Accelerating Convergence of the Frank-Wolfe Algorithm for Solving the Traffic Assignment Problem

Saïda Arrache† **and Rachid Ouafi**††,

† Mathematics Department, University Saad Dahleb, Blida, Algeria
†† Operational Research Department, University of Sciences and Technologies Houari Boumedien, Algiers, Algeria

Summary
The Frank-Wolfe method (FW) is one of the most widely used algorithm for solving routing problems in the telecom and traffic areas, its popularity is attributed to its simplicity and modest memory requirements. However, FW converges very slowly, that makes it less favourable to use without modifications. To improve its performance, different modifications of the Frank-Wolfe method have been suggested. In This paper we describe a new modified version (FWFλ), the algorithm consists to combine Fukushima direction (FWF), with a widened line search technique (FWλ). We also present preliminary computational studies in a C++Builder5, in these we apply (FW), (FWλ), (FWF), and (FWFλ) methods to some Traffic assignment problems. The computational results indicate that the proposed algorithm yield satisfactory results within reasonable computational time comparing to the other methods

Key words:
Algorithms, Convergence, Frank-Wolfe method, Traffic assignment.

1. Introduction

The Frank-Wolfe method [4] was first introduced in quadratic programming at once it proved very effective for the resolution of large scale flood problems, the Frank-Wolfe method, is famous for its advantages: it is easy to implement and it performs well far from the optimal solution. However, it has a property that makes it less favourable to use without modifications, namely that it shows very slow asymptotic convergence due to that the feasible solutions tend to zig-zag towards the optimal solution. To improve its performance, different modifications of the Frank-Wolfe method have been suggested, starting with the first "L. J. Leblanc" works [8] until the recent works of "Ziyou Gao & Al" [18]. One approach is to find better search directions e.g. combine the search direction with previous ones [6] [10], or make conjugate consecutive search directions to each other [1]. A second approach is to modify the line search e.g. take longer steps [14] or using predetermined step lengths [15]. A third type of modification is to modify the subproblems and thus avoid generating extreme point solutions [9].

On our behalf, in this paper we propose a new improvement of the FW method (FWFλ) for solving the traffic assignment problem, this modification consist to combine Fukushima direction [6] (FWF), with a widened line search technique [11] (FWλ).
The main proponent of this attempt is to avoid the zig-zagging in the path described by the solution points of the pure FW method.
For the remainder, Section 2 reviews the FW method and the modified versions: the FWF and FWλ. Section 3 is devoted to the new modified version, we present the possibility of combining the two alternatives presented above, to be able to imbricate them in the FW algorithm, we choose the best possible combination which can join together the maximum effectiveness to FW algorithm and preserve its convergence.
Section 4 establishes the convergences of the modifications cited. Finally, implementation issues and numerical results are discussed in section 5.

2. The Frank-Wolfe Method and modified versions

In this section, we consider the following non-linear programming problem (P):

$$(P)\begin{cases}\min\ f(x)\\ x\in X\end{cases}$$

Where $f : X \to IR$ is a twice continuously differentiable convex function, $X \subset IR^n$ is a compact convex polyhedral set.
By continuity of f, (P) has a solution, unique if f is strictly convex.

2.1 The Frank-Wolfe method

We can summarize the FW algorithm as follows:
At iteration k, FW approximates f by linearizing at the current iteration x^k, giving an affine minorant f_k to f:

$$f_k(y) \cong f(x^k) + \nabla f(x^k)(y - x^k) \qquad (1)$$

Manuscript received May 5, 2008
Manuscript revised May 20, 2008

First step: the determination of the search direction, which is done by a linear program:

$$LP(k)\begin{cases}\min\ f_k(y)\\ y \in X\end{cases}$$

Let y^k the optimal solution of $LP(k)$. The direction of Frank-Wolfe is then defined by:

$$d^k = y^k - x^k \qquad (2)$$

In the traffic application, $LP(k)$ decomposes into a set of shortest path problems [14].

Second step: In this phase, the objective function is minimized along the line segment passing by the point x^k and the direction d^k. the point obtained from the one-dimensional minimization is x^{k+1}.

Analytically this means that:

$$x^{k+1} = x^k + \alpha_k d^k \,/\, \alpha_k \in [0,1] \qquad (3)$$

With:

$$f(x^{k+1}) = \min_{0\le\alpha\le1} f(x^k + \alpha d^k) \qquad (4)$$

Convergence test: It is defined from the variation of the objective function between two successive iterations.

Let:

$LBD_k = f_k(y^k)$: a lower bound at iteration k.

$UBD_k = f(x^{k+1})$: a upper bound at iteration k.

LBD : is the best available lower bound.

The idea with FW is to iteratively shrink the gap between the so far best found lower bound and the upper bound from the last iteration, until they are sufficiently close to each other. The relative gap at iteration k can then be defined as [1]:

$$\frac{UBD_k - LBD}{LBD} \qquad (5)$$

So, the algorithm will stop if the relative gap is less then ε, $\varepsilon > 0$.

2.1.1 Convergence of FW method

Theorem1 The Frank-Wolfe method iterative process for the resolution of the problem (P), can be completely described by the following algorithmic map:

$$A : IR^n \to IR^n$$

$$x^k \to A(x^k) = x^{k+1} \text{ for } k \in IN^* \text{ with } A = MD$$

The FW algorithmic map A is composed of two maps:

The first one is the direction research map:

$$D : IR^n \to IR^{2n}$$

$$x^k \to D(x^k) = \left\{(x^k, d^k) / d^k = y^k - x^k\right\}$$

y^k the solution of the sub problem LP (k).

The other one is the line search map:

$$M : IR^{2n} \to IR^n$$

$$(x^k, d^k) \to M(x^k, d^k)$$

$$= \left\{x^{k+1} \,/\, f(x^{k+1}) = \min_{0\le\alpha\le1} f(x^k + \alpha d^k)\right\}$$

Lemma 1 The FW algorithmic map A is closed.

Lemma 2 Let suppose that the feasible area X, has p nods remotely finished: $y^1,\ldots,y^p$.

Let x^1 a starting feasible solution of the FW algorithm applied to the problem (P), then $\{x^k\}_1^\infty \subset C = \overline{Co(x^1, y^1,\ldots, y^p)}$. C is compact and $C \subset X$.

I.e. the sequence of the points generated by the FW algorithm is contained in the convex closing envelop of X.

To establish the FW convergence, we will apply the Convergence Theorem [17]:

Theorem2 (Convergence Theorem)

Let A be an algorithm on X, Z the evaluation function, $\Omega \subset X$ the solution set, let the sequence $\{x^k\}_1^\infty$ generated by the FW algorithm for the resolution of the problem (P), and suppose:

i) All points x^k are contained in a compact $S \subset X$.

ii) There is a continuous function Z on X such that:

if $x \notin \Omega$ then: $Z(y) < Z(x)$ for all points $y \in A(x)$.

if $x \in \Omega$ then: $Z(y) \le Z(x)$ for all points $y \in A(x)$.

iii) The mapping $A(x)$ is closed at points outside Ω.

Then the limit of any convergent subsequence of $\{x^k\}_1^\infty$ is a solution

2.2 The widened step technique

This technique consists in using, during a certain iteration count a descent step larger than the one obtained from one-dimensional optimization, while keeping the basic movement directions. It results a way in wider zigzag thus from it, where the algorithm points would advance more quickly towards the optimal solution (see Fig. 1):

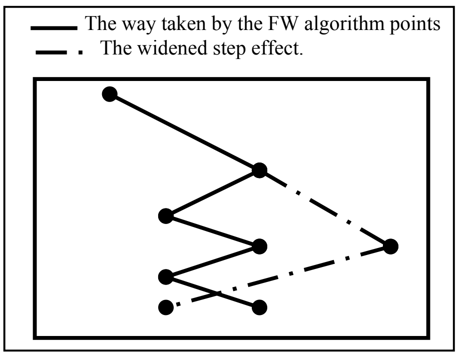

Fig. 1 The widened step effect.

2.2.1 Presentation of the FWλ method [11]

We will introduce this alternative in the form of mixed algorithm:
Let:
$A_k : IR^n \to IR^n$ the algorithmic map of the FW method.
$B : IR^n \to IR^n$ the modified version map.
With for $k \in K \backslash K^0$, $|K^0|$ finished $B = A_k$.
And for all $k \in K^0$, the following additional stages are required in the pure FW algorithm:
1- Set $\gamma_k = \lambda_k \alpha_k$; λ_k : is the step descent multiplier to the iteration k, with $\lambda_k > 1$.
2- Set $\beta_k = \min(\gamma_k, 1)$.
(This condition let us in the feasible area).
3- Let $x^{k+1} = x^k + \beta_k d^k$.
if $f(x^{k+1}) < f(x^k)$ set $x^{k+1} = x^{k+1}$.
else set $x^{k+1} = x^k + \alpha_k d^k$.
(If by modifying the descent step, we don't improve the objective function, we take $\lambda_k = 1$).

2.3 Descente direction modification

In this part, we present another modified version of the FW algorithm, it concerned the Fukushima method [6], we label it by FWF.

2.3.1 Presentation of the FWF method

We suppose that the optimal solution belongs to a face "S" of the feasible area (we show that the convergence is less slow if the optimum is an interior point of the feasible area [7]).
S is a convex polyhedron; therefore S can be represented by the convex envelope of a finished number of extreme points from the feasible area $y^1, \ldots y^n$.
Let x^* the optimal solution of the problem;
$x^* \in S$, thus there exists $\delta_i, i = 1,\ldots, n$ such as:
$x^* = \sum \delta_i y^i$ with $\sum \delta_i = 1, \ \delta_i \geq 0$ for $i = 1,\ldots, n$
These enable us to write x^* in the form:
$x^* = x^k + \alpha_k d^k \,/\, \alpha_k = 1$
And
$d^k = (\sum_{i=1}^{n} \delta_i y^i) - x^k$ ……………………….(*)
As it is practically impossible to determine with exactitude the y^i and the δ_i. however, we note that the FW algorithm generates y^i as solutions of sub-problems *LP*(i), which let us think that a convex combination of a certain number of the y^i previously generated by the algorithm could be used to approximate (*), and thus provide, a more refined direction than that of FW, which uses the y^i the most recently generated.
Moreover, in order to guarantee a maximum effectiveness of the modified algorithm, the two directions are compared in term of directional derivative, and those of the smallest value are selected to be the current descent direction.

The FWF can be summarized in the following steps:

Step0. Let x^1 a starting feasible solution, and let ℓ a positive integer. Set k=1.

Step1. Resolve the sub-problem *LP*(k), let y^k the optimal solution.
Step2. If $\nabla f(x^k)(y^k - x^k) = 0$ stop, else go to *Step3*
*Step3.*Let μ_i, $i = k$-q…,k chosen such as:

$$\sum_{i=k-q}^{k} \mu_i = 1, \ \mu_i \geq 0 \ \ \forall i = k-q,\ldots, k$$

Set:

$$v^k = (\sum_{i=k-q}^{k} \mu_i y^i) - x^k$$

Such as $q = \min\{k, \ell\}$-1,
And

$$w^k = y^k - x^k$$

*Step4.*Calculate the directional derivatives:

$$\gamma_1^k = \nabla f(x^k) v^k / \|v^k\| \tag{6}$$

$$\gamma_2^k = \nabla f(x^k) w^k / \|w^k\| \tag{7}$$

Put:

$$d^k = \begin{cases} v^k \text{ if } \gamma_1^k < \gamma_2^k \\ w^k \text{else} \end{cases} \tag{8}$$

Step5. Resolve the one-dimensional problem:

$$\min_{0 \leq \alpha \leq 1} f(x^k + \alpha d^k)$$

Let α_k its optimal solution.
Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k+1$, return to *Step1*.

2.4 The combined method, FWFλ

Now, we study the possibility of combining the two methods presented above, to be able to imbricate them in the FW algorithm.
We will choose the best possible combination which can join together the maximum effectiveness to FW algorithm; for a matter of convergence, this alternative must have the mixed algorithm form, which consequently leaves us two possibilities to consider:

- First possibility: Consider (FWF) as the basic algorithm, and the descent step technique as the modification to be incorporated.
- Second possibility: Taking (FWλ) as the basic algorithm, and Fukushima's modification like compensation measures.

As the (FWF) method acts on the descent directions which are responsible of the zigzag phenomenon; seen as this alternative equivalent to the FW method, in most unfavourable cases, which represents a certain guarantee of improvement, contrary to (FWλ) which rests on a heuristic technique. Thus, the first possibility is the best adapted to design a mixed algorithm.

2.4.1 Presentation of the FWFλ method

In this algorithm, the widened step technique will be apply on the « ℓ» first iterations, where the (FW) direction is preserved, considering the number « ℓ» of y^k used in Fukushima modification is not reached yet. Then, we use the (FWF) algorithm:

Step0. (Initialisation), Let x^1 a starting feasible solution, set k=1.

Step1. (FWFλ direction determination), Calculate the search direction of Fukushima " d^k ".

Step2. (Line search) Find the step α_k, which minimizes the objective: $\min_{0\le\alpha\le1} f(x^k + \alpha d^k)$

by application of the (FWλ) technique, if ($k \le \ell$)
by application of the (FWF) method, else.

Step3. if a stopping condition is satisfied stop,
else $k = k+1$; return to *Step1*.

3. Convergence of the modified versions

Consider that: $B : IR^n \to IR^n$ the algorithmic map of the modified versions. The convergence of the modified versions rises from the convergence theorem of the mixed algorithms.

3.1 The alternative (FWλ)

It is thus enough to check the following conditions of convergence:
1. For any $k \in K^0$; $f(x^{k+1}) \le f(x^k)$
2. For any $k \in K$; $x^k \in X$, where X is compact.
3. Let $x^* \in \Omega$, and let y a point of the algorithm; if $f(y) \le f(x^*)$, then $y \in \Omega$.

3.2 The alternative (FWF)

Suppose that the μ_i, $i = k\text{-}q\ldots,k$ are the continuous functions of $y^{k-q},\ldots y^k$ and x^k .
I.e. $\mu_i = \mu_i(y^{k-q},\ldots, y^k, x^k)$, for any k, and $v^k \neq 0$.
Then, the FWFλ algorithm is finishes in an optimal solution or generates an infinite continuation, of which any accumulation point is an optimal solution.
Thus, the Frank-Wolfe modifications proposed are globally convergent.

4. Computational Results

The main areas of the FW method applications and its modifications are traffic assignment problem; this last can be formulated as an optimization program with a nonlinear objective function and linear constraints [14]. Consider an urban traffic network represented as a graph G(N, A) where N and A are the sets of nodes and links, respectively. O is the set of origins and D is the set of destinations. The user-equilibrium traffic assignment problem can be stated as:

$$\min\ z = \sum_{a\in A} \int_0^{x_a} t_a(\omega)\,d\omega$$

$$\begin{cases} \sum_{p\in c_{od}} f_p \delta'_{p,od} = q_{od} & \forall o \in O, d \in D \\ f_p \ge 0 & \forall p \in C_{od}, \forall o \in O, d \in D \\ x_a = \sum_{od} \sum_{p\in C_{od}} f_p \delta_{a,p} & \forall a \in A \end{cases}$$

Where:
x_a the total flow on link a
$t_a(\omega)$ link cost function which is continuously differentiable and convex
q_{od} the total traffic demand between o and d
f_p the flow on path p
C_{od} the set of paths connecting o and d
$\delta_{a,p}$ the path-link incidence matrix.
$\delta'_{p,od}$ the path-pair(o, d) incidence matrix.

To demonstrate the efficiency of the new algorithm (FWFλ), in this section we present the results of the numerical examples in which the three algorithms: FWF, FWλ, FWFλ and the FW algorithm, are applied for solving some traffic assignment problems. In this study, the computer programs are coded in C++Builder5.

- Dijkstra's algorithm and Golden section method were used respectively for the resolution of the shortest path sub-problems and the one-dimensional problems.

- The starting feasible solution is obtained by application of the all-or-nothing method
- The strategies of the descent step modification were realized by repeated tests. The most representative strategies are:

For $N \leq 25$ $\lambda_k=1.6$ $K^0= \{k \in K / k \leq 5\}$
For $N > 25$ $\lambda_k=1.4$ or 1.5 $K^0= \{k \in K / k \leq 5\}$ or $= \{k \in K / k \leq 10\}$

- $\mu_i = 1/q, \forall i.$
- $\ell = |K^0|$.
- The convergence test is defined as:

$$\frac{UBD_k - LBD}{LBD} < 10^{-4} \tag{9}$$

4.1 First test

A first test was applied on a real problem suggested by L. J. Leblanc [8], witch considers the Sioux Falls network (South of Dakota; USA), composed of 24 nodes and 76 arcs.
We applied the following strategy:
$K^0= \{k \in K / k \leq 10\}$ and $\lambda_k=1.5$ for any $k \in K^0$.
On this example (see Table 1), we can note that the FWFλ algorithm is definitely higher than the FW method:

Table 1: Comparison between FW and FWFλ for the Sioux Falls network

Iteration k	FW	FWFλ
0	158.670	158.670
1	63.438	77.249
2	56.059	66.472
5	48.999	52.453
8	45.106	46.909
9	44.531	46.318
10	44.210	45.094
11	43.891	44.744
14	43.286	43.587
15	43.196	43.373
18	42.953	42.570
21	42.743	42.425
26	42.590	42.322
27	42.568	42.316
28	42.555	
32	42.491	
39	42.440	
40	42.438	
42	42.432	
48	42.404	
49	42.398	
55	42.317	

We remark that the FWλ technique did not make a rather good amelioration since the algorithm starts with a feasible solution very far of the optimum.
However, the effectiveness of the FWFλ method appears clearly during the introduction of the Fukushima direction. (see Table 1 for k>10)

4.2 Second test

Nine networks were randomly generated (i.e. networks configuration, total traffic demand between origins destinations and links parameters) for sensitivity analysis. The purpose is twofold. First, conduct more computational examples to examine FWFλ's performance furthermore. Second, make comparison between different FW modifications, especially on large scale networks. Some results are represented in Table 2: reports the objective function value, iteration number and CPU time of the different modifications for four networks:

Table 2: Comparison between FW, FWλ, FWF and FWFλ

(L, N)	method	optimum	iteration	Time (sec)	λ Value
(37,15)	FW	40335.25	114	0.485	K<5, λ=1.6
	FWλ	40341.15	50	0.221	
	FWF	40271.31	30	0.133	
	FWFλ	40276.78	14	0.065	
(90,35)	FW	87389.00	149	3.024	K<10, λ=1.4
	FWλ	86174.37	20	0.388	
	FWF	85784.06	31	0.645	
	FWFλ	85488.75	15	0.346	
(110,85)	FW	998809.94	91	2.458	K<5, λ=1.4
	FWλ	998201.00	60	1.459	
	FWF	992837.37	20	0.513	
	FWFλ	992837.37	18	0.458	
(120,40)	FW	204905.69	199	5.460	K<5, λ=1.4
	FWλ	203138.06	175	3.946	
	FWF	202362.12	75	2.011	
	FWFλ	200969.37	34	0.927	
(330,128)	FW	593610.39	251	8.032	K<10, λ=1.6
	FWλ	590869.97	234	7.997	
	FWF	587772.46	118	5.328	
	FWFλ	586931.82	67	3.151	

The global comparison is definitely favourable to FWFλ the results obtained indicate that:

- FWFλ gives a very good approximation of the optimum.
- The superiority of FWFλ is very remarkable. It reduces on the average more than 85% the iteration count required by the FW method, and more than 55% these required by FWF method.
- We also record, a good performance of FWF and FWλ.
- We also note that for some problems FWλ is higher in relation to FWF.

- According to the results obtained, we noted that the advantage of FWFλ becomes more significant in general, when the size of the network increases, and the required precision is finer, and for feasible starting solutions enough far from the optimum.

Fig 2 illustrates the improvement effect carried through the evaluation of the objective function during the first iterations of the first problem:

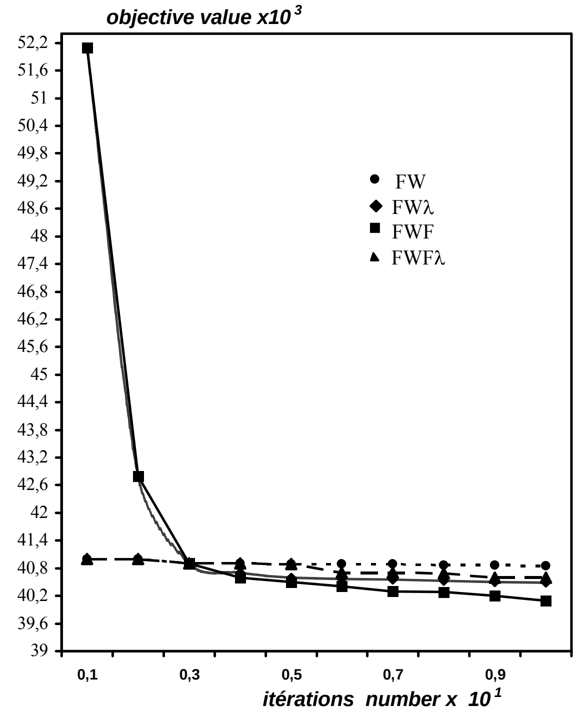

Fig. 2 Comparison for the first network

5. Conclusion

The FW method is widely applied for solving the urban traffic assignment problem, particularly for the user equilibrium assignment problem, which has a special structure that consists of an iterative process of linear programming. However, the FW algorithm converges very slowly when iterations are closing to the optimal solution. In order to expedite the rate of convergence of the Frank-Wolfe algorithm, the paper has specified a new algorithm for solving the traffic assignment problem.

A numerical example has been given to demonstrate the effectiveness of the proposed algorithm. The computational results indicate that the performance of the FWFλ is better than the pure FW and the other methods discussed.

Finally, we have here, at least, in the case of the FW method a thesis confirmation on which the new algorithm FWFλ was conceived and which is: The Combination of certain techniques of acceleration the convergence rate within an algorithm generates further improvement, even to define the good compatibilities conditions.

References

[1] M. Daneva, P.O. Lindberg. “The Stiff is Moving - Conjugate Direction Frank-Wolfe Methods with Applications to Traffic Assignment”. Submitted. (2004).

[2] C. Duhamel, Ph.Mahey. “Multicommodity flow problems with a bounded number of paths: a flow deviation approach”. March 22, 2005.

[3] M. Florian. “An improved linear approximation algorithm or the network equilibrium (packet witching) problem”, in: Proceedings of the 1977 IEEE Conference on Decision Control, 1997, pp. 812-818.

[4] M. Frank and PH. Wolfe. “An algorithm for quadratic programming”. Naval Res. Logist. Quart.3:95-110, 1956.

[5] L. Fratta, M. Gerla, and L. Kleinrock. “The flow deviation method: An approach to store-and-forward communication network design. Networks”, 3:97-133, 1973

[6] M. Fukushima. “A modified Frank-Wolfe algorithm for solving the traffic assignment problem”. Transportation Res. PartB, 18(2):169-177, 1984.

[7] C.C. Holloway. “An Extension to the Frank-Wolfe method of feasible directions”. Mathematical Programming. 6, 14-27, (1974).

[8] L.J. Leblanc. “Mathematical programming algorithms for large scale network equilibrium and network design problems”. PHD thesis, IE/MS Dept, Northwestern University, Evanston IL, 1973.

[9] T. Larsson, M. Patriksson, and C. Rydergren. “Application of simplicial decomposition with nonlinear column generation to nonlinear network flows”. In P. M. Pardalos, D.W. Hearn, and W. W. Hager, editors, Network Optimisation, pages 346-373.Springer, Berlin, 1997.

[10] L.J. Leblanc.,R.V. Helgason, and D.E. Boyce. “Improved efficiency of the Frank-Wolfe algorithm for convex network programs”. Transportation Sci.,19(4):445-462, 1985.

[11] R. Ouafi. “Analyse et contrôle des réseaux de trafic urbain par la méthode de Frank-Wolfe ”. Thèse de Doctorat 3ème cycle, Université Paris 6, 1988.

[12] R. Ouafi, S. Arrache. “Une version modifiée de la méthode de Frank-Wolfe pour la résolution du problème de l’affectation du trafic urbain”. Sur l’acte du colloque International sur l’optimisation et les systèmes d’information, COSI’05. Université de Béjaia, Algérie, juin 2005.

[13] R. Ouafi, S. Arrache. “Improved Frank-Wolfe method. Application to the traffic assignment problem”. International Congress of Mathematicians Madrid, 2006.

[14] M. Patriksson. *The traffic Assignment Problem–Models and Methods*. VSP, Utrecht, 1994.

[15] W.B. Powell and Y. Sheffi. “The convergence of equilibrium algorithms with predetermined step sizes”. Transportation Sci., 16(1):45-55, 1982.

[16] A. Weintraub, C. Ortiz, and J. González. “Accelerating convergence of the Frank-Wolfe algorithm”. Transportation Res. Part B, 19(2):113-122, 1985.

[17] Willard I. Zangwill. “Nonlinear programming: a unified approach”. Prentice-Hall Inc., Englewood Cliffs, N.J., 1969

[18] Ziyou Gao, W.H.K. Lam, S.C. Wrong, H. Yang. “The convergence of equilibrium algorithms with non-monotone line search technique”. Applied Mathematics and Computation, 148 pp1-13, 2004.