Fusion of Artificial Transportation Systems and the World Wide Web

Songhang Chen, Fenghua Zhu, Gang Xiong and Peijun Ye

Abstract—**Artificial Transportation Systems (ATS) is a promising way to study complex actual transportation systems. When using ATS to deal with actual transportation issues, a solid information foundation for modeling and simulation is usually required. Thus, this paper fuses ATS with the Word Wide Web from two aspects. First, desired information is collected from open resources and people on the Web to improve modeling and simulation of ATS. Second, with rich media technologies, ATS can be shown on the Web. Benefiting from this, ATS become open and have more potential applications. To validate the feasibility, a prototype has been developed. Its software architecture together with user interfaces and web services are described. Finally, three applications are conducted on the prototype to demonstrate how collected information is used in ATS.**

I. INTRODUCTION

Artificial Transportation Systems (ATS) provide a new perspective to study urban transportation systems, which are a kind of open and complex giant system [1]. However, at present, constructing ATS which qualify for reference or alternative versions of actual transportation systems is a time and labor consuming job, due to contradiction of supply and demand of actual traffic-related information: On one hand, ATS employ agent-based modeling and simulation to emerge traffic flow form bottom up [2]-[4]. This means that only if all agents are modeled well and their actions are simulated exactly, the final macroscopic traffic flow can be reasonable and effective. To realize this, abundant actual information is required for reference. On the other hand, though information level of traffic-related fields has been improved greatly, their information systems usually exist as isolated islands due to safety and law concerns, and it is difficult to access even for the purpose of traffic research.

Therefore, the World Wide Web, commonly known as the Web, is fused with ATS in this paper. As the world's largest open source information platform, the Web has been involved into people's daily life and had a profound effect on the development of human society. Information on the Web is all-inclusive and there are always a great number of people online creating or sharing information and knowledge. Thus, by collecting desired actual information from the Web, a solid information foundation for construction and simulation of ATS can be provided. With rich media technologies, ATS can be shown on the Web so that the public can use ATS and offer information just through web browsers. And by web services, ATS can interact with other modules of intelligent transportation systems even in heterogeneous environments. In short, with the help of the massive information and people's group intelligence, many information obstacles of ATS can be overcome and ATS will become open. This is the motivation and also objective of this paper.

We organized this paper as follows. An existing simulation engine of ATS is introduced in Section II, together with related research on web information access and some available open source software. In Section III, we present key models and actual data required in ATS, followed by web information collection for ATS in Section IV. Section V addresses the software architecture of ATS with fusion of the Web. Base on this, a prototype is developed and its user interfaces together with web services are demonstrated. Three applications conducted on the prototype are described in Section VI. Finally, conclusions are drawn and future work is discussed.

II. RELATED WORK

A. TransWorld

From the raise of ATS in 2003, there are already many researches on modeling, experiment and application of ATS [5]-[7]. Together with computational experiments [8] and parallel execution theory [9], ATS aims to achieve an integrated solution to transportation issues. To this end, our lab developed an ATS simulation engine called as TransWorld with agent-oriented programming method [10]. Working like a 3D render engine, TransWorld can compute and present a dynamic traffic scene automatically according to inputted model data and running parameters, and output traffic measures and process data specified by users into a database for later analysis and evaluation. Compared with traditional transportation simulation programs, TransWorld seals simulated objects as agents, which exchange messages and interact with each other by request mechanism rather than direct method calling. This can satisfy better the requirement of real-time and concurrent processing during traffic simulation. It can be said that "Agent" specializes "Object" but has higher performance [11].

Manuscript received August 15, 2011. This work was supported in part by NNSFC Grants 70890084, 60921061, 60974095l, 60904057, 90920305, MOST 2006CB705506, 2006AA010106; and CAS 2F09N05, 2F09N06.

Songhang Chen is a PhD student at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. (e-mail: chensohg@gmail.com).

Fenghua Zhu is an associate research scientist with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100080, China. (e-mail: fenghua.zhu@ia.ac.cn).

Gang Xiong is a research scientist at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. (e-mail: gang.xiong@ia.ac.cn).

Peijun Ye is a PhD student at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. (e-mail: dreamflight@163.com).

Base on TransWorld, a Parallel Management System for Public Transportation developed by Wang's team was formally put into use at the Traffic Information Command Center of Guangzhou to offer public transportation management and service guarantee for 2010 Asian Games [12]. In this paper, TransWorld is also utilized and extended to a web program.

B. *Web Information Access*

According to the statistics of Netcraft, there are nearly 346 million websites in the world up to June 2011 [13]. To help people find desired information effectively, varied web search engines have existed and research on information extraction and retrieval is emerging. In the traffic field, Sara Carvalho *et al.* use support vector machine to find out traffic-related information from massive Twitter messages [14]. Passos L. S. *et al.* analysis classifiers for web content related to traffic information [15]. Both are pioneering attempts to access traffic-related web information. However, accurate classification is just the first step to apply traffic-related information into transportation modeling and simulation, since there is still a long way for machine to extract desired data which can be used directly, especially when this information is unstructured like natural language.

To enable machine to access web information, a search engine is usually required. It is noteworthy that developing a search engine independently is a large software project, which refers to complex information extraction and retrieval process. Fortunately, there are some open source projects available for saving software development: 1) Heritrix [16] is a configurable web crawler which can crawl along specified hyperlinks and download webpages as local documents; 2) HTMLParser [17] is a famous HTML parser tool which can be used to extract structured information from webpage; 3) Lucene [18] provides a Java-based indexing and search implementation, as well as advanced analysis and tokenization capabilities. These open source tools are integrated in this paper to develop a vertical search engine.

III. Information Requirements of ATS

ATS is an extension of traditional microscopic traffic simulation (TMTS) from the perspective of complex systems [19]. Their microscopic presentations of traffic flow may be similar, but their principles for production of traffic flow are different. TMTS is based on initialed static or dynamic origin-destination (OD) matrix, while ATS is based on the theory of traffic demand generation based on activity (TDGA) [20]. Concretely, every traffic participant is modeled as an agent with certain autonomy in ATS. Production of their travel is bottom-up and time-driven, as shown in Fig.1.

Firstly, the person agent is enabled to check the system clock actively. If it is not the end time of the current activity, it has to stay at its current place. Otherwise, it will request the Travel Demand Model (TDM) to generate demand and destination of next travel. Then, through the Travel Solution Model (TSM), a travel solution comprised of one or multiple pairs of route and vehicle can be decided. Third, according to the solution, the person agent will start to move in the circumstance. Finally, after arriving at the destination, it will invoke the Activity Duration Model (ADM) to decide how long it should stay in the destination to finish its travel demand.

The circumstance in ATS is classified into two categories: physical transportation infrastructure and abstract natural or social process. The first is mainly composed of intersections, roads and places. The second is composed of different events, such as weather, road repair, traffic accident and large-scale activity. By rules, the circumstance has effect on actions of agents [19]. For example, if it is rainy or foggy, some types of travel demand will be weakened and vehicle agents should decrease driving speed with a certain probability for safety.

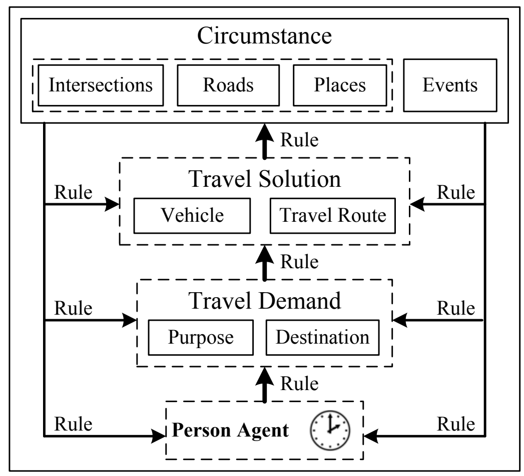

Fig. 1. Production of travel activity from bottom up.

The whole process is bottom-up and the circumstance adversely affects agents and models, so many factors may impact on the final macroscopic traffic flow of ATS. In order to model and simulate actual transportation systems exactly, all models and circumstance must be reasonable. Therefore, we think that ATS should require obtaining information from corresponding actual transportation systems from two aspects:

1) Models

As Fig.1 shows, the Person Agent is a core and bottom model of the whole ATS. Its attributes such as age, gender, household status are derived from an actual population structure. Place Agent is another key model requiring actual information. Because every travel of person agents has a place as departure and destination, place distribution impacts on the space-time distribution of traffic flow significantly. Only if all person agents synthesized and place distribution modeled are reasonable, TDM is able to arrange reasonable travel demand and travel destinations. Actual car ownership and proportion of different travel modes are important for TSM to decide travel mode and route. And duration distribution of different types of activity is important to ADM. Expect these models which impact on traffic flow directly, there are other models requiring actual information in ATS and see [21][22] for more details.

2) Events

Events that we pay attention to are also those that have effect on traffic flow significantly. This event information is usually required in computation experiments to evaluate some events and parallel execution of ATS. At present, TransWorld can take in and apply information of road repair, traffic accident, large-scale activity (e.g. sport game), and weather (e.g. rainy and foggy). We simulate their impacts on person agents' activity and movement of vehicles with IF-THEN rules.

It is costly to use traditional methods like questionnaire survey to get such information, if not impossible. By contrast, using the Web is a quite promising way. It does not cost much and can get a wide-range of traffic-related data and events which are usually open to the public and free. Furthermore, along with the coming of Web 2.0, once some emergencies happen, the Web as a social media can do quicker, more sensitive and exact reaction than traditional media [23]. In particular, even when some desired actual information does not exist on the Web, it is possible for net citizens from everywhere to contribute their knowledge. The emerging Q&A websites such as StackOverflow.Com, Quora.Com and Baidu Zhidao (zhidao.baidu.com) are successful examples in this aspect.

IV. WEB INFORMATION COLLECTION FOR ATS

For ATS, we design two kinds of web information collection modes: active and passive. The active mode is using machines to collect desired information actively from public websites and open third-party web interfaces, especially information of which are structured or semi-structured and can be transferred into workable data directly. The passive mode is relaying on a website to wait people to offer information, which usually needs manual or semi-automatic transformation before use.

A. Active Mode

To collection web information, a search engine is necessary. Considering the fact that traffic-related information just takes up a small part of the Web, a vertical search engine which focuses on a specific segment of online content is enough for us. Compared with general-purpose search engines such as Google and Baidu, a vertical search engine is lightweight and more flexible. Besides, unstructured information is still too complex for machine to process automatically at present, so we prefer to search structured or semi-structured information in this paper.

The process of our information extraction and retrieval is shown in Fig.2. Firstly, a web crawler is activated to crawl along hyperlinks which march a certain regular expression and download webpages as off-line webpage documents on a local disk system. To reduce data amount of downloads, only text information is requested. Secondly, if a webpage document is structured, mature DOM (Document Object Model) tree technologies [24] can be used to extract desired data easily. Otherwise, it is needed to use some Text Mining methods like Natural Language Processing (NLP) [25]. The data extracted is store into a relational database. Finally, information retrieval technology is used to build indexes for webpage documents in the file system and structured data in database. Based on this, user and program interfaces for quick information lookup can be offered.

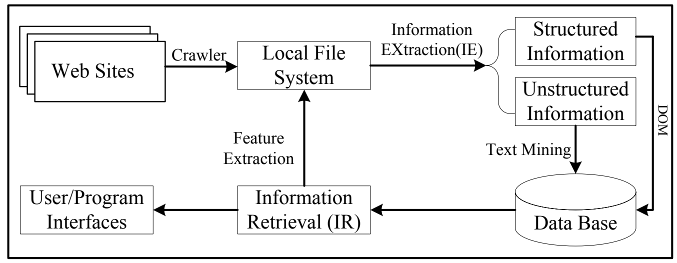

Fig. 2. Process of information extraction and retrieval.

With open source projects mentioned in Section II-B, we developed a tentative vertical search engine for traffic. Searching information of different types of place in Haidian District of Beijing is as a task to test the engine. First, some hyperlinks are as seeds for the crawler Heritrix to crawl and download webpages. To avoid download useless webpages, regular expressions are used as filters to march new hyperlinks. Then, HTMLParser is used to parse semi-structured webpages and extract desired data. Finally, we get information of 1587 residential areas, 396 office buildings, 9017 restaurants, 2305 entertainment places, 780 playgrounds and 6310 shops.

Data visualization can make information checking easier. Therefore, we get latitude and longitude of every place by its address with the open Google Geocoding Service [26]. And then through the open Google Maps JavaScript API (Application Program Interfaces) [27], show them in the form of WebGIS. Fig.3 (a)-(c) illustrate partial restaurants, residential areas and entertainment places within the Zhongguancun Area respectively, which is a part of Haidian District. As Fig.3 (d) shows, main roads of this area is extracted. An ATS based on this road network and called as the Zhongguancun ATS is constructed for applications in Section VI.

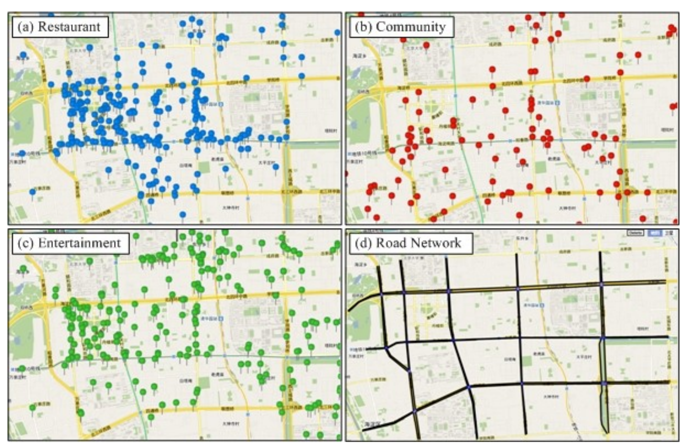

Fig. 3. Data visualization for places gathered and road network extracted.

In addition, if related third-party information systems publish their data on the Web, we will connect with them according to their protocols. For example, a Chinese website [28] offers real-time weather information in the form of standard web services. Its data is from authoritative China Meteorological Administration and updated about every one hour. We use it as an alternative information source for ATS to get real-time actual weather condition.

B. Passive Mode

As an important complement to active information collection from open resources, an interactive website with the topic of transportation is constructed to collect actual information from net citizens. Concretely, according to actual data that ATS needs, two types of information interactions are designed to make use of people's group intelligence.

(1) Model Information

The website opens interfaces for people to view and edit basic models in the ATS, such as places, roads and traffic lights. Because actual transportation systems refer to a great deal of engineered infrastructure and some changes with time easily, it is almost impossible for just several professional persons to cover all infrastructure in the real world and check all information collected by the vertical search engine. Therefore, we need the help of the public, who live in different places and can correct, consummate and update information of models in the ATS. Public participation also brings other benefits. For example, some business places which many people pay attention to would also cause the attention of ATS to adjust the TDM model, which is responsible to select a destination for a person agent's travel.

The website also welcomes people to submit their personal information such as age, gender, home address, family income, daily travels, and avocation. This information can be used as demographic samples to generate person agents and improve artificial population model of ATS [7][29].

(2) Event Information

Functions that allow people to release information about weather, traffic accident, road condition, and so on are also offered. Today, everybody can report news. As for some local events, it is more possible for the local to report and disseminate them by micro-blog or BBS with their mobile phones or PC at the earliest time. Thus, encouraging people to release information of events on the website is very helpful for ATS to feel changes in the actual world and touch off corresponding processing timely.

In additional, a Q&A module is necessary, because some information and statistics is beyond search engines. For example, an ATS needs information of a limited urban area, but there is only related information of the whole city on the Web. In this case, by the Q&A module, we can publish questions and encourage people to contribute their knowledge or launch a web survey to get related statistics. In return, a well-modeled ATS can offer free and available services for people's usual life. For example, people's daily travel solutions can be evaluated and some travel suggestion of every day can be given out.

V. PROTOTYPE

Based on existing work and some other open source projects for building a website, we made a software implementation for ATS with fusion of the Web (abbr. as Web-ATS). The layered software architecture is described. User interfaces and web services are demonstrated.

A. Software Architecture

The architecture is shown in Fig.4.The web client includes mainstream networked devices, such as traditional computers, mobile terminals and some automatic clients (e.g. third-party web servers). The web server adapts the classical three-layer structure composed of Presentation Layer, Business Logic Layer and Data Access Layer. Besides, Data Storage Layer and Computing Support Layer are also needed to support massive data storage and huge computations respectively. Each layer is introduced as follows.

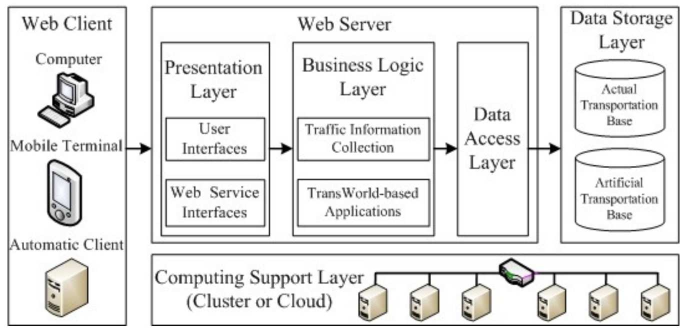

Fig. 4. Architecture diagram of Web-ATS.

Presentation Layer consists of user interfaces and web service interfaces. User interfaces offer user visual interaction with the ATS and web service interfaces offer ATS interoperation with third-party applications. Rich media technologies such as Flex and JavaScript are applied to make user interfaces and web presentation of ATS user-friendly.

Business Logical Layer is also known as the domain layer. It is not only responsible for the web information collection mentioned in previous section but also applications based on TransWorld. Concretely, Wang proposed three typical applications of ATS [30][31]: 1) Training and Learning; 2) Testing and Evaluation; 3) Control and Management. How to implement and present them on the Web is still needed to be designed systemically, but some exploratory work is done in this paper to validate the feasibility of existing technologies. TransWorld, which was a desktop program and cannot be accessed remotely, is sealed with ASP web service to be a web program.

Data Access Layer centralizes all data access interfaces. The setting of this layer is to decouple application and storage of data, and shield differences of Data Storage Layer, which is divided into Actual and Artificial Transportation Base for convenience use and maintenance. The Actual Transportation Base stores the data collected from the Web or some physical sensors, and the data used for construction of ATS and produced during simulation are stored in the Artificial Transportation Base.

Computing Support Layer takes charge of the running of ATS. Because we need Web-ATS to run different ATS or multi instances of an ATS at the same time, a distributed computing platform which can offer flexible computational ability is asked in this layer, such as computer clusters and cloud computing services.

B. User Interfaces

Friendly user interfaces are very important to a web system, so we use free Google Map as the carrier of ATS to provide intuitive interaction. For example, if users choose the Zhongguancun ATS to see, the corresponding web map will be loaded as the background and the navigation tool is offered, as shown in the left of Fig.5. Every tag on the web map represents a modeled place in the ATS. And roads and intersections covered by the ATS are shown with black lines and light-blue polygons. In the right of Fig.5, it is an operation panel for users to interact with ATS, including: 1) submit personal information; 2) view and edit places or roads, including three atomic operations—add, delete and update; 3) report real-time event information; 4) submit suggestions like strategies of traffic signal control.

Fig. 5. A webpage for users' interaction.

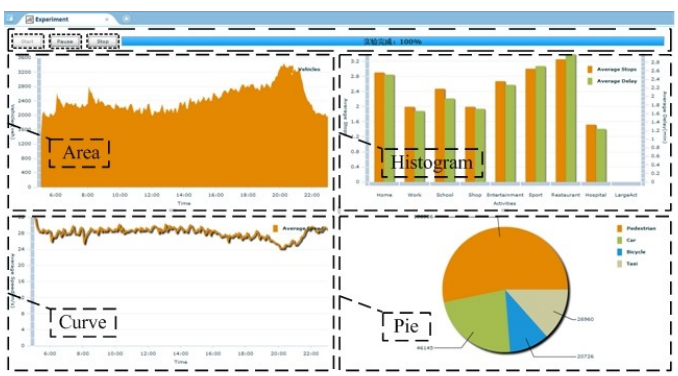

Fig. 6. A webpage of a remote experiment.

An ATS can be tread as a "guinea pig" on computer to conduct repeatable, controllable, and configurable experiments, which are called computational experiments and aim to analysis and evaluate transportation systems. Thus, the Web-ATS also allows users to carry out computational experiments remotely just with web browsers. As shown in the top of Fig.6, there are three buttons to start or pause or stop an experiment and a progress bar to indicate running of the experiment. When running is done, the Web-ATS will return a web report of some traffic measures, as shown in the bottom of Fig.6. The area diagram in the top left corner describes the changes of motor vehicle number in the road-network over time. The double y-axis histogram in the top right corner shows the average stop times and delay time of different travel activities by car. The curve in the bottom left corner presents the changes of average speed of motor vehicles over time. The pie chart in the bottom right corner shows the proportion of different travel modes. Until now, TransWorld supports 7 kinds of modes: car, bicycle, walk, taxi, subway, common bus and Bus Rapid Transit (BRT).

C. Web Services

In recent years, as requirement for flexibility and integration of IT business increases, the concept of Web Service is emerging and gets wide support in IT field. Because of web services, interoperation among different platforms in heterogeneous environments becomes easy. Thus, in order to extend application domains of ATS, we apply web service technologies to open some functions of ATS so that other systems can use them freely. To realize this, a XML file with Web Service Description Language (WSDL) is first and must. As the name suggests, WSDL is used to describe web services and can be understand as public function declaration informally. Take real-time road condition query of ATS as an example. A XML segment to indicate this function is as follow:

```
<operation name="QueryRoadCond">
  <input message="tns:InQueryRoadCondRequest"/>
  <output message="tns:OutQueryRoadCondResponse"/>
</operation>
```

The *QueryRoadCond* is the name of function. The input and output data are both sealed as message. And the input message *tns:InQueryRoadCondRequest* is declared further as follow:

```
<message name="InQueryRoadCondRequest">
  <part name="QueryRoadCond_roadName" type="xsd:string"/>
</message>
```

As we see, there is only one parameter, name and data type of which are specified in pairs. Similarly, the output message tns:OutQueryRoadCondResponse is as follow.

```
<message name="OutQueryRoadCondResponse">
  <part name="QueryRoadCond_Result" type="xsd:string"/>
</message>
```

The XML information is self-describing so that third-party programs are able to know how to invoke the function of ATS through the Web. With the same way, ATS can know and use web services of other systems.

VI. Application

Based on the prototype and Zhongguancun ATS, we conducted three applications, in which information gathered from the Web is used. Both real-time and non-real time information is demonstrated.

A. Population Synthesis

In ATS, Artificial population takes places of OD matrix to emerge traffic flow from bottom up, so it is important to synthesis reasonable population [29]. With the fifth China Population Census published on the Web by NBSC (National Bureau of Statistics of China), we can get related data of population structure to synthesize artificial population for the

Zhongguancun ATS. In China, the census is carried out by sub-district offices. And the area modeled in the Zhongguancun ATS is under the jurisdiction of Zhongguancun Sub-district Office politically. The key data of this sub-district is extracted and presented in the following two tables.

TABLE I
MACROCOSMIC STATISTICS

Item	Number		
	Total	Male	Female
Population	58367	32185	26182
Household Population	48768	24410	24358

TABLE II
AGE DISTRIBUTION

Age Interval	Number		
	Total	Male	Female
<=14	6390	3361	3029
15-64	46935	26136	20799
>=64	5042	2688	2354

In Table I, the first row shows the number of total population, male population and female population in turn. The second row shows the same three kinds of number of household population, which is a part of population and refers to people with households. According to the statistics, there are total 16511 households in the Zhongguancun Area. The rest population is composed of those who occupy group quarters. Table II lists the age distribution which is divided into three groups: less than 14, between 15 and 64 and more than 64. In each ground, total number together with number of male and female are also given out.

Based on this data, we can start the population synthesis. First of all, it is need to use place information collected before to create household and group quarter agents. For example, communities are treated as households and dormitories of collage or company are used as group quarters. Then, the process to make a person agent can be repeated to synthesize all population. Concretely, the process can be divided into three steps.

--Step 1: based on the rate among total number of three age intervals, use Monte Carlo method to decide an age interval and then select an integer in this interval as the age of the new person agent. The maximum of age is limited as 120.

--Step 2: according to the age in the Step1, find the interval which the age owing to in the Table II. Then use the male-to-female rate to decide a gender.

--Step 3: if there is not enough household population, choose an empty household or arrange one with some person agents already for this person agent. The male-to-female rate of household population should be kept. Otherwise, choose a group quarter for it.

Finally, when the synthesis is finished, the same statistics with Table I and II can be worked out to validate the structure of synthesized population. As shown in Table III and IV, the symbol '$(=)$' behind a data means that the data is the same with the census; '$(\uparrow\delta)$' means that the data is δ percent more than the census; '$(\downarrow\delta)$' means that the data is δ percent less than the census. The result shows that the max relative error is 1.06%, which is an acceptable error.

TABLE III
MACROCOSMIC STATISTICS of SYNTHESIZED POPULATION

Item	Number		
	Total	Male	Female
Population	58367 (=)	32057(↓0.40)	26310(↑0.49)
Family Population	48768 (=)	24410 (=)	24358 (=)

TABLE IV
AGE DISTRIBUTION of SYNTHESIZED POPULATION

Age Interval	Number		
	Total	Male	Female
<=14	6358(↓0.50)	3341(↓0.60)	3017(↓0.40)
15-64	47004(↑0.15)	26040(↓0.37)	20964(↑0.79)
>=64	5005(↓0.73)	2676(↓0.45)	2329(↓**1.06**)

TABLE V
WEATHER INFORMATION COLLECTED OF A DAY

Time Interval (HH:mm)	Temperature (℃)	Wind Direction	Wind Level	Humidity (%)
04:50-05:44	26	South	2	55
05:44-06:55	26	Southwest	2	55
06:55-07:45	26	South	2	57
07:45-08:52	26	South	2	58
08:52-09:58	27	Southwest	2	55
09:58-10:48	28	South	2	54
10:48-11:54	29	South	3	54
11:54-12:45	31	Southwest	3	41
12:45-13:54	31	Southwest	2	39
13:54-14:44	31	Southwest	2	40
14:44-15:50	31	Southwest	2	41
15:50-16:56	32	South	2	37
16:56-17:47	32	South	2	38
17:47-18:55	31	South	2	36
18:55-19:46	31	South	2	37
19:46-20:54	29	South	2	44
20:54-21:44	29	South	2	48

B. *Real-time Weather*

When ATS is used in Control and Management application, the actual and artificial systems must be connected in real time and online, and the artificial systems must replicate actual behaviors with high fidelity [31]. This means that the running of TransWorld should keep synchronization with the physical clock and take in changes of actual systems timely.

As a kind of event information that impacts on transportation directly, actual weather condition is monitored by Web-ATS and any change will be sent to corresponding ATS which is running synchronously. The Table V summaries the weather changes of Zhongguancun Area on July 10, 2011. It was a sunny day that even a little hot. This data is gained from open third-party web services every 30 minutes, which is a time granularity enough for weather changes in most cases.

Rainfall is another important weather factor, but we had not found available websites or services to get this information until finished this paper. The reason may be that rainfall is zero in most cases and when it is rainy, rainfall is usually measured by an hour or a day, which is far from real-time. It is expected that people release rain-related information on the Web-ATS so that TransWorld can feel the weather change.

Impacts of weather condition on ATS by IF-THEN rules can be concluded into three aspects.

--First, travel demand generated by the TDM model is affected by the current weather. For example:

If the current temperature is more than 35 ℃, then shopping is canceled with probability 0.8.

--Second, activity duration decided by the ADM model is affected by the current weather. For example:

If it is rainy, then duration of eating out is extended by half an hour with probability 0.5.

--Third, movement of vehicles in the road network is affected by the current weather. For example:

If it is rainstorm, then vehicles decrease their max speed to 80% with probability 0.7.

During synchronous simulation, number and average speed of motor vehicles in the ATS are recorded every 5 minutes by TransWorld. These two indicators are important to represent traffic congestion status and are widely used in urban traffic control and management. Thus, we can use these two indicators to observe impacts of weather condition on ATS. Fig.7 illustrates two group curves. The two solid curves are results of a rainy day, when there is rainstorm last from about 4:30am to 5:30am, and the two dotted curves are the results without respect to weather condition. Due to the rainstorm, there is a sharp increase of motor vehicles number and a corresponding decrease of average speed.

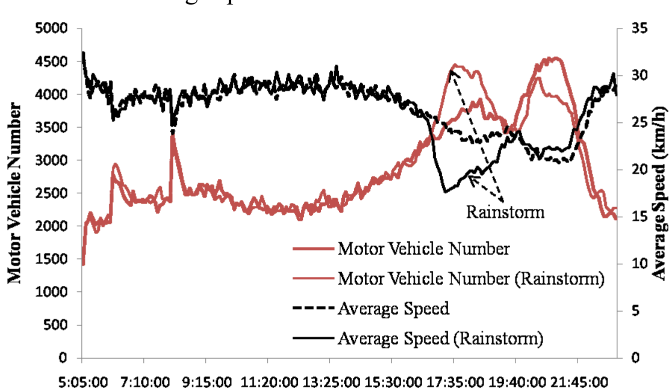

Fig. 7. Number and average speed of motor vehicles.

C. Road Condition Comparison

Production of traffic flow in ATS is based on bottom-up emergence rather than traditional OD matrix from practical investigation, so to what extent the traffic flow produced in an ATS is similar to the actual is a common concern of researchers. Thus, we carried out the experiment to compare two kinds of traffic flow by road condition.

At present, many websites offers real-time road condition information on the Web, like the official site of Beijing Traffic Management Bureau [32]. This site presents road condition of main roads in Beijing, which is updated every 2 minutes and classified into three levels by driving speed: smooth (more than 40km/h), slow (between 20km/h and 40km/h) and crowd (less than 20km/h). From this site, we use the Web-ATS to collect time-varying road condition of desired roads. To balance the performance of vertical search engine and the timeliness of information, the search cycle of engine is set as 5 minutes.

According to data collected on July 5, 2011, Fig.8 illustrates actual road condition of two roads—Jiyu Bridge to Park Road and P.E BNU to Dazhong Tempal. The horizontal axis is the time from 5:00 to 23:00 and the vertical axis is the degree of congestion. Degree 1, 2 and 3 represent smooth, slow and crowd road condition respectively. These two roads are also modeled in the Zhongguancun ATS, so by running TransWorld to simulate the ATS also from 5:00 to 23:00, we can get their artificial road condition at the same time point.

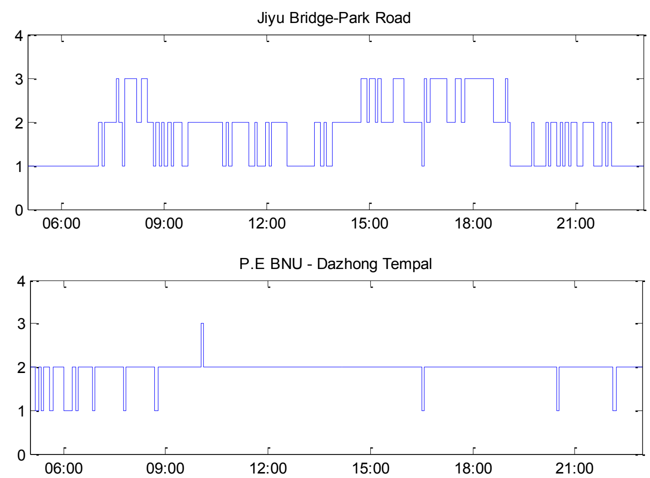

Fig. 8. Stair step chart of road condition of two roads.

To compare the actual and artificial road condition of a road, we treat them as time series denoted as $S'=\{s'_{t_1},s'_{t_2},...,s'_{t_k}\}$ and $S''=\{s''_{t_1},s''_{t_2},...,s''_{t_k}\}$ $(s'_{t_i},s''_{t_i} \in \{1,2,3\}, i \in \{1,2,...,k\})$ respectively. Then, based on the following Euclidean distance formula, we can calculate their Euclidean distance (Ed). To describe similarity by Ed, we classify degree of similarity into three levels: low (if the Ed is more than 20% of the maximum), middle (if the Ed is between 10% and 20% of the maximum), and high (if the Ed is less than 10% of the maximum). In this case, because $k=216$ and $\left|s'_{t_i}-s''_{t_i}\right| \le 2$, we can calculate that the max Euclidean distance is 29.39.

Euclidean distance formula:

$$d(S',S'')=\sqrt{(s'_{t_1}-s''_{t_1})^2+(s'_{t_2}-s''_{t_2})^2+....+(s'_{t_k}-s''_{t_k})^2} \quad (1)$$

Take those two roads in Fig.8 as examples. The result of comparison is shown in Table VI.

TABLE VI
CALCULATION RESULT

Road Name	Ed	Similarity
Jiyu Bridge - Park Road	15.33	low
P.E BNU - Dazhong Tempal	14.25	low

With the same method, we can compare other roads in the Zhongguancun ATS to get a whole understand. The reason of low similarity should be analyzed seriously. It may indicate weakness of models and rules in the ATS. More work is still needed to demonstrate the ability of ATS to generate traffic flow similar with the actual, though the aim of ATS is not to represent or approach the true state of actual traffic systems. We think an ATS which is proved effective to assist transportation engineer and managers in practice is enough.

VII. CONCLUSION

In this paper, to make construction, simulation and application of ATS more effective, ATS is fused with the Web on the meanings and forms. Web information collection mechanism for ATS from open resources and net citizens is designed and implemented. Based on existing work and open source projects, a prototype was developed to show ATS on the Web to collect information and offer services. Its layered architecture together with user interfaces and web services are also addressed. On this prototype, three applications were conducted to show how information collected is used in ATS.

Our future work will still focus on collection and application of web information for ATS. The traffic vertical search engine must be enhanced to access more kinds of information, such as some real-time public emergency events. Machine learning methods can be adapted to improve models and rules of ATS, for example, those involved in making person agents' activities. And data mining methods can be employed to analysis massive data produced by ATS and discovery interesting patterns.

ACKNOWLEDGMENT

We would like to thank Professor Fei-Yue Wang for his intensive instructions. We also would like to thank the Lab of Complex Adaptive Systems for Transportation (CAST), Prof. Shuming Tang, and all hard works of Yisheng Lv, Yan Ou and Peijun Ye. Without their works and supports, we cannot implement the introduced ATS platform.

REFERENCES

[1] Qian, X.S., Yu, J.Y., and Dai, R.W. A new discipline of science— the study of open complex giant system and its methodology [J]. Nature Journal, 1990, 13(1), 3-10.

[2] F.-Y. Wang, Shuming Tang. A Framework for Artificial Transportation Systems: From Computer Simulations to Computational Experiments [A]. Proceedings of 8th IEEE International Conference on Intelligent Transportation Systems [C]. Sep. 13-16 2005, Vienna, Austria: 1130-1134.

[3] Tang Shuming, Wang Feiyue, Liu Xiaoming, etc. A Preliminary Investigation on Artificial Transportation Systems [J]. Journal of System Simulation, 2005, 17(3):704-709.

[4] F.-Y. Wang, S. M. Tang. Concept and Framework of Artificial Transportation Systems [J]. Journal of Complex Systems and Complexity Science, 2004, 1(2):52-59.

[5] Tang Shuming. A Preliminary Study for Basic Approaches in Artificial Transportation System [D], Institute of Automation, Chinese Academy of Sciences, 2005.

[6] Zhu Fenghua. A Study on the Evaluation of Urban Traffic Signal Control System Based on Artificial Transportation Systems [D], Institute of Automation, Chinese Academy of Sciences, 2008.

[7] Zhao Hongxia. A Study on Computational Experiments of Relationship between Artificial Population and Travel Demand [D], Institute of Automation, Chinese Academy of Sciences, 2009.

[8] WANG Fei-yue. Computational Experiments for Behavior Analysis and Decision Evaluation of Complex Systems [J]. Journal of System Simulation, 2004, 16(5):893-897.

[9] WANG Fei-yue. Parallel system methods for management and control of complex systems [J]. Control and Decision, 2004, 19(5):485-514.

[10] Y. Shoham. Agent-oriented programming [J]. Artificial Intelligence, Vol.60, No.1, pp.51-92, 1993.

[11] Feng Shan, Huang Jingping, M N Jun, SHEN Chong. Agent Oriented Technology for Modeling and Simulation of Complex Systems [J], Journal of System Simulation, 1999, 11(3):161-166.

[12] Gang Xiong, Kunfeng Wang, Fenghua Zhu, etc. Parallel Traffic Management for the 2010 Asian Games [J]. IEEE Intelligent Systems, vol.25, no.3, 2010, pp.81-85.

[13] [Online]. Available: http://news.netcraft.com/archives/2011/06/07/june-2011-web-server-survey.html.

[14] Sara Carvalho, Lu´ıs Sarmento, Rosaldo J. F. Rossetti. Real-Time Sensing of Traffic Information in Twitter Messages [A]. Proceedings of the IEEE ITSC 2010 Workshop on Artificial Transportation Systems and Simulation.

[15] Passos L. S., Rosaldo Rosseti. Analyzing Classifiers for Web Content Related to Traffic Information [A], Proceedings of the 2010 IEEE Conference on Automation Science and Engineering, Toronto, Ontario, Canada,pp. 307-312.

[16] [Online]. Available: http://sourceforge.net/projects/archive-crawler/.

[17] [Online]. Available: http://htmlparser.sourceforge.net/.

[18] [Online]. Available: http://lucene.apache.org/.

[19] Jinyuan Li, Shuming Tang, Xiqin Wang, etc. Growing Artificial Transportation Systems: A Rule-Based Iterative Desgin Process [J]. IEEE Transaction on Intelligent Transportation System, 2011, 12(2):322-332.

[20] W. Davidson, R. Donnelly, et al., "Synthesis of first practices and operational research approaches in activity-based travel demand modeling," Transportation Research Part A: Policy and Practice, vol. 41, pp. 464-488, 2007.

[21] Jinyuan Li, Shuming Tang, Xiqin Wang, Fei-Yue Wang. A Software Architecture for Artificial Transportation Systems - principles and framework [A]. Proceedings of the 10th IEEE International Conference on Intelligent Transportation Systems[C], 2007, pp.229–234.

[22] Zhu Fenghua, Feiyue Wang, Dong Fan, etc. Modeling and Analysis of Transportation Systems Using ACP Approach. Proceedings of the 2011 IEEE Conference on Artificial Transportation Systems, accepted.

[23] ZENG Dajun, WANG Feiyue, CAO Zhidong. Open Source Information in Emergency Response [J]. Science & Technology Review, 2008, 16(6):27-33.

[24] Li Xiaodong, Gu Yuqing. DOM-based Information Extraction for the Web Sources. CHINESE J.COMPUTER, 2002, 25(5):1-8.

[25] Isahara, H.Resource-based Natural Language Processing. International Conference on Natural Language Processing and Knowledge Engineering, 2007, pp: 1-2.

[26] [Online]. Available: http://code.google.com/intl/zh-CN/apis/maps/documentation/javascript/services.html.

[27] [Online]. Available: http://code.google.com/intl/zh-CN/apis/maps /documentation/javascript/basics.html.

[28] [Online]. Available:http://www.webxml.com.cn/webservices/weatherwebservice.asmx.

[29] Miao Qinghai, Zhu Fenghua, Lv Yisheng, Cheng Changjian, Chen Cheng, Qiu Xiaogang, "A Game-Engine-Based Platform for Modeling and Computing of Artificial Transportation Systems", IEEE Transactions on Intelligent Transportation Systems, 2011, 12(2):343-353.

[30] Fei-Yue Wang. Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications [J]. IEEE Transactions on Intelligent Transportation System, Vol.11, No.3, September 2010.

[31] F.-Y. Wang. Toward a Revolution in Transportation Operations: AI for Complex Systems [J]. IEEE Intelligent Systems, vol. 23, no. 6, 2008, pp. 8–13.

[32] [Online]. Available: http://sslk.bjjtgl.gov.cn/roadpublish/Map /trafficOutNew1.jsp.

[33] Zhu Fenghua, Li Guoxi, Li Zhenjiang, Chen Cheng, Wen Ding, "A Case Study of Evaluating Traffic Signal Control Systems using Computational Experiments", IEEE transactions on Intelligent Transportation Systems, 2011, accepted.