**User Guide**

This is a simple user guide for the GMNS data format version of Big data-driven transportation computational graph framework (BTCG).

1. **Introduction**

This code is built on the GMNS data format version of BTCG (Wu et al., 2018). BTCG is a forward and backward propagation algorithmic framework on a layered computational graph which can achieve hierarchical travel demand estimation using multiple data source.

The specific relationship between the layers of HFN is shown in Fig1. HFN of Traffic Demand Flow Estimation (TDFE) model. The figure shows the input/output variables of each layer and the correlation calculation between layers.

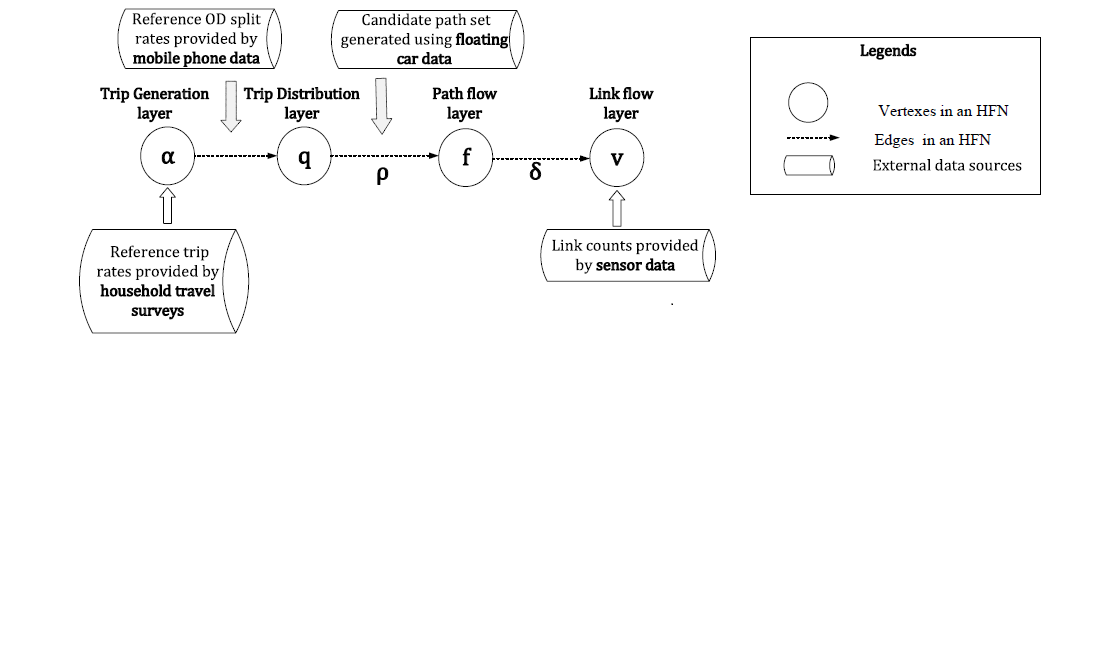


Fig. 1. HFN of TDFE model

BTCG is the implementation of Hierarchical Flow Network (HFN), by applying the back propagation (BP) algorithm, and by viewing each variable as a vertex and the edge between vertexes as calculation process between variables. To minimize the loss function, parameters like α, π and θ are updated during the training process. Finally, the traffic flow of each layer (e.g. α of ozone layer, γ of OD layer, v of link layer) are estimated jointly.

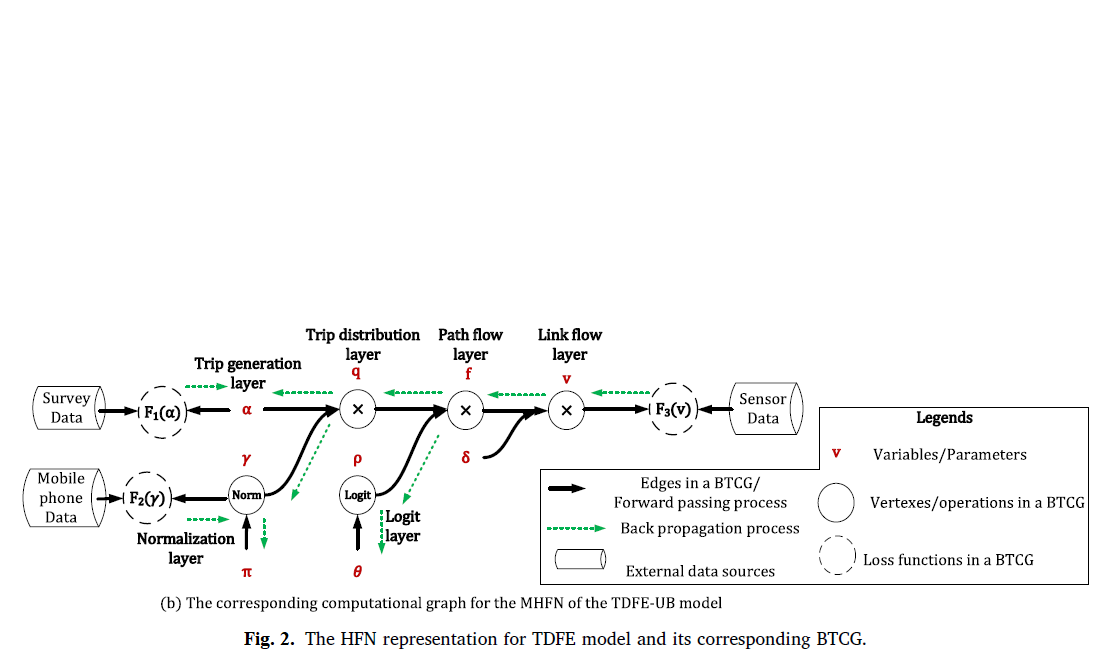


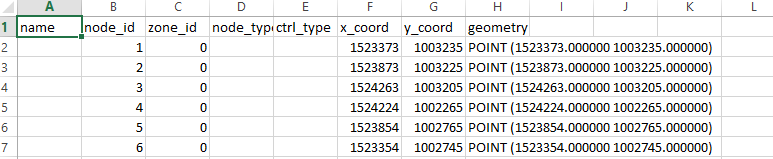
Fig. 2. Corresponding BTCG of HFN

1. **Data flow**

|  |  |
| --- | --- |
| **Input files** | **Output files** |
| node.csv | output\_ozone.csv  output\_od.csv  output\_path.csv  output\_link.csv |
| input\_agent.csv  agent\_type.csv |
| road\_link.csv |

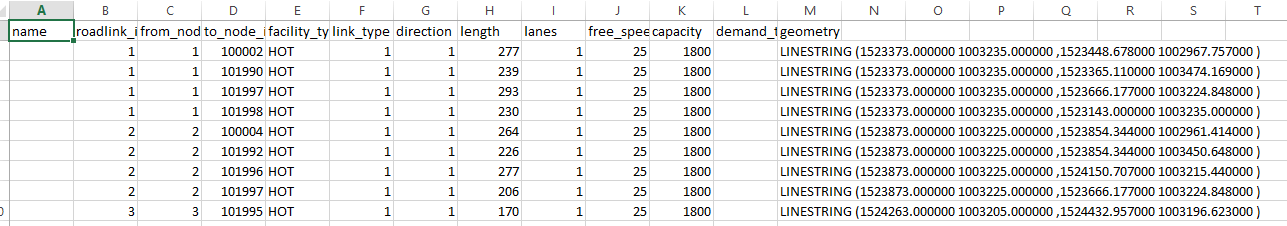
node.csv

This file node.csv may include the basic node information about the test network, such as name, node\_id, zone\_id, node\_type, ctrl\_type, x\_coord, y\_coord, and geometry.



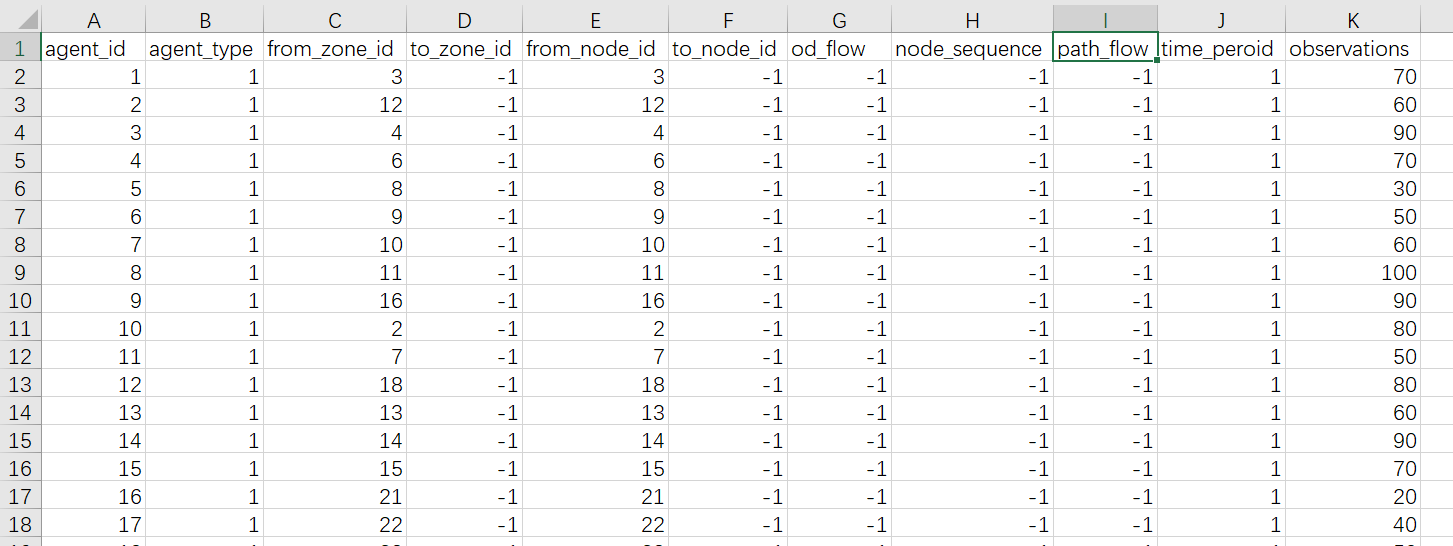
road\_link.csv

Road\_link.csv includes basic link-level information.



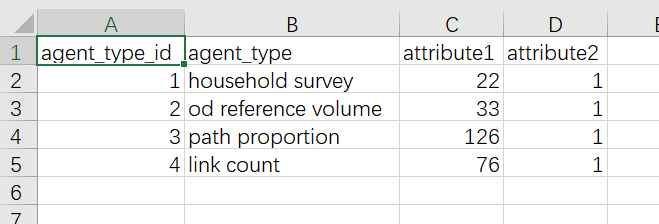
input\_agent.csv

All the measurements are currently stored in input\_agent.csv. Our code requires the household survey data (ozone data), the OD reference volume or the OD split rate (mobile phone data), the link count (sensor data), and the path information such as its node sequence.



agent\_type.csv

agent\_type.csv is used to explain the agent type in input\_agent.csv. Currently, it includes agent types :1. household survey; 2. OD reference volume; 3. path proportion; 4. link count, the attribute 1 means the number of observations, and attribute 2 means the number of time period. In our code, we consider one day of traffic volume, so the number of time periods is 1.



1. **Case study**

The Sioux Falls network is used as the test case, it has 22 zones, 24 nodes, 33 OD pairs, 76 links and 132 paths, the network structure is shown in Fig. 3. The GMNS format data is provided in the sub folder of /SiouxFalls network.

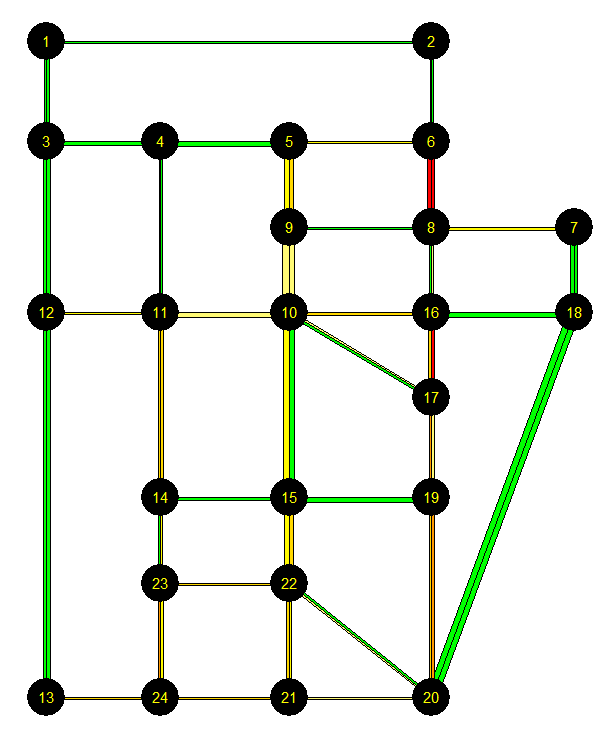
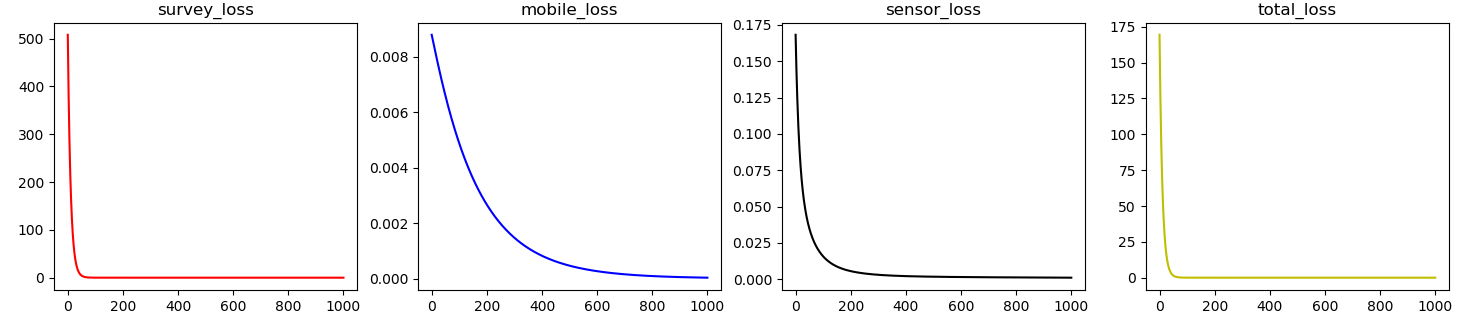


Fig. 3. Sioux Falls Network

Training results (1000 epochs)：

BTCGLite\_GMNS.py (this code don’t use logit model and class object, so it runs faster ):



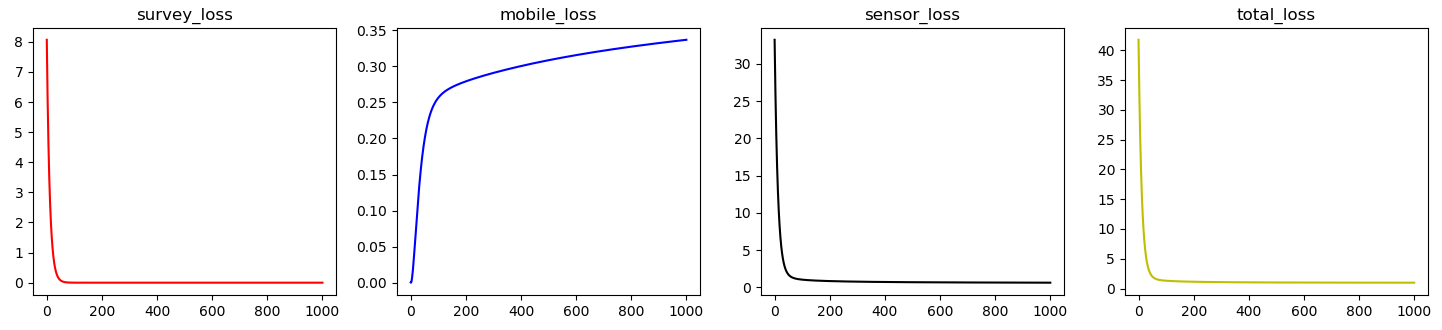
step 1000 :survey error= 6.983434152935154e-08

step 1000 :mobile error= 3.8845606447143905e-05

step 1000 :sensor error= 0.0010404423796513053

step 1000 :total error= 0.0010792072388590332

BTCG\_GMNS.py Although the mobile loss increases to a certain degree,the total loss has been reduced, to demonstrate how the BTCG works.



step 1000 :survey error= 3.1824046e-09

step 1000 :mobile error= 0.3365797

step 1000 :sensor error= 0.6412415

step 1000 :total error= 0.9778478

**References.**

Wu, X., Guo, J., Xian, K., Zhou, X., 2018. Hierarchical travel demand estimation using multiple data sources: A forward and backward propagation algorithmic framework on a layered computational graph. Transportation Research Part C: Emerging Technologies 96, 321-346.