Fig.1 (a) A shanghai metro map. There are 16 metro lines in the map. The size of each station represents inbound and outbound traffic for the day. (b) The trip frequency distribution of all users. Users in the data set made between 100 and 300 trips in four months. (c) The daily trips of all users. The dotted lines represent the average number of trips per day, and they're all integers. The solid lines represent the average number of trips per user per day, and they're all decimals.

Fig.2 (a) Various label trip distribution diagram of commuter. The solid blue line represents that home-based work trips of commuter. The dotted green line represents home-based other trips of commuter. The orange dotted line represents non-home based trips of commuter. The dotted pink line represents all trips of commuter. (b) Various label trip distribution diagram of non-commuter. The solid blue line represents that home-based work trips of non-commuter. But since non-commuter doesn't have work station, they don't have this kind of travel. The dotted green line represents home-based other trips of non-commuter. The orange dotted line represents non-home based trips of non-commuter. The dotted pink line represents all trips of non-commuter. (c) Trip displacement distribution diagram of three types users. The solid blue line represents commuters. The solid red line represents non-commuters. The green solid line represents non-home. Travel of about ten kilometers accounts for the majority.

Fig.3 (a) Modeling flow chart of state at the last moment and state at the next moment. (b) (c) The state transition matrix of a commuter in the morning and evening. The horizontal and vertical axes are the ten transition states. The vertical axis represents the state at the last time slot. The horizontal axis represents the state at the next time slot. The values and color gradients in the squares represent the transition probability. (d) (e) The state transition matrix of a non-commuter in the morning and evening. (f) The choice probability matrix for other station. Take commuter’s evening as an example. The vertical axis is the state at the next time slot. The horizontal axis is the other station selected at the next timeslot. (e) The choice probability matrix the state of other to other. The vertical axis is the state of other to other. The horizontal axis is the combination of other station selected at the next timeslot.

Fig.4 (a) The accuracy rate of next state prediction and daily trips prediction for three type users. The dotted red line represents next state prediction and dotted blue line represents daily trips prediction. (b)-(e) The distribution map of all users travel frequency throughout the day. (b) (c) The next state prediction result and the real data are compared to the travel volume of the two types of users. The blue bars represent predicted values and the red bars represent actual values. (d) (e) The daily trips prediction result and the real data are compared to the travel volume of the two types of users.

Fig.5 (a)-(e) The distribution map of the predicted and actual values of all day flow at six major station. They respectively represent People’s square station, Lujiazui station, Xinzhuang station, Sijing station, Caohejing Development station and Nanjing west road station. The solid blue line represents the predicted inbound flow. The solid orange line represents the predicted outbound flow. The solid green line represents actual inbound flow. The solid red line represents the actual outbound flow.