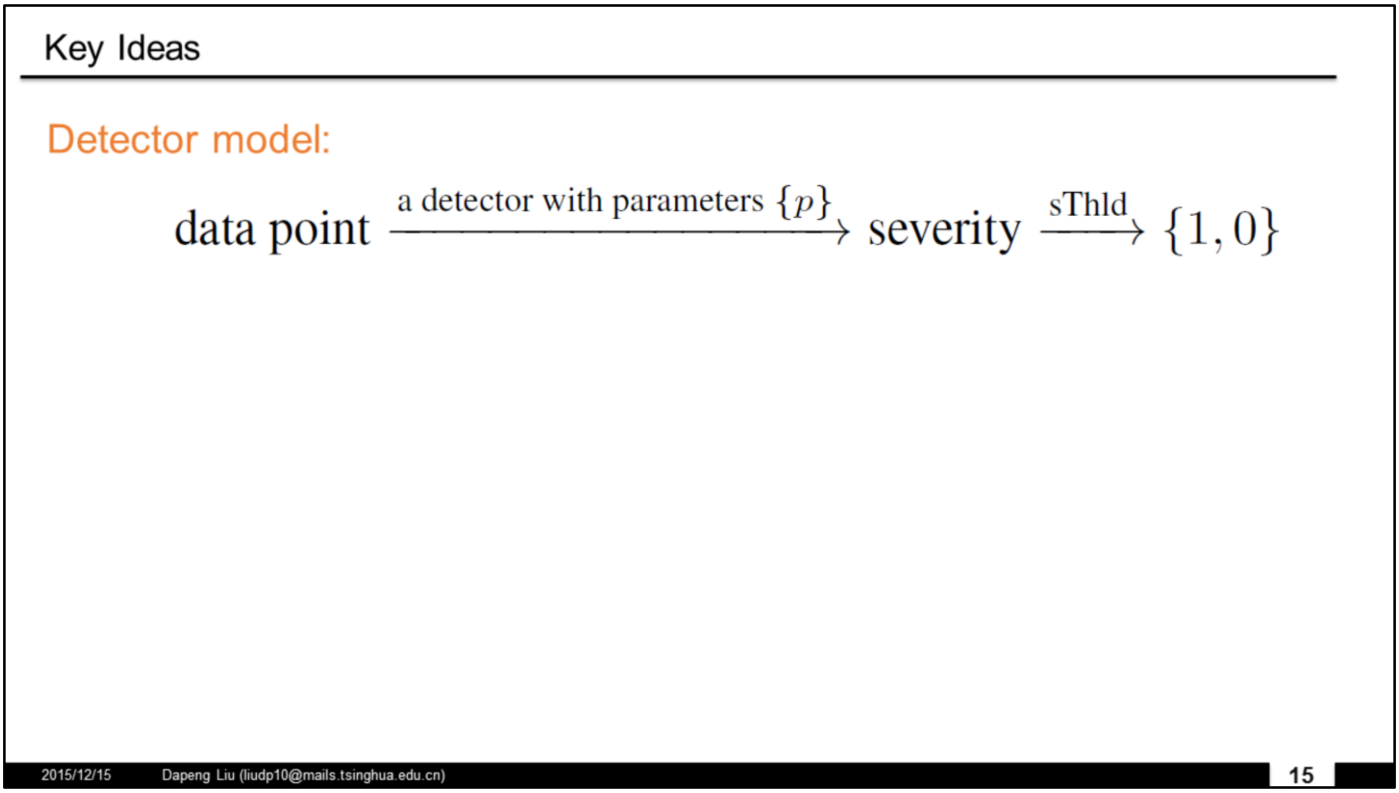


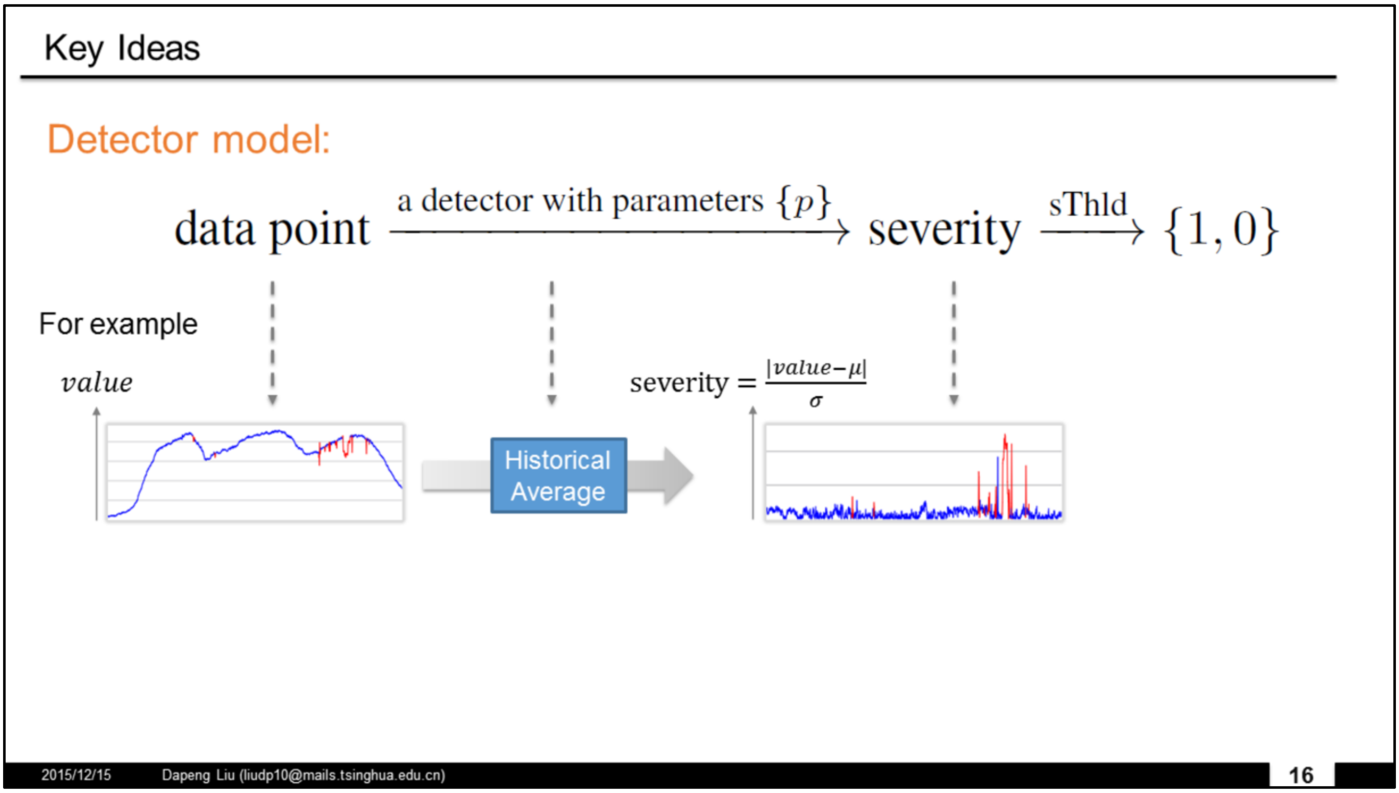
Key Ideas

Detector model:

$$\text{data point} \xrightarrow{\text{a detector with parameters } \{p\}} \text{severity} \xrightarrow{\text{sThld}} \{1,0\}$$

First, we use a uniform model to represent how different detectors work.

It is basically a two-step process.



Key Ideas

Detector model:

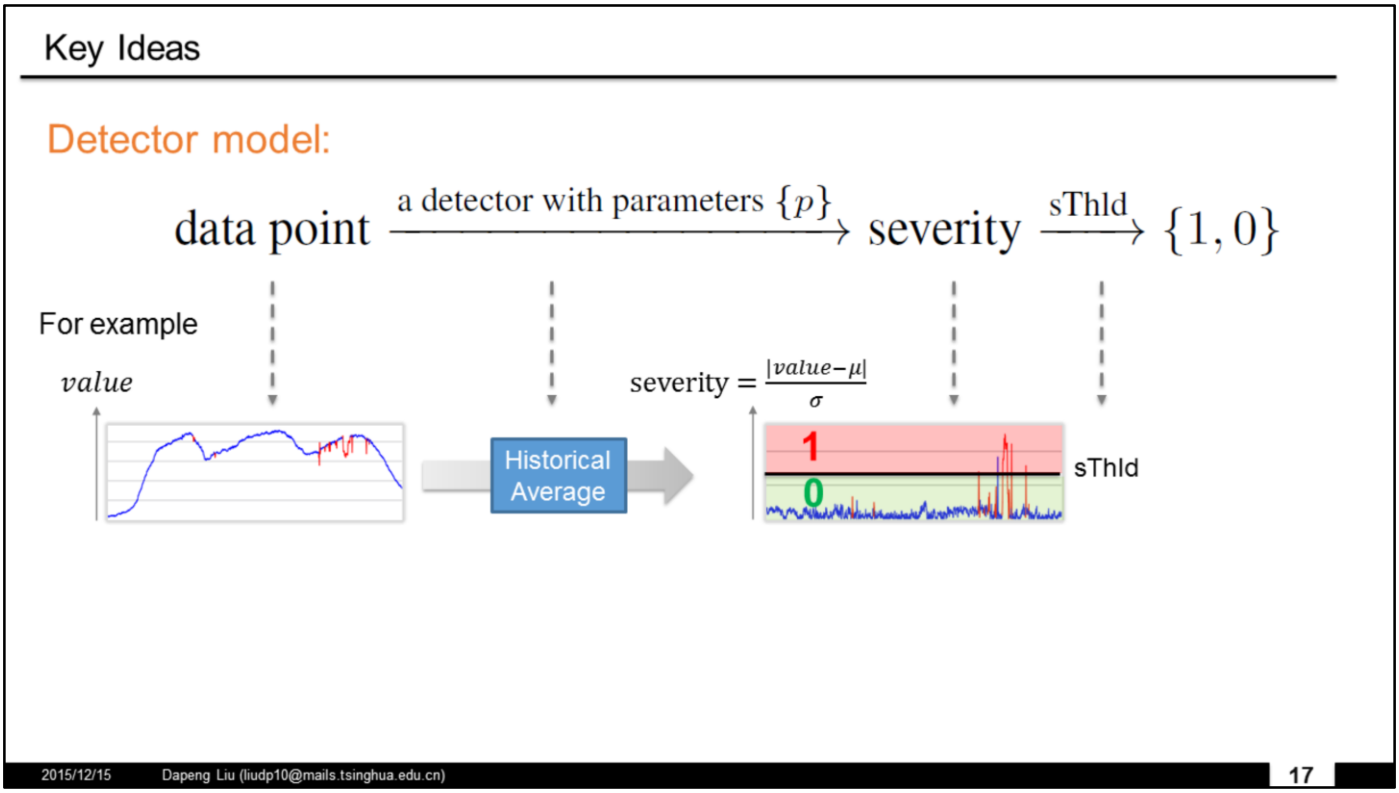
$$\text{data point} \xrightarrow{\text{a detector with parameters } \{p\}} \text{severity} \xrightarrow{\text{sThld}} \{1,0\}$$

For example

$$\text{severity} = \frac{|value - \mu|}{\sigma}$$

For example, this is our KPI data. The red parts are the anomalies labeled by the operators.

** Click
If we apply a detector, historical average for example, to this data, it will generate the anomaly severity for each data point. This detector measure the severity using how many times of the standard deviation each point is away from the average. The higher the severity is, the more anomalous the data point is.



Key Ideas

Detector model:

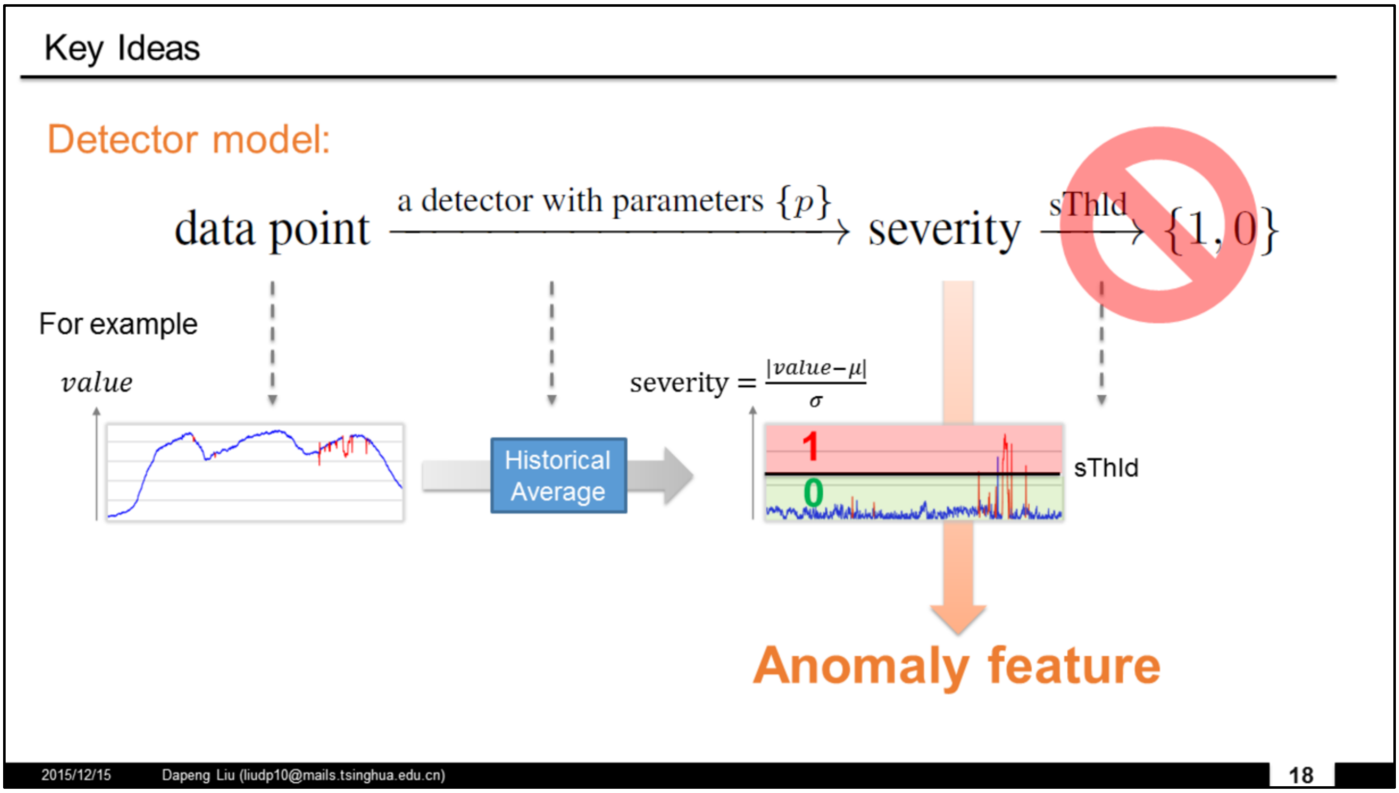
$$\text{data point} \xrightarrow{\text{a detector with parameters } \{p\}} \text{severity} \xrightarrow{\text{sThld}} \{1,0\}$$

For example

$$\text{severity} = \frac{|value - \mu|}{\sigma}$$

Then, we need to set a severity threshold to decide anomalies.

So, this is a general detector model. Different detectors basically work in these two steps, except that they use different techniques or algorithms to measure the severities.



Key Ideas

Detector model:

$$\text{data point} \xrightarrow{\text{a detector with parameters } \{p\}} \text{severity} \xrightarrow{\text{sThld}} \{1, 0\}$$

For example

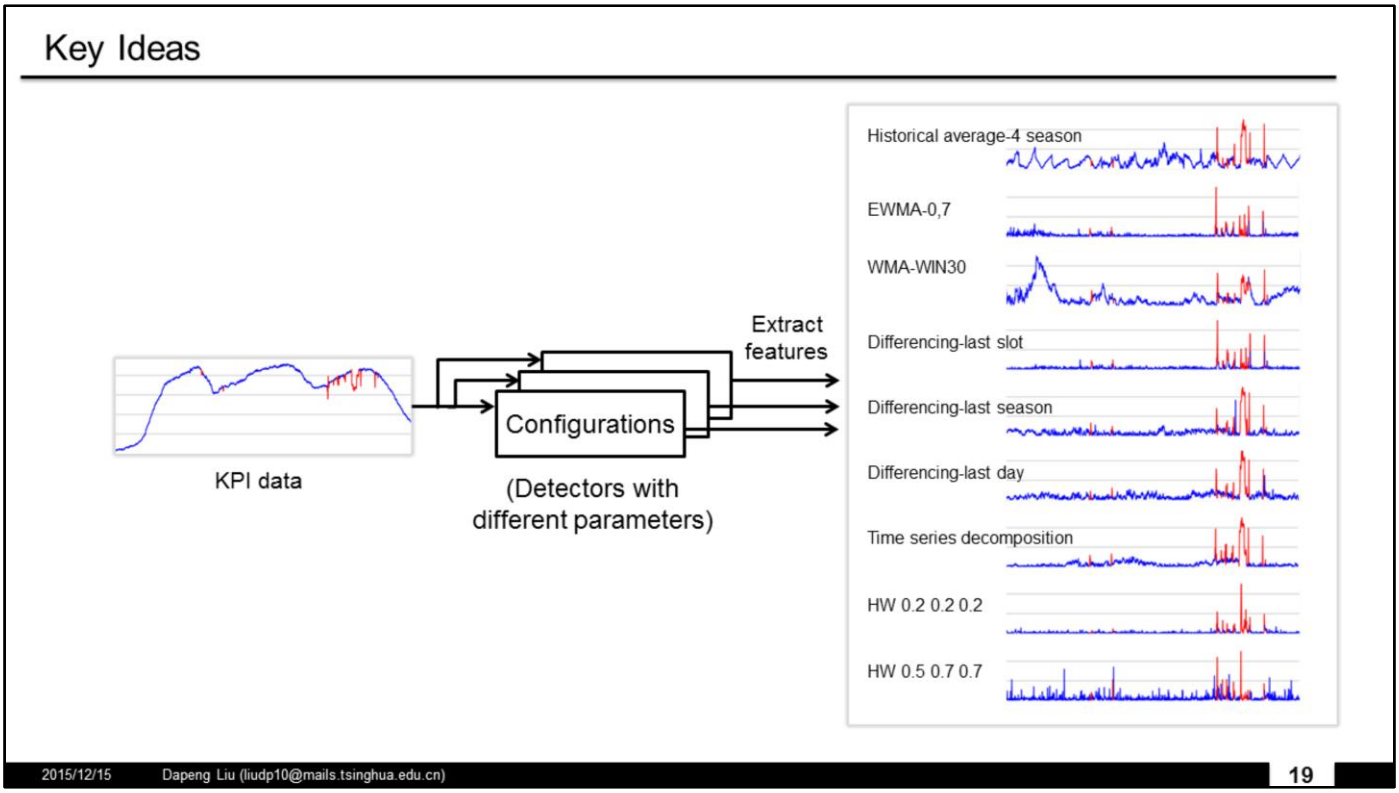
$value$

Historical Average

$$\text{severity} = \frac{|value - \mu|}{\sigma}$$

Anomaly feature

In Opprentice, we do not let each detector determine the anomaly by itself. Instead, we only use the severity as the anomaly feature.



Key Ideas

KPI data

Configurations

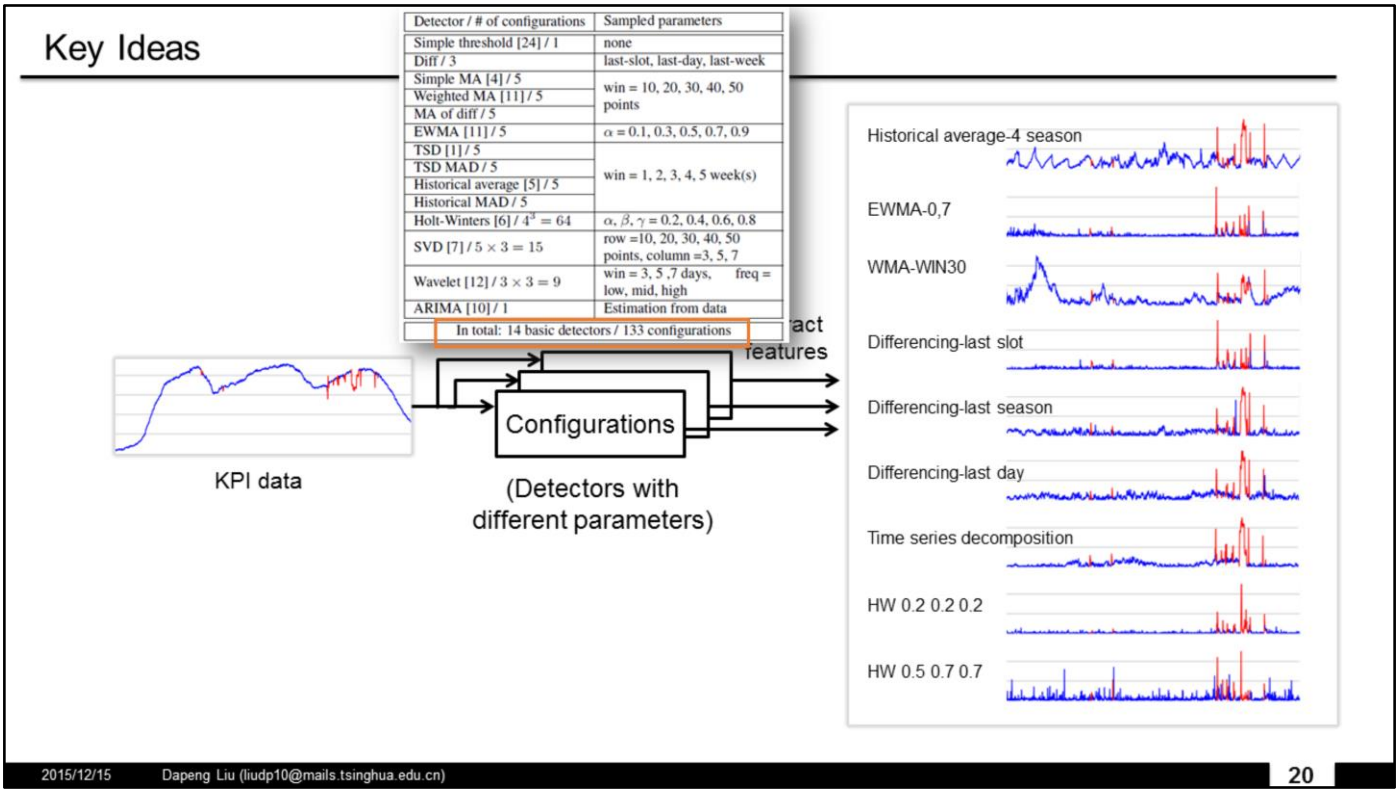
(Detectors with different parameters)

Extract features

2015/12/15 Dapeng Liu (liudp10@mails.tsinghua.edu.cn) 19

Multiple detectors with different parameters are used simultaneously to extract different anomaly features of the data.

We call a combination of a detector and its specific parameters a configuration.



Key Ideas

Detector / # of configurations	Sampled parameters
Simple threshold [24] / 1	none
Diff / 3	last-slot, last-day, last-week
Simple MA [4] / 5	win = 10, 20, 30, 40, 50 points
Weighted MA [11] / 5	
MA of diff / 5	
EWMA [11] / 5	$\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$
TSD [1] / 5	win = 1, 2, 3, 4, 5 week(s)
TSD MAD / 5	
Historical average [5] / 5	
Historical MAD / 5	
Holt-Winters [6] / $4^3 = 64$	$\alpha, \beta, \gamma = 0.2, 0.4, 0.6, 0.8$
SVD [7] / $5 \times 3 = 15$	row =10, 20, 30, 40, 50 points, column =3, 5, 7
Wavelet [12] / $3 \times 3 = 9$	win = 3, 5 ,7 days, freq = low, mid, high
ARIMA [10] / 1	Estimation from data
In total: 14 basic detectors / 133 configurations	

KPI data

Configurations

(Detectors with different parameters)

ract features

We broadly select 14 detectors and sample some parameters. Finally, we get 133 configurations. In other words, we have 133 anomaly features.