

## DIRECT OBSERVATION OF REROUTING PHENOMENA IN TRAFFIC NETWORKS

Rafał Kucharski<sup>1</sup>, Guido Gentile<sup>2</sup>

<sup>1</sup> Cracow University of Technology, Department of Transportation Systems, Cracow, Poland  
e-mail: rkucharski@pk.edu.pl

<sup>2</sup>DICEA, Sapienza University of Rome, Rome, Italy

---

**Abstract:** *In this paper we propose how available dataset can be used to estimate rerouting phenomena in traffic networks. We show how to look at set of paths observed during unexpected events to understand the rerouting phenomena. We use the information comply model [1] and propose its estimation method. We propose the likelihood formula and show how the theoretical and observed rerouting probabilities can be obtained. We conclude with illustrative example showing how a single observed path can be processes and what information it provides. Contrary to parallel paper [2] where rerouting phenomena is estimated using real traffic flow measures from Warsaw, here we use only synthetic data.*

*The paper is organized as follows. First we elaborate on rerouting phenomena and define the traffic network, then we summarize the literature behind rerouting phenomena. We follow with a synthetic definition of dynamic traffic assignment needed to introduce ICM model in subsequent section. Based on that introduction we define the observations and propose estimation method based on them followed by illustrative example. Paper is summarized with conclusions and pointing of future directions.*

**Key words:** *dynamic traffic assignment, rerouting phenomena, route choice model, information comply model*

---

### 1. Introduction

By rerouting we mean changing the currently chosen path in road network after either receiving some information or observing consequences of an unexpected traffic event. We broadly define unexpected event as any relevant traffic information that is not known in advance by at least some percentage of drivers and implies changes in the perception of the supply side. This can include: incident, road closure, sport event, demonstration, planned road work, etc. Rerouting phenomena originates in rational process when some reason leads an individual to alternate his plans. In case of rerouting it is either observation or information about unexpected situation in traffic network which leads to changing the path to avoid negative consequences. The process is latent as it takes place in mind of every individual and cannot be observed, however, what can be observed are a) reasons for rerouting b) rerouting itself (changing the path).

We follow definition of rerouting phenomena from information comply model - ICM [1]. This can be seen as an extension of dynamic traffic assignment covering rerouting phenomena. It models the phenomenon through calculating probability of

rerouting  $\alpha$  for given place and time in the network subject to current situation and information. ICM is parameterized by: a) how information reaches drivers, b) how do they react and c) how do they choose their new paths, which all influence the resulting probability. Such detailed representation yields realistic results as it covers cognitive process of rerouting, unfortunately it is hard to estimate and validate. Fortunately, thanks to recent developments, revealed preference (RP) data of how individuals travel through the network in time became more available. As we show in the parallel article [2] rerouting can be also (to some extent) indirectly observed by looking at the flows in the network. Here we focus on how the rerouting behavior can be estimated with set of paths observed during unexpected events.

In this paper we follow classical definition of road network represented by means of an oriented graph  $G(N, A)$ , where  $N$  is the set of nodes and  $A \subseteq N \times N$  is the set of arcs. Each arc  $a \in A$  is described through a vector of characteristics  $\delta_a(\tau)$  that allow to represent its performances (travel times  $t_a(\tau)$  and cost  $c_a(\tau)$ ). The initial node of the generic arc  $a \in A$  is referred to as tail and denoted  $a^- \in N$ , while the final node is referred to as head and denoted  $a^+ \in N$ .

The set of arcs with tail at the generic node  $i \in N$  is referred to as forward star and denoted  $i^+ = \{a \in A: a^- = i\}$ . Symmetrically, the set of arcs entering node  $i \in N$  is referred to as backward star and denoted  $i^- = \{a \in A: a^+ = i\}$ . We define generic path  $k$  in the network  $G$  as an ordered set of arcs  $a \in A$  and nodes  $i \in A$  reached at time  $t_a^k$ . Generic path  $k$  connects origin node  $o \in N$  with destination node  $d \in N$ .

## 2. Literature review

Valuable starting point to define rerouting phenomena is a comprehensive literature review on route-choice models by [3], where general definitions of routing behavior are summarized. Comprehensive definition of the rerouting problem can be found in [4] which clearly highlights the dynamic nature of rerouting phenomena and places it in traffic assignment.

We follow here classic definition of traffic assignment (TA) in terms of econometric equilibrium [5], and ICM model used in this paper follows fundamental traffic assignment assumptions, most importantly expected-utility concept [6] as an explanation of routing behavior. As the rerouting phenomena is strictly dynamic (takes place in space and time), we consider dynamic version of user equilibrium - DUE [7][8]. In fact we work with the stochastic macroscopic DTA model proposed in [9] with sequential route choices [10] at the demand side and General Link Transmission Model [11] on the supply side. The most up-to-date implementation of this framework can be found in [12].

[13] extended typical day-to-day TA model and allowed individuals for en-route rerouting. In their concept rerouting is driven by experienced delay while traversing the network. They model rerouting through agent-based simulation where agents everyday update knowledge on how perturbation on single arc influences the remainder of the network – this knowledge is then stored in correlation matrices. It is assumed that experiencing delay while travelling can lead to assumption that downstream arcs are also perturbed, which can, in turn, lead to rerouting.

While Snowden assumed rerouting is driven only by experience and actual observation [14] assumed that user have access to perfect knowledge about actual state of the whole network. They assume that

rerouting take place if the possible gains of changing the route are big, moreover it takes place only if the difference, both relative and absolute, is greater than so-called ‘indifference-band’. Mahmassani propose term ‘schedule-day’ for what we call here typical day.

Both Snowden and Mahmassani implicitly follow one of most valuable concepts for rerouting, namely the hybrid-model formulated by [15]. Hybrid-model originates from strong distinction between pre-trip and en-route route choice and ascertainment that rerouting takes place as a mixture of them, namely ‘hybrid-routing’: following pre-trip chosen route until there’s a good reason to deviate from it and follow new route up to destination. Hybrid-model addresses rerouting with a sequential procedure executed at each decision point using utility of rerouting. The utility of rerouting is function of: travel times, elasticity parameter (showing how easy drivers will deviate from pre-trip choices, handled in ICM within compliance model). Utility of rerouting is calculated strictly subject to destination. The final result of hybrid model is the new, recalculated route-choice probability at each node and is analogical to probability of rerouting proposed in ICM. Furthermore [15] provide valuable considerations about information and define it as a function which transform actual costs into perceived costs used by individuals for routing.

[4] defines rerouting problem while analyzing within-day re-planning of agents activity plans due to events, which can include i.e. altering destination, later departure time or even cancellation for non-obligatory trips. Within-day re-planning is based on strong distinction between iterative process adequate for recurrent conditions and single-shot simulation for special cases (i.e. re-planning, rerouting, evacuation), explaining why classic TA methods fail in modelling reaction of drivers to exceptional events. The agent structure proposed by Dobler uses the BDI structure (beliefs, desires, intentions, as defined in [16]) which is fruitful to define rerouting where beliefs are perception of the network states, desires are to avoid negative impact and finally intentions are to reroute or not. This BDI structure was origin of ICM structure of information, observation and compliance models. Dobler proposes the Rayleigh distribution [17] to define the information process.

Notable stream of papers by Gao, Frejinger and Ben-Akiva addressed rerouting phenomena through adaptive routing policies.[18] opposed routing policies to adaptive route choices and defined it as the mapping of the network ‘support-points’ (realization of travel times) to routing decisions. The realization of policy is a single path which can be adapted due to actual network conditions (the general concept is similar to hyperpath [19] and routing strategy [20]). Routing policy is built on assumption that probability of each ‘support point’ of the network is known. When driver observes actual state of the network (realization of travel times) he updates probability of downstream arcs travel times. So that if new paths become optimal the rerouting is observed. It is assumed that drivers use their experience to update policies, while here we focus on unexpected, non-recurrent events.

The information spread model of ICM originated from evacuation models [21] where information about event spreads in time reaching more people. Fortunately recent data from Twitter [22] led to way better understanding of information spread processes. Numerous researches were conducted using Twitter data (i.e. [23][24][25]) and provided valuable information on two aspects: dynamics and range of the information spread process.

The dynamics are observable through the ‘tweets’ posted after emergency situations: earthquake, hurricane, riots, etc. [26] made an outstanding research on how fast the information dissipates through the communication network of Twitter. They investigated the twitter traffic related to the false news (i.e. “Rioters released wild animals from London ZOO”) and showed how the society reacts, believe and deny the hoaxes. The analysis was made time dependent so that speed of the twitter is observable. This is probably the most evident and precise feedback on what is the process of informing in time. Several examples analyzed by Procter showed similar properties: The shape of the information spread curves resembles what researchers usually adopt while talking about information spread ([21], [27], [28]) –Rayleigh-like distribution (or any of similar shape). Twitter research revealed other important phenomena of information: virality. [29] showed that information in the communication network is either completely negligible and forgotten very fast, or completely opposite: it spreads like a virus through

communication network, reposted forward with exponential probability [25]. The observations on virality led us to parameterize both information spread speed and range with significance of the information (measured with the total delay it causes).

Finally we refer to stated- and revealed-preference studies of rerouting. [30] conducted a survey to see what is the impact of VMS and radio broadcasts on route choice. 60% respondents claimed that their route-choice is influenced by radio broadcasts, and 40% by VMS signs. [31] got much less optimistic results with the revealed-preference data from floating mobile data. Analysis on how VMS information affects route-choices showed ~30% of compliance to exact guidance provided through VMS. Thanks to recent solutions detailed data on paths in traffic networks is available and working on path data in various studies becomes more accessible. Paths are observed through GPS tracks are more broadly used in number of researches, i.e. to obtain OD matrices[32], to detect modes [33] and trip purposes ([34], to define trip generation [35].

### 3. Dynamic Traffic Assignment

Before we can introduce the ICM model we need to provide brief introduction to DTA. In general, DTA determines the traffic flows on the network satisfying the demand [36]. It is done through an assignment methods, typically following the ‘user-equilibrium’ concept of balancing travel costs of all drivers [5]. In dynamic context ‘user-equilibrium’ becomes a dynamic user equilibrium (DUE), defined as a traffic pattern at which no driver finds it convenient to (unilaterally) change his/her route and departure time (see i.e. [37]). DUE is obtained through an iterative process where at each iteration route choices are adjusted based on outcomes of decisions made in previous iterations. The process is converging to a fixed-point, where demand and supply are stabilized [38]. The results of DTA are the network performances (i.e. temporal profile of travel costs and times) and the demand pattern. Demand pattern of DTA is either set of OD paths defined with specific temporal profile of flows, or, alternatively, [39] set of local routing decisions (arc conditional probabilities) defined for each node which coupled with origin demand becomes equivalent to explicit paths.

The outcomes of DTA are travel costs and flows of the network obtained through equilibrium for typical case; ICM rerouting model provides the same outcomes, yet for atypical situation of unexpected event.

#### 4. Observations

In the remainder of the paper we look at set of paths  $K=\{k\}$  observed during the unexpected event. It can be obtained during a long term study of sample individuals recurrently travelling through the traffic network. If during this long term study it is possible to identify an atypical day resulting from some unexpected event we have satisfactory input for proposed estimation procedure.

For further analysis only a subset of paths which could possibly be affected by this event is needed. For those paths for which it is possible let's determine its typical realization  $\hat{k}$ , i.e. most probable path of this individual connecting the same origin and destination. Unfortunately  $\hat{k}$  can be easily defined only for typical, recurrent trips and for long-term studies.

Using results of DTA, for each arc  $a$  of path  $k$  we can define probability of choosing it while being at its tail node  $a$  at time  $\tau$  subject to travelling destination  $d$ , denoted as  $p_a^d(\tau)$ , further called arc conditional probability.  $p_a^d(\tau)$  is a result of sequential route choice model (RCM) with implicit path enumeration calculated for given set of travel costs  $c$  and times  $t$  [10]. In particular we introduce two RCMs; one calculated with typical travel costs and times:  $\hat{p}_a^d(\tau) = RCM(\hat{c}, \hat{t})$  and one calculated using actual costs and times:  $\tilde{p}_a^d(\tau) = RCM(\tilde{c}, \tilde{t})$ .

By typical costs (denoted throughout the paper with superscript  $\hat{\cdot}$ ) we mean conditions observed during a typical day (when no unexpected events are present), which coincide with costs and times of dynamic user equilibrium and are, at the same time, the conditions expected by individuals to occur when making route choices. Actual conditions (denoted throughout the paper with superscript  $\tilde{\cdot}$ ) in turn are those observed as a consequence of unexpected event. Actual travel times and costs will be used by individual to choose a new path to avoid consequences of unexpected event.

For explicit path  $k$  we can define its probability using arc conditional probabilities from implicit

RCM [9] being product of arc conditional probabilities along the path, as defined in (1).

$$p_k = \prod_{a \in k} p_a^d(t_a) \quad (1)$$

Let's further define for each path the rerouting point  $r$  as the point in time and space where individual reroutes, i.e. where he/she makes a decision to change path to destination to avoid consequences of the event. Technically we define  $r$  as a point at which individual starts using actual travel times and costs for routing, i.e. instead of using  $\hat{p}_a^d(\tau)$  he switches to  $\tilde{p}_a^d(\tau)$  to get to destination from  $r$ . Location of  $r$  is not seen directly when looking at path  $k$ , yet we propose following two schemas to determine it: one more general and another one utilizing typical realization  $\hat{k}$  of rerouting path  $k$ . In general case, using  $r$  as rerouting point we can redefine (1) for the case of rerouting as follows:

$$p_k^r = \prod_{a \in k} p_k^*(t_a) \quad , \quad p_k^*(t_a) = \begin{cases} \hat{p}_a^d(t_a) & \text{if } a < r \\ \tilde{p}_a^d(t_a) & \text{if } a \geq r \end{cases} \quad (2)$$

Where by  $a < r$  we mean subpath between origin  $o$  and rerouting point  $r$  and by  $a \geq r$  we mean subpath between rerouting point  $r$  and destination  $d$ . We define  $r$  as the point for which  $p_k^r$  is maximized:

$$r = \arg \max_{r \in k} \{p_k^r\} \quad (3)$$

From (3) we can express  $r$  as the point at which path  $k$  yields has highest probability (2), in other words  $r$  is chosen so that path  $k$  is consistent with  $\hat{p}_a^d(\tau)$  prior the rerouting and with  $\tilde{p}_a^d(\tau)$  after the rerouting. For boundary cases, if  $r=d$  the rerouting was not observed and whole path is more consistent with  $\hat{p}_a^d(\tau)$  than with any combination through  $p_k^r$ . Respectively if  $r=o$  the actual costs were already considered at the departure, i.e.  $p_k^r$  is maximized for  $\tilde{p}_a^d(\tau)$ .

To improve the procedure we add the following criteria for rerouting point  $r$ :

$$\hat{p}_r^d(t_r^k) \neq \tilde{p}_r^d(t_r^k) \quad (4)$$

$$\alpha_i^d(t_r^k) > 0 \quad (5)$$

Condition (4) guarantees that at the rerouting point actual arc conditional probability differs from

typical one, otherwise it could happen that  $p_k^r$  will be equal for number points before actual rerouting point. Condition (5) guarantees that ICM model (further elaborated below) evaluated at  $r$  yields positive rerouting probability, this way we guarantee the general consistency between observed point  $r$  and results of ICM model. Unfortunately (5) leads to circularity as  $\alpha_i^d(t_r^k)$  is a result of ICM model calibrated with data from paths  $K$  including observed rerouting point of each path calculated with (3). However condition (5) is intended to eliminate big errors (i.e. rerouting point placed before event took place).

Alternatively, for the paths for which we know their typical representation  $\hat{k}$  we can obtain  $r$  as the last node of overlapping part of  $k$  and  $\hat{k}$ . In this case  $r$  can be seen as the point at which individual first acts atypically (i.e. diverged from his typical path  $\hat{k}$ ).

Using (3) we extend data from  $K$  by including rerouting point  $r$  and rerouting time  $t_r^k$  for each of observed path  $k$ . This way we can say that the realized path  $k$  consisted of two parts: prior and after the rerouting decision (if such decision was made).

Let's define observed rerouting probability  $\bar{\alpha}_i^d(t_i^k)$  for each node of the path  $i$  with formula (6). Mind that (6) defines  $\alpha$  for all nodes prior rerouting point  $r$ , but it is not defined for subpath  $a > r$ , after the rerouting point, this is because the decision process is already over at  $r$  and  $\alpha$  has no particular meaning for  $a \geq r$ .

$$\bar{\alpha}_i^d(t_i^k) = \begin{cases} 0 & \text{if } i < r \\ 1 & \text{if } i = r \end{cases} \quad (6)$$

We further define probability of realization for path  $k$  with rerouting path  $\bar{\alpha}_k$  using (6) in a joint probability that rerouting decision was not taken before rerouting point  $r$  and it was taken at point  $r$ . Of course for each observed path  $k$  probability of observing this path equals 1 and is computed as in (7).

$$\bar{\alpha}_k = \prod_{i < r} (1 - \bar{\alpha}_i^d(t_i^k)) \cdot \bar{\alpha}_r^d(t_r^k) \quad (7)$$

## 5. ICM model

This section briefly summarizes information comply model (ICM) proposed in [1]. Here we provide only essential information about ICM model needed to understand how rerouting phenomena is modelled and what is the meaning of parameters.

ICM models probability of individual travelling to destination  $d$  to reroute  $\alpha_i^d(\tau)$  for each point in time  $\tau$  and space  $i$  based on current situation. It is designed to resemble reasoning process made by each individual who is represented through three sub models:

- information model  $t_i^d(\tau)$ , telling if individual has receive information about the event,
- observation model  $o_i^d(\tau)$ , telling if individual has observed atypical situation and linked it with the event, and:
- compliance model  $\kappa_i^d(\tau)$ , telling if individual reroutes to avoid negative impact of the event, each parameterized to fit to actual observed behavior.

For each node  $i$  in the network we can compute result of ICM model using input from DTA, most importantly typical and actual travel times and costs. ICM computes  $\alpha_i^d(\tau)$  as shown in formula (8) which links three submodels of ICM defined through formulas (9) to (11):

$$\alpha_i^d(\tau) = (1 - t_i^d(\tau))(1 - o_i^d(\tau))\kappa_i^d(\tau) \quad (8)$$

$$t_i^d(\tau) = I_0 \cdot \exp\left(-\frac{a_1}{M(\tau)}\right) \left(1 - \exp\left(-\frac{\tau^2}{2(a_2 / M(\tau))^2}\right)\right) \quad (9)$$

$$o_i^d(\tau) = a_3(\Delta t_i)^{a_4} \quad (10)$$

$$\kappa_i^d(\tau) = \left(1 + \exp\left(\frac{-\beta_{\Delta p_{id}} \Delta p_{id}(\tau) + \beta_{\Delta w_{id}} \Delta w_{id}(\tau)}{\eta}\right)\right)^{-1} \quad (11)$$

Submodels of ICM model ((9) to (11)) are defined using following terms:

$M(\tau)$  – global impact of the event calculated as total network delay at time  $\tau$ :

$$M(\tau) = \sum_{a \in A_0} \int_0^\tau \hat{t}_a(\theta) d\theta \Big/ \sum_{a \in A_0} \int_0^\tau \hat{t}_a(\theta) d\theta \quad (12)$$

$\Delta t_i(\tau)$  – cumulated delay at subpath from origin to node  $i$ :

$$\Delta t_i^d(\tau) = \sum_{a \leq i} (\tilde{t}_a(t_i^k) - \hat{t}_a(t_i^k)) \quad (13)$$

$\Delta p_i^d(\tau)$  – cosine similarity between  $\hat{p}_i^d(\tau)$  and  $\tilde{p}_i^d(\tau)$  vectors, showing how the demand pattern differs:

$$\Delta p_i^d(\tau) = \frac{\sum_{a \in i^d} \tilde{p}_a^d(\tau) \cdot \hat{p}_a^d(\tau)}{\sqrt{\sum_{a \in i^d} \tilde{p}_a^d(\tau)^2} \times \sqrt{\sum_{a \in i^d} \hat{p}_a^d(\tau)^2}} \quad (14)$$

$\Delta w_i^d(\tau)$  – relative gap between actual and typical node satisfaction  $w$  (expected cost to get to destination – [44]):

$$\Delta w_i^d(\tau) = \frac{\hat{w}_i^d(\tau)}{\tilde{w}_i^d(\tau)} = \frac{\sum_{a \in i^d} \hat{p}_a^d(\tau) \cdot \tilde{w}_{a^*}^d(\tau)}{\sum_{a \in i^d} \tilde{p}_a^d(\tau) \cdot \tilde{w}_{a^*}^d(\tau)} \quad (15)$$

ICM is parameterized through set of parameters  $\mathbf{a} = \{a_1, \dots, a_4\}$  plus parameters of logit model  $\beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta$  with following meaning:

$a_1, a_2$  – sensitivity of information spread to total severity of the event,  $a_1$  alters the total probability of receiving information and  $a_2$  the pace at which information is spread.

$a_3, a_4$  – probability of *guessing* the event from the total delay.

$\beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta$  – relative weight of  $\Delta p_{id}(\tau)$  and  $\Delta w_{id}(\tau)$  in the utility of rerouting used in binomial

logit (11) with logit parameter  $\eta$

The proposed structure of ICM model is, at this stage, only a hypothesis about actual rerouting behavior of individuals and should be further validated against observations. However based on literature review we can say which parts are more determined than the others. The Rayleigh shape of information spreading process  $t$  is verified through number twitter based observations (most notably [26]). Further manipulations of Rayleigh distribution with  $M(\tau)$  seem to be reasonable through virality analysis of information spread. On contrary functional form of observation model is not justified, although some ATIS [40] and day-to-day analysis [41] suggest that exponential smoothing rule based on cost difference is appropriate. Also the proposed binomial logit is recognized solution for modelling the discrete choices, while difference between costs represented

here with  $\Delta p_{id}(\tau)$  and  $\Delta w_{id}(\tau)$  can resemble the possible gains and losses which link to acknowledged prospect theory describing decision making process.

## 6. Calibrating ICM using observed paths

We propose the following method to estimate ICM model based on the observations of  $K$  paths. We define our estimation process with theoretical values being probabilities of rerouting for each element given by ICM model  $\alpha_i^d(\tau)$  calculated with (8) and realized values being observed rerouting probability  $\bar{\alpha}_i^d(t_a^k)$  calculated with (6).

We propose following optimization program with the main objective to model rerouting probabilities  $\alpha_i^d(\tau)$  matching observed  $\bar{\alpha}_i^d(t_a^k)$ . In general optimization program can be defined as in (16) where  $\|x\|$  is any distance measure. Yet for estimating models with categorical outcomes (in our case it is a dichotomous yes/no model) we follow most of researchers and propose the maximum log-likelihood formula using log of likelihood as shown in (17).

$$L(\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta) = \arg \min_{\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta} \left\{ \sum_{k \in K} \sum_{a \leq r} \left\| \bar{\alpha}_i^d(t_a^k) - \alpha_i^d(t_a^k) \right\| \right\} \quad (16)$$

$$L(\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta) = \sum_{k \in K} \sum_{a \leq r} \left( \bar{\alpha}_i^d(t_a^k) \cdot \ln(\alpha_a^d(t_a)) + (1 - \bar{\alpha}_i^d(t_a^k)) \cdot \ln(1 - \alpha_a^d(t_a)) \right) \quad (17)$$

The log-likelihood  $L$  of (17) is used as an objective function in the optimization problem. The parameters  $\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta$  are estimated by finding the maximum of  $L$ , using some numerical method or software package, i.e. [42].

Mind that the formulation of likelihood will sum over all elements of paths  $k \in K$  prior its rerouting points. And most of observations will appear to be for not rerouting (i.e. having zero value of  $\bar{\alpha}_i^d(t_a^k)$ ), as there are just  $|k|$  rerouting points and much more elements of the paths. This will result in optimization driven by second term of likelihood, and the best fit will be obtained at ICM producing zero probability. Therefore reformulation of (17) is

advocated, also because in likelihood estimation it is assumed that observations are independent. (17) uses each element of subpath  $a \leq r$  as an independent observation, while it is more adequate to treat each path  $k$  as independent observation.

Therefore we propose alternative formulation of  $L$  (17) using  $\bar{\alpha}_k$  as defined in (7). Thus we need to provide equivalent of  $\bar{\alpha}_k$  from ICM model, which can be defined with following:

$$\alpha_k = \prod_{i < r} (1 - \alpha_i^d(t_i^k)) \cdot \alpha_r^d(t_r^k) \tag{18}$$

Based on the above we can redefine  $L$  so that:

$$L(\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta) = \sum_{k \in K} \left( \bar{\alpha}_k \cdot \ln(\alpha_k) + (1 - \bar{\alpha}_k) \cdot \ln(1 - \alpha_k) \right) \tag{19}$$

Which, as we use only observed paths with  $\bar{\alpha}_k = 1$ , simplifies to:

$$L(\mathbf{a}, \beta_{\Delta p_{id}}, \beta_{\Delta w_{id}}, \eta) = \sum_{k \in K} \bar{\alpha}_k \cdot \ln(\alpha_k) \tag{20}$$

Above formulation of log-likelihood function reduces number of observations for calibration, yet it is more consistent with the actual correlations between the observations, which are not captured in (17). For single observed path  $k$  the  $L$  computed both formulations will be the same, however formulation of (20) will yield different structure of optimization program, supposedly more consistent with structure of the problem. Now the log-likelihood maximization can be seen as maximizing the probability of realization of the observed paths produced by ICM through (20). Sample can be extended to cover also paths for which rerouting was not observed (for which  $r$  computed with (3) is

placed at destination) by assuming in ICM model  $\alpha_a^d(\tau) \equiv 1$ .

### 7. Illustrative example

In this section we provide synthetic example with single path  $k$  observed during unexpected event. The path consists of 22 decision points  $i$  reached at respective times  $t_i^k$ . Using DTA we can obtain arc conditional probabilities, both typical  $\hat{p}_a^d(t_i^k)$  and actual  $\tilde{p}_a^d(t_i^k)$  of each arc  $a$  of the path and apply them through (3) to determine rerouting point. In our case the rerouting point was placed at node 19<sup>th</sup>.

The whole example is depicted at fig. 1. For consecutive decision points we can see values of theoretical and observed probability of rerouting  $\alpha$  as well as explanatory variables of ICM:  $\Delta t$  – delay  $\Delta p$ ,  $\Delta w$  – utility of rerouting,  $\iota$  – information spread,  $M$  – significance of the event.

Fig. 1 shows that information spread process increases along CDF of Rayleigh distribution according to (9), the severity of the event  $M$  reaches its maximum at time  $t_{13}^k$  and then decreases.

Individual experienced atypical delay only at nodes 14<sup>th</sup> and 15<sup>th</sup> but it did not make him reroute yet.  $\alpha$  is positive only if there is utility of rerouting, resulting from positive values of  $\Delta p$ , or  $\Delta w$ . That theoretical rerouting probability is positive only for 8<sup>th</sup>, 13<sup>th</sup>, 15<sup>th</sup> and 19<sup>th</sup> decision points with various combinations of  $\Delta p$  and  $\Delta w$ . The subpath from 20<sup>th</sup> to 22<sup>nd</sup> element (subpath after the rerouting decision was made) is not used as it does not bring information.

theoretical and observed rerouting probability and explanatory variables along path  $k$

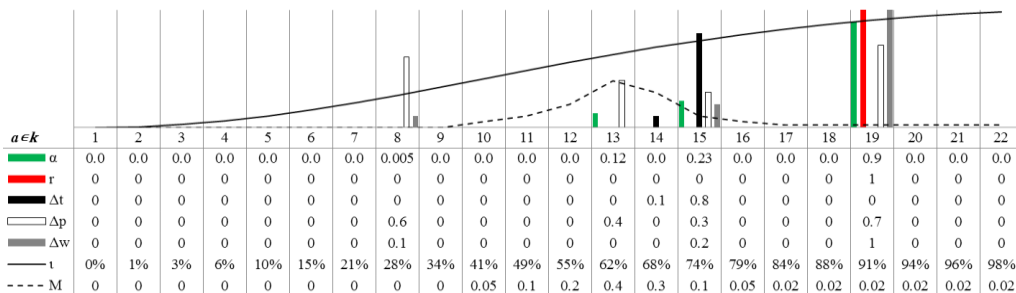


Fig. 1. Illustrative example – observation of single path  $k$

For numerical reasons we cannot estimate the model which produces null  $\alpha$  as  $\log(0)$  does not exist, therefore we transform it to a very small number, i.e.  $10^{-10}$ . The theoretical probability of realizing this path computed with (18), being probability of not rerouting prior 19<sup>th</sup> decision point and rerouting at 19<sup>th</sup> decision point is 0.606. The log-likelihood of the proposed model calculated both with (17) and (20) coincides and equals -0.4995. Model estimation for this single path do not make sense as this single path cannot make a sample, however the applied procedure resulted in  $L \approx 0$  by modifying information spread model  $\iota$  so that it is zero prior  $t_{19}^k$  and quickly grows to 1 for  $t_{19}^k$  which shows on one hand that the model can be estimated, but on the other hand that much bigger sample is needed. Also it shows that additional boundary conditions for parameters should be well thought to fit the reality and keep the cognitive process captured by ICM model and revealed by other researchers.

## 8. Perceived costs

Mind that ICM hides circularity as an individual makes a decision based on  $\alpha$  which is calculated using  $\tilde{p}_a^d(\tau)$  based on actual travel times, which in turn are function of flow inside DTA:  $c_a = f(q_a)$ . While flows  $q$  result from rerouting decisions modelled through  $\alpha$ :  $q_a = f(\alpha)$ . Therefore DTA is usually solved through iterative procedures, i.e. fixed-point problem leading to convergence through iterations. Unfortunately rerouting phenomena does not lead to convergence as drivers do not have perfect knowledge, therefore we propose to execute ICM for given number  $n_{DTA}$  of DTA iterations, which resemble the forecasting capabilities of individuals while deciding to reroute. If  $n_{DTA}=1$  individuals do not consider consequences of other decision makers, if  $n_{DTA} \gg 1$  individuals can fully predict consequences of decisions made by others.  $n_{DTA}$  can be seen as another parameter of ICM to be estimated; unfortunately it cannot be included in likelihood of (19). However estimation of  $n_{DTA}$  should be considered jointly with testing the hypothesis on what costs are used for rerouting part, namely what is  $\tilde{p}_i^d(\tau)$ . We initially assumed perfect knowledge restricted only with number of iterations  $n_{DTA}$ .

While reality can show that either: a) instantaneous actual, b) free flow or c) maximal expected travel times will be utilized. Such hypothesis can be tested by applying any formula generating perceived costs  $c$  as a functional of observed, typical and instantaneous costs. Such formula can be validated through maximizing the consistency of route-choices using the typical route-choice estimation methods (i.e.[43]) coupled with placing the rerouting point  $r$  with formula (3).

## 9. Conclusion

The paper shows how data set available nowadays can be used to estimate rerouting model and to understand the rerouting behavior. Thanks to the method proposed here through statistical tests, i.e. log-likelihood ratio tests it will be possible to understand ICM parameters, most importantly we can estimate significance of information and observation, also we can identify points at which rerouting takes place. The method proposed here can be applied in order to define magnitude of rerouting phenomena and its impact on traffic patterns. It can be then seen if this phenomena is significant enough and if how important it is to cover it in traffic management and ITS solutions. This paper should be further extended with the real dataset of observed paths  $K$  for a city with established DTA real-time model available.

## References

- [1] Kucharski, R., Gentile, G. & Meschini, L. 2014, Information Comply Model – new model to represent rerouting phenomena in Dynamic Traffic Assignment, 5th International Symposium on Dynamic Traffic Assignment, pages 1-32, Salerno, 2014.
- [2] Kucharski, R., Gentile, G., (in press) Indirect observation of rerouting phenomena in traffic networks – case study of Warsaw bridges. Archives of Transport, vol. 32/Issue 4, 2014.
- [3] Gao, S., Chabini, I., (2006) Optimal routing policy problems in stochastic time-dependent networks. Transportation Research Part B: Methodological, 40(2), pp.93–122.
- [4] Dobler C. (2013) Travel behaviour modelling for scenarios with exceptional events - methods and implementations, PhD Dissertation ETH Zurich.



- [5] Wardrop J. G. (1952) Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers*, 1 (3) pp. 325–362.
- [6] Bernoulli D. (1738) Exposition of a new theory on the measurement of risk (*Comentarii Academiae Scientiarum Imperialis Petropolitanae*). *Econometrica* 22, pp.22–36 (translated and reprinted in 1954).
- [7] Merchant D.K., Nemhauser G.C. (1978) A model and an algorithm for the dynamic traffic assignment problems, *Transportation Science* 12, pp.183–199.
- [8] Merchant D.K., Nemhauser G.C. (1978) Optimality conditions for a dynamic traffic assignment model, *Transportation Science* 12, pp.200–217.
- [9] Bellei, G., Gentile, G., Papola, N., (2005). A within-day dynamic traffic assignment model for urban road networks, *Transportation Research Part B* 39 (2005), pp. 1-29.
- [10] Gentile G., Papola A. (2006) An alternative approach to route choice simulation: the sequential models, *Proceedings of the European Transport Conference 2006 – ETC 2006*.
- [11] Gentile G. (2010) The General Link Transmission Model for Dynamic Network Loading and a comparison with the DUE algorithm. *New developments in transport planning: advances in Dynamic Traffic Assignment* (selected papers from the DTA 2008 Conference, Leuven), ed.s L.G.H. Immers, C.M.J. Tampere, F. Viti, *Transport Economics, Management and Policy Series*, Edward Elgar Publishing, MA, USA.
- [12] Gentile G., Kucharski R., Tiddi D. (2013) Splitting rates equilibrium in Dynamic Traffic Assignment using the General Link Transmission Model to simulate network performances, *hEART 2013 - 2nd Symposium of the European Association for Research in Transportation*.
- [13] Snowdon, J., Waterson, B., Fangohr H. (2012) Evolution of adaptive route choice behaviour in drivers. *UTSG: 44th Annual Conference of the Universities' Transport Study Group*, Aberdeen.
- [14] Mahmassani H.S., Hu, T. (1997) Day-to-Day Evolution of network flows under real-time information and reactive signal control, *Transportation Research Part C: Emerging Technologies*, 5, pp.51-69.
- [15] Pel A.J., Bliemer M.C.J., Hoogendoorn, S.P. (2009) Hybrid Route Choice Modeling in Dynamic Traffic Assignment, *Transportation Research Record: Journal of the Transportation Research Board*, 2091(-1), pp.100–107.
- [16] Wooldridge M. (2000) *Reasoning about Rational Agents*, MIT Press, Cambridge.
- [17] Papoulis, A. (1984) *Probability, Random Variables, and Stochastic Processes*, 2nd ed. McGraw-Hill, pp. 104 and 148.
- [18] Gao S., Frejinger E., Ben-Akiva M. (2008) Adaptive route choice models in stochastic time-dependent networks. *Transportation Research Record: Journal of TRB*.
- [19] Trozzi V., Hosseinloo S.H., Gentile G., Bell M.G.H. (2010) Dynamic hyperpaths: the stop model, *Proceedings of Models and Technologies for Intelligent Transportation Systems, International Conference Rome2009*, ed.s G. Fusco, Aracne, 334-343.
- [20] Hamdouch Y., Lawphongpanich, S. (2008) Schedule-based transit assignment model with travel strategies and capacity constraints. *Transportation Research Part B: Methodological*, 42.
- [21] Tweedie, S.W., J. R. Rowland, S. J. Walsh, R. P. Rhoten and P. I. Hagle (1986) A methodology for estimating emergency evacuation times, *The Social Science Journal*, 23 (2) 189–204.
- [22] Milstein S., Chowdhury A., Hochmuth G., Lorica B., Magoulas R. (2008) Twitter and the micro-messaging revolution: Communication, connections, and immediacy. 140 characters at a time. O'Reilly Media.
- [23] Earle P., Bowden D., Guy M. (2012) Twitter earthquake detection: earthquake monitoring in a social world, *Annals of Geophysics*, v. 54/6
- [24] Leskovec J., Adamic L., Huberman, B. (2007) The dynamics of viral marketing, *ACM Transactions on the Web*, 1(1).
- [25] Sakaki T., Okazaki M., Matsuo Y. (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. *WWW '10 Proceedings of the 19th international conference on World wide web*, pp 851-860.

- [26]Procter R., Vis F., Voss A. (2013) Reading the riots on Twitter: methodological innovation for the analysis of big data, *International Journal of Social Research Methodology*, Vol. 16, Iss. 3, 2013.
- [27]Cova T.J., Johnson, J.P. (2003) A network flow model for lane-based evacuation routing. *Transportation Research Part A: Policy and Practice*, 37(7), pp.579–604.
- [28]Lindell M. K. (2008) EMBLEM2: An empirically based large scale evacuation time estimate model, *Transportation Research Part A: Policy and Practice*, 42 (1) 140–15.
- [29]Ghosh R., Kuo T.; Hsu C.; Lin S.; Lerman, K. (2011) Time-Aware Ranking in Dynamic Citation Networks, *Data Mining Workshops (ICDMW), IEEE 11th International Conference on*, pp.373-380.
- [30]Emmerink R. H. M, Nijkamp P., Rietveld P., Van Ommeren J. N. (1996) Variable message signs and radio traffic information: An integrated empirical analysis of drivers' route choice behaviour. *Transportation Research Part A: Policy and Practice*, vol. 30, pp. 135–153.
- [31]Schlaich J. (2010). Analyzing Route Choice Behavior with Mobile Phone Trajectories. *Transportation Research Record: Journal of the Transportation Research Board*, 2157, pp.78–85.
- [32]Friedrich, M., Immisch, K., Jehlicka, P., Otterstätter, T., & Schlaich, J. (2010). Generating Origin-Destination Matrices from Mobile Phone Trajectories. *Transportation Research Record: Journal of the Transportation Research Board*, 2196(-1), 93–101. doi:10.3141/2196-10.
- [33]Gong, H., Chen, C., Bialostozky, E., & Lawson, C. T. (2012). A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems*, 36(2), 131–139. doi:10.1016/j.compenvurbsys.2011.05.003.
- [34]Bohte, W., & Maat, K. (2009). Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*, 17(3), 285–297.
- [35]Choudhury and Iqbal (2014). Developing Trip Generation Models Using Mobile Phone Data, hEART Conference, Leeds, 2014.
- [36]Cascetta E. (2009) *Transportation Systems Analysis*, Springer.
- [37]Friesz, T.L., Kwon, C., Bernstein, D. (2000) Analytical dynamic traffic assignment models, Hensher D., Button K.J. *Handbook of Transport Modeling*, Elsevier.
- [38]Banach S. (1922) Sur les opérations dans les ensembles abstraits et leur application aux équations intégrals, *Fund. Math*, 3133–181
- [39]Meschini, L., Bellei G., Gentile G., Papola N. (1999) An implicit path enumeration model and algorithm for dynamic traffic assignment with congestion spillback.
- [40]Bifulco G. N., Cantarella G. E., Simonelli F., Velonà P. (2013) Equilibrium and Day-to-Day Stability in Traffic Networks under ATIS, *Proceedings of 3rd International Conference on Models and Technologies for Intelligent Transportation Systems*, pp.503-512.
- [41]Watling D. P. (1999) Stability of the stochastic equilibrium assignment problem: a dynamical systems approach, *Transport Research B* 1999;33(4):pp.281–312.
- [42]Bierlaire, M. (2003) BIOGEME: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch.
- [43]Ramming S. (2002) Network knowledge and route choice. PhD thesis, Massachusetts Institute of Technology.
- [44]Dial R. B. (1971) A Probabilistic Multipath Traffic Assignment Algorithm which Obviates Path Enumeration, *Transportation Research* 5, pp.83-111.