Failure Diagnosis with Incomplete Information in Cable Networks

Penn

Yun Mao† Hani Jamjoom‡ Shu Tao‡

† University of Pennsylvania ‡ IBM T.J. Watson Research

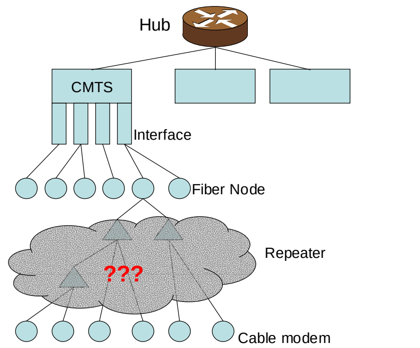
IBM

Background

- **Broadband cable Internet service is a huge market**
  - **"In 2004 there are 90,600 average number of cable modem installs per week in the US". *(Kagan Research, LLC, March 2004)***
  - **"By 2007, 90% (or 46.4M households) of all broadband households in the US will have home networking services available to them". *(Parks Associates, 4Q 2003)***
- **Failures are frequent in cable networks**
  - **Large distribution networks with hundred of thousands of devices in each administered area**
  - **Confirmed by measurement from a major US cable provider**
- **Failure management is critical to service providers, or…**
  - **Customers may go away, i.e. less revenue**
  - **Increase management cost**

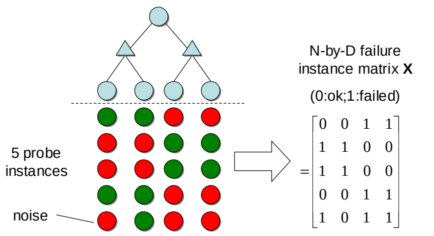
Challenges

- **Missing topology between fiber nodes and cable modems**
  - **Most of the failures are due to failures at that level (based on observations of ~2 month measurement)**
  - **Unknown status of repeaters**
- **Misconfigurations**
  - **Typo in address, wrong mapping from address to geocode, etc**
- **Question: can we infer the missing topology or the failure groups of nodes that share the same risks?**



Design

- **Intuition: explore temporal and spatial correlation of the failures**
- **Combine passive monitoring and active probing to retrieve failure data**



▪**Goal: decompose X into two *smaller* binary matrices U, V such that X ≈ U *V**

$$\begin{bmatrix}0&0&1&1\\1&1&0&0\\1&1&0&0\\0&0&1&1\\1&0&1&1\end{bmatrix} \approx \begin{bmatrix}0&1\\1&0\\1&0\\0&1\\0&1\end{bmatrix} * \begin{bmatrix}1&1&0&0\\0&0&1&1\end{bmatrix}$$

- ▪**V represents the compositions of the failure groups**
- ▪**U indicates the failure status of each failure group**

▪**Many well-known algorithms (e.g. PCA/SVD) don't fit**

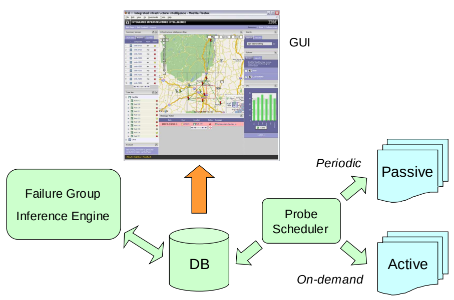
- ▪**They can do the decomposition, but into real value matrices**
- ▪**Negative values in their output are inexplicable**

▪**Our core algorithm: Non-negative matrix factorization (NMF)**

- ▪**NMF does the decomposition, with non-negative constraints**
- ▪**Try to minimize** $\sum_i \sum_j (X_{ij} - (UV)_{ij})^2$
- ▪**Converge to local minima quickly**

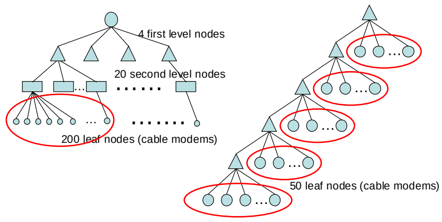
▪**Further normalize the results from NMF and apply threshold to obtain the binary matrices**

Implementation

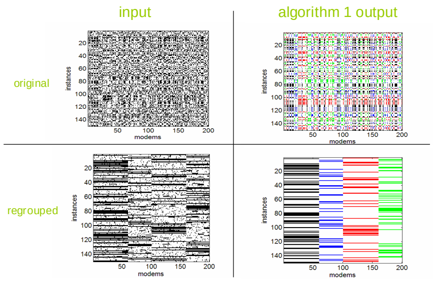


Experimental Results

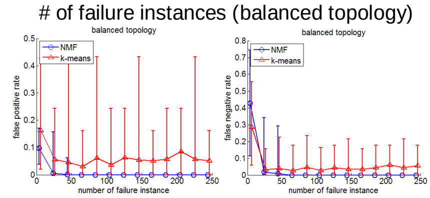
balanced topology vs. cascaded topology



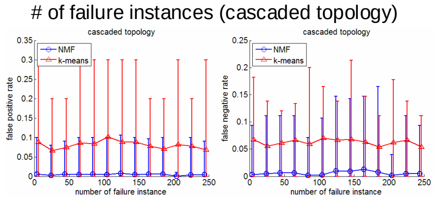
An example (balanced topology)



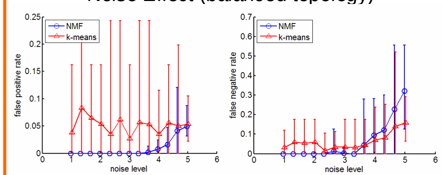
\# of failure instances (balanced topology)



\# of failure instances (cascaded topology)



Noise Effect (balanced topology)



Noise Level = P(individual failure) / P(group failure)