Hybrid Calibration of Agent-Based Travel Model using Traffic Counts and AVI Data

2 authors, including:

Fei-Yue Wang
Qingdao academy of intelligent indistries
1,101 PUBLICATIONS   **11,368** CITATIONS

Some of the authors of this publication are also working on these related projects:

Parallel Vision for Perception and Understanding of Complex Scenes View project

Social manufacturing View project

Hybrid Calibration of Agent-Based Travel Model using Traffic Counts and AVI Data

Peijun Ye and Fei-Yue Wang, *Fellow, IEEE*

Abstract—Artificial transportation system is critical to parallel transportation management. At its kernel is the agent-based computation which simulates individual's travel behaviors via disaggregated models and "grows" complex traffic scenarios for computational experiments. However, as a common problem, there still lacks a general calibration method for its agent travel behavioral models. Motivated by this, the paper proposes a calibration method for agent-based travel model in artificial system, which correlates macro traffic data with micro behavioral model parameters. The behavioral model is calibrated by two data sources from real urban transportation: link traffic counts and Automatic Vehicle Identification data. Our results indicate that the proposed method can help receive a reasonable model and be applied in general calibration problems.

Keywords—Calibration; Agent-Based Model; Parallel Transportation Management

I. INTRODUCTION

Parallel transportation control is an emerging technology for complex urban transportation management [1] [2]. From the perspective of complex system, it aims to generate various traffic scenarios in artificial transportation system (ATS) and impose the most appropriate control strategy on actual traffic via diverse computational experiments. By connecting and simultaneously controlling the artificial and actual systems, parallel management can guide the actual transportation to a statistical "optimal" direction, thus to some extent, solve the systemic complexity through a bottom-up approach. Such process is called the ACP (Artificial Society, Computational Experiment, Parallel Execution) method [3]– [11]. Evidently, as a fundamental part of the parallel transportation management, ATS needs to generate and simulate various scenarios for the traffic policy test and validation.

At the core of ATS is the agent-based computation that adopts the basic idea of artificial society and social computing [12]– [14]. At the beginning, ATS is initialized by synthesizing an artificial population [15] [16]. The artificial population is generated in spatial distributions according to census, traffic survey, labor force survey, tax record, and other urban survey

data [17]. Then disaggregate travel behavioral models are introduced to construct software-defined agents [18]. These heterogeneous agents will autonomously perform reasoning and planning to complete their own routine activities. Therefore, when such activities need to be conducted in different places, travel takes place. Obviously, it is required to calibrate the agent-based travel models so that ATS can provide reasonable and reliable traffic demand for the successive computational experiments. This calibration, at least, should endow the ABM with the ability that can both retain heterogeneity of individual decisions and reproduce historical traffic flows. However, although some research has been conducted about the issue, there still lacks a general agent calibration method to our knowledge. Motivated by this, the main contribution of this paper is to propose a solution by relating agent micro decision parameters to overall traffic flows. The proposed method is tested by calibrating one of ABMs in ATS via actual traffic data from Chengdu, China. And results show that it can help receive a reasonable model and reproduce a satisfactory traffic flow.

The remainder of this paper is organized as follows. Section II reviews some representative calibration methods and models in agent-based traffic simulation, specifically the calibration of public transit path choice and pedestrian dynamics. Section III elucidates our calibration methods with some theoretical analysis. The actual traffic dataset used in our calibration is introduced in Section IV. The evaluation criterion and the calibrated results are also presented in the section. Finally, the paper concludes with some further discussions in Section V.

II. RELATED WORKS

Agent-based traffic simulation has become an indispensable approach for travel demand analysis. Many simulation systems or platforms have been developed during the past decades such as FAMOS [19], CEMDAP [20], ALBATROSS [21], ILUTE [22] [23], TASHA [24], TRANSIMS [25] [26] and so on. In the last few years, many studies have concentrated on the calibration of ABM in this field.

In general, the mechanisms of ABM calibration in transportation simulation can be roughly categorized into two types—manual adjust by experience and automated calibration. Though the latter is much demanding on the technical side and requires certain level of expertise in the use of sensors and software, it is less time and participation-consuming thus is becoming more prevalent. Tang and Jia pointed out that it is necessary to calibrate and validate the pedestrian model before using it to mimic the crowd dynamics [27]. They adopted a regression approach based on least square method to calibrate the social force pedestrian model. Real

pedestrian tracks in a Beijing massive transit railway were used to test the calibration on Netlogo multi-agent simulation platform. The outcome show the calibrated model was able to reproduce the characteristics of pedestrian flow. Zhao and Sadek reported their continued efforts toward the development, calibration, validation and application of a large-scale, agent-based model based on TRANSIMS platform [28]. In their study, the calibration focused on high level and a total of 162 hourly count station volume data from the Buffalo-Niagara metropolitan area were used. Total traffic demand as well as its diurnal distribution was calculated through arbitrary trials, which seems not operable in large scale scenarios. Voloshin et al. used generic algorithms to calibrate a pedestrian behavioral model through a rather frugal optimization of parameters [29]. The behavioral model was implemented by a social-force inspired algorithm and the mutual repulsion among agents as well as among the agents and obstacles was calculated. According to the video data from a busy metro station in the center of St. Petersburg, the calibration focused on the agent count, flow rate, passage time, etc.. Unlike the above three approaches where calibration more or less relies on the developers experience, Oliveros and Nagel developed an autonomous calibration procedure for public transport path choice [30]. The procedure was embedded in each simulation iteration, updating the posterior plan choice probability in a multinomial logit model. The day-based counts of Berlin public transport system was collected and used as a test scenario. At start, the fixed choice set and randomized routing had 128.5% and 100% mean relative errors. After the calibration, the parameters were gradually adjusted and the errors reduced to 42% and 15% respectively around iteration 600. Recently, Zeng et al. proposed a hybrid approach for modeling of pedestrian movement at signalized crosswalks by the combining route search and social force-based method [31]. Their model calibration adopted a generic algorithm, specifically the evolutionary computing process involving reproduction, mutation, crossover, and selection. The fitness function was set to be the combination of the relative distance error and the relative angle error. Crosswalks captured by an optical camera at an intersection in Beijing, China, were collected to be the benchmark, which consisted of the trajectories of 494 pedestrians and 156 turning vehicles. The results demonstrated that the simulated trajectories are consistent with the true trajectories within the acceptable deviation.

Though much effort has been made in the calibration of ABM, some difficulties still remain. In essence, the problem lies in the ambiguity of replication in modeling. From the view of transportation management center, it is the systemic efficiency of the road network rather than individual travel convenience is always the pursuit. Such objective is usually achieved by measuring and optimizing aggregate flows. In the context of models created to describe realistic traffic phenomena, emphasis must be placed on the reproducibility of computational experiments, so that management policies and control strategies are validated scientifically. But unfortunately, ABM makes it impossible to think in an aggregate way. Because in agent-based simulations, aggregate metrics are only measured from the bottom up by summing the individual quantities. As a consequence, the interpretation of the mechanism of traffic flow fluctuation, for instance, is somehow arbitrary. The difficulty arises due to the fact that the models contain several

parameters and these parameters are microscopic. Moreover, the parameters cannot be directly measured but they must be correctly tuned. Thus as an alternative, most scholars have to use a real traffic dataset to determine the parameter values as reviewed before. And it is increasingly accepted that the simulated traffic flow from ABM should closely match the actual data. Based on this, a general calibration method is proposed in the next section.

III. CALIBRATION OF AGENT-BASED MODEL IN ATS

This section will give a brief introduction of agent travel model in ATS at first, and then present our calibration method.

A. Agent-Based Travel Models in ATS

Based on basic idea of artificial society, ATS considers comprehensive facets of transportation, namely population distribution, individual internal state, activity chain, traffic infrastructure, social network, etc. [32] [33] [34]. To make this paper self-sustained, agent travel models, specifically the activity place selection and travel path determination, are introduced concisely.

Since daily traffic demand stems from the travel behaviors of urban residents, ATS starts the traffic simulation by synthesizing artificial population using census data. Initially, the population generator distributes the basic population to residential places of the network where other types of places such as schools, hospitals, parks, malls, business centers, hotels are also contained. Each person is modeled as an agent and has his own heterogeneous attributes. The agent will arrange his daily activities as an activity chain and perform each of them sequentially. Currently, two kinds of activities are considered, one of which is fixed activities whose start time, end time and duration cannot be altered easily while the other is flexible activities that can be adjusted according specific environment. Places and travel path of fixed activities like working, going to school are usually deterministic. And agents may probably select his travel mode, destination and travel path before each flexible activity. To limit the scope, this paper will only calibrate the destination and path selection models, which are reviewed as follows.

When the simulation clock reaches his expected departure time, each agent will choose his destination to perform his next activity according to his activity plan. This selection process is conducted among the place candidates that match the activity type in a certain range. Each eligible place will be endowed with a probability that is calculated via the maximum entropy method:

$$p_{(j|i)} = \frac{exp(\alpha \cdot D_{ij} + \beta \cdot log(C_j) + \gamma)}{\sum_r exp(\alpha \cdot D_{ir} + \beta \cdot log(C_r) + \gamma)} \quad (1)$$

Where $p_{(j|i)}$ is the conditional probability that the agent selects the j-th place when his current place is i. D_{ij} means the distance between the i-th and j-th places. C_j is the capacity of j-th place and (α, β, γ) are coefficients.

After the destination selected, the agent will further determine its travel route to that place. By default, the agent adopts his habitual path according to his experience or the shortest path preliminarily computed according to the network topology. But most probably, the agent will use the minimum
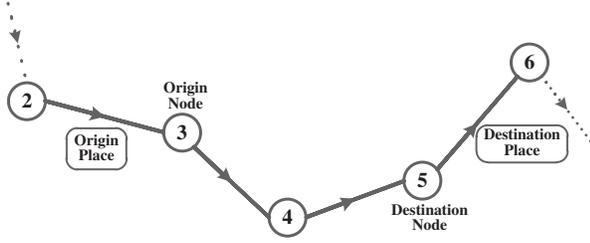
Fig. 1.   Relationship between Place and Node

cost path, which is updated based on real-time data and broadcasted at regular simulation time. Thus the minimum cost path is globally visible to every agent in ATS. The probability that the agent selects a particular path is determined by

$$p_l^e = \frac{exp(c_e/L_l + d_e \cdot F_{el})}{\sum_t exp(c_e/L_t + d_e \cdot F_{et})} \tag{2}$$

Where p_l^e stands for the probability that agent e selects the l-th path. L_l is the length of the path. F_{el}, a fuzzy variable ranging from 0 to 10, represents the agent's familiarity to the l-th path. c_e and d_e are coefficients.

B. Calibration of ABM

In the travel behavioral model, except the coefficients, other properties are deterministic given a particular road network and personal state. Therefore, the coefficients are to be calibrated. Let us conveniently denote $p_{(j|i)} = p_{(j|i)}(\alpha, \beta, \gamma)$ and $p_l^e = p_l^e(c_e, d_e)$. Consider an abstract road network $G(N, L)$ where N and L represent nodes and links, respectively. Since one's place selection cannot be directly measured, we need to consider the nearest downstream node of an origin place and the nearest upstream node of a destination place instead. For example, if an agent is now in the origin place shown in Figure 1 and he is going to the destination place, node 3 and 5 are used to represent the origin and destination, thus his route is $3 \rightarrow 4 \rightarrow 5$.

To limit the research scope, here we only consider the traveling by vehicle. In ATS, the total traffic flow of an OD pair is the sum of all the driving agents between the two nodes. In actual transportation systems, traffic data are usually collected during a period of time (called the detection cycle) and sent to the management center. Based on this mechanism, assume $t_0, t_1, \cdots, t_T$ are the time points that a detection cycle ends. Let $\Delta t_1 = t_1 - t_0, \cdots, \Delta t_T = t_T - t_{(T-1)}$ be the detection intervals. For the first interval, we have

$$\begin{bmatrix} p_{1|1}(\Delta t_1) & \cdots & p_{n|1}(\Delta t_1) \\ \vdots & \ddots & \vdots \\ p_{1|n}(\Delta t_1) & \cdots & p_{n|n}(\Delta t_1) \end{bmatrix} \cdot \begin{bmatrix} a_1(t_0) \\ \vdots \\ a_n(t_0) \end{bmatrix} = \begin{bmatrix} a_1(t_1) \\ \vdots \\ a_n(t_1) \end{bmatrix}$$

Where $a_i(t_0)$ and $a_i(t_1)$ are the agent numbers of the i-th node in t_0 and t_1, and $p_{(j|i)}(\Delta t_1)$ is the transferring probability in Δt_1 which has been defined before. $n = \|N\|$ is the number of nodes. Denoting the formula in a matrix form, there are

$$P(\Delta t_k) \cdot a(t_{(k-1)}) = a(t_k), 1 \le k \le T$$

Furthermore, we have the transferring agent number from the deduction of two formulas with adjacent time points:

$$\begin{aligned} a(\Delta t_k) &= a(t_k) - a(t_{(k-1)}) \\ &= [P(\Delta t_k) - I] \cdot a(t_{(k-1)}), 1 \le k \le T \end{aligned}$$

By introducing a coefficient matrix that maps the agent numbers to the traffic counts, there is

$$\begin{aligned} v(\Delta t_k) &= \Lambda \cdot a(\Delta t_k) \\ &= \Lambda \cdot [P(\Delta t_k) - I] \cdot a(t_{(k-1)}), 1 \le k \le T \end{aligned} \tag{3}$$

Where $v(\Delta t_k) = [v_1(\Delta t_k), \cdots, v_n(\Delta t_k)]^T$ are the net traffic counts flowed into each node and $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ are the preliminarily determined mapping constants from agent number to vehicular number. Note that entries of $a(\Delta t_k)$ and $v(\Delta t_k)$ can be negative which means agents in those places are reducing. According to the typical activity chain, agents start their activities at their homes at the beginning of each day. Thus the initial agent distribution can be set as the initial urban population distribution. This is achieve by population synthesis though various methods [15] [16] [26], which indicates $a(t_0)$ is known. $v(\Delta t_k)$ can be acquired from the actual transportation surveillance. Therefore, Eq. (3) is a recursive linear system and (α, β, γ) can be calibrated in each cycle.

After the destination calibrated, each OD demand of agent number is calculated through

$$a_{ij}(\Delta t_k) = p_{(j|i)}(\Delta t_k) \cdot a_i(t_{(k-1)}) \tag{4}$$

The traffic assignment proportions are computed by

$$p_l^{(i,j)}(\Delta t_k) = \frac{u_l^{(i,j)}(\Delta t_k)}{\sum_l u_l^{(i,j)}(\Delta t_k)} \tag{5}$$

Where $p_l^{(i,j)}(\Delta t_k)$ means the expectation of path l between OD pair (i, j). $u_l^{(i,j)}(\Delta t_k)$ is the AVI traffic flow of path l between OD pair (i, j). Assume that c_e and d_e obey Gaussian distributions $N(\mu_c, \sigma_c)$ and $N(\mu_d, \sigma_d)$ respectively, and $\sigma_c = 2\mu_c, \sigma_d = 2\mu_d$. There is

$$\frac{exp(\mu_c/L_l + \mu_d \cdot F_{el})}{\sum_t exp(\mu_c/L_t + \mu_d \cdot F_{et})} = p_l^{(i,j)}(\Delta t_k) \tag{6}$$

Since the denominator is a normalized parameter and the path selection probability is relative, we can arbitrarily set

$$exp(\mu_c/L_l + \mu_d \cdot F_{el}) = R \cdot p_l^{(i,j)}(\Delta t_k)$$

Where R is a normalization constant. Thus μ_c and μ_d are determined through Eq. (6). Using Monte Carlo sampling, path selection distributions of each agent from node i to node j can be calculated. The overall calibration process is shown in Figure 2.

IV.   DATA SOURCE AND NUMERICAL EXPERIMENTS

In order to validate our estimation model outlined in the previous part, this section will carry out a series of numerical experiments. These experiments are based on the actual detected data from Chengdu, one of the largest cities in the west of China. Our studied area is the central district of the city. As shown in Figure 3, the abstracted road network covers about 19.2 km^2, containing 37 nodes and 112 links.
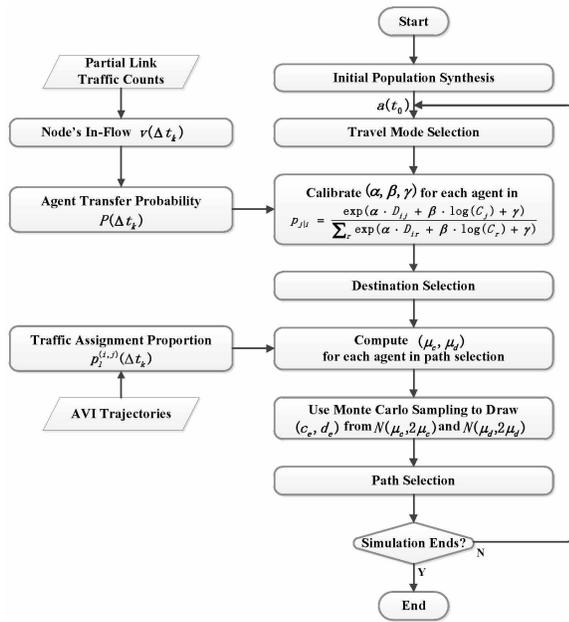
Fig. 2.   Overall Calibration Process
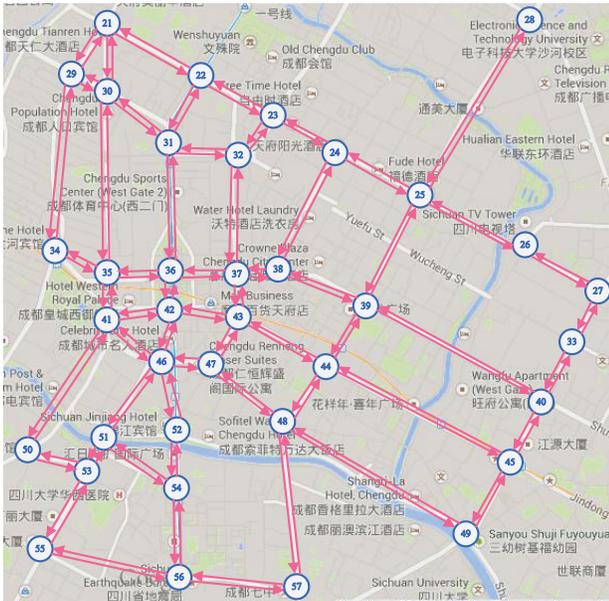
Fig. 3.   The abstract road network for experiments

The input data comes from two sources: partial link traffic counts and the active taxi information. These two types of data are collected by 225 loop detectors embedded in the road surface and the GPS location devices installed in 13,608 taxies. The loop detectors send their detected traffic counts to the transportation management center every minute. While GPS devices transmit the location coordinates to the management center at intervals of 10 seconds. In our experiments, the entire datasets collected in August 1st, 2014, are used, which contains 3,202,442 records of traffic counts and 77,645,666 records from taxis. The structures of the two datasets are illustrated in Table I and Table II. As can be seen, the loop detector records mainly provide volume, speed, occupancy and headway. The

Fig. 4.   Route with a loop

GPS devices give latitude, longitude, direction, and speed.

The original datasets cannot be directly used since they have much noise. For example, it is believed that the GPS device malfunctioned when the longitude and latitude are both 0 in a particular record. And when the loop detector provides a too large speed value (such as 200km/h), we can safely infer it was out of order (because according to local regulations, most of the urban area have a strict speed limit of 60km/h). Generally, three types of noise are tackled before the experiment execution:

1. Abnormal records. These records contain obvious outliers (like 200km/h speed) and are removed from the datasets.
2. Deficient records. In these records, some data items are missing and we retained the records as long as the items we concentrate on are complete.
3. Duplicate records. Duplicate results may arise due to some transmission problems. For this kind, only one record would be retained and other duplicates were removed.

After the data clean process, 3,200,580 traffic count records and 75,722,620 taxi records are filtered preliminarily. The proportions that the deleted records accounted for the whole datasets are 0.06% and 2.48%, which indicates the noise ratio stays at an extremely low level. For taxi data, a further step must be taken to convert the longitude-latitude into the actual roads. This operation is called map match. We adopt the Geocoding service provided by Baidu Corp. to complete this task [35]. After that, the trajectories of each taxi are calculated and the right side of Eq. (5) can be computed based on these results. For each trajectory, the origin and destination nodes are determined according to the longest path without loops.

TABLE I. TRAFFIC COUNT RECORDS

Link ID	Detector ID	Time	Lane Number	Volume	Speed	Occupancy	Headway
103111700364001	3014-200326	01-08-14 12.36.00pm	1	60	34	1	200
103111700364001	3014-200326	01-08-14 12.36.00pm	2	300	27	7	154
103142104041001	3014-290240	01-08-14 12.23.00pm	1	660	37	10	48
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE II. TAXI LOCATION RECORDS

Terminal ID	Vehicle Number	Longitude	Latitude	Direction	Speed	Time
100200102181	CA-TC408	104.048715	30.611158	0.000000	9.5	2014/8/1 13:21:30
100101180078	CA-TZ602	104.076188	30.635200	356.000000	8.3	2014/8/1 21:07:08
100101360016	CA-TC321	104.100890	30.625727	100.000000	17.7	2014/8/1 15:37:51
⋮	⋮	⋮	⋮	⋮	⋮	⋮

When a particular route contains one or more loops, they will be removed and calculated separately. For example, if a route of a taxi is like Figure 4, which manifests the vehicle went through the nodes $26 \to 25 \to 24 \to 38 \to 39 \to 25 \to 28$ in turn. Then OD pairs should be $(26, 28)$, $(25, 39)$ and $(24, 25)$. The latter two are from the longest paths in the loop: $25 \to 24 \to 38 \to 39$ and $24 \to 38 \to 39 \to 25$. The net traffic counts flowed into each node can be estimated via L1-minimization [18] [36]. Thus the left side of Eq. (3) is determined.

To test our proposed calibration method, half of the traffic count data and the whole AVI data are adopted as the input, whereas the remaining traffic counts are set to be the test criterion. The whole day is split into 5 time intervals: [0:00, 7:00], [7:00, 10:00], [10:00, 17:00], [17:00, 20:00], [20:00, 24:00] and the ABM is calibrated for each interval. The calibrated ABM is used to "reproduce" the traffic flow. Absolute Percent Deviation (APD) is calculated to investigate the errors:

$$APD_i = \frac{|f_{i,gen} - f_{i,act}|}{f_{i,act}} \times 100\%$$

Where f means the link flow and the subscripts i, gen, act represent the i-th observed link, the generated flow and the actual flow, respectively.

TABLE III. THREE INDICATORS OF EXPERIMENTS

Indicator	MAPD	RMSE	RMSPE
00:00-07:00	0.2117	835.9	1.079
07:00-10:00	0.2389	1369.1	1.218
10:00-17:00	0.7810	6290.5	3.982
17:00-20:00	0.1754	1128.8	0.895
20:00-24:00	0.2516	1317.2	1.283

Figure 5 shows the APDs of the experiment results in five intervals. Clearly, all of the deviations are below 100% and can be classified into two groups. One group is scattered below 20% which consists of 82 links among 130. The other group is mostly above 70% which consists of 46 links. Apart from the [10:00, 17:00] period, the results are good overall. This indicates that our proposed method gives better performance towards shorter time intervals. Further, we find that the number of links in the first group is relevant to the rank of assignment matrix. In other words, if observed set includes more linearly independent links, the calibration will become more accurate. Table III gives another three performance indicators: Mean Absolute Percent Deviation (MAPD), Root Mean Square Error (RMSE) and Root Mean Square Percent Error (RMSPE). The

calculation methods are:

$$MAPD = \frac{1}{m} \sum_i \frac{|f_{i,gen} - f_{i,act}|}{f_{i,act}}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_i (f_{i,gen} - f_{i,act})^2}$$

$$RMSPE = \sqrt{\frac{1}{m} \sum_i \left(\frac{f_{i,gen} - f_{i,act}}{f_{i,act}} \right)^2}$$

It can be seen that the results have large variances. This phenomenon can be also reflected in Figure 5. The errors of most link flows stay at extremely low levels, while the others are almost beyond 80%. The dispersal of errors undoubtedly brings large variances. Since the L1-minimization reconstructs the primary entries as accurate as possible from actual traffic counts, it can be expected that if the number of independent observed links increased, the calibration would be better.
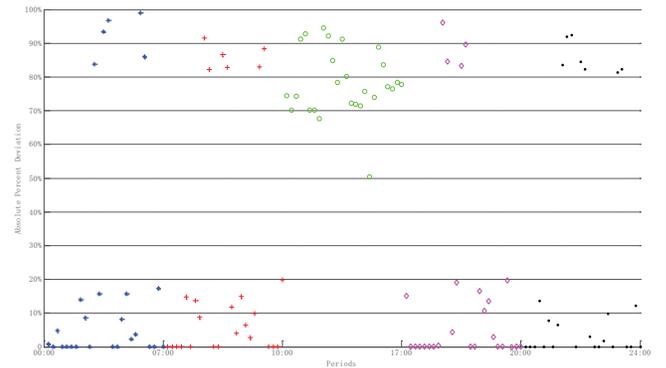

Fig. 5. Absolute Percent Deviations of the observed link flows

V. CONCLUSIONS AND DISCUSSIONS

This paper proposes a calibration method for agent-based travel model in ATS by correlating actual traffic counts and AVI data with micro behavioral model parameters. The new method is tested using the actual transportation data from the city of Chengdu, China. The results indicate that the calibrated ABM can "reproduce" traffic flow very well, and the proposed method can be used in general cases.

Although the current calibration is conducted in a real transportation network, the scale is relatively small. In future,

it requires to be tested in larger abstract road networks. Meanwhile, a part from traffic counts and AVI data, other types of transportation data also need to be used in the calibration. Thus the accuracy of calibrated ABM is expected to increase. However, it will be much more computational intensive as well. Such problems might be solved via distributed computing and more efficient algorithms.

REFERENCES

[1] F.-Y. Wang. *Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications.* IEEE Transactions on Intelligent Transportation Systems, 2010, 11(3): 630–638.

[2] F.-Y. Wang. *Toward A Revolution in Transportation Operations: AI for Complex Systems.* IEEE Intelligent Systems, 2008, 23(6): 8–13.

[3] F. Zhu, S. Chen, Z.-H. Mao and Q. Miao. *Parallel Public Transportation System and Its Application in Evaluating Evacuation Plans for Large-Scale Activities.* IEEE Transactions on Intelligent Transportation Systems, 2014, 15(4): 1728–1733.

[4] F. Zhu, Z. Li, S. Chen and G. Xiong. *Parallel Transportation Management and Control System and Its Applications in Building Smart Cities.* IEEE Transactions on Intelligent Transportation Systems, 2016, 17(6): 1576–1585.

[5] Y. Lv, X. Zhang, W. Kang and Y. Duan. *Managing Emergency Traffic Evacuation with a Partially Random Destination Allocation Strategy: A Computational-Experiment-Based Optimization Approach.* IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4): 2182–2191.

[6] G. Xiong, D. Shen, X. Dong, B. Hu, D. Fan, H. Dong and F. Zhu. *Parallel Transportation Management and Control System for Subways.* IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1974–1979.

[7] Y. Lv, Y. Duan, W. Kang, Z. Li and F.-Y. Wang. *Traffic Flow Prediction With Big Data: A Deep Learning Approach.* IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 865–873.

[8] F.-Y. Wang, X. Wang, L. Li and L. Li. *Steps toward Parallel Intelligence.* IEEE/CAA Journal of Automatica Sinica, 2016, 3(4): 345–348.

[9] L. Li, Y.-L. Lin, D.-P. Cao N.-N. Zheng and F.-Y. Wang. *Parallel Learning—A New Framework for Machine Learning.* Acta Automatica Sinica, 2017, 43(1): 1–8.

[10] X. Liu, X. Wang, W. Zhang, J. Wang and F.-Y. Wang. *Parallel Data: From Big Data to Data Intelligence.* Pattern Recognition and Artificial Intelligence, 2017, 30(8): 673-682.

[11] W. Zhu and F.-Y. Wang. *Reduction and Axiomization of Covering Generalized Rough Sets.* Information Sciences, 2003, 152: 217–230.

[12] Y. Lv, Y. Chen, X. Zhang, Y. Duan and N. Li. *Social Media Based Transportation Research: The State of The Work And The Networking.* IEEE/CAA Journal of Automatica Sinica, 2017, 4(1): 19–26.

[13] X. Wang, X. Zhang, K. Zeng and F.-Y. Wang. *Analysis of Cyber Interactive Behaviors Using Artificial Communities and Computational Experiments.* IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(6): 995–1006.

[14] X. Zheng, W. Chen, P. Wang, D. Shen, S. Chen, X. Wang, Q. Zhang and L. Yang. *Big Data for Social Transportation.* IEEE Transactions on Intelligent Transportation Systems, 2016, 17(3): 620–630.

[15] H. Zhao, S. Tang and Y. Lv. *Generating Artificial Population for Traffic Micro-simulation.* IEEE Intelligent Transportation Systems Magazine, 2009, 1(3): 22–28.

[16] P. Ye, X. Wang, C. Chen, Y. Lin and F.-Y. Wang. *Hybrid Agent Modeling in Population Simulation: Current Approaches and Future Directions.* Journal of Artificial Societies and Social Simulation, 2016, 19(1): 12.

[17] P. Ye, X. Hu, Y. Yuan and F.-Y. Wang. *Population Synthesis Based on Joint Distribution Inference without Disaggregate Samples.* Journal of Artificial Societies and Social Simulation, 2017. (Accepted)

[18] P. Ye and D. Wen. *A Study of Destination Selection Model Based on Link Flows.* IEEE Transactions on Intelligent Transportation Systems, 2013, 14(1): 428–437.

[19] R. M. Pendyala, R. Kitamura, A. Kikuchi, T. Yamamoto, and S. Fujii. *Florida Activity Mobility Simulator: Overview and Preliminary Validation Results.* Transportation Research Record, 2005, 1921: 123–130.

[20] C. R. Bhat, J. Y. Guo, S. Srinivasan and A. Sivakumar. *Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns.* Transportation Research Record, 2004, 1984: 57–66.

[21] T. A. Arentze and H. J. P. Timmermans. *A Learning-Based Transportation Oriented Simulation System.* Transportation Research Part B: Methodological, 2004, 38(7):613–633.

[22] P. Salvini and E. J. Miller. *ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems.* Networks and Spatial Economics, 2005, 5(2): 217–234.

[23] D. R. Pritchard. *Synthesizing Agents and Relationships for Land Use/Transportation Modeling.* Master's Thesis, Department of Civil Engineering, University of Toronto, 2008.

[24] M. J. Roorda, E. J. Miller and K. M.N. Habib. *Validation of TASHA: A 24-h Activity Scheduling Microsimulation Model.* Transportation Research Part A: Policy and Practice, 2008, 42(2): 360–375.

[25] L. Smith, R. Beckman and K. Baggerly. *TRANSIMS: TRansportation ANalysis and SIMulation System.* Los Alamos National Laboratory Report, LA-UR-95-1641, 1995.

[26] R. J. Beckman, K. A. Baggerly and M. D. McKay. *Creating Synthetic Baseline Populations.* Transportation Research Part A: Policy and Practice, 1996, 30(6): 415–429.

[27] M. Tang and H. Jia. *An Approach for Calibration and Validation of the Social Force Pedestrian Model.* In Proceedings of IEEE International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), Changchun, China, December 16-18, 2011: 2026–2031.

[28] Y. Zhao and A. W. Sadek. *Large-scale Agent-based Traffic Microsimulation: Experiences with Model Refinement, Calibration, Validation and Application.* Procedia Computer Science, 2012, 10: 815–820.

[29] D. Voloshin, D. Rybokonenko and V. Karbovskii. *Optimization-based Calibration for Micro-level Agent-based Simulation of Pedestrian Behavior in Public Spaces.* Procedia Computer Science, 2015, 66: 372–381.

[30] M. M. Oliveros and K. Nagel. *Automatic calibration of agent-based public transit assignment path choice to count data.* Transportation Research Part C, 2016, 64: 58–71.

[31] W. Zeng, P. Chen, G. Yu and Y. Wang. *Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach.* Transportation Research Part C, 2017, 80: 37–70.

[32] F. Zhu, D. Wen and S. Chen. *Computational Traffic Experiments Based on Artificial Transportation Systems: An Application of ACP Approach.* IEEE Transactions on Intelligent Transportation Systems, 2013, 14(1): 189–198.

[33] S. Chen and F. Zhu. *Artificial Transportation System construction based on ontology-based knowledge base systems.* In Proceedings of IEEE International Conference on Service Operations and Logistics, and Informatics, Suzhou, China, 2012: 207–212.

[34] S. Chen, F. Zhu and J. Cao. *Growing Spatially Embedded Social Networks for Activity-Travel Analysis Based on Artificial Transportation Systems.* IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2111–2120.

[35] Web site: http://developer.baidu.com/map/index.php?title=webapi/guide/webservice-geocoding.

[36] P. Ye and D. Wen. *Optimal Traffic Sensor Location for Origin-Destination Estimation using Compressed Sensing Framework.* IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1857-1866.