

Proactive Network Anomaly Detection Using A Statistical Approach

Amar Agrawal

MTech - 1, KReSIT, IIT Bombay.

Seminar Guide : Prof. Varsha Apte

Dept. of CSE, IIT Bombay

The problem

- The prediction of network anomalies before they cause catastrophic network faults.
- Monitoring the network health continuously, detect and diagnose potential network problems.
- Initiate appropriate recovery actions.

Approaches

- Rule based approaches
- Finite state machines
- Pattern Matching
- Network probes
- Statistical Analysis

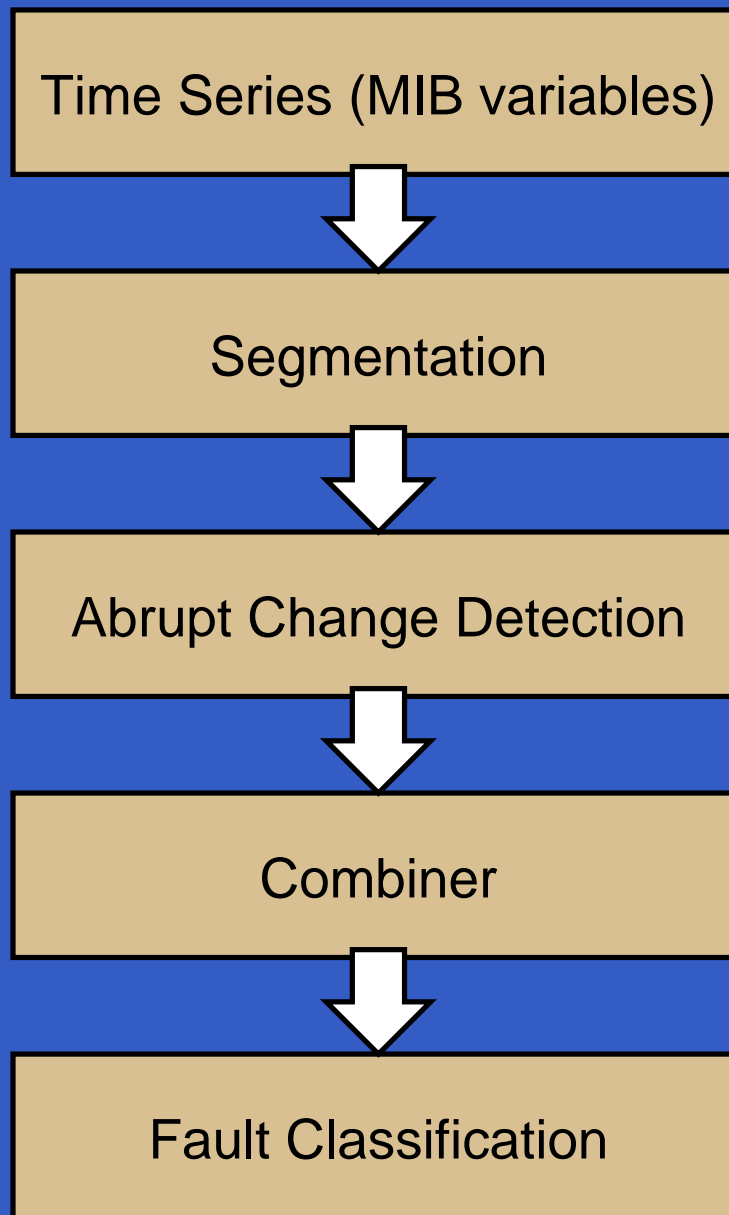
An intuitive solution

- Model (quantify) the normal network behavior.
- Use signal processing techniques for measuring, analyzing and synthesizing the network information.
- Detect deviation from the normal behavior due to an anomalous event.

The statistical approach

- Data source - MIB variables for individual network devices.
- Choice of appropriate variables
- Time series for each MIB variable by sampling.
- Detect abrupt and correlated changes in the values of these variables.
- Generate a network health function and use it to detect anomalous behavior.

The process



Segmentation

- Non-stationary time series.
- Division into piecewise stationary segments.
- Statistical properties assumed to be stationary within each segment.
- AR modelling of the segments.

Abrupt change detection

- Learning window $L(t)$, Test window $T(t)$
- Quantify the change in the statistical properties of the segments.
- Hypothesis test using Generalized Likelihood Ratio (GLR).
- Abnormality indicator $\vec{\psi}(t)$ obtained for each MIB variable.

Hypothesis test

The joint likelihood l of the residual errors in the two windows $L(t)$ and $T(t)$

$$l = \left(\frac{1}{\sqrt{2\pi\sigma_L^2}} \right)^{N_L} \left(\frac{1}{\sqrt{2\pi\sigma_T^2}} \right)^{N_T} \exp\left(\frac{-N_L \hat{\sigma}_L^2}{2\sigma_L^2} \right) \exp\left(\frac{-N_T \hat{\sigma}_T^2}{2\sigma_T^2} \right) \quad (1)$$

Hypothesis test

- H_0 : No change is observed between the two windows

$$l_0 = \left(\frac{1}{\sqrt{2\pi\sigma_P^2}} \right)^{N_L+N_T} \exp\left(\frac{-(N_L + N_T)\hat{\sigma}_P^2}{2\sigma_P^2} \right)$$

- H_1 : Change is observed between the two windows

$$l_1 = l$$

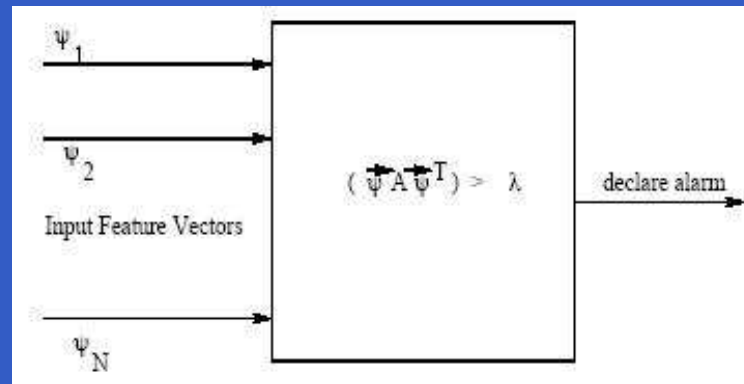
$$\text{Likelihood Ratio}(\eta) = \frac{l_1}{l_1 + l_0} = \frac{\hat{\sigma}_L^{-N_L} \hat{\sigma}_T^{-N_T}}{\hat{\sigma}_L^{-N_L} \hat{\sigma}_T^{-N_T} + \hat{\sigma}_P^{-(N_L+N_T)}} \quad 0 \leq \eta \leq 1$$

Spatial correlation using combiner

- Individual abnormality indicators combined to form abnormality vector $\vec{\psi}(t)$.
- Quadratic functional

$$f(\vec{\psi}(t)) = \vec{\psi}(t)A\vec{\psi}(t)$$

used to generate scalar network health indicator.



- Linear operator A based on the correlation matrix used.

Combiner (contd.)

$\vec{\psi}(t)$ can be represented as linear combination of the eigenvectors of A

$$\begin{aligned} A\vec{\phi} &= \lambda\vec{\phi} \\ \vec{\psi}(t) &= \sum_{i=1}^M c_i \vec{\phi}_i \\ A\vec{\psi}(t) &= \sum_{i=1}^M c_i \lambda \vec{\phi}_i \end{aligned}$$

An average scalar value of the transformation is obtained by using the function

$$\begin{aligned} f(\vec{\psi}(t)) &= \vec{\psi}(t) A \vec{\psi}(t) \\ &= \sum_{i=1}^M c_i^2 \lambda_i \\ &= E(\lambda) \end{aligned}$$

where $E(\lambda)$ is the measure of the average abnormality in the network as perceived by the node.

Node Level Alarms

A bound for $E(\lambda)$ can be obtained

$$\min_{\lambda_i \in \{\lambda_1, \dots, \lambda_N\}} (\lambda_i) \leq E(\lambda) \leq \max_{\lambda_i \in \{\lambda_1, \dots, \lambda_N\}} (\lambda_i)$$

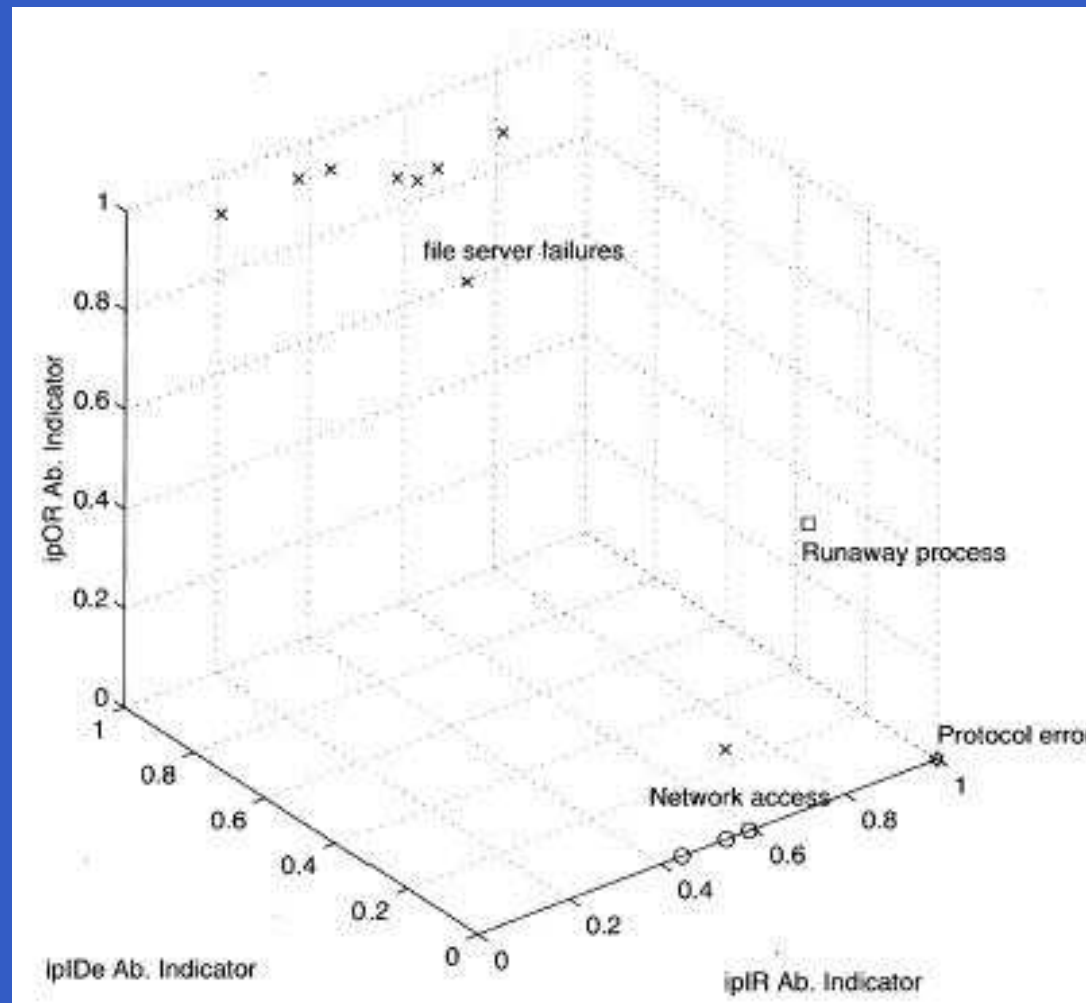
Using these bounds on the functional $f(\vec{\psi}(t))$ an alarm is declared when

$$E(\lambda) > \min_{\lambda_i \in \{\lambda_1, \dots, \lambda_N\}} (\lambda_i)$$

Classification Of Faults

- Faults tend to be clustered in the fault space.
- Euclidean distance between two fault clusters is sufficient to distinguish between the fault types.
- Abnormality vectors for known faults monitored and corresponding area in the fault space identified.
- Use this prior knowledge to identify known faults.

Fault Classification : An example



References

1. Chuanyi Ji Marina Thottan. Anomaly detection in ip networks. *IEEE Transactions on Signal Processing*, 51(8):2191-2204, 2003.
2. M. Thottan and C. Ji. Using network fault predictions to enable ip traffic management. *J. Netw. Syst. Manage.*, 9(3):327-346, 2001.
3. Chuanyi Ji Marina Thottan. Fault prediction at the network layer using intelligent agents. In *IFIP/IEEE International Symposium on Integrated Network Management*, pages 745-759. IEEE, 2000.
4. M. Thottan and C. Ji. Adaptive thresholding for proactive network problem detection. In *Proceedings of the IEEE Third International Workshop on Systems Management*, page 108. IEEE Computer Society, 1998.
5. Kishor S. Trivedi. *Probability and statistics with reliability, queuing and computer science applications*. John Wiley and Sons Ltd., New York, 2001.
6. William Hines et al. *Probability and Statistics in Engineering, 4th edition*. John Wiley and Sons Ltd., Singapore, 2003.