Spillback Effects from Traffic Accidents

Jacobien Carstens, Bart Litjens, Verena Schamboeck, Tineke School, Veerle Timmermans, Jacob Turner

Contents

1 Introduction

Understanding how traffic incidents affect the flow of traffic is a problem of great importance. It is of interest to both the general public as well as to many private companies that rely heavily on transportation. Disturbances to traffic come in an enormous number of varieties and are very sensitive to a large number of different situational and environmental factors. A complete understanding is therefore far from feasible at this present time as many individual aspects are still not well described.

In this note, we focus on a very specific problem inside of the general theory. Given a traffic incident that inhibits vehicles to travel freely, as in normal conditions, how does the gradual build up of slow moving vehicles congesting the road behave? There are several basic questions to investigate in relation to this situation:

(a) How quickly does traffic congestion build upstream from the incident given the nature of the disruption?

(b) If this back up progresses all the way back to an intersection, will it cause back up on other roads? (When traffic backs up onto another road, it is called *spillback*)

(c) If this is the case, can the roads that will also suffer significant congestion be predicted?

(d) How many vehicles can be expected to avoid the road on which the disturbance has occurred and which alternative roads will thus see an inflow of traffic as rerouting begins to occur?

(e) How will an individual driver make the decision to wait in traffic versus finding an alternative route?

As evidenced by the large number of complex questions above, even this simple situation is unlikely to have a single consistent pattern. We now explicitly draw attention to two distinct effects that occur when a traffic disturbance is present, as can already be seen in the questions above. Namely:

Effect 1. Those drivers who will remain in the congested area until they can proceed along their originally intended route, contributing to a back up on the affected motorway.

Effect 2. Those drivers who will seek to avoid the affected road altogether and deviate from the initial route onto different routes.

The latter phenomenon is very dynamic and difficult to predict. The former effect can be studied easily with some simplifying assumptions. This effect will be present if there is essentially no choice for the drivers in their route given their origin and destination. For example, one might expect a long stretch of road connecting different cities to be more prone to backup. If such a road becomes affected by a traffic incident, it is not uncommon that any other route linking the two cities will be a significant deviation in time and distance, likely involving travel to a completely different city. This is because of the relative sparsity of roads between cities in contrast to roads within a city area that makes long traffic jams more likely.

However, such intuition may not be reflected in reality and the aforementioned roads need not be the only ones on which drivers will feel that any alternative route would be such a deviation that the only realistic choice is to wait in traffic. Roads with this property will be called *vulnerable*. Leaving out circumstantial causes that may affect the drivers choice, it should be clear that *vulnerability* of a road is a property of the road network itself.

The problem then becomes how to identify such roads and to formulate some measure of road vulnerability. If an incident occurs on a very vulnerable road, then we should expect Effect 2 to be negligible. In this case, understanding how traffic behaves becomes less complex. The follow-up problem is then to describe how traffic behaves in this simpler type of scenario.

This paper is organized as follows. In Section 2 we address the notion of link vulnerability. In order to do so, we first describe the model of the road network that we use in Section 2.1. Then, in Section 2.2, we define several vulnerability measures for roads in the network. Some definitions of road vulnerability have been considered before in Freeman et al. (1991); Jenelius (2009, 2010); Knoop et al. (2008). Their primary focus was on roads on which an incident causes the maximum disruption of traffic in the whole network. Our notion of vulnerability however is orthogonal to the amount of traffic flow on the road. It captures how much choice a driver taking that road has in choosing an alternative route.

Subsequently, in Section 3, we use vulnerability to make some actual predictions. In Section 3.1 vulnerability as well as some additional time-dependent parameters are used to estimate the rate of people rerouting in case of an event on a fixed road. Lastly, in Section 3.2 local order-destination information is used to predict spillback on highly vulnerable roads.

2 Link vulnerability

2.1 The model

We consider the Dutch road network to be a weighted undirected graph $G = (V, E)$, where each edge (or link) represents a part of the motorway and each vertex (or node) represents a junction of motorways. Only the motorways, which in the Netherlands are indicated by the letter A followed by a number, and a few provincial roads, which are important for the global structure of the road network, are taken into account. In this paper, we will refer to the chosen network as the 'motorway network'. A more comprehensive model would also include all provincial and city roads. We assume that at each node one has the possibility to move to any motorway incident with that node. For $e \in E$, let $\ell(e)$ denote the time it takes to travel from one endpoint of e to the other. In this paper these times are computed using Google Maps at a specific time (2pm on a weekday without traffic incidents). A more accurate weight is obtained by averaging over different times on several days. The weighted graph is shown in Figure 1.

The reason for restricting the network to motorways and a few important roads is that we have access to detailed data on the traffic on these roads. There are thousands of sensors throughout this part of the Dutch road network, recording the number of cars passing and their velocity every minute of the day. In Section 3.2 we use this data to analyse how quickly traffic backs up after an incident occurs.

For any path $P \subseteq E$, let $\ell(P)$ be *length* of P, i.e., $\ell(P) = \sum_{e \in P} \ell(e)$. Whenever we speak of a path, it is assumed to be simple, i.e, without repeated edges or vertices. For $i, j \in V$, we define $P(i,j)$ to be the set of paths connecting the vertices i and j. Then we define the length $c(i,j)$ of a *shortest path* between i and j as

$$c(i,j) := \min\{\ell(P) \mid P \in P(i,j)\}.$$

As we are also interested in alternative routes, for any $e \in E$ we furthermore define $c(i,j,e)$ to be the length of the shortest path from i to j in the graph G when the edge e is missing,

$$c(i,j,e) := \min\{\ell(P) \mid P \text{ a path from } i \text{ to } j \text{ in the graph obtained by removing } e\}.$$

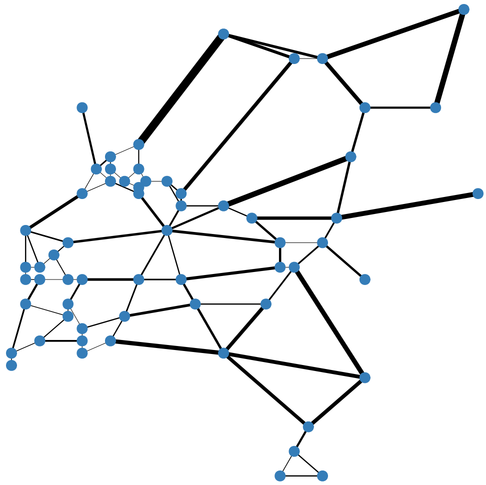

Figure 1: The Dutch motorway network.

Let $i, j \in V$ and $k \in \mathbb{R}_{\geqslant 1}$. We define $P(i,j,k)$ as the set of all paths from i to j whose length is at most k times the length of the shortest path between i and j,

$$P(i,j,k) := \{P \in P(i,j) \mid l(P) \leqslant k \cdot c(i,j)\}. \tag{1}$$

For $e \in E$, we are also interested in the subset of $P(i,j,k)$ consisting of paths that contain e,

$$P(i,j,k,e) := \{P \in P(i,j,k) \mid e \in P\}. \tag{2}$$

We define the set of order-destination pairs that suffer from the fact that the link e becomes inaccessible,

$$S(e) := \{(i,j) \in V^2 \mid \text{ there exists a shortest path from } i \text{ to } j \text{ that contains } e\}. \tag{3}$$

Lastly, whenever we will speak of *free flow* on an edge e, we mean that all lanes at e are open and that the average speed of the cars on e is at least 10 km/hr.

2.2 Vulnerability measures

In this section a vulnerability measure is assigned to each link that indicates whether or not there are good alternative routes available if a link becomes inaccessible. We define these measures to satisfy the following properties:

1. The vulnerability should be a number between 0 and 1. A rate of 0 implies that many alternative routes are available. A rate of 1 implies that no alternative roads are available.

2. The vulnerability can be computed using the network topology, taking into account the travel time on each link assuming free flow. Hence, this rate does not depend on the current traffic situation, and only needs to be computed once using the graph G defined in Section 2.1.

We will define two different vulnerability measures that all satisfy the above properties. For definitions and notation, see Section 2.1. The first measure generalizes the notion of *(edge-)betweenness centrality*, a network theoretic concept that has been formally defined first by Freeman (1977).

Definition 2.1 (Vulnerability measure 1, based on edge-betweenness centrality)**.** Let $e \in E$ and $k \in \mathbb{R}_{\geqslant 1}$. Then we define

$$V_1(k,e) := \frac{1}{|V|(|V|-1)} \sum_{i \neq j \in V} \frac{|P(i,j,k,e)|}{|P(i,j,k)|}, \tag{4}$$

where the sum runs over all distinct vertices i and j.

The factor in front of the sum normalizes the sum of the ratios to ensure the measure V_1 is a number between zero and one. In practice, only the cases $1 \leqslant k \leqslant 2$ are interesting, as we do not expect drivers to reroute if the alternative route would take more than twice as long as usual.

The second vulnerability measure that we define considers drivers that suffer from the closing of link e. We compute the average fraction of time that is lost by closing link e. Note that when e is not on any shortest path, we define $V_2(e) = 0$, as no one suffers from deleting this link. On the other hand, if deleting e would disconnect the network (i.e, if e is a *bridge*), we set $V_2(e) = 1$.

Definition 2.2 (Vulnerability measure 2, based on edge deletion I)**.** Let $e \in E$. Then we define

$$V_2(e) = \begin{cases} 0 & \text{if } e \text{ is not in any shortest path,} \\ 1 & \text{if } e \text{ is a bridge,} \\ \frac{1}{|S(e)|} \sum_{(i,j) \in S(e)} \frac{c(i,j,e) - c(i,j)}{c(i,j,e)} & \text{otherwise.} \end{cases} \tag{5}$$

Note that $V_2(e)$ is well-defined as both $|S(e)|$ and $c(i,j,e)$ are nonzero if e is on a shortest path and not a bridge. The two different vulnerability rates are depicted in Figure 2.

Notice that the vulnerability measures V_1 and V_2 take on completely different values on bridges that are on the fringe of the network. Our current implementation of V_1 is too slow to compute the vulnerability rates of the complete motorway network of the Netherlands. Figure 3 depicts the vulnerability rate V_2 for this network.

One could improve these measures by considering weighted sums, where the weights are defined as the number of times an order-destination pair (i,j) is traveled. One difficulty there is that these order-destination pairs are hard to determine from data. Locally, however, this can be done and this method is exploited in Section 3.2. An easier approach is to define the weights proportionally to the travel time, assuming more people drive shorter routes.

Edge	Unweighted V_1	Unweighted V_2	Weighted V_1	Weighted V_2
$\{1,2\}$	0.104	0.271	0.087	0.043
$\{1,3\}$	0.104	0.500	0.101	0.645
$\{2,3\}$	0.104	0.271	0.106	0.353
$\{2,4\}$	0.250	1.000	0.250	1.000
$\{4,5\}$	0.180	0.327	0.168	0.258
$\{4,7\}$	0.180	0.252	0.164	0.347
$\{5,6\}$	0.180	0.194	0.178	0.218
$\{6,7\}$	0.180	0.172	0.178	0.323
$\{6,8\}$	0.111	1.000	0.111	1.000
$\{7,9\}$	0.111	1.000	0.111	1.000

Table 1: The values of the vulnerability measures V_1 with $k = 2$ and V_2 for the network in Figure 2. The unweighted values are computed for the network with all edge weights equal to 1, the weighted values are computed using the edge weights as shown in Figure 2.

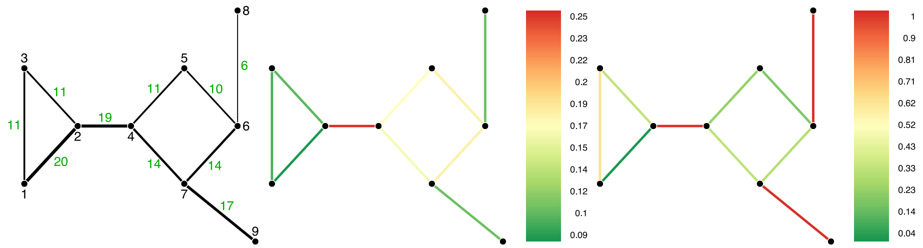

Figure 2: From left to right: weighted example network where edge thickness corresponds to edge weights, the edges of the network are colour coded by the vulnerability rate V_1 for $k = 2$, the edges of the network are colour coded by the vulnerability rate V_2.

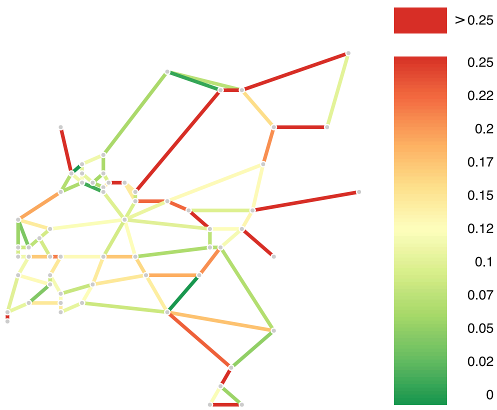

Figure 3: The vulnerability rate V_2 of the motorways network of the Netherlands.

3 Vulnerability in practice

3.1 Predicting rerouting

In this section we estimate the amount of people that will take an alternative route, and the number of commuters that will stick with their initial route in case of a traffic incident. In order to do so, we need more than only the vulnerability measure. The percentage of drivers deviating from their original routes is also affected by the current flow in comparison with the capacity of the link, and the size of the accident (in terms of the number of lanes that are closed).

Let $e \in E$. Then by $f_t(e)$ we denote the number of cars on e that at time t are in free flow. We call $f_t(e)$ the *flow of* e at time t. The maximum number of cars in free flow on e is called the *capacity of* e and is denoted by $c(e)$. By $\text{lanes}(e)$ we denote the number of lanes on e. We write $\text{open}_t(e)$ for the number of open lanes on e at time t.

Using the notions defined above we will now describe a function F that is an estimate of the percentage of people on a road e that will reroute in case at time t an incident happens and causes $\text{lanes}(e) - \text{open}_t(e)$ lanes to close, given a flow that equals $f_t(e)$ at that time. The function F furthermore depends on the capacity and the vulnerability measure. Fix a $k \in \mathbb{R}_{\geqslant 1}$ and set $V_1(e) := V_1(k, e)$ for $e \in E$. Given $e \in E$ and a time t, we first define the function

$$h_t(e) := \begin{cases} 1 \text{ if } f_t(e) \leqslant c(e), \\ 0 \text{ else.} \end{cases}$$

Then we define F as follows

$$F_t(e, i) := \frac{\alpha_1 \cdot V_i(e) + \alpha_2 \cdot (\text{open}_t(e)/\text{lanes}(e)) + \alpha_3 \cdot h_t(e)}{\alpha_1 + \alpha_2 + \alpha_3}, \tag{6}$$

where $i \in \{1, 2, 3\}$ and where $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}_{>0}$. The parameters α_1, α_2 and α_3, which at present do not depend on time, correlate the variables involved in the function F and will need to be determined from the actual data. The function F then returns the rate of people that will stay on their original route. In Section 4 we discuss ways of refining equation (6).

3.2 Predicting spillback

In this section, we assume that we have correctly identified a link as highly vulnerable. How do we expect back-up and spillback to happen? This is the question we now seek to address.

The first problem is to identify how quickly traffic will back up once an incident happens. The detection of a disturbance can occur within a matter of minutes from the induction loop detectors placed on the motorway. A sudden and significant drop in average speed is sufficient for this purpose.

If one plots a graph with axes corresponding to position and time, and each point color-coded based on the average speed of traffic as given by the induction loop detectors, it is a well known empirical phenomenon that traffic incidents cause a parallelogram shape to appear in the colors of low speeds. This tells us that backup on a road accumulates linearly. As such, the rate of backup can be quickly determined using the first few minutes of incoming data after the accident has been detected.

An example of such a parallelogram is shown in Figure 4. The slope of side of the parallelogram running roughly in the vertical direction indicates the rate of backup of the traffic.

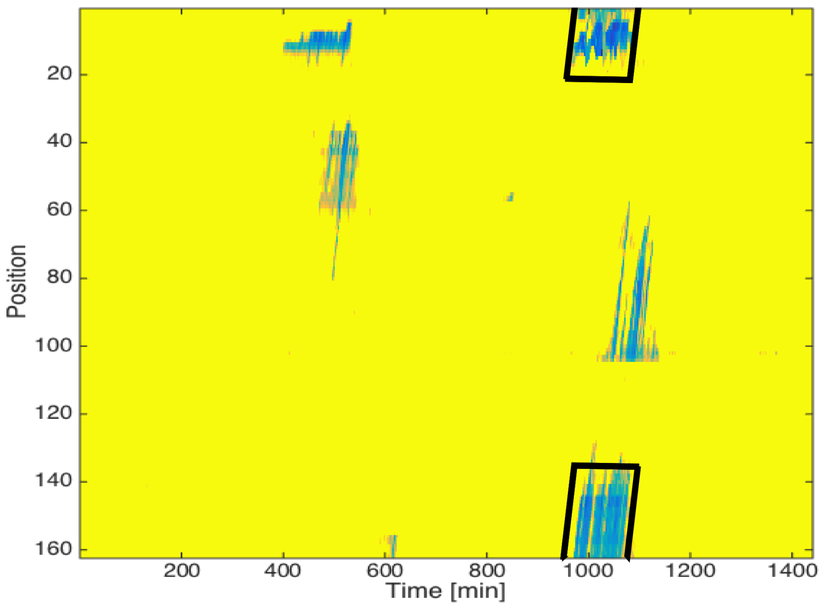

Figure 4: In this picture, the color blue represents a reduced speed. The parallelogram corresponding to the traffic congestion resulting from an incident has been framed. In this particular example, the road in question was a ring and so the graph should be viewed as on a cylinder, hence why the parallelogram is split in the picture.

Of course, it is an altogether different question to try and predict how long such an incident will take place and whether or not backup will reach an intersection. On this point, we make no comment. Instead, let us focus on what we expect to happen in the event that the backup does reach the nearest intersection. To make this prediction, we will define the *local traffic matrix* for an intersection which will depend on empirical data concerning the typical traffic behavior. In fact, one should have many matrices associated to an intersection, one for each time of day/week/year as these conditions can greatly affect what would be considered “normal traffic”.

Now let vertex v denote the intersection in question. Let I be the set of those edges incoming to v and O those that are outgoing. Then the local traffic matrix, $M(v)$, will be a matrix with rows indexed by I and columns indexed by O. For $i \in I$ and $o \in O$, the entry in $M(v)_{i,o}$ will be the average percentage of traffic that turns from road i to road o at v. Such a matrix should be computable given sufficient data and if there are sensors placed on entry and exit ramps between motorways.

Given that traffic on a vulnerable road o backs up to the intersection v, we would expect those roads i such that $M(v)_{i,o}$ is large to also experience congestion. Indeed, one would expext that if $x\%$ of traffic on i turns onto o at v, it will experience $x\%$ of the rate of backup that road o is experiencing.

If the traffic spills back onto another vulnerable road, then after enough time it may reach another intersection and the same method of prediction is possible. However, it seems naive to expect this spillback to continue indefinitely given enough time. One would expect that eventually people would begin canceling trips altogether as news of such a major accident spread. Additionally, if a truly large amount of spillback is occuring, officials may close the road altogether, again forcing people to cancel their trip. These effects would mitigate spillback onto more motorways even if the motorway on which the incident occured was very vulnerable.

4 Discussion

In this section we address the assumptions that were made throughout the paper and discuss ways to verify them from the data. Furthermore, we investigate how to improve the vulnerability measures defined in Section 2.2 and the function (6) defined in Section 3.1.

4.1 Improvements on the vulnerability measures

In Section 2 we defined the graph that represents the Dutch road network. As mentioned there, all motorways are included but the provincial and city roads have not been included. This results in the fact that for instance the motorway $A2$ between Weert and Maastricht (in the southern part of the Netherlands) is a bridge, and therefore maximally vulnerable in our model. However, in practice there is a very good alternative for that piece of road in the event of a traffic accident, namely the $N276$ (this is a provincial road).

In order to account for these kind of alternatives, we strongly recommend to include the provincial and city roads in further research that uses our model. From a graph theoretic point of view, this would make the graph far more complex. The vulnerability measure 2, as defined in equation (5), can still be computed efficiently, as there exists a fast algorithm to compute shortest paths in graphs. The time needed to compute vulnerability measure 1 (see equation (4)), would increase exponentially. However, the vulnerability measure only need to be computed once (for every edge). Therefore, we consider it still worthwile to explore this extended graph.

With respect to the function $F_t(e, i)$, defined in equation (6), computationally nothing changes. The only real-time data it depends on is the number of open lanes and the current flow, both of which can be computed quickly from the data. There are however some refinements that we want to address. For instance, note that in computing $F_t(e,i)$, preferably one would also take the time after the accident in consideration. Drivers will only reroute if they are aware of the accident and if there is still time to take the alternative route. A more sophisticated approach would be to consider a dynamical system in which the value of the function F at a specific time depends on the value F at an earlier time and, in turn, serves as input for computations of F at later times.

Another way to improve upon the function $F_t(e,i)$ is to investigate quadratic or higher-order dependencies. In the current formula, the function depends on $V_i(e)$, $\text{open}_t(e)$ and $h_t(e)$ linearly. This may be a good first-order approximation but higher-order terms certainly will make the function more accurate.

Another potential issue is the disparity between objective understanding of the Dutch road network and the perception of drivers. While the measure of vulnerability should be solely a network measure, its definition fundamentally hinges on the notion of driver choice. As such, there are subjective factors at play and the network that should be measured should be, in some sense, the network as people imagine it, as opposed to how it actually is. If this difference is great, then it seems unlikely to craft a measure simply from geospatial information and a closer investigation of driver behavior will have to be taken into account.

4.2 Implicit assumptions and testable hypthoses

The analysis in Section 3.2 rested on some silent assumptions that should not be simply taken as axioms. We outline these assumptions here as testable hypotheses, to be confirmed or denied using available empiral data.

(a) We have tried to divorce our notion of vulnerability from the amount of traffic flow typical on a given motorway. While it seems clear that the ability to reroute is indeed independent of such considerations, the perceived ability to reroute may not be. It may be that roads most susceptible to back up are very short stretches of road that are very heavily traveled. Even though there may be many alternative routes, the shortness of the stretch of road could make people believe that they can push through in a short amount of time.

(b) We expect that incidents are 1) more common at intersections and 2) the accidents occuring near intersections will cause the greatest amount of spillback because of their proximity to other roads in the network. If this is true, then instead of focusing on the vulnerability of links, it may be more prudent to consider the vulnerability of intersections.

(c) If spillback occurs, how frequently does it occur across two or more upstream intersections? Our hypothesis is that this is an incredibly rare occurence and that after traffic has spilled back across one intersection, the knowledge and increased visibility of the accident will cause significant rerouting, mitigating the upstream backup. If this is the case, it makes predicting spillback much simpler, although the question of rerouting related congestion remains complicated.

(d) Can the severity of spillback be dichotomized according to intercity versus intracity incidents? Or do highly vulnerable roads exist in both situations?

References

L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

L. C. Freeman, S. P. Borgatti, and D. R. White. Centrality in valued graphs: A measure of betweenness based on network flow. *Social networks*, 13(2):141–154, 1991.

E. Jenelius. Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*, 17(3):234–244, 2009.

E. Jenelius. Redundancy importance: Links as rerouting alternatives during road network disruptions. *Procedia Engineering*, 3:129–137, 2010.

V. Knoop, H. van Zuylen, and S. Hoogendoorn. The influence of spillback modelling when assessing consequences of blockings in a road network. *EJTIR*, 4(8):287–300, 2008.