



Complex Event Processing Synergies with Predictive Analytics

Introduction

CEP deals with detecting real-time complex events PA deals with different kinds of prediction

Motivation

CEP reaction should be triggered earlier, even before the occurrence of the primary event, events should be predicted.

Solution

- > By the connection of CEP and PA, complex events can be predicted
- > Initial CEP application is represented by the white boxes and ellipses
- > A certain complex event can be detected with the help of CEP
- > This event is called primary complex event (PCE), while the related query is called primary CEP query
- > The CEP engine also produces the predictors (secondary complex events, SCE)
- > The SCEs are transmitted physically by the event handler
- > The PCE is the target of the prediction

- > Training set is extended based on the earlier defined predictors (SCEs) and the prediction target (PCE)
- > Learning model is periodically refreshed based on the currently extended training set
- > PA gives a prediction for the PCE at a later time
- > Prediction is sent back to the CEP engine
- > The CEP engine generates the new (predicted) PCE event

CEP-PA Interaction Sequence

- > How does it work?
- > A snapshot at measurement point 7
- > Training set
 - > is extended with data coming from measurement point 5
 - > predictors are calculated from SCEs on interval 1-5
 - > learning labels are calculated from PCE based on interval 3-7
- > Prediction is
 - > based on the calculation of SCEs on interval 3-7
 - > a forecast to the future point 8 and 9

Proof of Concept

- > Entry system: the CEP engine is listening the traffic data
- > Incomings are greater than 25 during the last 1.5 hours → alert (complex event)
- > The event is predicted 0.5 or 1 hour before it really occurs

```

2005.07.26. 14:00:00,70,11,ms
2005.07.26. 14:00:00,71,11,ms
2005.07.26. 14:00:00,72,11,ms
2005.07.26. 14:00:00,73,11,ms
2005.07.26. 14:00:00,74,10,ms
2005.07.26. 14:00:00,75,10,ms
2005.07.26. 14:00:00,76,10,ms
2005.07.26. 14:00:00,77,10,ms
2005.07.26. 14:00:00,78,10,ms
2005.07.26. 14:00:00,79,10,ms
2005.07.26. 14:00:00,80,10,ms
2005.07.26. 14:00:00,81,10,ms
    
```

```

2005.08.04 7:30:00 - Predicting: no
2005.08.04 8:00:00 - Predicting: no
2005.08.04 8:30:00 - Predicting: no
2005.08.04 9:00:00 - Predicting: yes
2005.08.04 9:30:00 - Predicting: yes
2005.08.04 10:00:00 - PVPNT (in the trainingset this time is shifted back)
2005.08.04 10:30:00 - Predicting: yes
2005.08.04 10:50:00 - EVENT (in the trainingset this time is shifted back)
2005.08.04 11:00:00 - Predicting: yes
2005.08.04 11:30:00 - LVL2-1 (in the trainingset this time is shifted back)
2005.08.04 11:00:00 - Predicting: no
    
```

Results

- > Success = correct prediction in 1 hour
- > 93.15% precision
- > 92.23% recall

Reference

Lajos Jenő Fülöp, Gabriella Tóth, Róbert Rác, János Pánczél, Tamás Gergely, Árpád Beszédés and Lóránt Farkas: Survey on Complex Event Processing and Predictive Analytics. Technical report, 2010
www.inf.u-szeged.hu/~gtoth/research/cep_pa_tech2010.pdf