procedure and are then optimized using a hybrid of the ant colony optimization and the variable neighborhood search. This method was applied to a real world example in Great Tunis. They conclude that their algorithm can provide improvements over the current system. A similar investigation by Shi, Zhao and Zhang (2010) [3] on a public bus network in Cheng guan District, Lanzhou also suggested that improvements to the system could be made using ant colony optimization. The bus network problems cover many of the key points such as passenger requirements and optimization of bus and driver usage. However they are still simplified models in that the distances between nodes represent the distances between the stops and not the time. In many road networks the shortest distance is not always the quickest due to traffic levels and other road restrictions. Later in this report we will look into how ACO is used to

find the quickest path between when taking into account road and traffic factors.
The problem is very difficult to solve exactly even for small numbers of cities and therefore alternative methods such as heuristic algorithms are used to find near optimal results. Ghoseiri and Morshedsolouk [2] give an overview of the different methods that have been used to solve the problem in 2006. Ant colony optimization is particularly suited to this type of problem and easily out performs the other methods, finding the optimum solution on a significantly lower number of iterations. For this reason it is often used in other related problems such as timetabling. This is a solution of a single problem in train scheduling, that of a single track line. Real train networks are much more complicated with double tracks, other branching lines and trains that will not stop at every station. Also this model does not consider the requirements of the passengers such as more trains at rush hour or regularly spaced trains for convenience. There are also trains and driver issues as the trains do not return to their point of origin. The dynamic basis of Ant Colony Optimization makes it well suited to optimizing road traffic networks which are continuously changing and unpredictable. Knowing the route from one place to another is not enough on its own to provide good navigation and management of a road network.
This is because there numerous factors which can affect the journey time and efficiency on a given section of road. Such factors include: road accidents and break downs, adverse weather conditions, speed limits, and slow moving vehicles. The variance and unpredictability of traffic means that a dynamic system is needed to manage it. Traffic routing is a complex task and can have several objectives that can be in conflict with each other. Each road user is interested in traveling the shortest and quickest route to their destination with minimum route changes. This goal is known as User Equilibrium. From the interest of managing the road network as a whole, the goal is to maximize the traffic flow; this could be for certain important sections of the road network or for the entire network as a whole. This goal is called the System Optimum and aims to minimize the average travel time in the network, which may mean that some users travel along suboptimal routes or at suboptimal speeds for that user's interest. These two contrasting objectives are hard to reach separately in its dynamic environment and the choosing the right balance between them is also difficult; when a good balance is used between the two objectives the network will be more stable.
Why use ACO? Modern algorithms tend to use much precompiled data, which leads to determine their usage for static route planning. Application in dynamic networks is inefficient, because of the need to recalculate at least the part of pre-processed information. [8] ACO may not always find the most optimal solution, but is very good at quickly producing good solutions.
Basic route planning is based on solving the shortest path problem for a road network. When ACO is used for traffic planning a graph is used to represent the road network and this graph is used to find the shortest paths between points. The edges in the graph represent different roads and the vertices represent turning, crossings and other important road features, which separation the sections of road.
The weighting of the edges are usually the length of the edge and is always positive. Basic graphs are static and do not model changes in the network such as closure of a road or reduction in the maximum speed limit. A dynamic road network is almost always be more accurate provide that enough information about the roads it is modeling is collected and used correctly. For finding the shortest path in static graphs ACO work by depositing pheromones based on how many ants traverse an edge, thus this edge will become be attractive as more pheromones are add.
However, when managing traffic flow the amount of pheromone deposited could be used to represent congestions so edges become less attractive with increasing pheromones. Also as there are many objectives that a system may be trying to achieve the function for pheromone depositing will be much more complicated than just the amount of ants which traversed the edge. Such functions for pheromone depositing will need to take into account the dynamic features of the road network



Ant Colony algorithm methodology

V.CONCLUSION

Current traffic systems intermittently make use of research, and it may be that in future designs, Nature inspired design techniques could be employed to design road systems, to develop a transport infrastructure that could be beneficial to overall efficiency and safety. The de-centralized control of an ant colony allows it to easily work around problems in a network. Given a model of a working system, changes and restrictions can be applied to help simulate different styles of problems. For example, in a road network system a lane closure could be simulated through decreasing the maximum capacity of that path or by increasing the estimated time taken to traverse down it. In the case of a major road traffic collision it could be represented in a similar way or by completely removing a connection in network depending on the severity of the accident.

VI.REFERENCES

[1]. Euchi, Jalel and Mraihi, Rafaa: The Urban Bus Routing Problem in the Tunisian case by the Hybrid

Artificial Ant Colony Algorithm. In: Swarm and Evolutionary Computation, Vol. 2, Pages 15-24, Elsevier, 2012.

[2]. Ghoseiri, Keivan and Morshedsolouk, Fahimed: ACS-TS: Train Scheduling using Ant Colony System. In: Advances in Decision Sciences, Hindawi, New York, 2006.

[3]. Shi, Baizhan, Zhao, Chunyan and Zhang, Yong: Route Optimisation for Bus Dispatching Based of Improved Ant Colony Algorithm. In: The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010. Vol. 2, Pages 807-810, IEEE, 2010.

[4]. Weisstein, Eric W. "Traveling Salesman Problem." From MathWorld—A Wolfram Web Resource. http://mathworld.wolfram.com/TravelingSalesmanProblem.html. Accessed 16/02/2012

[5]. Jau-Ming, Su and Jen-Yu, Huang: Using Ant Colony Optimization to Solve Train Timetabling Problem of Mass Rapid Transit. In: Proceedings of the 2006 Joint Conference on Information Sciences, JCIS, Atlantis Press, 2006.

[6]. Transport for London. Facts and Figures. Accessed from ww.tfl.gov.uk/corporate/modesoftransport/londonunderground/1608.aspx

[7]. Rachel Lockwood. Applying Ant Colony Optimisation Techniques to the London Underground Network. University of Bath. 2008. Accessed from http://www.cs.bath.ac.uk/~mdv/courses/CM30082/projects.bho/2007-8/Lockwood-RL-dissertation-2007-8.pdf

[8]. Pavel Kr¨omer, Jan Martinoviˇc, Michal Radeck´y, Radek Tomis, V´aclav Sn´aˇsel. Ant Colony Inspired Algorithm for Adaptive Traffic Routing. 2011.

[9]. TomTom IQ Routes. Accessed from http://www.tomtom.com/page/iq-routes

[10]. B. Tatomir and L. Rothkrantz. Hierarchical routing in traffic using swarm-intelligence. 2006.

[11]. D. Alves, J. van Ast, Z. Cong, B. De Schutter, and R. Babuˇska. Ant colony optimization for traffic dispersion routing. 2010.

[12]. Punam Bedi, Neha Mediratta, Silky Dhand, Ravish Sharma, Archana Singhal. Avoiding traffic jam using Ant Colony Optimization – A novel approach. 2007.

[13]. An Efficient Ant Colony System Based on Receding Horizon Control for the Aircraft Arrival Sequencing and Scheduling Problem - Zhi-Hui Zhan, Student Member, IEEE, Jun Zhang, Senior Member, IEEE, Yun Li, Member, IEEE, Ou Liu, S. K. Kwok, W. H. Ip, and Okyay Kaynak, Fellow, IEEE

[14]. Aircraft Taxi Path Optimization Based on Ant Colony Algorithm - Changyou Liu, Yongcheng Wang - Civil Aviation University of China Tianjin 300300, China Ant Colony Optimization for Aircraft Landings. Ihsen Farah, Ali Kansou. Adnan Yassine, Thierry Galinho. University of Le Havre Ant Colony Optimization. Marco Dorigo, Mauro Birattari, and Thomas St¨utzle Universit´e Libre de Bruxelles, Belgium Ant Colony Optimization lecture Notes. Dirk Sudholt. University Of Birmingham

[15]. Aircraft Taxi Path Optimization Based on Ant Colony Algorithm - Changyou Liu, Yongcheng Wang - Civil Aviation University of China Tianjin 300300, China S. Isaacman, R. Becker, R. Caceres, S. Kobourov, M. Martonosi, J. Rowland, and A. Varshavsky. Identifying important places in people's lives from cellular network data. In Pervasive Computing, pages 133{151. Springer, 2011.

[16]. C. Kang, S. Sobolevsky, Y. Liu, and C. Ratti. Exploring human movements in Singapore: a comparative analysis based on mobile phone and taxicab usages. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp 2013), 2013.

[17]. L. Liao, D. Patterson, D. Fox, and H. Kautz. Learning and inferring transportation routines. Articial Intelligence, 171(5-6):311{331, 2007.

[18]. C. Pan, J. Lu, S. Di, and B. Ran. Cellular-based data-extracting method for trip distribution. In Transportation Research Record: Journal of the Transportation Research Board, volume 1945, pages 33{39, 2006.