Intelligent Classification and Visualization of Network Scans

Chris Muelder,

Lei Chen, Russell Thomason, Kwan-Liu Ma
VIDi group at University of California, Davis

Tony Bartoletti
Lawrence Livermore National Laboratory

Motivation

- Counterintelligence efforts
  - Want to learn about attackers
  - Tools/OS/Hardware/Internet location
  - ID them if/when they return
- Commonly source address is used for ID
  - Source does not indicate tool/os/hardware

Motivation

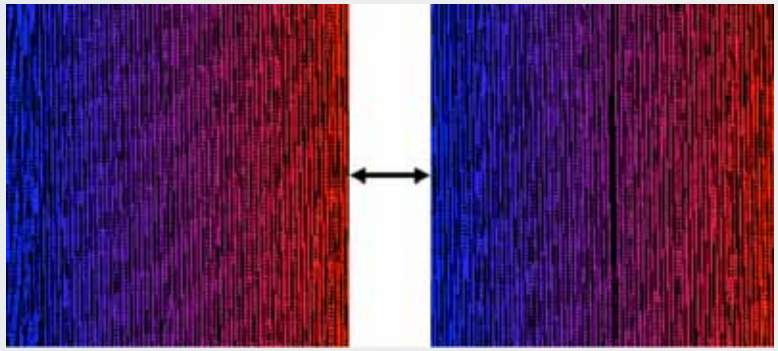
- Sources can be dynamic, spoofed, etc…
- Want to ID an attacker based on unalterable properties
- Timing is fairly unalterable and very difficult to spoof
  - Hardware factors
  - Software factors
  - Routing factors

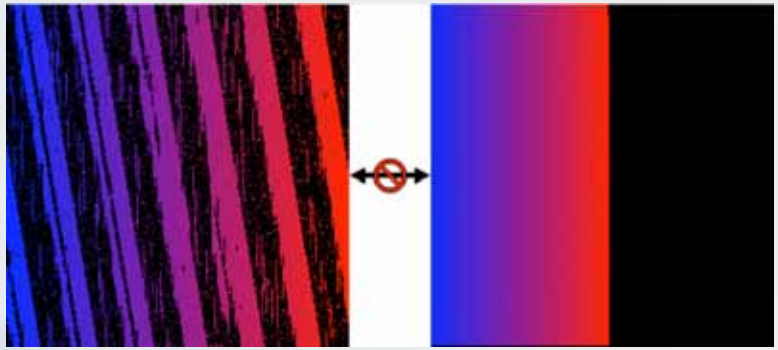
Network Scans

- Good source of timing information
- Probe every possible address
  - Find out what is there
  - Often followed by more serious attack
  - Could contain an attack (Worms)
  - Could be benign (Web spiders)

Previous Approaches

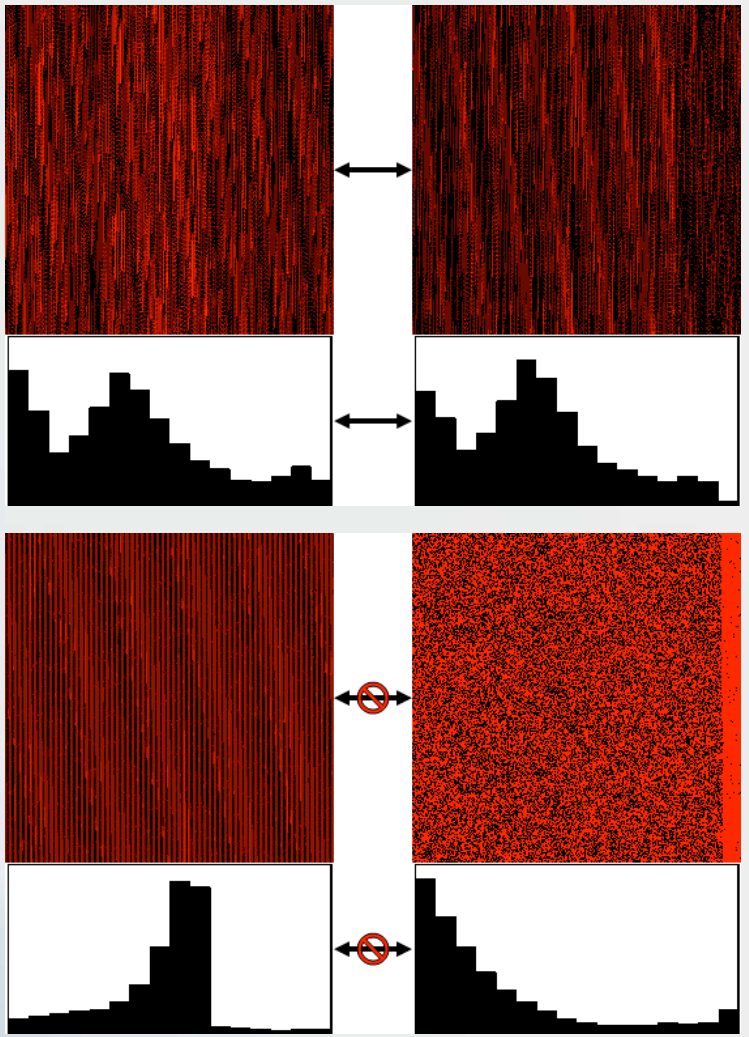
- Direct visual inspection
  - Class B network (A.B.0.0/16)
  - 3rd and 4th bytes are axes
  - Color is a time based metric
    - We use a deviation from a linear expectation
  - Effectiveness
    - Pattern matching easy for human eye
    - Can not scale well



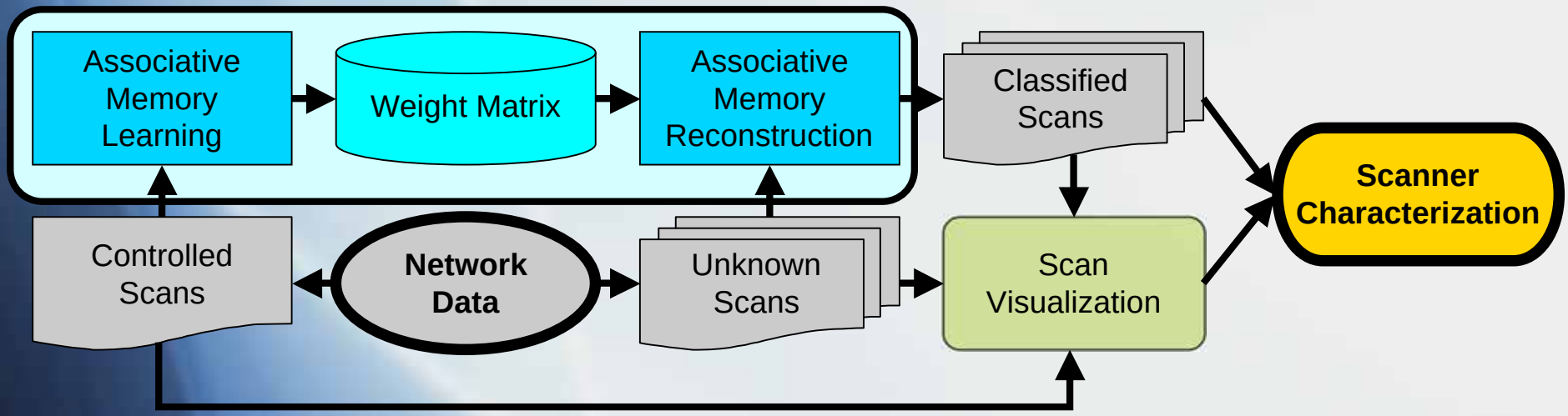


Previous Approaches

- Wavelet Analysis
  - Reduce 65,536 values to 16-dimensional scalogram
  - Captures frequency properties
  - Effectiveness
    - Scalograms can be automatically compared
    - Loses data
    - Can match somewhat dissimilar patterns



Overview



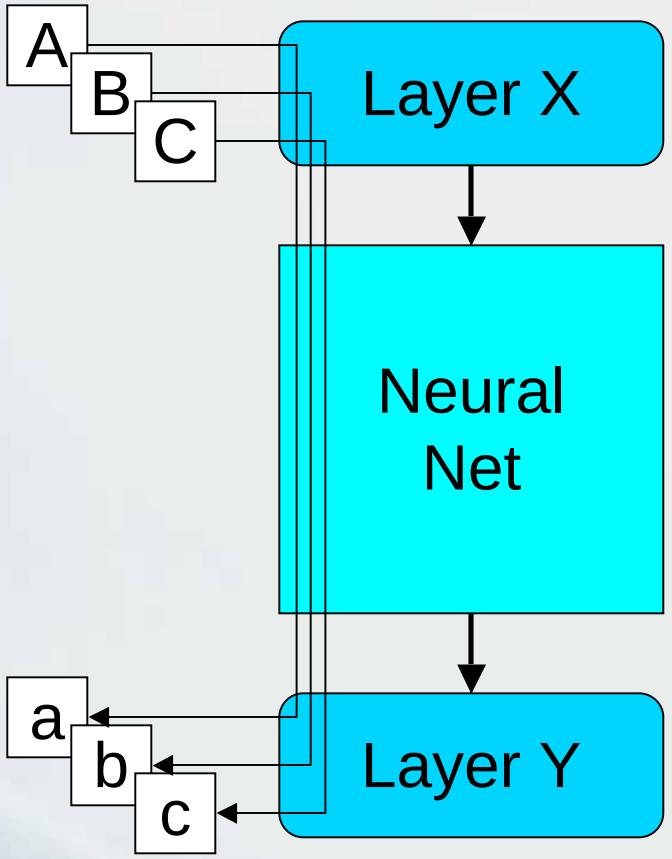
- Extract timing information from scans
- Analyze with an intelligent approach
- Combine with existing visual approach

Scan Data

- Collected by Computer Incident Advisory Capability (CIAC) at Lawrence Livermore National Laboratory (LLNL)
- Controlled scans generated by running various common tools on an isolated LAN
- Unknown scans collected at LLNL border
- Detected by rate threshold (probes/second)

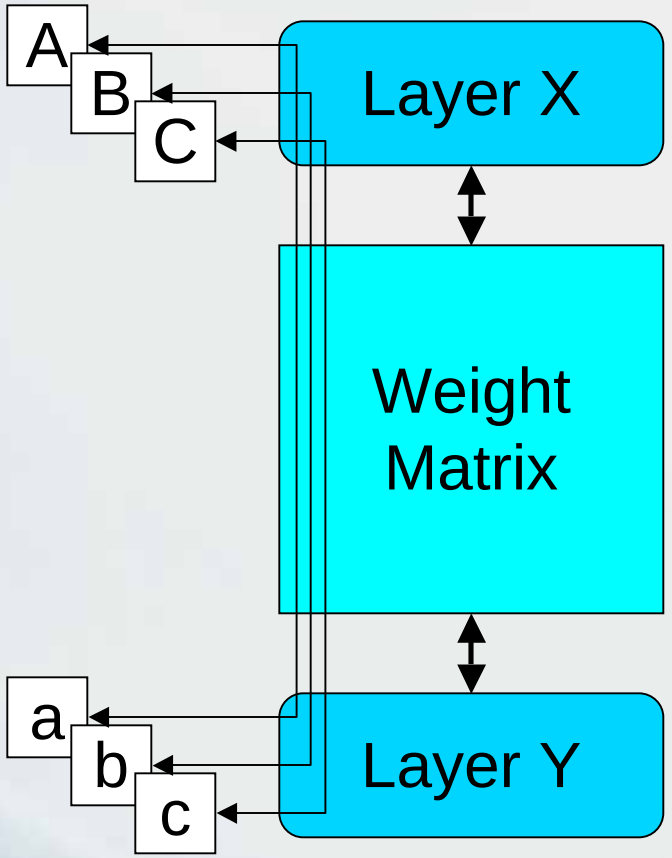
Associative Memory

- Maps from one pattern space to another
  - Map pattern in layer X to pattern in layer Y
  - Goes through a neural net of some sort
- Good for working with noisy patterns
- Several variants
  - BAM, Hopfield, etc…



Bidirectional Associative Memory

- BAM (Bidirectional Associative Memory) maps patterns in both directions
  - X --> Y and Y-->X
- Neural net is a matrix of weights
- Iterates back and forth until equilibrium
- Discrete, bipolar layers and patterns



BAM Training

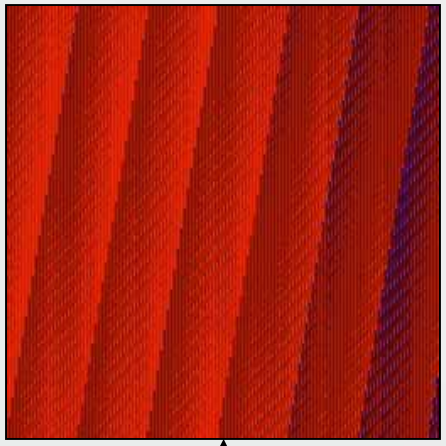
- Calculate weight matrix
  - Let:
    - $\mathbf{W} = \{W_{ij} \mid 0 \leq i < |\mathbf{X}|, 0 \leq j < |\mathbf{Y}|\}$
    - $\mathbf{X}_k = \{x_{ki} \mid 0 \leq i < |\mathbf{X}|\}$ for the k^{th} pattern
    - $\mathbf{Y}_k = \{y_{kj} \mid 0 \leq j < |\mathbf{Y}|\}$ for the k^{th} pattern
  - Then:
    - $W_{ij} = \Sigma_k x_{ki} * y_{kj}$

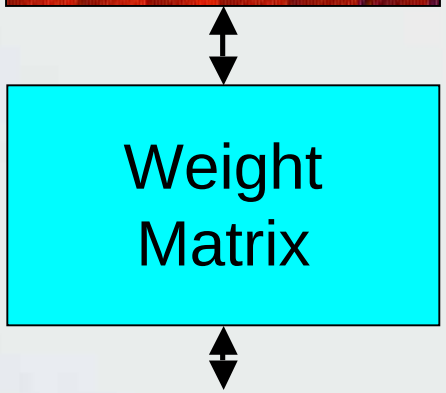
BAM Iteration

- Each iteration t from **X** layer to **Y** layer
  - Let
    - $x_i`(t) = \Sigma_j\ y_j(t-1)*w_{ij}$
  - Then
    - $x_i(t) = +1$ if $x_i`(t) > 0$
    - $x_i(t) = x_i(t-1)$ if $x_i`(t) = 0$
    - $x_i(t) = -1$ if $x_i`(t) < 0$
- Each iteration t from **Y** layer to **X** layer
  - Let
    - $y_j`(t) = \Sigma_i\ x_i(t-1)*w_{ij}$
  - Then
    - $y_j(t) = +1$ if $y_j`(t) > 0$
    - $y_j(t) = y_j(t-1)$ if $y_j`(t) = 0$
    - $y_j(t) = -1$ if $y_j`(t) < 0$

Application to Network Scans

- Map scans to ID patterns
  - Scans are converted to bipolar patterns
  - ID patterns are unique and randomly generated
  - Size of ID's proportional to number of training scans
- Classification
  - Train on known scans
  - Classify unknown scans

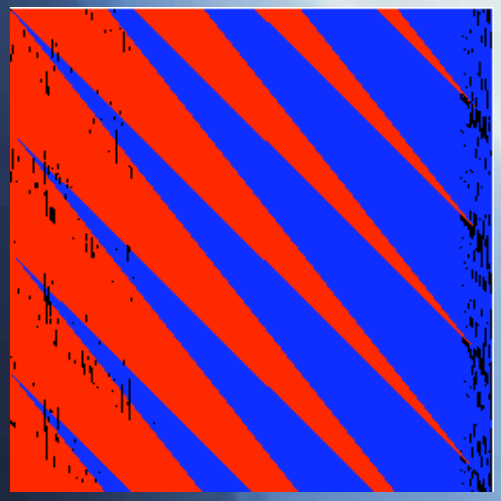




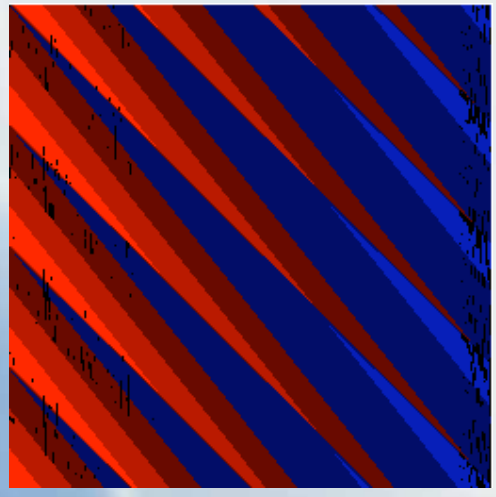
{1,1,-1,…,1,-1}

Bipolar Encoding

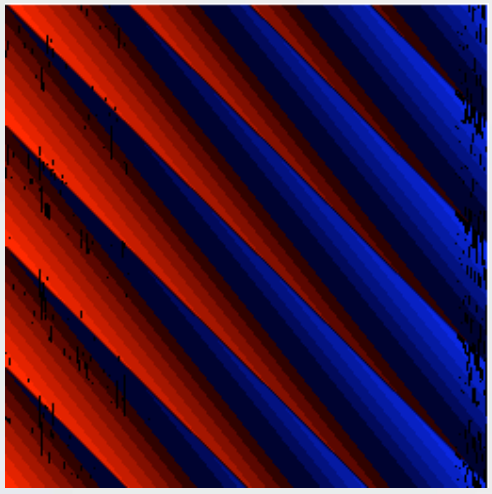
- Reduce float data to bipolar/binary patterns
- User adjustable numbers of bits
  - More bits = more resources



2 bits

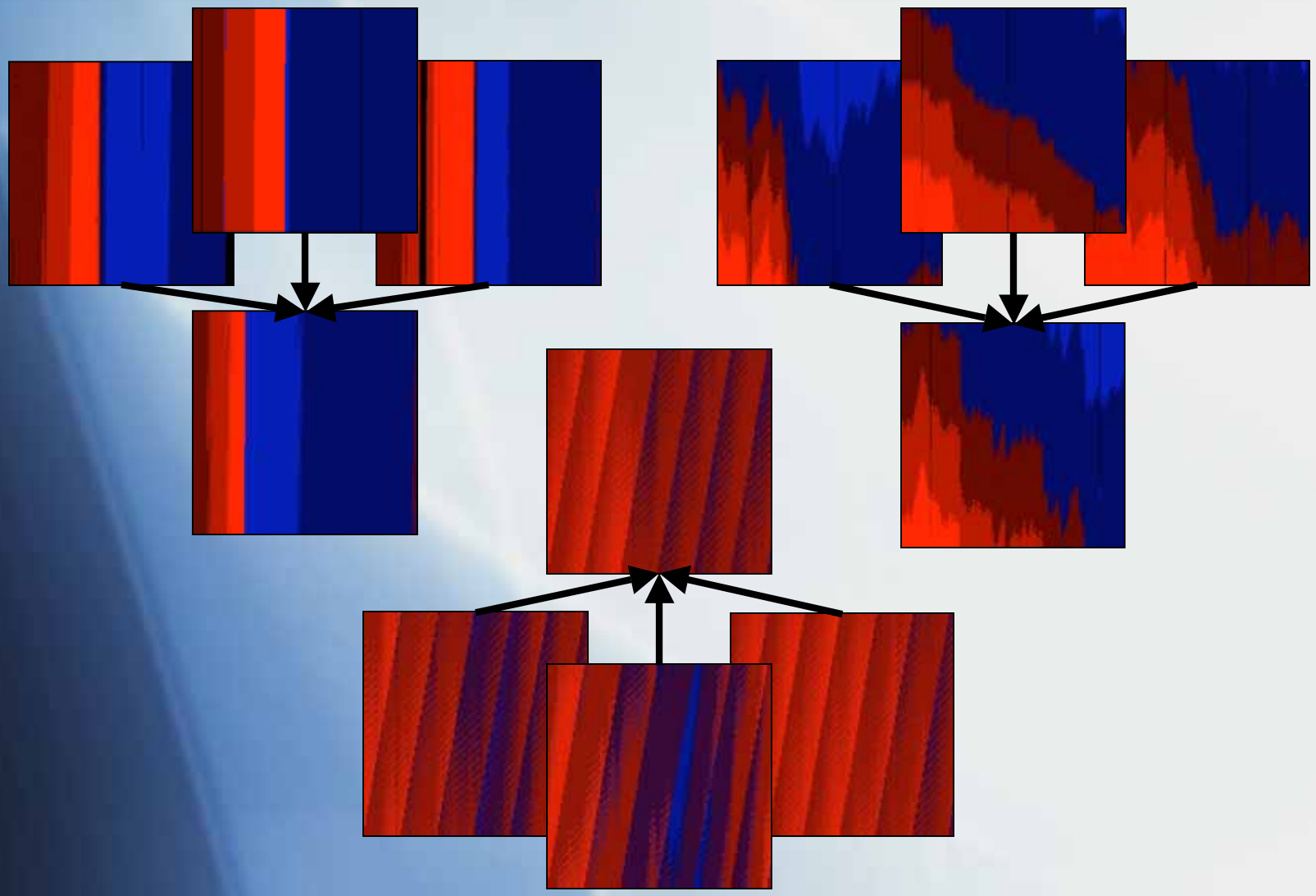


3 bits



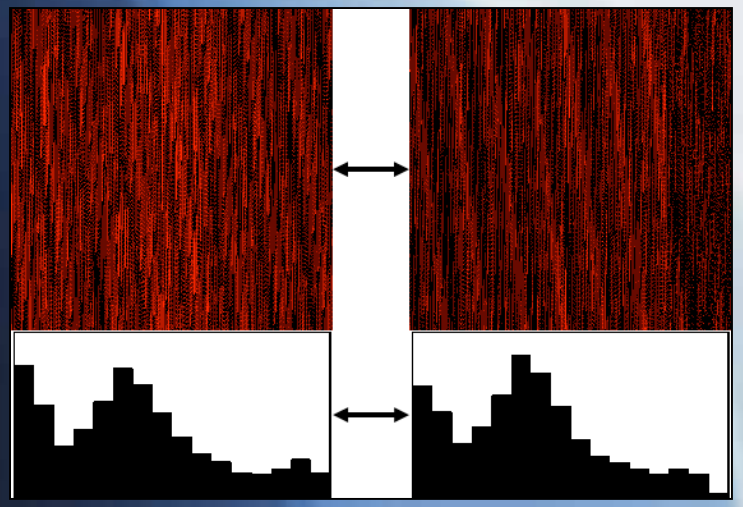
4 bits

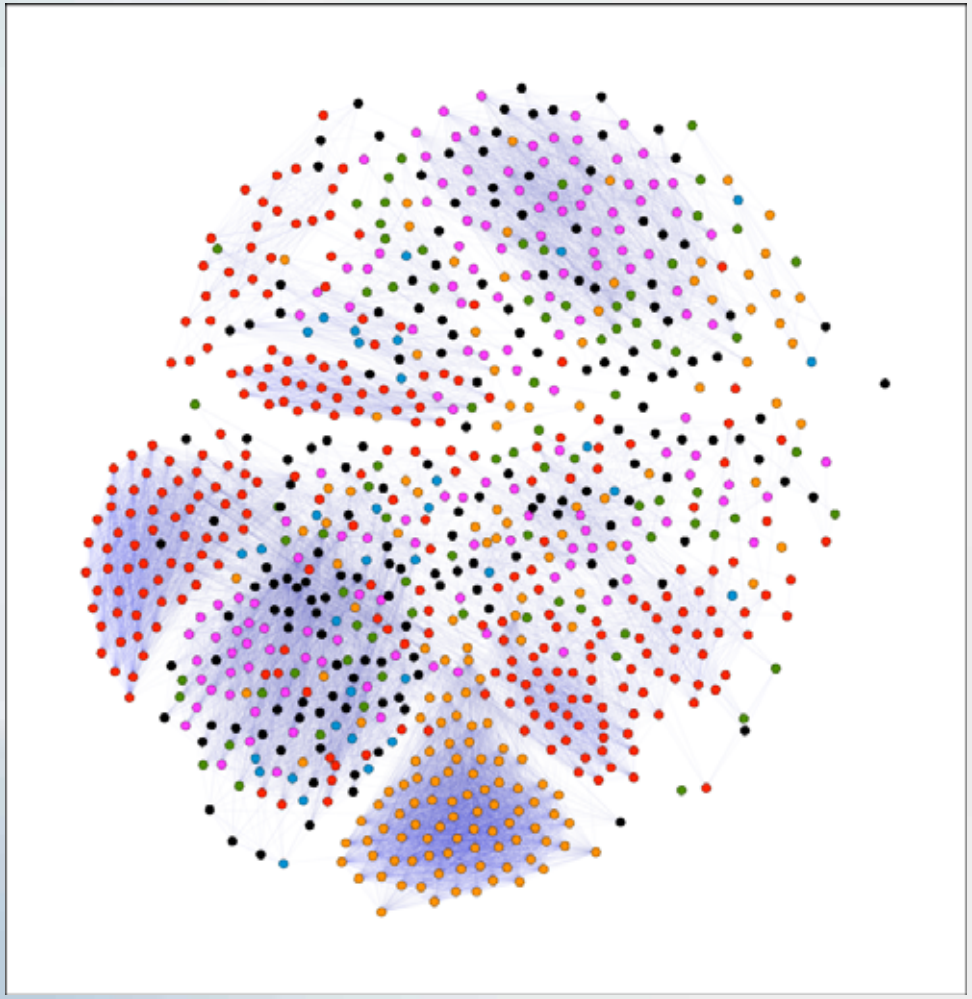
Results



Visualization Integration

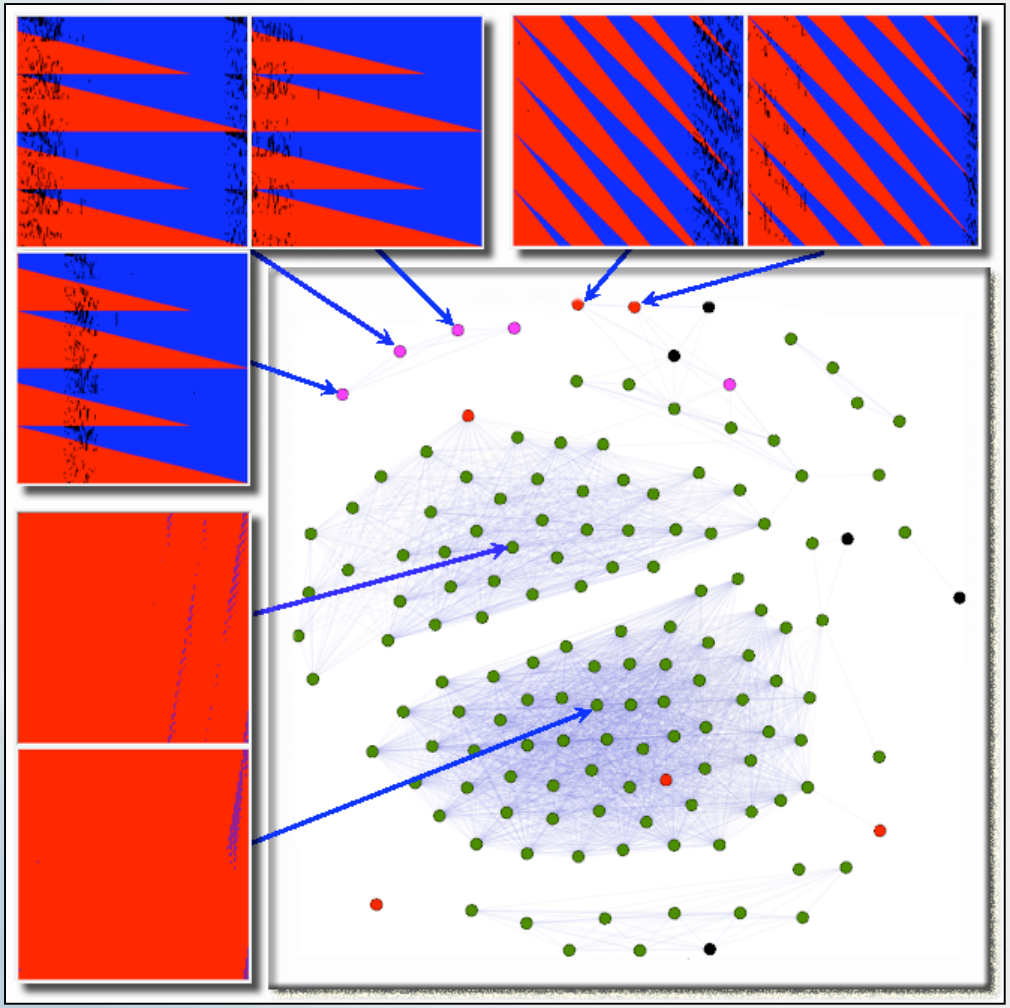
- ScanVis
  - VizSec 2005
  - Wavelet analysis
  - Graph overview



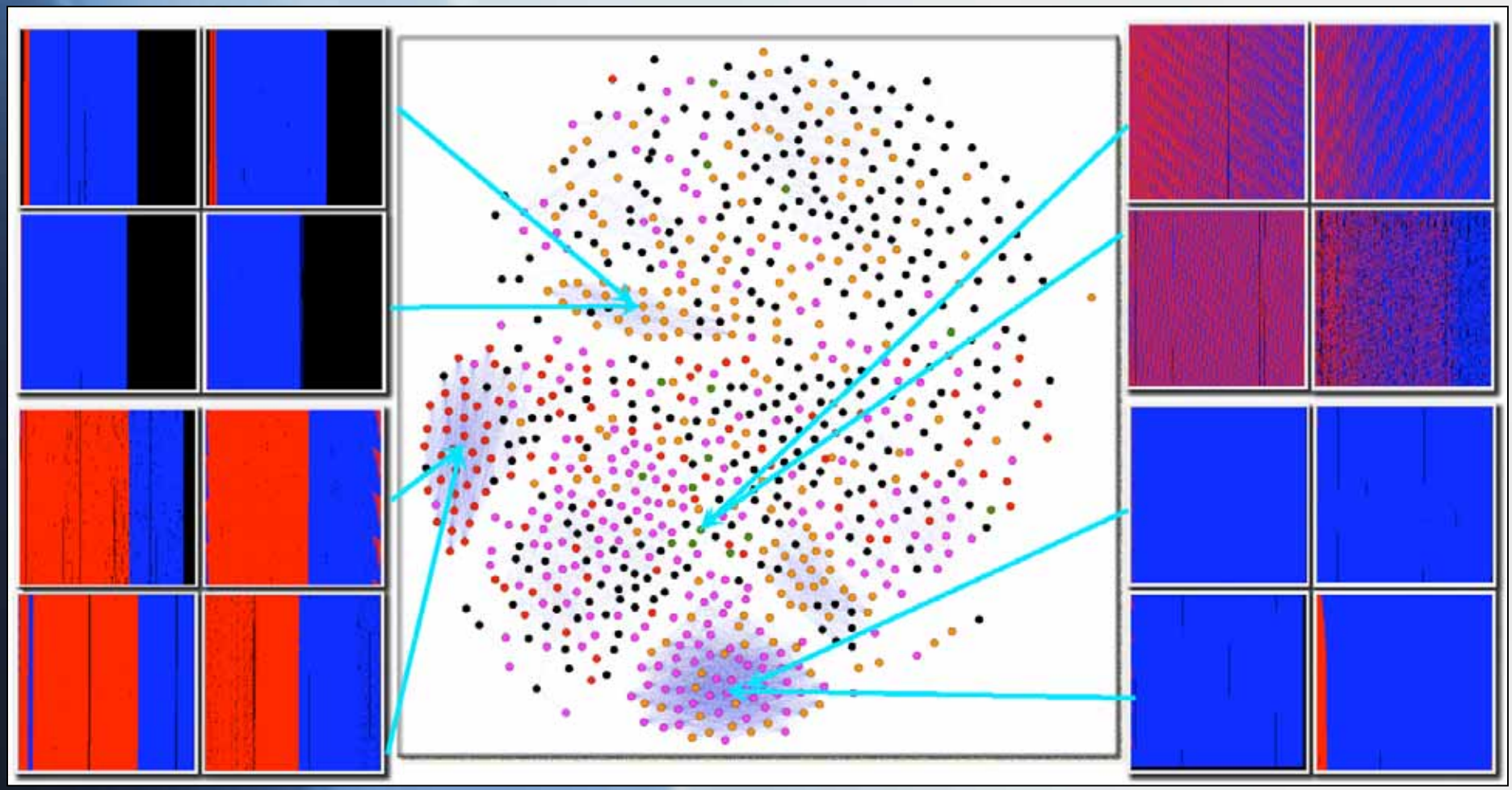


Visualization Integration

- BAM results can be integrated through color
  - Nodes colored according to classification
  - Control data results very good

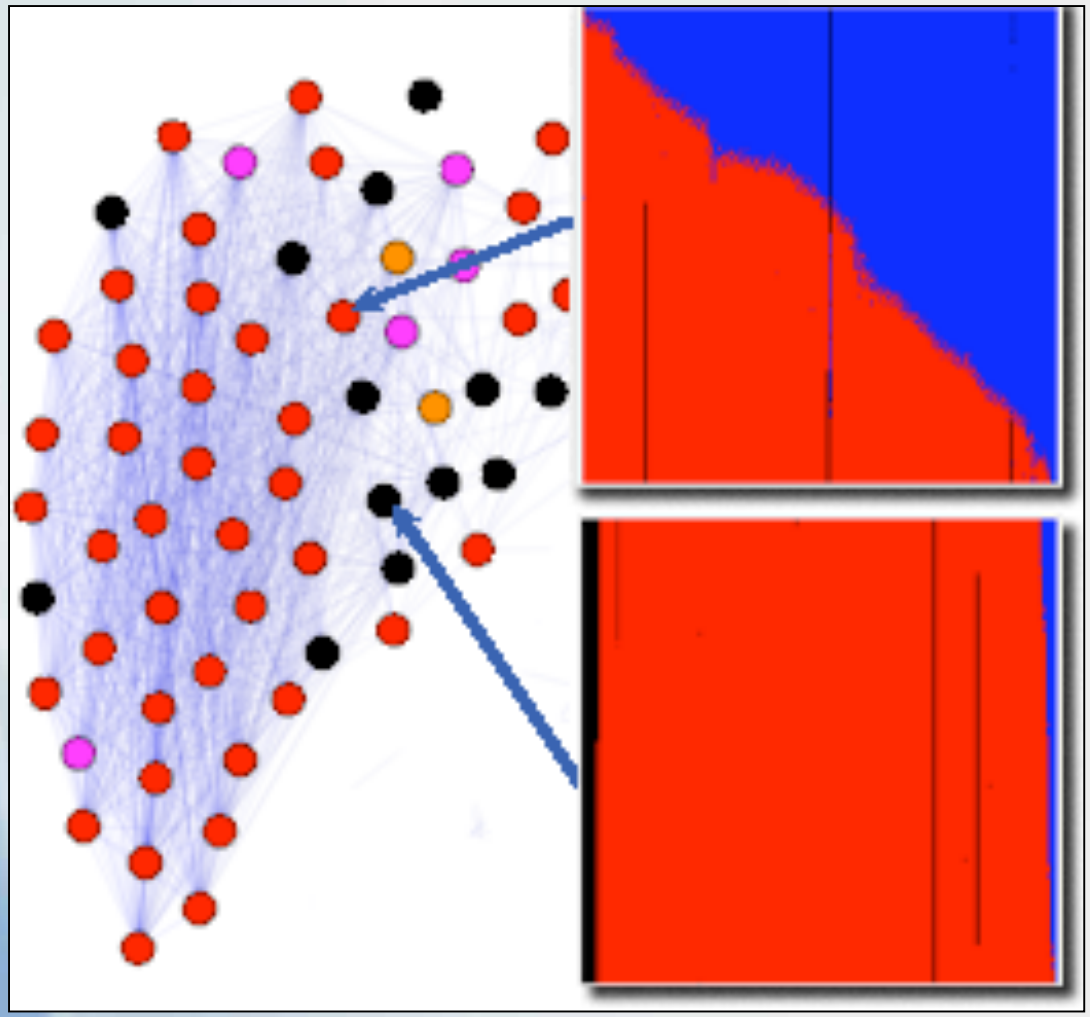


Application



Usefulness

- Discrepancies between BAM and wavelets
  - Wavelets say similar
  - BAM says dissimilar
  - Visual inspection confirms difference



Conclusion

- Visual and intelligent approaches
  - Capture different aspects of the data
  - Complement each other well in combination
- Bidirectional Associative Memory
  - Effectively classifies scans
  - Requires good controlled data

Future Work

- Other intelligent algorithms?
  - Continuous BAM
  - Unsupervised approaches
- Other metrics
- More controlled data
- Tighter visualization integration
  - Select training scans from graph
  - Modify graph layout according to classification

Questions?

Thanks for listening