**A Framework for Systematic Traffic Count Balance in Regional Networks**

Hongbo Chi1, Ph.D., Chunyu Lu2, P.E., and Jason Learned3

1AECOM Technology, 7650 Corporate Center Dr. Suite 400, Miami, FL 33126; e-mail: [chihongbo@gmail.com](file:///C:\Users\chih\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\H00P7PNK\chihongbo@gmail.com)

2AECOM Technology, 7650 West Courtney Campbell Causeway Suite 700 Tampa, FL 33607; e-mail: [chunyu.lu@aecom.com](file:///C:\Users\chih\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\H00P7PNK\chunyu.lu@aecom.com)

3Planning & Environmental Management Office, FDOT-D5, 719 S Woodland Blvd, DeLand, FL 32720; e-mail: [jason.learned@dot.state.fl.us](file:///C:\Users\chih\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\H00P7PNK\jason.learned@dot.state.fl.us)

# ABSTRACT

In this paper, a methodology framework is developed on systematic traffic count balancing for regional road networks. The framework includes three major components: two intercept link extractors, and one traffic count balancing model. The intercept link extractors are developed based on breadth search algorithm (BFS) designed to traverse networks to extract intercept links. The traffic count balancing model is a constrained weighted least square optimization model with constraints of traffic conservation equations based on intercept links. The proposed framework has the following features: 1. It can automatically traverse networks to identify and extract the intercept link counts for traffic conservation ; 2. It can balance multiperiod traffic counts in regional level; 3. Uncertainty of traffic counts can be considered by weight factors; and 4. the traffic balancing model can be solved efficiently due to its convex structure.

# INTRODUCTION

Consistent and accurate traffic counts are critical for calibration and validation of travel demand model and meso/micro-scopic traffic simulation model. However, a significant amount of inconsistencies and noises in traffic counts may exist due to the following reasons:

1. Different data collection date;
2. Deduced traffic counts from historical trend;
3. Different local conversion factors applied to raw traffic counts, such as seasonal factors;
4. Malfunction of the traffic count collection devices;
5. Erroneous post-processing for traffic counts, such as misplacement of count locations.

In order to identify and minimize those inconsistencies, traffic data balancing/reconciliation efforts are often needed. In practice, those efforts are very time-consuming and labor-intensive when performed manually. There are several studies on the traffic balancing/estimation for intersections and access-controlled facilities. Hauer, *et al* (1981) proposed a bi-proportional model to estimate the turn volumes for an intersection based on the total entering and leaving traffic volumes and turn proportions derived from historic data set. Xin, *et al* (2006) adjusted the existing intersection turning volumes by proportionally distributing volume difference between the total inflow and total outflow among any movements that can be changed (pro-rata adjustment method). Lee, *et al* (2014) developed rule- and model- based approach for real-time estimation of lane-to-lane turning flows for isolated intersections. In the rule-based approach, the entrance lane used by the vehicle detected in the exit lane is identified according to a set of specified rules. The model-based approach is based on utility maximization and is used to identify the most probable turns from a set of potential upstream entrance lanes.

As for traffic balancing for access-controlled facilities, FHWA (2001) recommended using a pro-rata adjustment procedure for freeway traffic count balancing. This method proportionally allocates the total traffic count discrepancy for a freeway segment to each section between the two anchor points in the mainline. Xin, *et al* (2006) proposed a constrained least square model for the traffic count balancing for freeways with minimizing the sum of squares of the traffic count adjustment under traffic conservation constraints. Zhao, *et al* (1997) developed a traffic data balancing/reconciliation system TCAS for freeways. A constrained generalized least square model (GLS) was developed to resolve the discrepancies. The objective function is a squared Mahalanobis length of count adjustment vector, and the constraints represent the traffic conservation condition along the corridor. The model uses the inverse of covariance matrix of traffic counts as the weighting factors to assess the noise level of traffic counts. An important assumption is that the error in traffic counts follows independent normal distribution. Kwon, *et al* (2007) proposed a similar constrained weighted least square regression (WLSR) traffic count balancing model for an access‐controlled linear corridor. The model uses predefined confidence levels (excellent/good/fair/poor) for each count detector as the weight matrix. In order to obtain the close form solution, the model also assumes that traffic count error follows independent normal distribution. Shaw, *et al* (2014) developed a constrained least GEH traffic balance model. the GEH measurement was claimed to be a more robust measure for traffic adjustment. The model was employed to balance the three-leg freeway-to-freeway interchange near Bellevue, Wisconsin. It should be noted that not like the GLS or WLSR model, the GEH model cannot guarantee a global solution since its non-convexity property.

According to the literature review, we have the following summaries on the current studies on traffic count balancing:

1. There are various models on traffic balancing for isolated intersections or access-controlled facilities.
2. Uncertainty in the traffic counts has been considered as the standard deviation or predefined weight factors as in the studies by Zhao, *et al* (1997) and by Kwon, *et al* (2007).
3. The Robustness of the measurement of the count adjustment is emphasized in the study by Shaw, *et al* (2014).
4. Relatively few studies exist for extracting and organizing the traffic counts from networks (intercept link counts) to inform the traffic balancing model, which could be very time consuming for a long corridor or a large network if performed manually.
5. Studies on the regional level traffic balancing are rare.

This study proposes a methodology framework for balancing traffic counts in network level with the following characteristics:

1. Efficient method to extract the intercept links from network automatically;
2. Simple and expandable structure for integrating multiperiod traffic counts from intercept links to perform network level traffic balancing;
3. Flexibility to assign the weight factors for traffic counts to account for uncertainty level of different measurements.
4. Global solution guaranteed due to its convex structure;

The paper is organized as follows. The proposed method framework is described in Section 2. A case study is presented in Section 3. Section 4 concludes the paper.

# METHODOLOGY

The proposed methodology framework can be illustrated in Figure 1. The framework is comprised of three components, namely subarea intercept link extractor, corridor intercept link extractor and traffic count balancing model.

**Figure 1. Proposed methodology framework for traffic balancing**

# The two intercept link extractors are designed to traverse the network and identify the intercept links with traffic counts for any subarea of the network and for any subsegment of a corridor. The intercept links are defined as the links with traffic counts crossing the study area boundary, so that all traffic entering or exiting the subarea or subsegment must go through those links. In Figure 2, the green and red colored links stands for the inbound and outbound intercept links for a subarea and a subsegment. With those identified intercept links, the traffic conservation equation can be constructed, which is the sum of inbound traffic counts should be equal to the sum of outbound traffic counts.

|  |  |
| --- | --- |
| Intercept links for a subarea | Intercept links for a subsegment |
|  |  |

**Figure 2. Intercept links for a subarea and a subsegment of a freeway**

# The subarea intercept link extractor is based on the breadth first search algorithm (BFS). BFS is an algorithm for traversing graph data structures. It starts at some arbitrary node and explores all the neighboring nodes at the present depth prior to moving on to the nodes at the next depth level. BFS uses a queue data structure with First In First Out (FIFO) rule. The queue contains the frontier along which the algorithm is currently search. BFS algorithm is an efficient search algorithm with time complexity of O(V+E), since every vertex and every edge will be explored only once in the worst scenario. The classic BFS works with the following steps:

1. Putting any one of the graph’s vertices at the end of a queue.
2. Taking the front item of the queue and adding it to the list of the visited.
3. Creating a list of that vertex’s adjacent nodes. Adding the ones which aren’t in the list of the visited to the end of the queue.
4. Repeating steps 2 and 3 until the queue is empty.

# For the subarea intercept link extraction, a modified breadth first search algorithm (MBFS1) is developed with some additional termination conditions. It should be noted that the original input graph needs to be augmented to guarantee its strong connectivity property to facilitate MBFS1. The strong connectivity means that the any node can reach all its neighbor nodes. Additionally, the search needs to report warning message when reaching centroid or internal intercept links (the two nodes of a link with count can be reached by search). The followings are the procedure for the MBFS1:

1. Augmenting the input directed graph G by adding the opposite links for these one-way links.
2. Adding the starting node into a queue Q as the search frontier.
3. Taking the front item of the queue and adding it the list of the visited
4. Creating a list of that vertex’s adjacent nodes, adding the ones which aren’t in the list of the visited to the end of the queue.
5. Terminating the search at a new node if the new links have counts in their both directions and reporting this link as the intercept link with correct direction.
6. Reporting a warning message if the new node is a centroid, since the centroids make the traffic flow non-conservative.
7. Reporting a warning message if a search link with counts for both directions can be reached by the algorithm from its both end nodes since the link is an internal link instead of an intercept link.
8. Adding the new node into the end of the queue if the new link does not contain traffic counts.
9. Repeating steps 3 to 8 until Q is empty.

Some simple postprocessing procedures are needed for the intercept links with warning message. If the warning message is zone centroid related, then the whole set of the intercept links for the subarea should be canceled out. If the wanting message is for the internal links, then only those internal intercept links need to be removed. It should be mentioned the above algorithm is a prototype to extract the intercept links for one subarea. The MBFS1 can be easily customized to perform intercept links search across the whole regional network.

# The corridor intercept link extractor is also based on the BFS. In this study, a modified breadth first search algorithm (MBFS2) is developed. In the same manner, the input direct graph needs to be augmented. It should be mentioned that in this application, ramps are still treated as the intercept links even there are no traffic counts on them. The search on ramps should be terminated if it reaches any non-ramp links or it has observed traffic count on itself. The following is the procedure for the MBFS2:

1. Augmenting the input directed graph G by adding the opposite links for these one-way links.
2. Obtaining a mainline node search list L by using a classic BFS only along the mainline of the corridor.
3. Popping the starting node of L and adding into a queue Q as the search frontier.
4. Taking the front item of the queue an adding it the visited list.
5. Creating a list of that vertex’s adjacent nodes, adding the ones which aren’t in the visited list to the back of the queue.
6. Terminating the search at a node if the node is the ending node of the corridor and reporting its parent link as intercept link with correct direction.
7. Terminating the search at a new node if the new link has counts and reporting the new link as the intercept links with correct direction.
8. Terminating the search at a node if its parent link is a ramp link and there are non-ramp adjacent links and reporting the parent link as the intercept links with correct direction.
9. Adding the new node at the end of the queue if the adjacent link does not contain traffic counts.
10. Repeating steps 4 to 9 until Q is empty
11. Repeating steps 3 to 9 until L are empty.

It should be mentioned that users should find the traffic counts from other data sources for the intercept links without traffic count identified by MBFS2. Otherwise, the model could generate very strange results without those traffic counts information.

For the traffic balancing model, traffic counts are assumed to reflect real traffic information; therefore, the sum of squares of traffic count adjustment percentage should be minimized. The weighted least square formulation of traffic count balancing model is shown in Eq.1:

(Eq. 1)

Subject to:

count spatial conservation constraints

count temporal conservation constraints

count non-negative constraints

In Eq. 1, CountBal stands for balanced traffic counts, which is the decision variable of the model. CountObsrepresents observed traffic counts from intercept links. *W* is the weighting factor standing for user’s confidence in traffic counts. For example, the traffic counts in mainline of a freeway corridor are usually more reliable with less oscillation than those collected from ramps. In that case, user can assign a larger weighting factor value to the mainline counts in the objective function in Eq. 1, so that they will be less likely to be adjusted or are adjusted to a less extent. As for the constraints, the first one addresses the fundamental traffic conservation condition for subareas or subsegments. The second one makes sure that the TOD traffic counts amount to the balanced daily traffic counts. The third constraint simply states the non-negative nature of traffic counts.

Eq.1 has a simple structure and can easily integrate the traffic data from intercept links identified by the two extractors. It should be noted that in order to perform the traffic balancing in regional level, the shared intercept links (for example, same ramps shared by two freeway systems) should be coded with same link number. Minimal efforts are required to assemble the data for balancing regional traffic balance.

Eq. 1 is a quadratic mathematic programming problem with linear constraints. The objective function has positive definite Hessians, so theoretically, the global optimum solution is guaranteed. In this study, the solver MINOS from GAMS software package was employed to solve the problem using gradient descent algorithm.

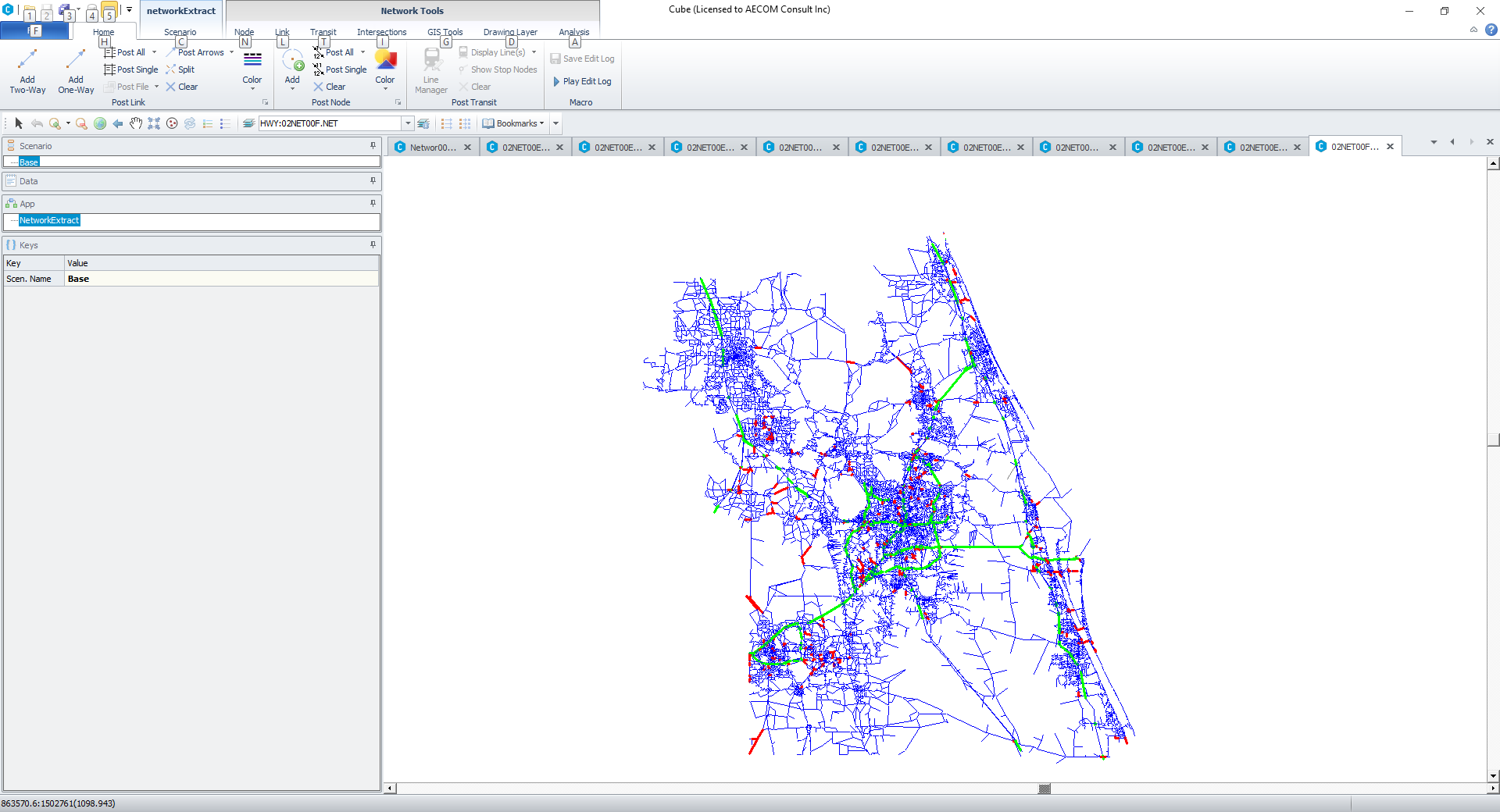
# CASE STUDY

In this case study, the 2015 highway network from Central Florida Regional Planning Model (CFRPM) is used. The network includes 5,406 zones and 72,750 links, of which 10,376 carry one daily and four TOD traffic counts. The proposed extractors can efficiently identify the intercept links. Figure 3 illustrates the extracted intercept links by MBFS1. The red colored links are the intercept links for three selected subareas in the CFRPM network. It could be very time consuming to manually identify those intercept links for a large network like CFRPM.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Figure 3. Intercept links extracted for three selected subareas**

The intercept link extraction for corridors is very similar to the extraction for subarea. The only difference is that the MBFS2 also reports the ramps without traffic counts as intercept links. Figure 4 shows the all the intercept links identified from the two extractors for CFRPM network. The red colored links are the intercept links identified by MBFS1 and the green colored links by MBFS2. It only takes a few seconds for the extractors to identify those intercept links.



**Figure 4. Intercept links extracted for CFRPM network**

Table 1 shows the intercept link summary for the two extractors. Totally there are 2,385 unique links identified as the intercept links for traffic balancing, comprising 884 basic road segments (subareas and subsegments). The spatial and temporal traffic conservation condition will be applied on each of those basic segments. It should be noted that there are some duplicate intercept links identified by the two extractors, driving the number of unique intercept links not equal to the total.

**Table 1. Intercept link summary**

|  |  |  |
| --- | --- | --- |
| Extractor | # of unique intercept links | # of basic segment |
| Subarea | 2,112 | 535 |
| Corridor | 1,182 | 349 |
| Total | 2,385 | 884 |

With intercept links extracted, the multi-period traffic counts data can be organized to feed into the traffic balancing model. Before performing the balancing work, the weighting factorst need to be defined. Without any priori information on the traffic counts, the weighting factors are assumed to an equation of the facilities type and time period as defined in Table 2. It is reasonable to assume that traffic counts from freeways especially from toll roads are more accurate than other facilities, and daily counts are more reliable than TOD counts. Users also have the flexibility to specify those weighting factors based on their educated judgement. For the link counts with large weighting factors, the balancing model will unlikely adjust them much.

**Table 2. Weighting factors by facility type and time period**

|  |  |  |
| --- | --- | --- |
| Facility | Daily traffic count | TOD traffic count |
| Freeway main line | 10 | 5 |
| Freeway ramp | 8 | 4 |
| Other facilities | 4 | 2 |

With well prepared traffic counts and weighting factors, the traffic balancing model is ready to run. According to Table 1, the model has 2,385\*5=11,925 decision variables, 884\*5=4,420 spatial constraints, 2,385\*1=2,385 temporal constraints and 11,925 non negative constraints. The model can be solved with MINOS within 10 minutes to reach its optimal solution. The objective value declines with number of iterations very smoothly as shown in Figure 5.

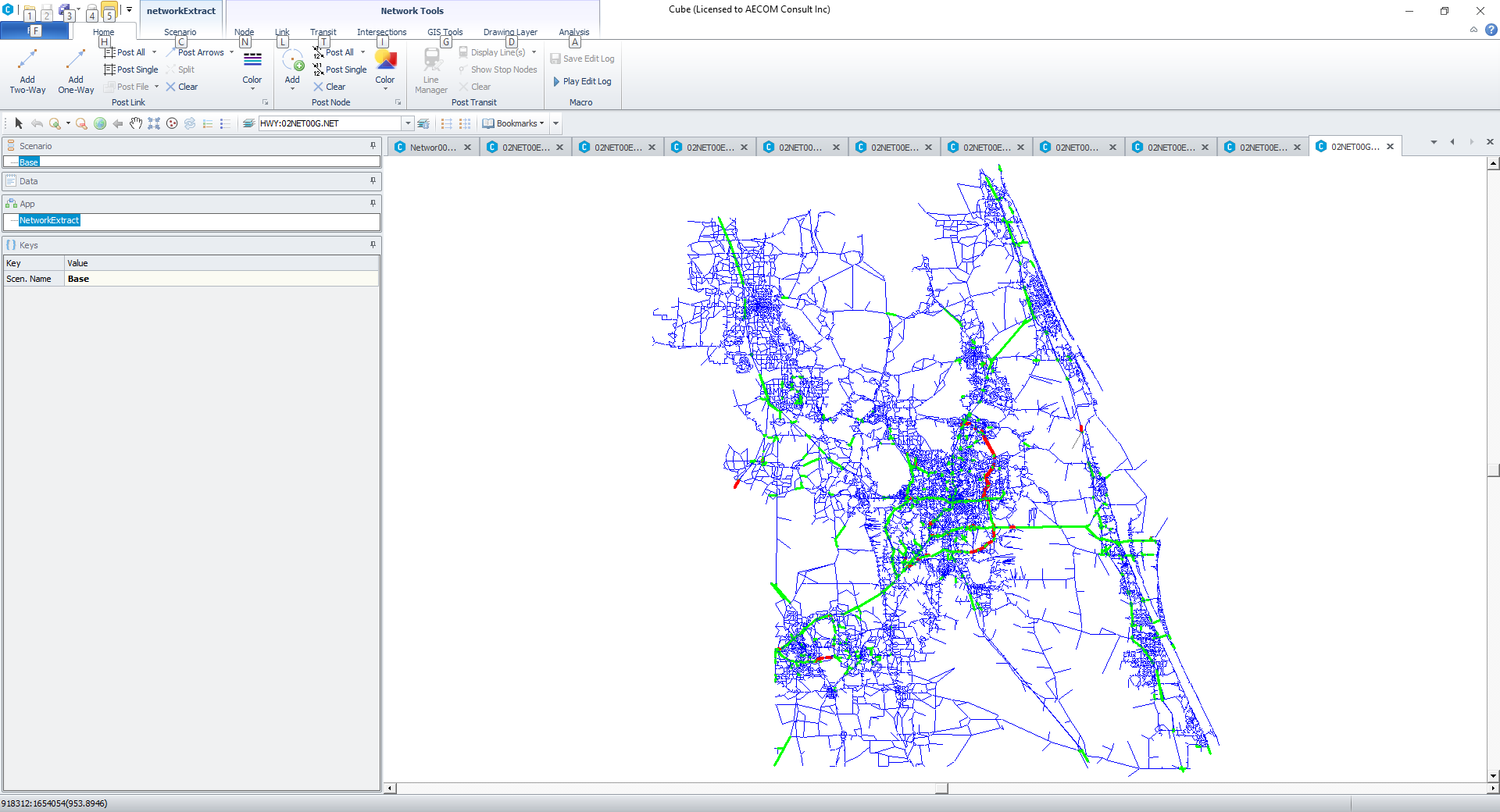
**Figure 5. Objective function values with iterations**

Table 3 shows the adjustment percentage summary information for daily traffic count after the balancing work. The average of adjustment percentage looks acceptable within 5% of range. However, some adjustments are greater than 25%, mostly in freeway mainlines and ramps.

**Table 3. Daily traffic count adjustment percentage summary**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Facility | # of Links | AVG | SD | Min | Max |
| Freeway main | 373 | -4% | 17% | -68% | 100% |
| Freeway ramp | 864 | -2% | 13% | -100% | 75% |
| Other | 1,148 | 0% | 4% | -32% | 18% |
| Total | 2,358 | -1% | 11% | -100% | 100% |

In Figure 6, the links with green color stand for the extracted intercept links, and the links with red color for the intercept links with daily traffic count adjustment percentage over +/-25%. The Figure indicates that some major adjustment happens in freeway corridor SR 417 in the center part of the network, indicating that large inconsistencies exist in the collected traffic counts along this corridor. Some sanity checks are required to confirm: 1. the counts are coded correctly in the network and 2. the counts are comparable to their historical records. In order to obtain a reasonable balancing result, users need to check and correct the traffic counts with large adjustment percentage first.



**Figure 6. Intercept Links with large daily traffic count adjustment percentage**

# CONCLUSION

In this study, a traffic count balancing methodology framework is developed to resolve traffic count discrepancies in regional level. The framework consists of three components: two intercept link extractors for subarea and corridor respectively and one constrained weighted least square based traffic balancing model. The first two components can significantly reduce manual work to construct spatial traffic conservation equations for traffic balancing model. The third component can efficiently balance multiperiod traffic counts in regional level. The model also has the functionality to integrate the confidence factors with traffic counts. The case study for balancing traffic counts on CFRPM network demonstrates the strength of the proposed methodology.

It should be noted that, networks from travel demand models are usually simplified with some local roads missing. Efforts might need to be taken to review the intercept links to make sure that there is no potential traffic leaking in the subarea level or subsegment level of network.

In the traffic balancing model, traffic count distribution is not considered, and the overall method solely reports a set of balanced counts dictated by the objective function and selected constraints. In the future, traffic count distribution may be explored in order to get some descriptive statistics on the balanced traffic counts, such as confidence interval. In addition, some extra studies are needed to design a more sophisticate method to determine the weighting factors.

# REFERENCES

Hauer, E., Pagitsas, E., and Shin, B. T. (1981). “Estimation of Turning Flows from Automatic Counts.” *Transportation Research Record*,795, pp.1-7*.*

Zhao, M., Garrick, N., Achenie, L., and Paladagu, S. (1997). *Data Reconciliation Based Traffic Count Analysis System*. University of Connecticut Transportation Institute, Storrs, CT.

Federal Highway Administration (FHWA). (2001), *Traffic Monitoring Guide*, Section 3, pp.3‐42 ­­- 3‐49. Washington, DC.

Xin, W., Hourdos, J., and Michalopoulos, P. (2006). *Streamlining the Traffic Modeling Process for Implementation in the Twin Cities Freeway Network – Phase II.* Minnesota Department of Transportation. St. Paul. MN.

Kwon, J., Petty, K., Shieh, E., Kopelias, P., and Papandreou, K. (2007). “An Automatic Method for Imputing and Balancing Link Traffic Counts.” *Transportation Research Board 87th Annual Meeting Compendium of Papers CD,* Transportation Research Board of the National Academies, Washington, DC.

Shaw, J.W., and Noyce, D.A. (2014). “Automated Optimal Balancing of Traffic Volume Data for Large Access-Controlled Highway Networks and Freeway-to-Freeway Interchanges.” *Transportation Research Board 93rd Annual Meeting Compendium of Papers CD,* Transportation Research Board of the National Academies, Washington, DC.

Lee, S., Wong, S., Pang, C., and Choi, K. (2014). “Real-Time Estimation of Lane-to-Lane Turning Flows at Isolated Signalized Junctions.” *IEEE Transactions on Intelligent Transportation Systems*,16(3), pp.1549-1558*.*