

# Canadian travelers minimize the traversed distance: definitions, bounds and heuristics

March 28, 2018

## 1 Introduction

The *Canadian traveler Problem* (CTP) was introduced by Papadimitriou and Yannakakis [5]. This is a generalization of the *Shortest Path Problem*. Given an undirected weighted graph  $G = (V, E, \omega)$  and two nodes  $s, t \in V$ , the objective is to design a strategy in order to make a traveler walk from  $s$  to  $t$  through the graph  $G$ , knowing that some edges can be blocked. The traveler ignores which edges are blocked when he begins and discover them when he visits an adjacent node. The *k-Canadian traveler Problem* (*k*-CTP) is the parameterized variant of CTP where we specify an upper bound for the total number of blocked edges. Both CTP and *k*-CTP are PSPACE-complete [1, 5].

**State-of-the-art:** Several strategies have been designed and studied through the competitive analysis, which is a way to assess the quality of an online algorithm. A first class of strategies to be considered are deterministic. Westphal proved that there is no deterministic algorithm that can achieve a better competitive ratio than  $2k + 1$  where  $k$  is an upper bound of the number of blockages and that this ratio is achieved by REPOSITION algorithm [6]. However, in practice (for example in the case of an urban network), returning to node  $s$  everytime the traveler is blocked does not seem to be an efficient strategy. This is why Xu et al. introduced the GREEDY algorithm for the CTP which achieves a  $2^{k+1} - 1$  ratio [7]. For grids, they showed that the GREEDY strategy achieves a  $\mathcal{O}(1)$  ratio, independent of  $k$ , under realistic hypotheses. Both GREEDY and REPOSITION strategies are executed in polynomial time.

A second class of strategies are randomized. We evaluate these strategies by supposing that an oblivious adversary is setting the blocked edges. Westphal proved that there is no randomized algorithm that can achieve a ratio lower than  $k + 1$  [6]. Recently, two randomized algorithms have been proposed. Demaine et al. designed a strategy for arbitrary graphs with a ratio  $\left(1 + \frac{\sqrt{2}}{2}\right)k + 1$ . This is executed in time  $\mathcal{O}\left(k\mu^2 |E|^2\right)$  where  $\mu$  is a parameter that can potentially be exponential [4]. However, for a rather large class of graph (graphs with a reasonable number of groups of path, where a group of path is a set of paths which have the same length), this strategy offers better results than deterministic methods and is executed in polynomial time. Furthermore, Bender et al. studied the specific case of node-disjoint-paths graphs and proposed a strategy of ratio  $(k + 1)$  and with a polynomial running time for this kind of graphs [2]. Their method consists in a randomized REPOSITION: we assign a probability  $p_i$  to each path  $P_i$  and execute a draw: the traveler crosses the path and, if he is blocked, returns to  $s$  and restarts the process.

## 2 Preliminaries

### 2.1 Notations and definitions with a single traveler

The traveler traverses an undirected weighted graph  $G = (V, E, \omega)$ ,  $n = |V|$  and  $m = |E|$ . He starts his walk at source  $s \in V$ . His objective is to reach target  $t \in V$  with a minimum cost (also called distance), which is the sum of the weights of edges traversed. Set  $E_*$  contains blocked edges, which means that when the traveler reaches an endpoint of one of these edges, he discovers that he cannot pass through it. A pair  $(G, E_*)$  is called a *road map*. From now on, we suppose that any road map  $(G, E_*)$  is feasible, *i.e.*  $s$  and  $t$  are always connected in graph  $G \setminus E_*$ .

We remind the definition of the competitive ratio introduced in [3]. Let  $\omega_A(G, E_*)$  be the distance traversed by the traveler guided by a given strategy  $A$  on graph  $G$  from source  $s$  to target  $t$  with blocked edges  $E_*$ . The shortest  $(s, t)$ -path in  $G \setminus E_*$  is called the *optimal offline path* of map  $(G, E_*)$  and its cost, noted  $\omega_{\min}(G, E_*)$ , is the optimal offline cost of map  $(G, E_*)$ . Strategy  $A$  is  $c_A$ -competitive if, for any road map  $(G, E_*)$ :

$$\omega_A(G, E_*) \leq c_A \omega_{\min}(G, E_*) + \eta,$$

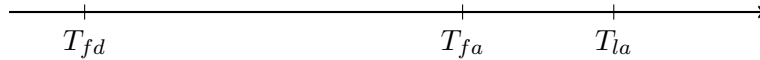
where  $\eta$  is constant. For randomized strategies, it becomes:

$$\mathbb{E}[\omega_A(G, E_*)] \leq c_A \omega_{\min}(G, E_*) + \eta.$$

### 2.2 Notations and definitions with $L$ travelers

The  $L$  travelers traverses the same undirected weighted graph  $G = (V, E, \omega)$ ,  $n = |V|$  and  $m = |E|$ . They all start their walk at source  $s \in V$ . Their objective is to reach target  $t \in V$  with a minimum cost (also called distance), which is the sum of the weights of edges traversed. We call  $d_i$  the distance traveled by the  $i^{\text{th}}$  traveler. We consider that the  $L$  travelers can communicate with each other. Set  $E_*$  contains blocked edges, which means that when a traveler reaches an endpoint of one of these edges, he discovers that he cannot pass through it and communicate this information to the other travelers. A pair  $(G, E_*)$  is called a *road map*. From now on, we suppose that any road map  $(G, E_*)$  is feasible, *i.e.*  $s$  and  $t$  are always connected in graph  $G \setminus E_*$ .

We consider the following timeline which brings forth three times we will need to study this problem.



$T_{fd}$  : time of the first departure

$T_{fa}$  : time of the first arrived

$T_{la}$  : time of the last arrived

We want to study the minimization of the distances traversed by all the travelers from the departure point  $s$  and the arrival point  $t$ . We consider  $d_i$  the distance traversed by the  $i^{\text{th}}$  traveler. We have two main problems to consider, we will try to minimize the distances traversed by all travelers:

- Before the first arrived :

$$\sum_{i=1}^L d_i \quad \text{between } T_{fd} \text{ and } T_{fa}$$

- After the last arrived :

$$\sum_{i=1}^L d_i \quad \text{between } T_{fd} \text{ and } T_{la}$$

### 3 Bounds of competitiveness: deterministic algorithms

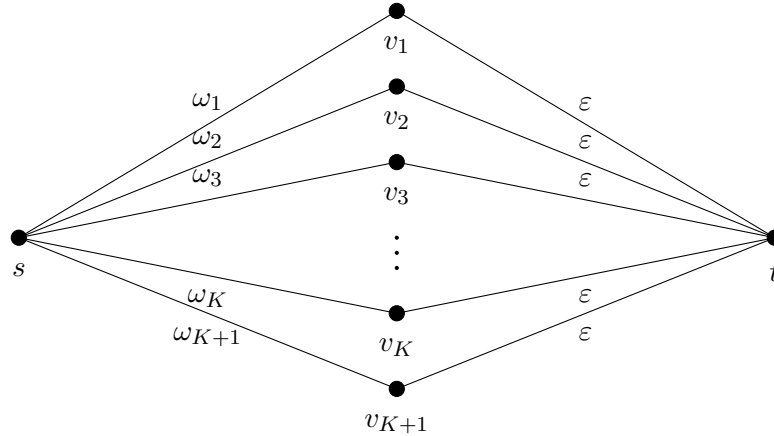
We will study this problem with three different types of communication. We will note  $P_0$  the proposition when we don't have any communication,  $P_1$  when we have partial communication and  $P_2$  for complete communication.

#### 3.1 Complete communication

We consider the following deterministic and online algorithm for  $K$  blocked edges and  $L$  travelers with complete communication: we launch the travelers one at a time. If the traveler meets a blocked edge, it has to stay there and we launch another traveler who will have to take the second shortest path possible. Once again when this one meets a blocked edge, it stays there and another is launched. Let's call this strategy the ABANDONMENT strategy.

**Lemma 1** *The cost of the abandonment strategy on the Westphal graph is  $2(K+1) - \min(K+1, L)$  between  $T_{fd}$  and  $T_{fa}$  and  $2K+L$  between  $T_{fd}$  and  $T_{la}$ .*

**Proof.** We consider the following Westphal graph, where all the ways from the departure point to the arrival point have the same cost  $1 + \varepsilon$ . We are searching the cost for the worst case scenario: the only unblocked route is the one we choose last  $(s, v_{K+1}, t)$ .



We consider that  $w_1 = w_2 = \dots = w_{K+1} = 1$ , that  $\varepsilon$  is very small, that for  $1 \leq i \leq K$ ,  $(v_i, t)$  is blocked and that  $t_{i+1}$  is a time after the  $i^{th}$  movement.

We have in fact 4 problems to consider: for each of the 2 main problems (between  $T_{fd}$  and  $T_{fa}$  and between  $T_{fd}$  and  $T_{la}$ ) we have two secondary problems, when  $K < L$  and  $K \geq L$ .

- Case 1: between  $T_{fd}$  and  $T_{fa}$  and  $K < L$

At the time  $t_1$ , we send a first traveler through the edge  $(s, v_1)$ , who discovers that the edge  $(v_1, t)$  is blocked and communicates it to the other travelers.

Then we send at the times  $t_i$  with  $1 < i \leq K$  a traveler through the edge  $(s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and communicates it to the other travelers.

At the time  $t_{K+1}$ , we send a traveler through the edges  $(s, v_{K+1})$  and then reaches the destination by the edge  $(v_{K+1}, t)$ .

Therefore we have:

$$\sum_{i=1}^K \omega_i + \omega_{K+1} + \varepsilon = \sum_{i=1}^K 1 + 1 + \varepsilon = K + 1 + \varepsilon.$$

- Case 2: between  $T_{fd}$  and  $T_{la}$  and  $K < L$

We redo the three first steps of the 1st case and at  $T_{pa}$  the first traveler communicate the unblocked route to the target and therefore that the edge  $(v_{K+1}, t)$  is not blocked.

Each traveler at the point  $v_i$  will then arrive at the destination with a cost of  $w_i + w_{K+1} + \varepsilon$ . And each traveler still at the departure point will arrive at destination with a cost of  $w_{K+1} + \varepsilon$ .

Therefore, we have:

$$\begin{aligned} \sum_{i=1}^K \omega_i &+ \omega_{K+1} + \varepsilon + \sum_{i=1}^K (\omega_i + \omega_{K+1} + \varepsilon) + (L - K)(\omega_{K+1} + \varepsilon) \\ &= 2 \sum_{i=1}^K \omega_i + L(\omega_{K+1} + \varepsilon) = 2 \sum_{i=1}^K 1 + L(1 + \varepsilon) \\ &= 2K + L(1 + \varepsilon). \end{aligned}$$

- Case 3: between  $T_{fd}$  and  $T_{fa}$  and  $K \geq L$

At the time  $t_1$ , we send a first traveler through the edge  $(s, v_1)$ , who discovers that the edge  $(v_1, t)$  is blocked and communicates it to the other travelers.

Then we send at the times  $t_i$  with  $1 < i \leq L$  a traveler through the edge  $(s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and communicates it to the other travelers.

At the time  $t_{L+1}$ , we send the traveler which is at the point  $v_L$  through the edges  $(v_L, s)$  and then  $(s, v_{L+1})$ . He discovers that the edge  $(v_{L+1}, t)$  is blocked and communicates it to the other travelers.

Then we send at the times  $t_i$  with  $L + 1 < i \leq K$  a traveler through the route  $(v_{i-1}, s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and communicates it to the other travelers.

At the time  $t_{K+1}$ , we send a traveler through the route  $(v_K, s, v_{K+1})$  and then reaches the destination by the edge  $(v_{K+1}, t)$ .

Therefore we have:

$$\begin{aligned} \sum_{i=1}^L \omega_i &+ \sum_{i=L+1}^K (\omega_{i-1} + \omega_i) + \omega_K + \omega_{K+1} + \varepsilon \\ &= \sum_{i=1}^L 1 + \sum_{i=1}^{K-L} 2 + 1 + 1 + \varepsilon \\ &= L + 2(K - L) + 2 + \varepsilon \\ &= 2(K + 1) - L + \varepsilon. \end{aligned}$$

- Case 4: between  $T_{fd}$  and  $T_{la}$  and  $K \geq L$

We redo the five first steps of the third case and at  $T_{pa}$  the first traveler will communicate the unblocked route to the target and therefore that the edge  $(v_{K+1}, t)$  is not blocked.

At the time  $t_{K+2}$  the  $L-1$  travelers blocked on the nodes  $v_1 \dots v_{L-1}$  come back to the point  $s$  and then take the unblocked route  $(s, v_{K+1}, t)$ .

Each traveler at the point  $v_i$  will then arrive at the destination with a cost of  $w_i + w_{K+1} + \varepsilon$ .

Therefore we have:

$$\begin{aligned} 2(K+1) &= L + \varepsilon + \sum_{i=1}^{L-1} (\omega_i + \omega_{K+1} + \varepsilon) \\ &= 2K + L(1 + \varepsilon). \end{aligned}$$

■

Therefore we have the following costs for the Westphal graph:

	between $T_{fd}$ and $T_{fa}$	between $T_{fd}$ and $T_{la}$
$K < L$	$K + 1$	$2K + L$
$K \geq L$	$2(K + 1) - L$	$2K + L$

**Lemma 2** *The abandonment strategy has the optimal cost on the Westphal graph.*

**Proof.** This Lemma is proved using recurrence.

**Recurrence hypothesis:** For  $K$  (blocked edges), for all  $L$  the optimal cost is :  $K + 1$  if  $L > K$  and  $2K - L + 1$  if  $L < K$ .

**Initialization:** Let us take the example of  $L = 1$  traveler. As shown in the existing literature we already have the cost equal to  $K + 1$ .

**Recurrence:** We know that there is a natural integer  $n$  number of travelers for which the cost is: if  $n > K$ , the cost is  $K + 1$ , and if  $n < K$ , the cost is  $2K - L + 1$ . Let's prove it for  $n + 1$ . If  $n + 1 < K$ , since we are in the worst case scenario, the  $n$  travelers have to discover the  $K$  blocked edges before reaching the end. To do so, the cost is at least of  $K$ . Then 1 is added to reach the end. However, our strategy has a cost of  $K + 1$  also, therefore once again, for  $n + 1 < K$ , the optimal cost is  $K + 1$ .

Now if  $n + 1 > K$ , we launch the first traveler. He discovers one blocked edge and stays there. We are now facing a problem of  $n$  travelers facing  $K - 1$  blockage. By recurrence hypothesis we know that for that case the cost is  $2(K - 1) - n + 1$  because  $n > K - 1$ . To that cost we add the cost of discovering the first blocked edge. We have  $2(K + 1) - n + 2$  which is equal to  $2K - (n + 1) - 1$ . This shows the second part of the recurrence.

We can therefore conclude that the optimal cost for the Westphal graph is :  $K + 1$  if  $L > K$  and  $2K - L + 1$  if  $L < K$ . This is the cost of the abandonment strategy. ■

**Lemma 3** *The competitive ratios found in the Lemma 3.1.1 are the minimum we can have for any graph. Therefore the best competitive ratio for the abandonment algorithm is  $2(K + 1) - \min(K + 1, L)$  between  $T_{fd}$  and  $T_{fa}$  and  $2K + L$  between  $T_{fd}$  and  $T_{la}$ .*

**Proof.** We consider the same graph as in the proof of the Lemma 3.1, but this time we consider that we have  $w_1 < w_2 < \dots < w_{K+1}$ . We still consider that  $\varepsilon$  is very small, that for  $1 \leq i \leq K$ ,  $(v_i, t)$  is blocked and that  $t_{i+1}$  is a time after the  $i^{th}$  movement. We are still searching for the cost of the worst case scenario: the only unblocked route is the one we choose last, in this case the one with the biggest cost  $(s, v_{K+1}, t)$ .

We still have the same 4 problems to study as in the proof of the Lemma 3.1.1, for each case, we follow the same procedure as before.

The true optimum route in each case will be the last one taken with a cost of  $\omega_{K+1} + \varepsilon$ .

- Case 1: between  $T_{fd}$  and  $T_{fa}$  and  $K < L$

We have:

$$\begin{aligned} \frac{\sum_{i=1}^K \omega_i + \omega_{K+1} + \varepsilon}{\omega_{K+1} + \varepsilon} &\leq \frac{\sum_{i=1}^K K + 1\omega_{K+1} + \varepsilon}{\omega_{K+1}} + o(1) \\ &\leq \frac{(K+1)\omega_{K+1}}{\omega_{K+1}} + o(1) \leq K + 1 + o(1). \end{aligned}$$

- Case 2: between  $T_{fd}$  and  $T_{la}$  and  $K < L$

We have:

$$\begin{aligned} &\frac{\sum_{i=1}^K \omega_i + \omega_{K+1} + \varepsilon + \sum_{i=1}^K (\omega_i + \omega_{K+1} + \varepsilon) + (L-K)(\omega_{K+1} + \varepsilon)}{\omega_{K+1} + \varepsilon} \\ &\leq \frac{2\sum_{i=1}^K \omega_{K+1} + L\omega_{K+1}}{\omega_{K+1}} + o(1) \leq \frac{2(K+L)\omega_{K+1}}{\omega_{K+1}} + o(1) \\ &\leq 2K + L + o(1). \end{aligned}$$

- Case 3: between  $T_{fd}$  and  $T_{fa}$  and  $K \geq L$

We have:

$$\begin{aligned} &\frac{\sum_{i=1}^L \omega_i + \sum_{i=L+1}^K (\omega_{i-1} + \omega_i) + \omega_K + \omega_{K+1} + \varepsilon}{\omega_{K+1} + \varepsilon} \leq \frac{\sum_{i=1}^L \omega_{K+1} + \sum_{i=1}^{K-L} 2\omega_{K+1} + 2\omega_{K+1}}{\omega_{K+1}} + o(1) \\ &\leq \frac{(L + 2(K-L) + 2)\omega_{K+1}}{\omega_{K+1}} + o(1) \leq 2(K+1) - L + o(1). \end{aligned}$$

- Case 4: between  $T_{fd}$  and  $T_{la}$  and  $K \geq L$

We have:

$$\begin{aligned} &\frac{\sum_{i=1}^L \omega_i + \sum_{i=L+1}^K (\omega_{i-1} + \omega_i) + \omega_K + \omega_{K+1} + \varepsilon + \sum_{i=1}^{L-1} (\omega_i + \omega_{K+1} + \varepsilon)}{\omega_{K+1} + \varepsilon} \\ &\leq \frac{\sum_{i=1}^L \omega_{K+1} + \sum_{i=1}^{K-L} 2\omega_{K+1} + 2\omega_{K+1} + \sum_{i=1}^{L-1} (\omega_{K+1} + \omega_{K+1})}{\omega_{K+1}} + o(1) \\ &\leq \frac{(L + 2(K-L) + 2 + 2(L-1))\omega_{K+1}}{\omega_{K+1}} + o(1) \leq 2K + L + o(1). \end{aligned}$$

■

Therefore we have the following competitive ratio for any graph:

	between $T_{fd}$ and $T_{fa}$	between $T_{fd}$ and $T_{la}$
$K < L$	$K + 1$	$2K + L$
$K \geq L$	$2(K + 1) - L$	$2K + L$

### 3.2 No communication

We now consider the same following deterministic and online algorithm for  $K$  blocked edges and  $L$  travelers with no communication: they don't even know that they are a part of a fleet, so they will all leave at the first moment and will all arrive at the same time at the end point. We will call this strategy the multiple single traveler strategy.

**Lemma 4** *The cost of the multiple single traveler strategy on the Westphal graph is  $(2K + 1)L$  between  $T_{fd}$  and  $T_{fa}$  and between  $T_{fd}$  and  $T_{la}$ .*

**Proof.** None of the  $L$  travelers know that they are a fleet. All of them will depart at the first time. We are in the worst case scenario, so all of them will go through all the blocked paths before finding the opened one. They will all be independent and act as a single traveler. They will all arrive at the destination at the same time. Therefore no matter if we take into account the distance between  $T_{fd}$  and  $T_{fa}$  or between  $T_{fd}$  and  $T_{la}$ , we will have a cost of  $(2K + 1)$  for each traveler, meaning  $(2K + 1)L$  total. ■

**Lemma 5** *The multiple single traveler strategy has the optimal cost on the Westphal graph.*

**Proof.** All of the travelers aren't communicating, therefore they are independent. We are in the worst case scenario, so each traveler will be as a single traveler and have a cost of  $(2K + 1)$  as seen in the existence literature. So we have a total cost of  $(2K + 1)L$ . ■

**Lemma 6** *The competitive ratios found in the Lemma 3.2.1 are the minimum we can have for any graph. Therefore the best competitive ratio for the multiple single traveler algorithm is  $(2K + 1)L$  between  $T_{fd}$  and  $T_{fa}$  and between  $T_{fd}$  and  $T_{la}$ .*

**Proof.** We consider the same graph as in the proof of the Lemma 3.2.1, but this time we consider that we have  $w_1 < w_2 < \dots < w_{K+1}$ . We still consider that  $\varepsilon$  is very small, that for  $1 \leq i \leq K$ ,  $(v_i, t)$  is blocked. We are still searching for the cost of the worst case scenario: the only unblocked route is the one we choose last, in this case the one with the biggest cost  $(s, v_{K+1}, t)$ .

The optimum route which is the only unblocked route in each case will be the last one taken with a cost of  $\omega_{K+1} + \varepsilon$ .

In both cases, we will have:

$$\begin{aligned} \frac{L(\sum_{i=1}^K 2 * \omega_i + \omega_{K+1} + \varepsilon)}{\omega_{K+1} + \varepsilon} &\leq \frac{L(\sum_{i=1}^K 2 * \omega_{K+1} + \omega_{K+1})}{\omega_{K+1}} + o(1) \\ &\leq \frac{L(2K + 1)\omega_{K+1}}{\omega_{K+1}} + o(1) \leq L(2K + 1) + o(1). \end{aligned}$$

For any graph, the best competitive ratio for the multiple single traveler strategy is  $L(2K + 1)$  for both problems: between  $T_{fd}$  and  $T_{fa}$  and between  $T_{fd}$  and  $T_{la}$ . ■

### 3.3 Partial communication

We now consider the same following deterministic and online algorithm for  $K$  blocked edges and  $L$  travelers with partial communication: the travelers can communicate at the beginning and decide what each traveler will do, they don't talk later on. We will have a rolling strategy where the first traveler takes the first route, the second one will wait a time a bit bigger than it would have taken the first one to arrive at destination if the first path was unblocked, and will take the second path and so on. In the cases where  $K > L$  the last traveler will continue to take each following path when we are in the first arrival problem. For the last arrival problem, we will minimize the worst case scenario by eliminating some paths for travelers as they will take different paths at the beginning and take the following one afterwards. Therefore they will not all be traveling all paths.

**Lemma 7** *The cost of the rolling travelers strategy on the Westphal graph is:*

	between $T_{fd}$ and $T_{fa}$	between $T_{fd}$ and $T_{la}$
$K < L$	$K + 1$	$2K + L - \lfloor \frac{L}{K+1} \rfloor (K + 1)(K + 2) - (L - \lfloor \frac{L}{K+1} \rfloor)(L - \lfloor \frac{L}{K+1} \rfloor + 1)$
$K \geq L$	$2(K + 1) - L$	$2K - L^2$

**Proof.** We consider the same Westphal graph as in the proof of the Lemma 3.1.1 and the same 4 problems. We consider that the times  $t_i$  is  $i$  times the time needed to go from  $s$  to  $t$  with an unblocked path (in this case time needed to travel  $1 + \varepsilon$ ).

- Case 1: between  $T_{fd}$  and  $T_{fa}$  and  $K < L$

At the time  $t_1$ , we send a first traveler through the edge  $(s, v_1)$ , who discovers that the edge  $(v_1, t)$  is blocked and stays there.

Then we send at the times  $t_i$  with  $1 < i \leq K$  a traveler through the edge  $(s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and stays there.

At the time  $t_{K+1}$ , we send a traveler through the edges  $(s, v_{K+1})$  and then reaches the destination by the edge  $(v_{K+1}, t)$ .

Therefore we have:

$$\sum_{i=1}^K \omega_i + \omega_{K+1} + \varepsilon = \sum_{i=1}^K 1 + 1 + \varepsilon = K + 1 + \varepsilon.$$

- Case 2: between  $T_{fd}$  and  $T_{fa}$  and  $K \geq L$

At the time  $t_1$ , we send a first traveler through the edge  $(s, v_1)$ , who discovers that the edge  $(v_1, t)$  is blocked and stays there.

Then we send at the times  $t_i$  with  $1 < i \leq L$  a traveler through the edge  $(s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and stays there.

At the time  $t_{L+1}$ , we send the traveler which is at the point  $v_L$  through the edges  $(v_L, s)$  and then  $(s, v_{L+1})$ . He discovers that the edge  $(v_{L+1}, t)$  is blocked and continues to the next path.

Then we send at the times  $t_i$  with  $L + 1 < i \leq K$  the same last traveler through the route  $(v_{i-1}, s, v_i)$  and discovers that the edge  $(v_i, t)$  is blocked and continues.



At the time  $t_{K+1}$ , we send the last traveler through the route  $(v_K, s, v_{K+1})$  and then reaches the destination by the edge  $(v_{K+1}, t)$ .

Therefore we have:

$$\begin{aligned} \sum_{i=1}^L \omega_i + \sum_{i=L+1}^K (\omega_{i-1} + \omega_i) + \omega_K + \omega_{K+1} + \varepsilon &= \sum_{i=1}^L 1 + \sum_{i=1}^{K-L} 2 + 1 + 1 + \varepsilon \\ &= L + 2(K-L) + 2 + \varepsilon = 2(K+1) - L + \varepsilon. \end{aligned}$$

- Case 3: between  $T_{fd}$  and  $T_{la}$  and  $K \geq L$

At the time  $t_1$ , we send the  $i^{th}$  traveler through the edges  $(s, v_i)$  for  $1 \leq i \leq L$  they all discover that the edge  $(v_i, t)$  is blocked.

For each time  $t_j$ , with  $1 \leq j \leq L - K$  we will send all the travelers through the following path. The  $i^{th}$  traveler through the edges  $(s, v_{i+j-1})$  for  $1 \leq i \leq L$ , all discover that the edge  $(v_{i+1}, t)$  is blocked.

For each following time  $t_j$ , with  $L - K < j \leq L$  we will send all the travelers through the following path. The  $i^{th}$  traveler through the edges  $(s, v_{i+j-1})$  for  $1 \leq i \leq L - K - j$ , all discover that the edge  $(v_{i+1}, t)$  is blocked except the last one who arrives at destination and stays there.

The first traveler will then have a cost of  $2K + 1$ , the second of  $2K$ . For  $1 \leq i \leq L$ , the  $i^{th}$  traveler will have a cost of  $2(K - i) + 1$

Therefore we have:

$$\begin{aligned} \sum_{i=1}^L (2(K - i) + 1) &= 2KL + L - 2 \sum_{i=1}^L i = L(2K + 1) - 2 \frac{L(L+1)}{2} \\ &= L(2K + 1 - (L + 1)) = L(2K - L). \end{aligned}$$

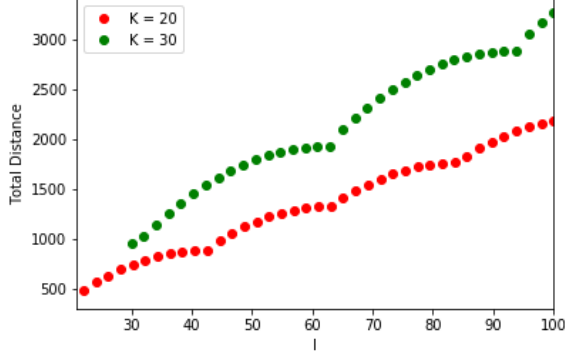
- Case 4: between  $T_{fd}$  and  $T_{la}$  and  $K < L$

We redo the same pattern as in the 3rd case, but there will be more than one traveler on the same road. The  $i^{th}$  traveler will take the  $i \bmod (K + 1)^{th}$  path.

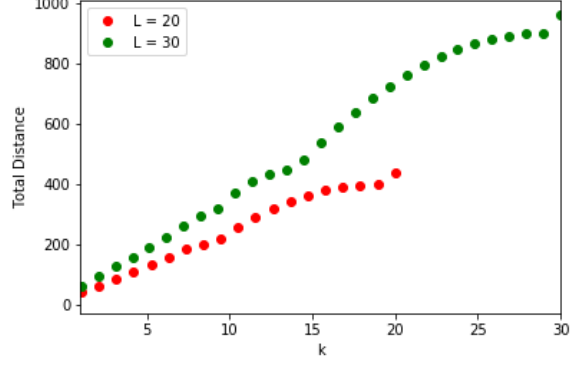
The first traveler will then have a cost of  $2K + 1$ , the second of  $2K$ . For  $1 \leq i \leq L$ , the  $i^{th}$  traveler will have a cost of  $2(K - i \bmod (K + 1)) + 1$

Therefore we have:

CEI/courbeNewKdeter.png

(a)  $K$  set to 20 and 30 for  $K < L$  in the deterministic approach

CEI/courbeNewLdeter.png

(b)  $L$  set to 20 and 30 for  $K < L$  in the deterministic approachFigure 1: Total distance for  $K < L$  in the deterministic approach

$$\begin{aligned}
\phi(K, L) &= \sum_{i=1}^L (2(K - (i - 1) \bmod (K + 1)) + 1) \\
&= L(2K + 1) - 2 \sum_{i=0}^{L-1} i \bmod (K + 1) \\
&= L(2K + 1) - 2 \left( \left\lfloor \frac{L-1}{K+1} \right\rfloor \sum_{i=0}^K i + \sum_{i=0}^{(L-1) \bmod (K+1)} i \right) \\
&= L(2K + 1) - 2 \left( \left\lfloor \frac{L-1}{K+1} \right\rfloor \frac{K(K+1)}{2} + \frac{((L-1) \bmod (K+1)) ((L-1) \bmod (K+1) + 1)}{2} \right) \\
&= L(2K + 1) - \left\lfloor \frac{L-1}{K+1} \right\rfloor K(K+1) - ((L-1) \bmod (K+1)) ((L-1) \bmod (K+1) + 1).
\end{aligned}$$

■

### 3.4 Conclusion for the deterministic approach

We note  $\phi(K, L) = L(2K + 1) - \left\lfloor \frac{L-1}{K+1} \right\rfloor K(K+1) - ((L-1) \bmod (K+1)) ((L-1) \bmod (K+1) + 1)$ .

	$P_0$	$P_1$		$P_2$	
		$K < L$	$K \geq L$	$K < L$	$K \geq L$
between $T_{fd}$ and $T_{fa}$	$2K + 1$	$K + 1$	$2(K + 1) - L$	$K + 1$	$2(K + 1) - L$
between $T_{fd}$ and $T_{la}$	$2K + 1$	$\phi(K, L)$	$L(2K - L)$	$2K + L$	$2K + L$

In order to have a better view of what  $\phi(K, L)$  represents, we propose to the reader to look at [Figure 1]:

First we set  $K$  and make  $L$  vary and then we set  $L$  and make  $K$  vary.

As one can expect, those curves are growing curves. Therefore the more travelers, the more blockages, the higher the cost. Nevertheless, one can notice that the tangent flattens out when

the number of car approaches a multiple of  $K + 1$ . Therefore, we can conclude that it is more interesting to use a multiple of  $K + 1$  number of car in this case.

## 4 Bounds of competitiveness: randomized algorithms

### 4.1 Complete communication

Now we consider the randomized online algorithms for  $K$  blocked edges and  $L$  travelers with complete communication: we assign a probability  $p_i$  to each path  $P_i$  and execute a draw: let the travelers crosses the path one by one, if the first traveler is blocked, leave him at the blocked point and let the second traveler to start the process; if all the travelers are blocked, then let the first traveler return to  $s$  and restarted the process. We can get:

**Lemma 8** *If we have more travelers than blocked edges, there is no randomized online algorithm with competitive ratio less than  $\frac{k+2}{2}$ .*

**Proof.** With  $K$  smaller than  $L$ , it is certain that we do not have to let a traveler return to  $s$  before we find all the blocked edges. If the algorithm is successful on its  $(l + 1)^{\text{th}}$  try, only the  $(l + 1)^{\text{th}}$  try cost  $(1 + \eta)$  and so the total cost will be:

$$(l - 1) \cdot 1 + (1 + \eta) = l + \eta$$

and the probability will be:

$$\frac{1}{l + 1}$$

Hence, its expected cost is at least:

$$\sum_{l=1}^{k+1} (l + \eta) \cdot \frac{1}{l + 1} = \frac{K + 2}{2} + \eta$$

As we know, the expected optimal offline cost is  $1 + \eta$ , so the competitive ratio is

$$\frac{K + 2}{2}.$$

■

**Lemma 9** *If we have less travelers than blocked edges ( $K > L$ ), there is no randomized online algorithm with competitive ratio less than  $K + 2 - L + \frac{L(L-1)}{2(K+1)}$ .*

**Proof.** Similar with the proof of Lemma 8, if the algorithm is successful on its  $(l + 1)^{\text{th}}$  try, the probability will be:

$$\frac{1}{l + 1}$$

But to calculus the cost, we have to consider two possibilities:

if  $l < L$ , which means we have find all the blocked edges without letting any traveler to return to  $s$ , the cost will be

$$(l - 1) \cdot 1 + 1 + \eta = l + \eta;$$

if  $l > L$ , it means that we have not find all the blocked edges after letting all the travelers depart  $s$ , so we have to let  $(l - L)$  travelers return to  $s$  and restart the process. In this case, there will

be  $L$  travelers with cost 1,  $(l - L - 1)$  travelers with cost 2 and the last traveler with cost  $(2 + \eta)$ . So the total cost will be:

$$L \cdot 1 + (l - L - 1) \cdot 2 + 2 + \eta = 2l - L + \eta.$$

So, its expected cost is at least:

$$\sum_{l=1}^L (l + \eta) \cdot \frac{1}{l+1} + \sum_{L+1}^{K+1} (2l - L + \eta) \cdot \frac{1}{l+1} = K + 2 - L + \frac{L(L-1)}{2(K+1)} + \eta$$

As we know, the expected optimal off-line cost is  $1 + \eta$ , so the competitive ratio is

$$K + 2 - L + \frac{L(L-1)}{2(K+1)}$$

■

If we consider the special case  $L = 1$  which means only one traveler, the competitive ratio becomes:

$$K + 2 - L + \frac{L(L-1)}{2(K+1)} = K + 2 - 1 = K + 1$$

This result corresponds to the randomized case of Bender et al. [2].

## 4.2 No communication

In the last part, we have discussed randomized online algorithms with total communication, where all the travelers can both send and receive messages. Now we will begin to study randomized cases without communication.

**Lemma 10** *Considering the distance taken by all travelers after the last arrived (distance between  $T_{fd}$  and  $T_{la}$ ), there is no randomized online algorithm with competitive ratio less than  $L(K+1)$ .*

**Proof.** According to the result of Bender et al. [2], in randomized case, the competitive ratio for only one traveler is  $K+1$ . When there is no communication between the travelers, all the travelers make choices independently. In this case, all the  $L$  travelers have to reach the target, so the obviously the result becomes  $L$  times  $K+1$ . ■

**Lemma 11** *Considering the distance taken by all travelers before the first arrived (distance between  $T_{fd}$  and  $T_{fa}$ ), there is no randomized online algorithm with competitive ratio less than:*

$$\sum_{l=1}^{K+1} \frac{(K-l+2)^L - (K-l+1)^L}{(K+1)^L} \cdot \left(\frac{2l-2}{L} + 1\right) \quad (1)$$

In this case, we have made an assumption which is different from the cases with complete communication: in the beginning, instead of setting off one by one, all the travelers set off together, which is more meaningful in the real world. So when the first travelers reached the target, the travelers, who had set off together with these lucky travelers but were blocked, are on the way back to the source  $s$ .

**Proof.**

The probability that the algorithm is successful on its 1<sup>th</sup> try is:  $1 - (\frac{K}{K+1})^L$ ;

The probability that the algorithm is successful on its  $2^{\text{th}}$  try is:  $(\frac{K}{K+1})^L \cdot [1 - (\frac{K-1}{K})^L]$ ;  
The probability that the algorithm is successful on its  $l^{\text{th}}$  try is:

$$(\frac{K}{K+1})^L \cdot (\frac{K-1}{K})^L \dots (\frac{K-l+2}{K-l+3})^L \cdot [1 - (\frac{K-l+1}{K-l+2})^L] = \frac{(K-l+2)^L - (K-l+1)^L}{(K+1)^L}$$

And according to the assumption, the cost will be:  $(1 + \eta) \cdot L + 2(l - 1)$ , so the total expected cost is:

$$\sum_{l=1}^{K+1} \frac{(K-l+2)^L - (K-l+1)^L}{(K+1)^L} \cdot ((1 + \eta) \cdot L + 2(l - 1))$$

As the expected optimal off-line cost is  $L(1 + \eta)$ , so we can find the result (1). ■

### 4.3 Conclusion for randomized approaches

	$P_0$	$P_2$	
		$K < L$	$K \geq L$
between $T_{fd}$ and $T_{fa}$	$\sum_{l=1}^{K+1} \frac{(K-l+2)^L - (K-l+1)^L}{(K+1)^L} \cdot (\frac{2l-2}{L} + 1)$	$\frac{K+2}{2}$	$K + 2 - L + \frac{L(L-1)}{2(K+1)}$
between $T_{fd}$ and $T_{la}$	$L(K+1)$		

When we study the distances between  $T_{fd}$  and  $T_{fa}$ , we have to notice that there is no meaning to compare the results of  $P_0$  and  $P_2$ , because these two cases are based on different assumptions. For  $P_0$ , travelers sets off one by one while in  $P_2$  all the travelers set off together.

Furthermore, in order to have a better view of the long formula for  $T_{fd}$  and  $T_{fa}$  with  $P_0$ , we can look into [Figure 2].

We can see from [Figure 2a] that if we fix  $K$ , the performance improves quickly with the growth of  $L$ , and tends to optimal. And from [Figure 2b] we know that if we set  $L=1$  for just one traveler, the result will be  $K + 1$ , which correspond with the result of Bender et al. [2]; when there are multiple travellers ( $L=20$  for example), the performance depends on the number of blockages, and usually when the number of blockages is less than or similar with the number of travellers, the performance is nearly optimal. But we also find that with the increase of  $K$ , the performance begin to drop faster and faster. So when  $K$  reaches a number which is big enough, the drop of the performance will no longer be negligible.

## 5 Conclusion

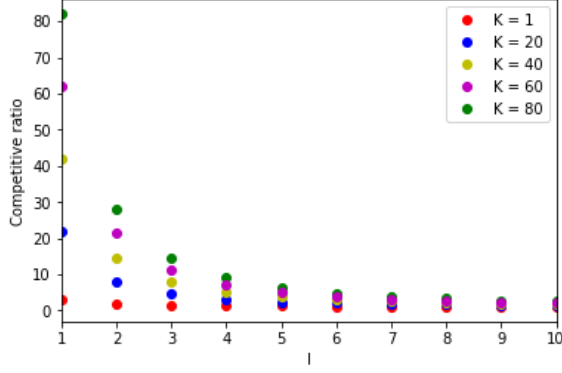
In this article, we have discussed the cost of distance for  $K$ -Canadian problem with deterministic approach and randomized approach, on considering different modes of communication.

The numerical results could be useful in many applications. For example, by studying the distance between  $T_{fd}$  and  $T_{fa}$ , we can judge whether a team of travellers can be more efficient in finding the target:

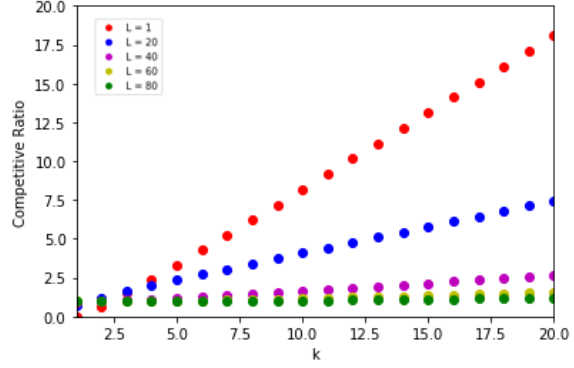
(1) With deterministic approach, the cost is  $2K + 1$  with one traveller, so it makes no difference to use multiple travellers when there is no communication ( $P_0$ ). However, with partial communication or complete communication, the more travellers we have, the better result we will get. Because the formula of the cost is  $2(K + 1) - L$  which concerns about the number of travellers, so if we want this result less than  $2K + 1$ , we should have  $L > 1$ .

(2) With randomized approach, the cost becomes  $K + 1$  for only one traveller. When there is no communication, we can compare the formula with  $K + 1$ . Unless  $K$  is enormously big,

CEI/courbeNewKrandom.png

(a) Competitive ratio for  $K$  set to 1, 20, 40, 60 in randomized approach

CEI/courbeNewLrandom.png

(b) Competitive ratio for  $L$  set to 1, 20, 40, 60 in randomized approachFigure 2: Competitive ratio for  $K$  and  $L$  set to different values in randomized approach

multiple travellers will be a better choice. With complete communication, if we want the result better than  $K + 1$  which is the case of one traveller, we should have the following condition:

$$(K + 1) - (K + 2 - L + \frac{L(L - 1)}{2(K + 1)}) > 0$$

so when  $L > 1$  and  $2K + 2 > L$ , multiple travellers can be better than only one traveller.

For future work, we recommend to explore the possibility of establishing a bi-criteria that would include both distance and time. This would expand the direct applications of the K-Canadian travellers problem for a fleet of vehicles that wants to reach destination quickly without spending too much on gas. Moreover, we recommend to apply the different communication systems (especially P0 et P1) with a time criteria. This would complete the previous work done by Zhang [8].

## References

- [1] Amotz Bar-Noy and Baruch Schieber. The Canadian Traveller Problem. In *Proc. of ACM/SIAM SODA*, pages 261–270, 1991.
- [2] Marco Bender and Stephan Westphal. An optimal randomized online algorithm for the k-Canadian Traveller Problem on node-disjoint paths. *J. Comb. Optim.*, 30(1):87–96, 2015.
- [3] Allan Borodin and Ran El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, 1998.
- [4] Erik D. Demaine, Yamming Huang, Chung-Shou Liao, and Kunihiro Sadakane. Canadians Should Travel Randomly. *Proc. of ICALP*, pages 380–391, 2014.
- [5] Christos Papadimitriou and Mihalis Yannakakis. Shortest paths without a map. *Theor. Comput. Sci.*, 84(1):127–150, 1991.
- [6] Stephan Westphal. A note on the k-Canadian Traveller Problem. *Inform. Proces. Lett.*, 106(3):87–89, 2008.

- [7] Yinfeng Xu, Maolin Hu, Bing Su, Binhai Zhu, and Zhijun Zhu. The Canadian traveller problem and its competitive analysis. *J. Comb. Optim.*, 18(2):195–205, 2009.
- [8] Huili Zhang, Yinfeng Xu, and Lan Qin. The k-Canadian Travelers Problem with communication. *J. of Comb. Opt.*, pages 251–265, 2011.