Improved solution quality through suboptimal subproblem solving in the FW method for traffic assignment

Johan Holmgren
School of Computing, Blekinge Institute of Technology, Karlshamn, Sweden

P O Lindberg

Static road network assignment (see, Patriksson, 1994) is a commonly occurring problem, e.g., in long-term traffic and transport planning. It appears, e.g., when assessing the impact of various types of transport policy and infrastructure measures. Given demands for traffic flow between pairs of nodes (so-called OD-pairs) in a network, and increasing travel time delay functions for each of the links, the objective is to assign traffic flows to the network in such a way that a traffic equilibrium is obtained. In the equilibrium state, which defines the optimal assignment, all routes that are actually in use for an OD-pair have the same travel time, not greater than unused routes.

The Frank-Wolfe (FW) method has for decades been the dominating method in practical use for static road network assignment. It operates by keeping track of the total flow assigned to each of the links in the network. During the solution process the flows on the links are iteratively updated by reassigning the traffic flow between the different routes connecting each of the OD pairs. The reassignment of flows are determined by solving a number of shortest path problems (one for each origin) that in each iteration appear as subproblems, and by generating a search direction, which is based on the link flows in the previous iteration and the solutions to the shortest path subproblems.

The FW method is known to show very fast initial convergence, but due to its slow asymptotic convergence it has been considered dead by methods oriented researchers. However, recent methodological improvements have significantly improved the time performance of the method. Mitradjieva & Lindberg (2013) improved the convergence of the method by applying conjugation of the search directions, thereby reducing the number of iterations to achieve required relative gaps. Holmgren & Lindberg (2012) improved the time per iteration by considering recursive sub-problem updating. Instead of solving each shortest path problem from scratch, we utilized the fact that two consecutive shortest path problems for the same origin in most cases are very similar (typically they differ only in a few links). For a number of standard test problems, we have achieved computation time improvements in the range of 93-99%, by applying conjugate search directions and recursive updating to the FW method.

In order to further improve the performance we have implemented a suboptimal approach to solving the shortest path subproblems, which can be used together with recursive updating and conjugate search directions. The idea is, in most FW iterations, to only consider those links that have been observed to appear often in the optimal solutions of the shortest path sub-problems, for a specific origin. Occasionally, all subproblems are solved to optimality, which is necessary in
order to be able to get lower bounds and terminate the method. By combining recursive subproblem updating with our suboptimal approach, we have for the test-problems referred to above, observed additional time reductions of approximately 30-45%, compared to recursive sub-problem updating.

In each iteration \( k \) of the FW method, a current feasible (pessimistic) solution is obtained, which gives an upper bound \( UBD_k \) to the optimal value. An infeasible (optimistic) solution is also obtained in each of the iterations, which gives a lower bound \( LBD_k \). The best so-far found lower bound, i.e., the maximum of all \( LBD_k \) up to and including iteration \( k \), is referred to as \( BLB_k \). The termination criterion typically used is the relative gap, which for iteration \( k \) is defined as \( \gamma_k = \frac{UBD_k - BLB_k}{BLB_k} \). The solution process proceeds until \( \gamma_k \) is smaller than some small value \( \epsilon \).

The obtained upper bounds will improve for each of the FW iterations (i.e., \( UBD_{k+1} < UBD_k \)), while \( BLB_k \) will improve stepwise (i.e., it is not necessarily the case that \( LBD_{k+1} > LBD_k \)). In particular, this behavior is interesting in our suboptimal approach to subproblem solving, because \( UBD_k \) will still improve in each iteration. A valid lower bound will only be achieved, and \( BLB_k \) can only be updated, in those iterations where all subproblems are solved to optimality. Still it cannot be guaranteed that the \( BLB_k \) is updated in those iterations where all subproblems are solved to optimality, due to the fact that \( LBDs \) are not always improving. Hence, the \( BLB_k \) will converge slower towards the optimal value when using suboptimal subproblem solving, than in approaches where the subproblems are always solved to optimality.

Due to the slower converge of \( BLB_k \), suboptimal approaches will typically require more iterations until termination, than is needed by approaches where subproblems are solved to optimality. Still they typically are able to significantly reduce the solution finding time, since each of the iterations in general is much faster. In addition, we have observed that the relative gap that is achieved at termination in general is significantly smaller than indicated by the termination criterion (\( \gamma_k < \epsilon \)), when using suboptimal subproblem solving. Moreover, for the same termination criterion, the best feasible solution at termination (i.e., UBD) is better when using suboptimal subproblem solving compared to approaches where the subproblems are always solved to optimality.

Our main conclusion is that suboptimal approaches to subproblem solving in the FW method do not only reduce the solution finding time, they also generate solutions that are significantly better in terms of relative gap and the best feasible solution.

References

