From Train Control System Log Files to Train Traffic Summaries

Using anomaly detection and NLG to translate Train Control System log files

TU Delft

CGI

About me

- MSc student in Data Science and Technology
- Finalizing my master thesis at a CGI

CGI

- IT consultancy company
- Transport, Post and Logistics department
- Rail Infra Unit

ProRail

Port of Rotterdam

The project

- ProRail
- Train control system in The Netherlands
- ASTRIS - Log files

Goal

Generating human readable reports about anomalies (unusual behavior) detected in the log files

```
2018-07-02T13:26:34.4012Z|IDCR_ZL01|niemand|Astris08|EMSCorrelationID='geen-
id'EMSBericht='<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<soapenv:Envelope xmlns:soapenv="http://www.w3.org/2003/05/soap-envelope"
xmlns="ESBConsumerEnvelope" xmlns:b="http://docs.oasis-open.org/wsn/b-2"
xmlns:esb="ESBConsumerWSDL" xmlns:wsa="http://www.w3.org/2005/08/addressing" xmlns:wsrf-
bf="http://docs.oasis-open.org/wsrf/bf-2">
<soapenv:Header>
<wsa:To>niemand</wsa:To>
<wsa:From>
<wsa:Address>ABS</wsa:Address>
</wsa:From>
<wsa:ReplyTo>
<wsa:Address>ABS</wsa:Address>
</wsa:ReplyTo>
<wsa:MessageID>1530537994401214</wsa:MessageID>
<wsa:Action>NotifyConsumer</wsa:Action>
</soapenv:Header>
<soapenv:Body>
<b:Notify>
<b:NotificationMessage>
<b:SubscriptionReference>
<wsa:Address>geen-ref</wsa:Address>
</b:SubscriptionReference>
<b:Topic Dialect="Simple">
<esb:topicName>Rijweg</esb:topicName>
</b:Topic>
<b:ProducerReference>
<wsa:Address>ASTRIS-IDCR</wsa:Address>
</b:ProducerReference>
<b:Message>
<esb:msgContainer>
<esb:msgHeader>
<esb:berichtVolgnr>13</esb:berichtVolgnr>
<esb:volgendeIsSnapshot>false</esb:volgendeIsSnapshot>
</esb:msgHeader>
<esb:msgBody>
<esb:ID>SEIN:110:PPLG:ZL:1:SEIN:86:PPLG:ZL:1:LL</esb:ID>
<RIM:Notification xmlns:RIM="RailinfraMelden" xmlns:RIT="RailinfraTypes">
 <RIM:wijzigingRijweg_1>
  <RIT:SafetyCode>
   <RIT:Checksum>2705496246</RIT:Checksum>
   <RIT:Volgnummer>0</RIT:Volgnummer>
   <RIT:Tijdstempel>
    <RIT:Seconden>1530537993</RIT:Seconden>
    <RIT:Microseconden>50473</RIT:Microseconden>
   </RIT:Tijdstempel>
  </RIT:SafetyCode>
  <RIT:RijwegIdentificatie>
```


The following unusual behaviors were detected after analyzing the given log file.

1. Unusual behavior was detected when setting the route : **108-202X-LLR**

One or more elements did not reach the requested position for the route. Consequently, state *Wordt_voorbereid* transitioned back to *Gereserveerd*, instead of transitioning to the expected state *Voorbereid*. The table below shows all the states visited and the attributes that were active in a particular state. Additionally, the time when a state was visited is indicated in the second column.

Message number	Timestamp (Seconds, Microseconds)	Checksum	State	Active attributes
1	1551963071,980126	566029257	Rust	[ ]
2	1551963079,199275	3455454404	Rust	[RR, VBR, IR]
3	1551963079, 213520	2612494063	Gereserveerd	[ ]
4	1551963079, 218601	3528715720	Wordt_voorbereid	[IR]
5	1551963246, 76928	3318898854	Gereserveerd	[IR]
6	1551963250, 37129	3853741571	Gereserveerd	[IR, H]
7	1551963250, 436923	3313027229	Rust	[ ]

2. Unusual behavior was detected when setting the route : **40-40X-L**

BEVNL denied an 'InstellenRijweg' request sent by ASTRIS. Therefore, ASTRIS transitioned from state *Instelopdracht_verstuurd* to *Instelopdracht_afgekurd*, instead of transitioning to state *Wordt_ingesteld*. The table below shows all the states visited and the attributes that were active in a particular state. Additionally, the time when a state was visited is indicated in the second

DataSet

- Log files provided by the company

Methodology

Natural Language Generation

- Content Determination task - **graph based anomaly detection** (State Machine and Adjacency Matrix) - Python
- SimpleNLG - JAVA

My experience

- Enjoyed getting domain knowledge
- The freedom given by the company
- I got to decide the techniques I use
- Got familiarized with being in a big company
- I like my topic - NLG

Questions

Email address : bojana.urumovska@yahoo.com

Thank you for your attention!