

Semestral work: Checkpoint 0

Ontologies and Semantic Web

Ondřej Borovec

October 2017

1 Motivation

Public transportation system plays an important role in regular city life. According to the DPP¹ more than 1,2 million passengers used Prague metro one day² so the total number of passengers of Prague public transportation system after including also all kinds of surface transportation is much higher. And all these people may be unpleasantly affected by any incident which cause delays because they may miss a meeting, miss their next connection or just stay longer time during a hot day in an overheated bus.

There are many possible causalities which may cause an irregularity of public transportation system, starting with a broken engine or just a driver who overslept. But the main reasons are, in general, related to some traffic accidents, traffic jams or road closures. Out of the last mentioned, you cannot easily predict such situation and if we would try to, it is a task for ML approach instead of ontology. But using an ontology we can put together more information which can help us to understand why such problem happened. In your case we would like to know what are the root reasons of delays in public transportation.

It is possible to find many source of information which are related to particular public transportation system of a city. For this semestral work we decided to focus at Prague and Brno and in section 2 are listed data sources which contain useful information related to this topic. Then in section 4 we describe like such combined information can be used to study the transportation systems or how it can be useful to regular passengers. The last section 5 is dedicated to the state of the art of ontologies related to public transportation and research which has been done in this field.

2 Data sources

In this section, there are listed data sources of information related to Prague public transportation system. Every source has a short description and also

¹Dopravní podnik hlavního města Prahy - <http://www.dpp.cz/>

²<http://www.dpp.cz/prazske-metro-v-den-prepravniho-pruzkumu-prepravilo-1-272-143-cestujicich/>

information content description, which for spacing reason placed to appendix A.

All data sources may be combined using common fields as timestamp and location of an event/incident/update. Out of these common features each of them contains more information which are highlighted in related subsections.

2.1 Dopravniinfo.cz³

This is a project of Czech Department of transportation⁴ to inform about problems related to any kind of transportation in the Czech Republic. It has user friendly web application and also mobile application which is convenient to users. But, more importantly, there is a machine readable feed⁵ for purposes of CTU students which contains the same information as the applications. Every feed has a pre-defined structure in xml format (schema descriptions⁶).

Each record contains only general information about a limitation on Czech roads. Useful information is clear classification of such situations 18 different classes (e.g. accidents, road closure, oversize-load transportation, ...).

2.2 Policie ČR - dopravní nehody⁷

Police of the Czech Republic provides information about every incident that happened on our roads. There is a poor form to filter events which does not show any incidents if the result list is longer then 100 records and also records are in form of pdf so special mining technique is needed. But every pdf contains a lot of information related to traffic incidents. Such information can be used for incidents classifications and severity estimation.

2.3 Prague Open Data⁸

Prague is publishing a lot of data related to daily life. One of such data sources is rss feed related to public transportation system and its singularity⁹, changes¹⁰ and planned changes¹¹. Another important source of information for this work is actual geo of each line of Prague integrated public transportation system¹².

³<http://www.dopravniinfo.cz>

⁴<http://www.rsd.cz/>

⁵[urlhttp://kbss.felk.cvut.cz/dopravni-info.zip](http://kbss.felk.cvut.cz/dopravni-info.zip)

⁶http://registr.dopravniinfo.cz/docs/x-format/cz-ndic_ddr-common-v3.2.5-en-html/format.html

⁷<http://pcr.jdvm.cz/pcr/>

⁸<http://opendata.praha.eu/>

⁹<http://opendata.praha.eu/dataset/dpp-mimoradne-udalosti>

¹⁰<http://opendata.praha.eu/dataset/dpp-zmeny-v-doprave>

¹¹<http://opendata.praha.eu/dataset/ropid-planovane-vyluky>

¹²http://opendata.praha.eu/dataset/ipr-prazska-integrovana-doprava-_trasy_aktualni_stav

2.4 Dopravní podnik města Brna¹³

Information about public transportation irregularities are not so rich as from Prague, but still some are published on this site¹⁴.

3 Integration options

The most straight forward integration option is based on time stamps and location. It is the easiest for combination of data source 2.1 and data source 2.2 where the both source contain the exact timing and location description and also the first mentioned contains a classification. Then we know, that only events classified as a type of accident may have an adequate record in the second source. To use the same features to link these data sources to data source 2.3 about Prague public transportation system we would have to use more complicated solution using standard routes of each line and analyze if an event location lies on that.

But for the purpose of this semestral work we need a different integration features then just time and location. The best candidate is public transportation irregularity classification - both, Prague and Brno, are using similar classification but not exactly the same.

4 Use cases

Information is powerful tool and you can use it in many way. We would like to connect information about traffic accidents with Prague public transportation lines which can be affected by them. After this first step it is possible to link also information about made changes due to such irregularities. As a result we have an information what types of traffic accidents have what kind of disruptive effect on public transportation system.

Below are listed some use cases how could be such information used.

4.1 Common public irregularity information system

When we are able to access all irregularities from one point it is much easier then to remember all the sites when it can be found based on location (city) in which we are.

4.2 Research about reasons of public transportation adjustments

We know if a bus or tram line had to change, but using only data source 2.3 there is not way to find out why that happened. The knowledge what are the

¹³<http://dpmb.cz/cs>

¹⁴<http://dpmb.cz/cs/vsechna-omezeni-dopravy>

reasons for such changed can be useful, but for such research we would need data.

4.3 Delay reasoning

Everybody hates when a public transportation in which he is or which he wants to take has a delay. Mostly we do not know why it faces such complications but for may people it would be easier to accept such situation if they know more information and reason behind that. There may be a mobile application which after searching of a connection can return a reason why it was delayed.

5 State of the art

There is a lot of potential for ontologies in the field of transportation. We can points out a solution from [3] which aim to help users with their travel planning giving them information about special offers, possible combinations of different transport modes and potential advantages on their path. The is also [2] which is focused at ontology to generate personalized user interfaces for transportation interactive systems.

The study about transport disruption ontology [1] is focused at the very similar problematic as this semestral work. As so it will be useful source of inspiration and idea for the best solution.

A Data source content

A.1 Dopravniinfo.cz

- Timing
 - Time of the message generation
 - Start of the event
 - End of the event
- Location
 - Country name
 - Region name, region code
 - Town ship name, town ship code
 - Town name, town code
 - Town district name, town district code
 - Street name, street code
 - Road number, Road class
 - Coordinate system (Altitude, Latitude)
- Description
 - Full text description
 - Free text description
- Event related data (list)
 - Textual description of the Event and its code
 - Textual description of the Update Class of the Event and its code

A.2 Policie ČR - dopravní nehody

- Timing
 - Date of an incident
 - Time of an incident
- Location
 - Road class, road number
 - Road type
- Incident related data
 - Type of collision

- Incident foreplay (cause, alcohol, road surface, visibility, road conditions, wind, other inputs, ...)
- Incident aftermath (deaths, injury, damage in money, leakage of fuel, ...)
- Description (Number of vehicles, driver conditions, driver info, vehicle types, vehicle info, ...)

A.3 Prague Open Data - all options

- Timing
 - Date of publishing
 - Start time
 - End time
 - Final end time
- Location
- Description
 - Affected type of transportation
 - Affected lines
 - Integrated rescue system flag
 - Effect

A.4 Dopravni podnik mesta Brna

- Timing
 - Start time
 - End time
- Location
- Description
 - Affected lines
 - Classification
 - suggested solution
 - Reasoning
 - Delay
 - Dirrection

References

- [1] David Corsar et al. “The transport disruption ontology”. In: *International Semantic Web Conference*. Springer. 2015, pp. 329–336.
- [2] KáThia MarçAl De Oliveira et al. “Transportation ontology definition and application for the content personalization of user interfaces”. In: *Expert Systems with Applications* 40.8 (2013), pp. 3145–3159.
- [3] Mnasser Houda et al. “A public transportation ontology to support user travel planning”. In: *Research Challenges in Information Science (RCIS), 2010 Fourth International Conference on*. IEEE. 2010, pp. 127–136.