NetClinic: Interactive Visualization to Enhance Network Fault Diagnosis

Zhicheng Liu, Bongshin Lee, Srikanth Kandula and Ratul Mahajan

Georgia Institute of Technology

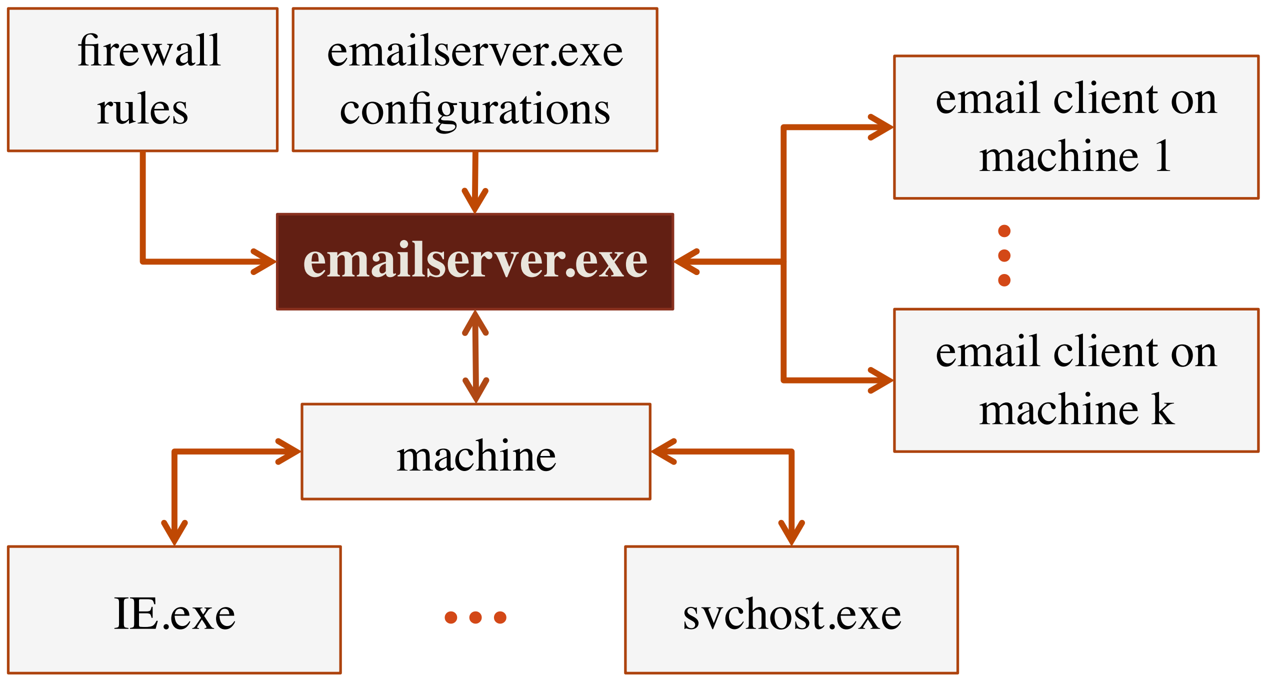
Microsoft Research

The Problem: Diagnosing Enterprise Networks

- Faults: Anomalies in application behavior
  - Cannot send email, browser extremely slow, network connectivity down …

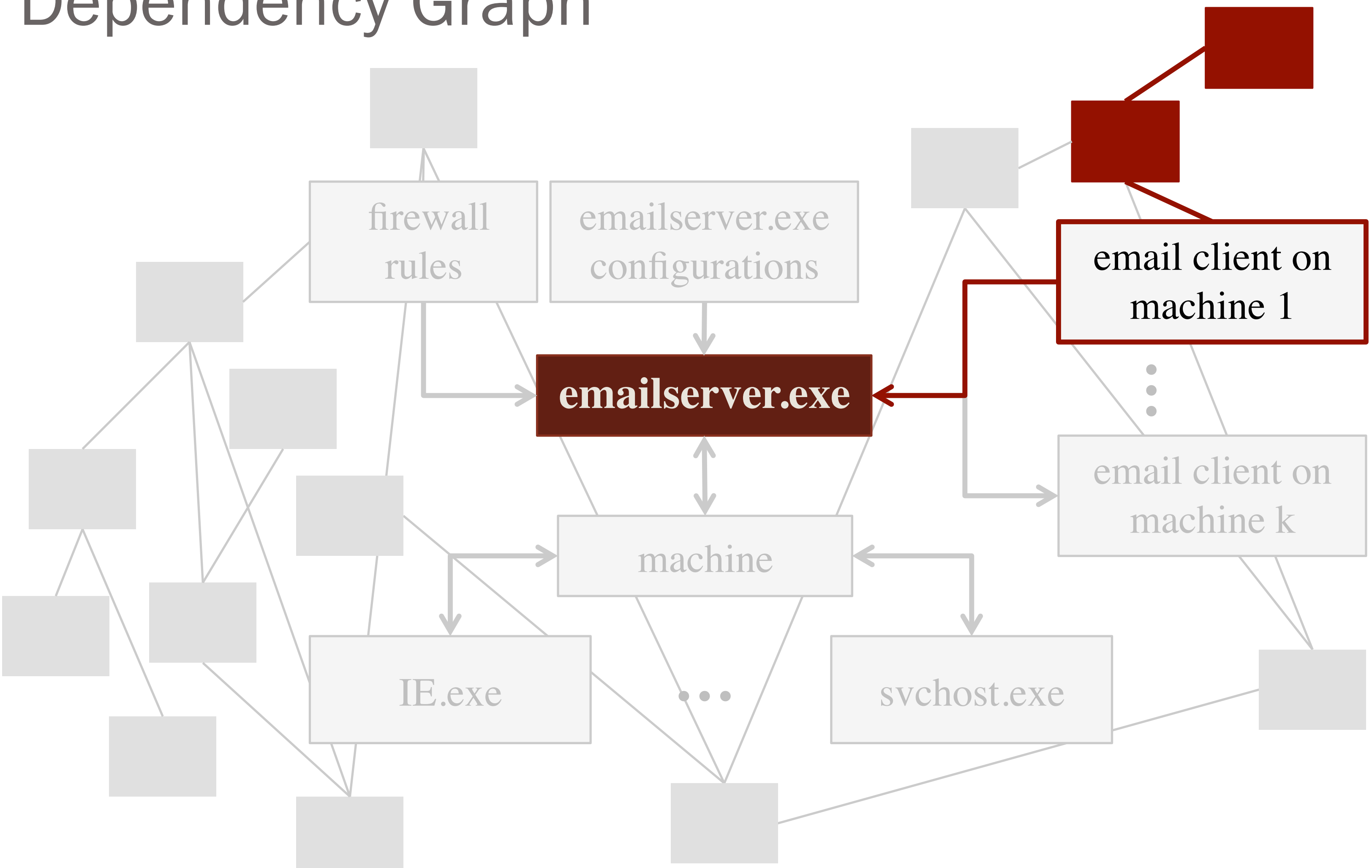
- Difficulty in identifying culprits / root causes
  - Network components interact in complex ways
  - Information overloading: too many variables


Modeling Complex Interaction as a Dependency Graph



Not diagnosing at the level of network path and individual switches

Modeling Complex Interaction as a Dependency Graph



Not diagnosing at the level of network path and individual switches

Motivation for Visual Analytics

- Automated diagnosis tools are not always accurate
  - Rely on minimal application specific semantic knowledge
  - Mostly statistical

- Even when true culprit is identified
  - Need for exploration and verification
    - Ground truth is not known before-hand

An ideal visual analytics problem

Automated Diagnosis

NetMedic [Kandula et al., SIGCOMM 2009]

- Variable Level: Performance Counters
- Component Level: Statistical Abnormality
- Edge Level: Potentiality of Impact
  - Statistical analysis of joint behavior of neighbors
- Network Level: Given a faulty component, Identifying Culprits
  - Rank edge weights to order likely causes

The Reciprocal Nature of Human Sensemaking

DATA
recognize/construct
a frame

FRAME
define, connect and filter
the data

[Klein et al. 2006]

Dynamic mixture of top-down and bottom-up processes

Data-Frame Interaction in Network Diagnosis

Problem: a SQL client cannot talk to the server

Network Level	other clients are overloading the server	the fault is due to the change in firewall rule
Edge Level		client using a port number blocked by the firewall change
Component Level	check server machine application state	firewall rules of the server machine changed
Variable Level	i/o operations normal CPU / memory normal	

Frame

Data

Design Considerations

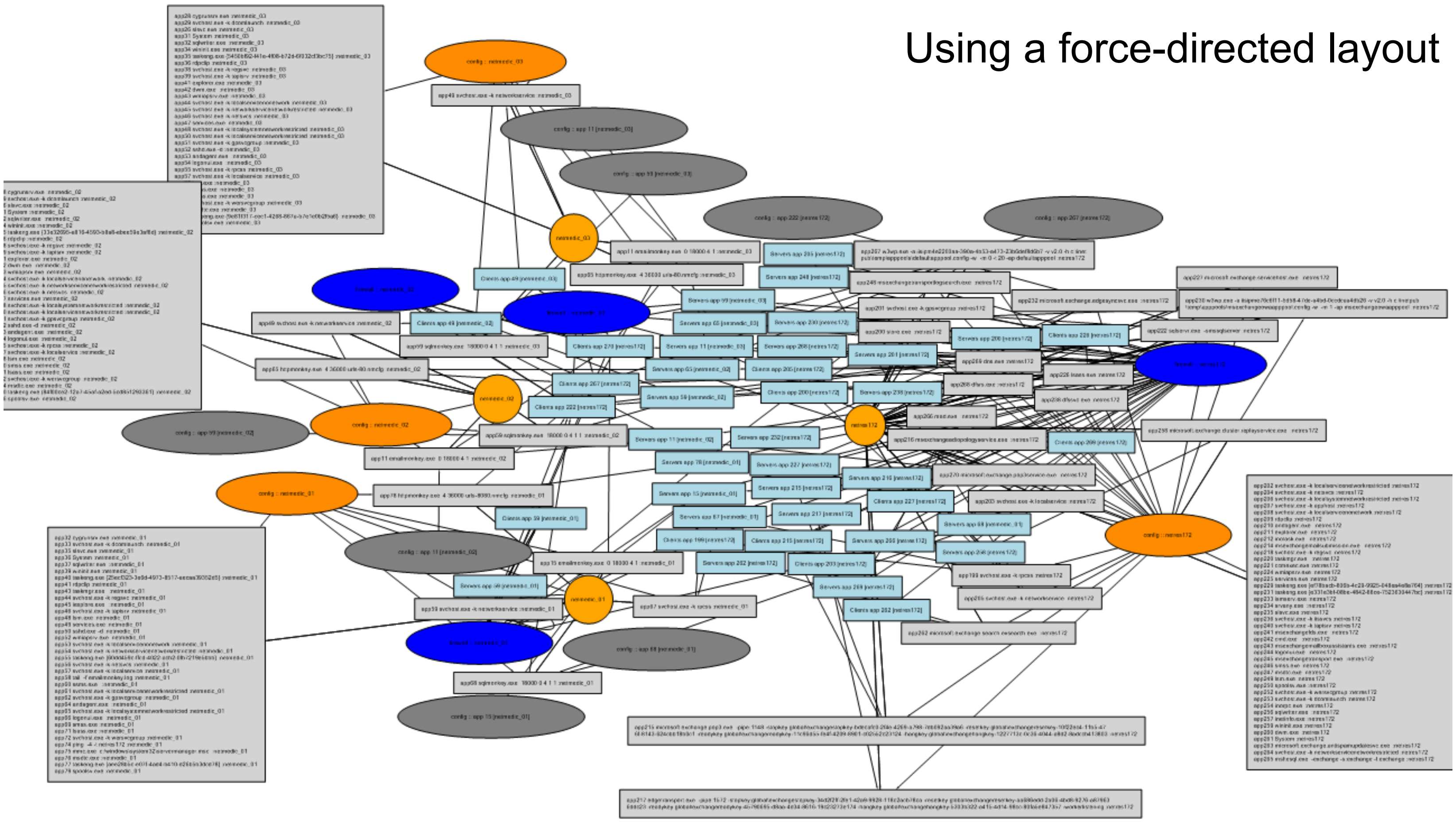
- Output of automated engines can be used as useful frames

- Show outputs at all levels of abstraction
  - Minimal constraints on navigation across levels of abstraction

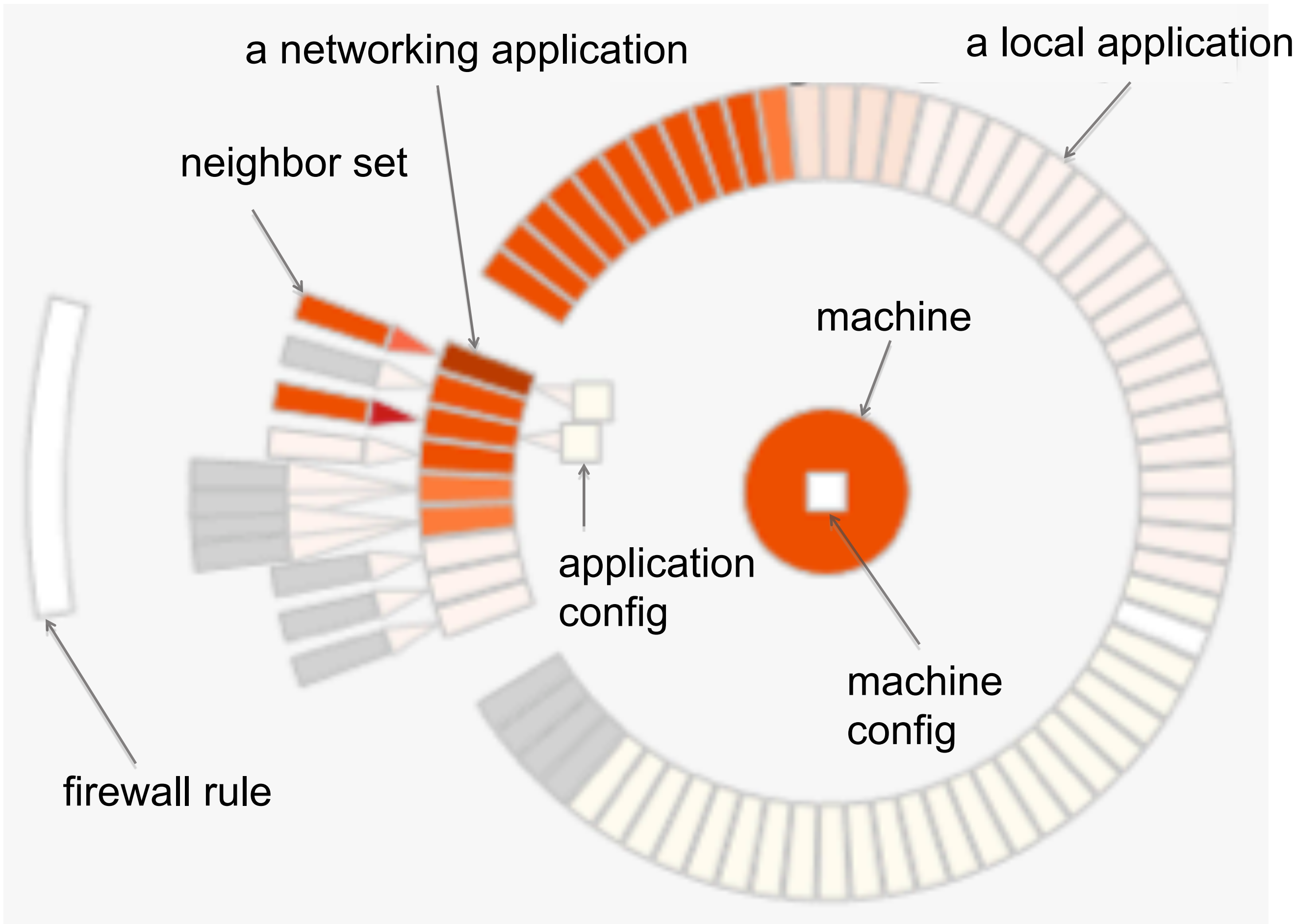
- Flexible exploration
  - Top-down exploration: verify the output of automated analysis
  - Bottom-up exploration: form and evaluate own hypothesis

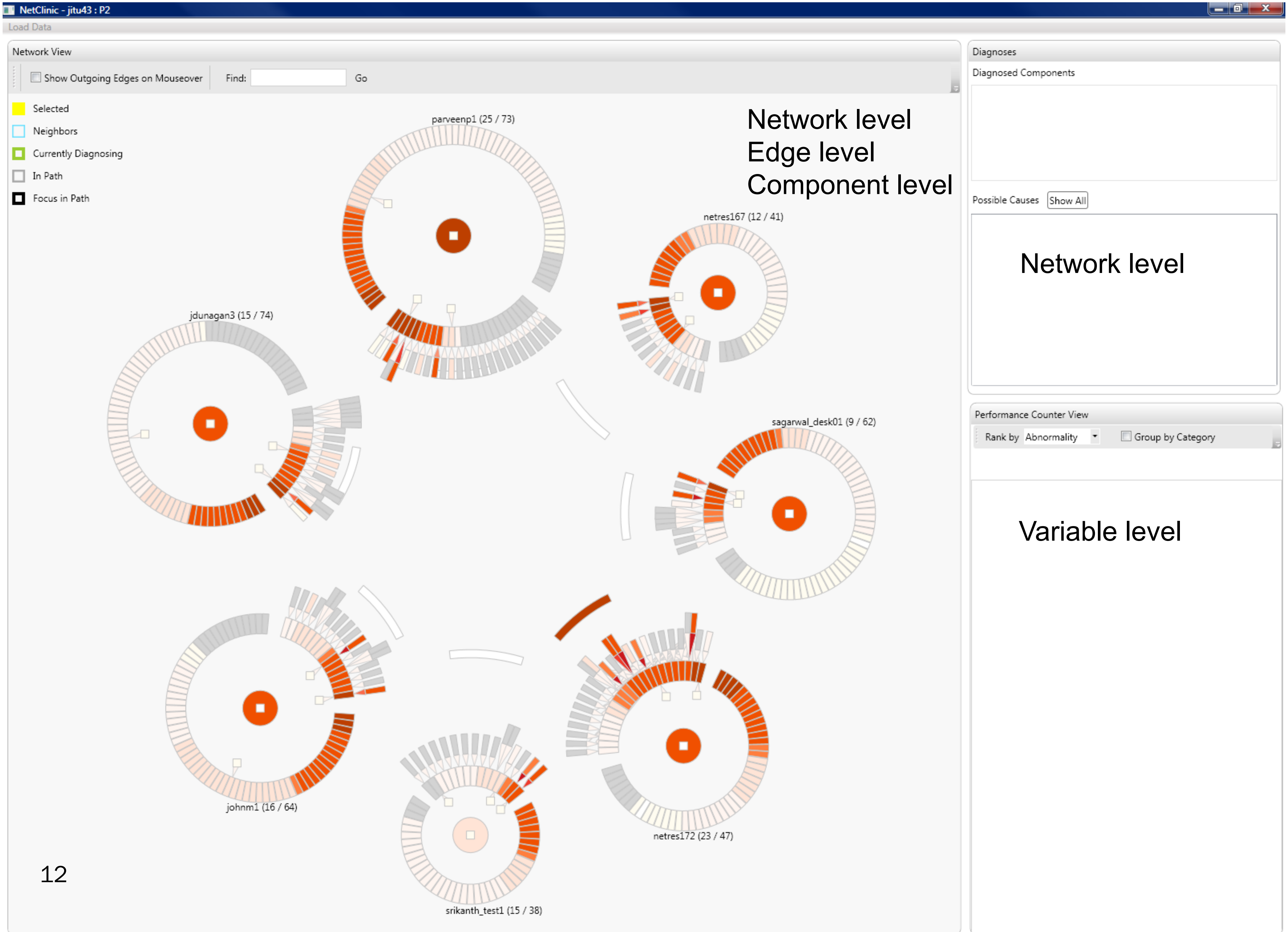
Main Design Challenge: Graph Layout

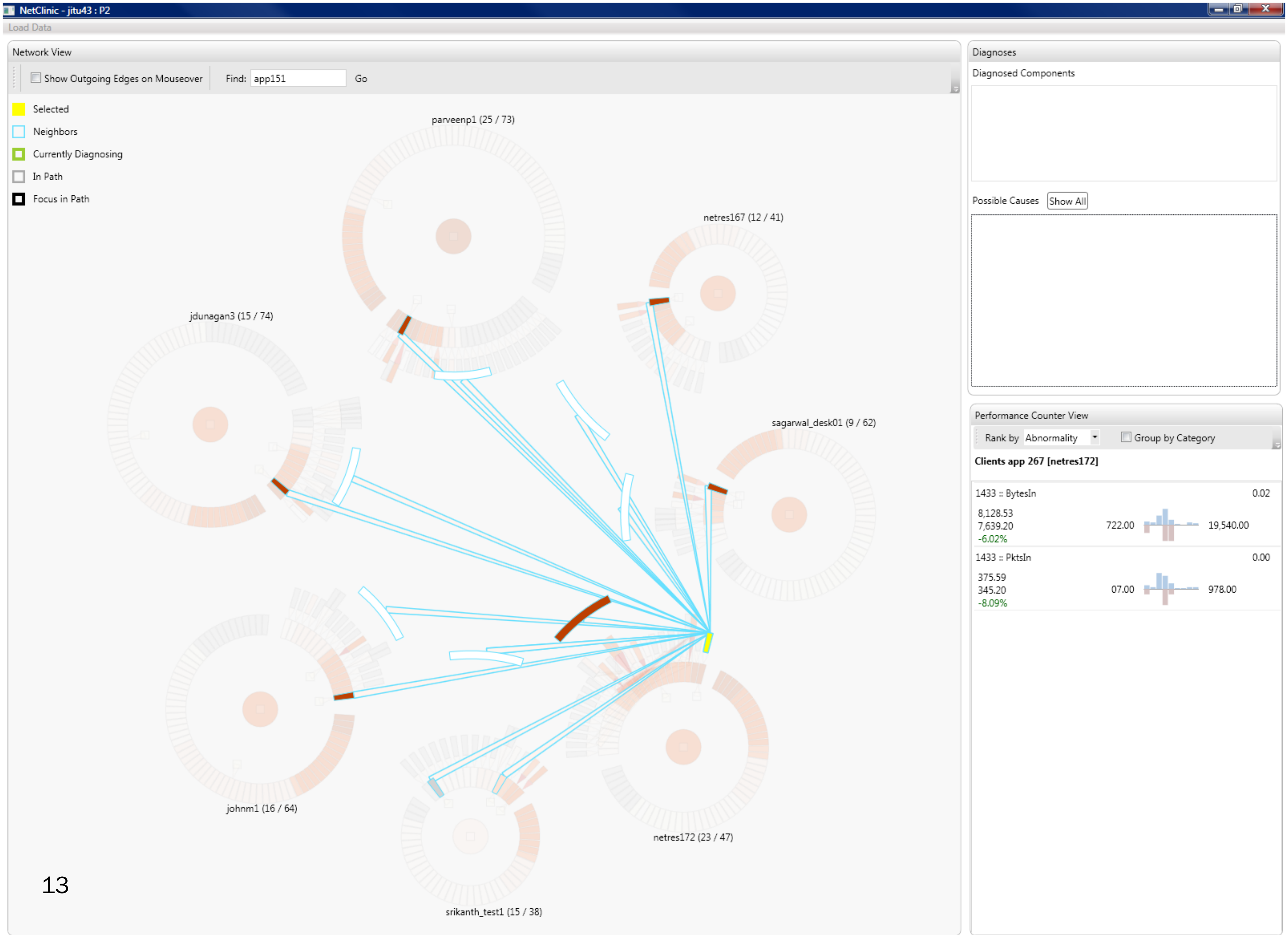
Using a force-directed layout

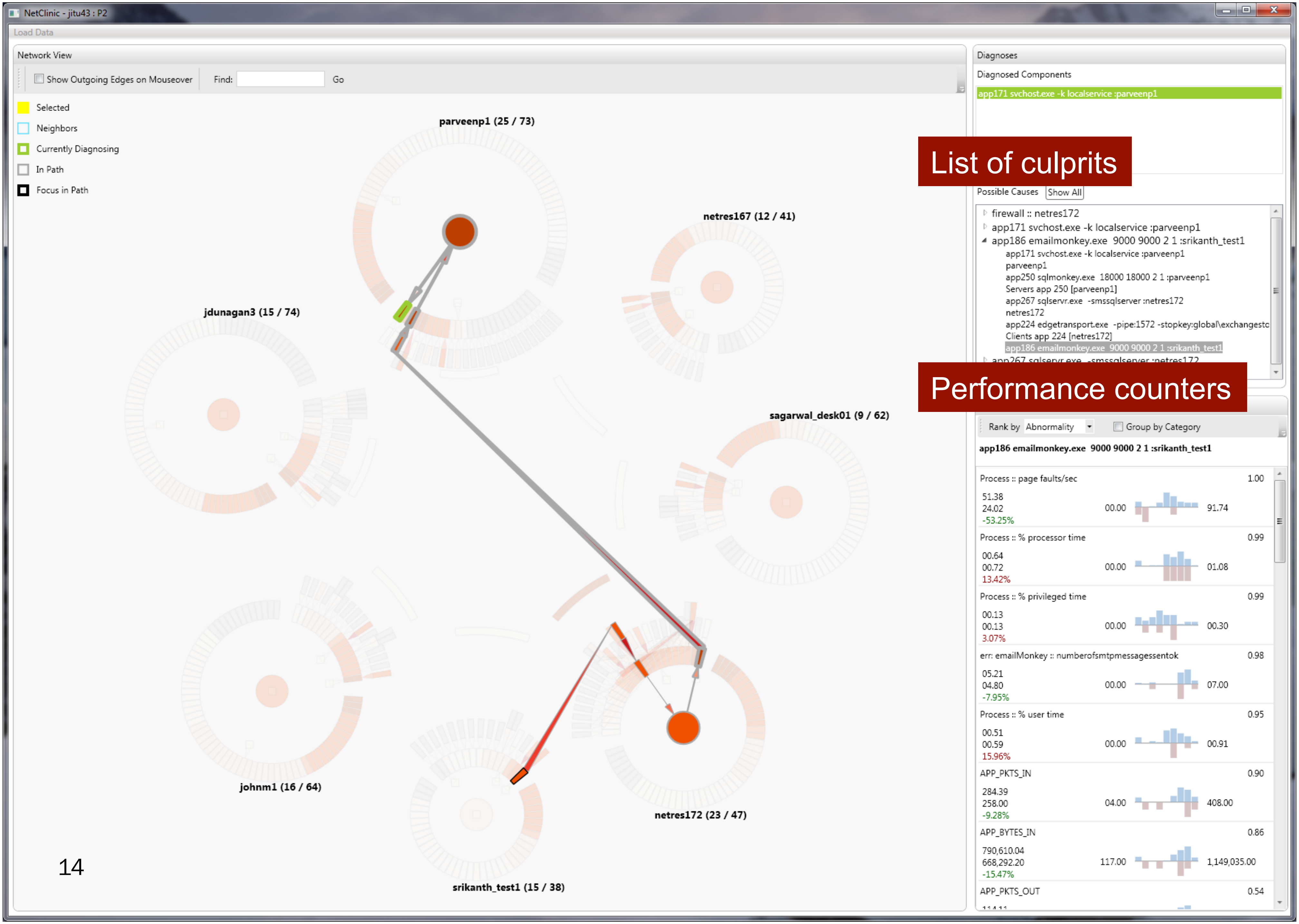


Machine-based clustering




NetClinic - jitu43 : P2
Load Data
Network View
Show Outgoing Edges on Mouseover
Find:
Go
Selected
Neighbors
Currently Diagnosing
In Path
Focus in Path
parveenp1 (25 / 73)
netres167 (12 / 41)
jdunagan3 (15 / 74)
sagarwal_desk01 (9 / 62)
johnm1 (16 / 64)
netres172 (23 / 47)
srikanth_test1 (15 / 38)
Network level
Edge level
Component level
Diagnoses
Diagnosed Components
Possible Causes
Show All
Network level
Performance Counter View
Rank by
Abnormality
Group by Category
Variable level


NetClinic - jitu43 : P2
Load Data
Network View
Show Outgoing Edges on Mouseover
Find: app151
Go
Selected
Neighbors
Currently Diagnosing
In Path
Focus in Path
parveenp1 (25 / 73)
netres167 (12 / 41)
jdunagan3 (15 / 74)
sagarwal_desk01 (9 / 62)
johnm1 (16 / 64)
netres172 (23 / 47)
srikanth_test1 (15 / 38)
Diagnoses
Diagnosed Components
Possible Causes
Show All
Performance Counter View
Rank by Abnormality
Group by Category
Clients app 267 [netres172]
1433 :: BytesIn
0.02
8,128.53
7,639.20
-6.02%
722.00
19,540.00
1433 :: PktsIn
0.00
375.59
345.20
-8.09%
07.00
978.00


NetClinic - jitu43 : P2
Load Data
Network View
Show Outgoing Edges on Mouseover
Find:
Go
Selected
Neighbors
Currently Diagnosing
In Path
Focus in Path
parveenp1 (25 / 73)
netres167 (12 / 41)
jdunagan3 (15 / 74)
sagarwal_desk01 (9 / 62)
johnm1 (16 / 64)
netres172 (23 / 47)
srikanth_test1 (15 / 38)
Diagnoses
Diagnosed Components
app171 svchost.exe -k localservice :parveenp1
List of culprits
Possible Causes
Show All
firewall :: netres172
app171 svchost.exe -k localservice :parveenp1
app186 emailmonkey.exe 9000 9000 2 1 :srikanth_test1
app171 svchost.exe -k localservice :parveenp1
parveenp1
app250 sqlmonkey.exe 18000 18000 2 1 :parveenp1
Servers app 250 [parveenp1]
app267 sqlservr.exe -smssqlserver :netres172
netres172
app224 edgetransport.exe -pipe:1572 -stopkey:global\exchangesto
Clients app 224 [netres172]
app186 emailmonkey.exe 9000 9000 2 1 :srikanth_test1
Performance counters
Rank by Abnormality
Group by Category
app186 emailmonkey.exe 9000 9000 2 1 :srikanth_test1
Process :: page faults/sec 1.00
51.38
24.02
-53.25%
00.00
91.74
Process :: % processor time 0.99
00.64
00.72
13.42%
00.00
01.08
Process :: % privileged time 0.99
00.13
00.13
3.07%
00.00
00.30
err: emailMonkey :: numberofsmtpmessagessentok 0.98
05.21
04.80
-7.95%
00.00
07.00
Process :: % user time 0.95
00.51
00.59
15.96%
00.00
00.91
APP_PKTS_IN 0.90
284.39
258.00
-9.28%
04.00
408.00
APP_BYTES_IN 0.86
790,610.04
668,292.20
-15.47%
117.00
1,149,035.00
APP_PKTS_OUT 0.54

Performance Counter View
Rank by Abnormality
Group by Category
app186 emailmonkey.exe 9000 9000 2 1 :srikanth_test1
Process :: page faults/sec 1.00
51.38
24.02
-53.25%
00.00 91.74
Process :: % processor time 0.99
00.64
00.72
13.42%
00.00 01.08
Process :: % privileged time 0.99
00.13
00.13
3.07%
00.00 00.30
err: emailMonkey :: numberofsmtpmessagessentok 0.98
05.21
04.80
-7.95%
00.00 07.00
Process :: % user time 0.95
00.51
00.59
15.96%
00.00 00.91
APP_PKTS_IN 0.90
284.39
258.00
-9.28%
04.00 408.00
APP_BYTES_IN 0.86
790,610.04
668,292.20
-15.47%
117.00 1,149,035.00
Historical avg
Current avg
% change
Statistical abnormality
Historical "training" values
Current values

Qualitative User Study

- Participants: 10 graduate students + 1 system engineer working on computer networks or operating systems
- Data: real environment with faults injected
  - Ground truths known
- NetClinic: suggest top 5 most likely causes
  - True culprit inside these five 50% of the time
- Training: 4 machines, 243 nodes, 683 links
- Test: 7 machines, 682 nodes, 2045 edges
- Video-taped, think-aloud protocols, semi-structured interviews

Tasks

- Given a reported problem, use NetClinic to find out the network component that most likely caused the problem.

	Symptom of Fault	Causes
Training	The email client on a machine is experiencing some errors	The client's configuration is broken
	Some SQL clients are experiencing poor performance	Another client is overloading the server
Testing	An email client cant get up-to-date data from server	The remote drive is dismounted
	Some users were unable to access a specific feature of a Web-based application	The firewall along the path was blocking https traffic
	Some clients cannot connect to the database serve	A port used by the problematic clients had been blocked by a change in firewall rules on the server machine

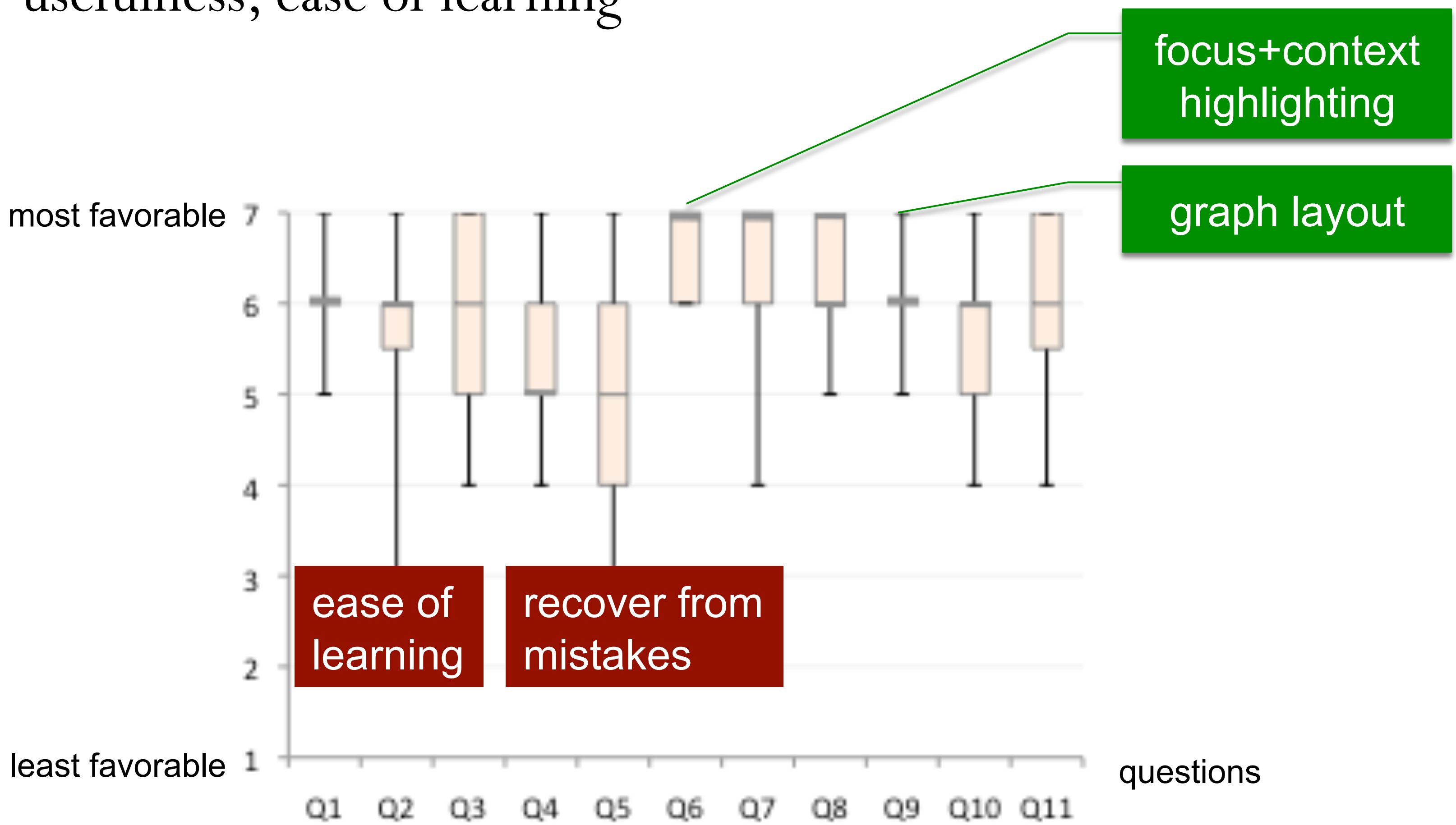
Results

- True culprits correctly identified in 29 out of 33 tasks (88%)
  - Culprits in top five suggestions 50% of the time

- Completed all 3 tasks within 1 hour

Survey

- subjective opinions on graph layout, visual design, usefulness, ease of learning



Flexibility in Exploration Strategies

- Most did not adopt a "least-effort" strategy
  - Verify all five suggestions before start self-exploration

- Using one diagnosis as entry point to learn about the problem

- Generate and verify frames, use automated diagnoses to make sure nothing was overlooked

- Not using network level diagnosis at all

Related Work

- Security monitoring / intrusion detection in computer networks
  - [Erbacher et al. 2002, Mansmann et al. 2007]
  - Tasks are different from fault diagnosis
- Visualization-based network diagnosis
  - SCUBA, nCompass, and MTreeDX
  - Mostly visualizing raw data
- Visual analytics in relationship networks
  - E.g., social networks [Social Action, 2006]

Contributions

- Coupling visualizations with a *sophisticated* reasoning engine
  - Integrated automated analyses across multiple levels
  - Explicit design consideration of sensemaking processes
  - A novel semantic graph layout design

Future Directions

- Scalability
  - Integrating machine-level diagnosis
- More evaluation
  - Long term study with professional administrators

Thank you

Questions?