Interoperability of Future Information Systems

ONR Grant Number: N00014-02-1-0499

Daniel Siewiorek, Katia Sycara (PI’s)
Joseph Giampapa, Ritika Sanghi,
Aaron Steinfeld

School of Computer Science
Carnegie Mellon University
Outline

• Motivation
• Research Approach
• Taxonomy Findings
• Agent Development Process
• What’s Next
Motivation

• Resolving network interoperability problems is difficult and time consuming
  – heterogeneity, admin policies, etc
• Advances in network flexibility will improve underlying performance
  – New HCI methods and tools will be required to enhance user awareness and problem resolution
Outline

- Motivation
- **Research Approach**
- Taxonomy Findings
- Agent Development Process
- What’s Next
Research Questions

• How does the user diagnose and remedy network interoperability problems?

• What options exist given the obstacles imposed by intermediary policies?
Research Plan

• Generate taxonomy of remote access interoperability problems
• Define agent interactions with existing network tools and formulate service profiles for future tools
• Develop agents for the resolution of network interoperability problems
Interoperability Problem Resolution Model (IPRM)
## Interoperability Problem Resolution Model - 1

<table>
<thead>
<tr>
<th>Model</th>
<th>Current</th>
<th>Research Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify the problem</td>
<td>Recognition of a real problem rather than a temporary event</td>
<td>Proactive monitoring and problem avoidance, periodic state maintenance, and formulation of constraining hypotheses</td>
</tr>
<tr>
<td>Collect Parameters and Symptoms</td>
<td>Manual, often requiring the user and expert to negotiate common definitions, mental models, and actions</td>
<td>Automatic and proactive, the ability to provide an intelligently structured “parameter dump”</td>
</tr>
</tbody>
</table>
# Interoperability Problem Resolution Model - 2

<table>
<thead>
<tr>
<th>Model</th>
<th>Current</th>
<th>Research Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve the Problem</td>
<td>a) Human reasoning, often reduced to previously solved problem, <em>or</em></td>
<td>A suite of autonomous and semi-autonomous actions:</td>
</tr>
<tr>
<td></td>
<td>b) Human reasoning augmented with Case Based Reasoning tools</td>
<td>a) Agent negotiation and suggestions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Middle agent interactions with network tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Agent-guided user actions</td>
</tr>
<tr>
<td>Learn</td>
<td>Cheat sheets and, rarely, failure analysis</td>
<td>Sharing of solutions between end-user and help desk agents</td>
</tr>
<tr>
<td>Re-use Knowledge</td>
<td>Knowledge base, FAQs, bug reports</td>
<td>Dissemination, indexing of problems and solutions</td>
</tr>
</tbody>
</table>
Outline

• Motivation
• Research Approach
• **Taxonomy Findings**
• Agent Development Process
• What’s Next
Taxonomy Findings

- SCS remote access trouble ticket case data for 6/5/2000 - 1/15/2003
- 528 Cases
  - Help only: 414, of these...
    - Single configuration events ("Single"): 88
    - Requests for modem numbers: 137
Case Flow

![Graph showing case flow over time]

- **X-axis**: Arrival Date
- **Y-axis**: Case (x 1,000)

The graph illustrates the increase in case flow over time, with a significant rise towards the end of the period.
Analysis Set 1

- 414 Help cases without outliers:
  - Zero or null Hours to Resolve: 12
  - Over 1,000 Hours to Resolve (notes): 4
By Type

- Phone number queries (significant)
- Network problems consume time fast

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>N</th>
<th>Hours to Resolve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Core</td>
<td>69</td>
<td>58</td>
</tr>
<tr>
<td>Network</td>
<td>45</td>
<td>77</td>
</tr>
<tr>
<td>Leaf</td>
<td>66</td>
<td>60</td>
</tr>
<tr>
<td>Single</td>
<td>86</td>
<td>52</td>
</tr>
<tr>
<td>Phone Number</td>
<td>132</td>
<td>27</td>
</tr>
<tr>
<td>Overall</td>
<td>398</td>
<td>49</td>
</tr>
</tbody>
</table>
By Operating System

- Macs: quicker & less variable
- Mixed OS & Unknown (significant)

<table>
<thead>
<tr>
<th>Operating System</th>
<th>N</th>
<th>Hours to Resolve</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
<td>Sum</td>
</tr>
<tr>
<td>Win</td>
<td>123</td>
<td>52</td>
<td>113</td>
<td></td>
<td>6,450</td>
</tr>
<tr>
<td>Mac</td>
<td>31</td>
<td>47</td>
<td>62</td>
<td></td>
<td>1,444</td>
</tr>
<tr>
<td>Li/unix (no Mac)</td>
<td>25</td>
<td>56</td>
<td>97</td>
<td></td>
<td>1,393</td>
</tr>
<tr>
<td>Unknown</td>
<td>211</td>
<td>43</td>
<td>101</td>
<td></td>
<td>9,036</td>
</tr>
<tr>
<td>Mixed</td>
<td>8</td>
<td>139</td>
<td>304</td>
<td></td>
<td>1,109</td>
</tr>
<tr>
<td>Overall</td>
<td>398</td>
<td>49</td>
<td>110</td>
<td></td>
<td>19,431</td>
</tr>
</tbody>
</table>
Duration, by Operating System

- Mac
- Li/unix (no Mac)
- Win
- Unknown
- Mixed

Hours to resolve

Percent Unresolved

Carnegie Mellon
Analysis Set 2

- 414 Help cases without outliers or phone number requests:
  - Zero or null Hours to Resolve: 12
  - Over 1,000 Hours to Resolve: 4
  - Requests for modem numbers: 132
Duration, by Type

- **Core**
- **Leaf**
- **Network**
- **Single**

- Hours to resolve:
  - 0
  - 168
  - 336
  - 504
  - 672
  - 840
  - 1008

- Percent Unresolved:
  - 0.0
  - 0.2
  - 0.4
  - 0.6
  - 0.8
  - 1.0

Carnegie Mellon
Modes: DSL, Modem, Wireless

- Combined are usually requests for same IP # in both modes
- No significant effects

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Hours to Resolve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Other</td>
<td>70</td>
<td>73</td>
</tr>
<tr>
<td>Modem</td>
<td>119</td>
<td>54</td>
</tr>
<tr>
<td>Wireless</td>
<td>26</td>
<td>50</td>
</tr>
<tr>
<td>DSL</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>DSL, Modem</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>DSL, Wireless</td>
<td>3</td>
<td>314</td>
</tr>
<tr>
<td>Overall</td>
<td>266</td>
<td>60</td>
</tr>
</tbody>
</table>
Security Policies: VPN, Realm

- 41% cases & 47% time involved either VPN or other security, authentication, or registration issues
- VPN and VPN*Realm (significant)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Hours to Resolve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>None</td>
<td>158</td>
<td>53</td>
</tr>
<tr>
<td>Realm</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>VPN</td>
<td>42</td>
<td>104</td>
</tr>
<tr>
<td>VPN, Realm</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>Overall</td>
<td>266</td>
<td>60</td>
</tr>
</tbody>
</table>
Very Little Knowledge Re-use

- Root Cause or Solution either
  - Not found
  - Not documented
- No significant effects

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Hours to Resolve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean Std Dev Sum</td>
</tr>
<tr>
<td>Fully Documented</td>
<td>102</td>
<td>56 108 5,681</td>
</tr>
<tr>
<td>Unknown Solution</td>
<td>35</td>
<td>41  99 1,435</td>
</tr>
<tr>
<td>Unknown Root Cause</td>
<td>34</td>
<td>58  93 1,972</td>
</tr>
<tr>
<td>Both Unknown</td>
<td>95</td>
<td>72 141 6,820</td>
</tr>
<tr>
<td>Overall</td>
<td>266</td>
<td>60 118 15,908</td>
</tr>
</tbody>
</table>

Carnegie Mellon
Taxonomy Findings Summary

• 22% from configuration changes
• 49 hrs/case for all help related
  – 60 hrs/case for subset not including phone number requests
• Security policy issues are frequent
• Very little knowledge sharing/re-use
  – Extracted by hand, rarely in existing database fields
Outline

- Motivation
- Research Approach
- Taxonomy Findings
- **Agent Development Process**
- What’s Next
Agent Development Process

- Model the Problem Domain
- Map Agents and Service Descriptions to:
  - Interoperability Problem Resolution Model, and
  - Problem Domain
- Implement, Deploy, Test, Evaluate
- Automatic Process Refinement
The Problem Domain Model

• Multiple Views and Options Intersect
  – Connectivity Model
  – Connectivity Security Model
  – Security Programs and Features
    • VPN, SSH, SCP, Kerberos

• Typical Motivating Applications
  – Interact with the above 3 models
  – Multiple ways to achieve application goals
  – Users get lost in the intersections
Typical Motivating Applications

• E-mail
  – Send and receive
  – From: on- or off- campus

• Intranet Quality of Service (QoS)
  – Institute-wide access
    • Printing, e-service subscriptions
    • Software licenses, downloads and updates
  – Bandwidth/speed

• File Transfer & File System Access
Mappings

• Each Connection between nodes:
  – Indicates a possibly new authentication step
  – Puts the user in a new application and access rights context
  – Potentially grants the user new privileges
  – Can be modeled as a service description

• Each Node:
  – Has (potentially) verifiable configuration parameters
  – Can be modeled as an agent

• A User’s Connectivity Problem
  – Possibly parameterized by goals of using certain applications
  – Resolved automatically by agent / service description matching
System Architecture

- Agents will have models of application, connectivity, and security tasks
- Agents will shadow local and remote applications
- Agents will also interact with SysAd agents for updates and policy changes
Service Descriptions

- Service profile: represents what a service does
- Service model: describes how a service works
- Service grounding: specifies service access information
Functional Architecture

User 1 > Interface Agent 1

Goal and Task Specifications

Tasks

Interface Agent 1 > Task Agent 1

Info & Service Requests

Advertisements

Info Agent 1 > Info Source 1

Queries

User 2 > Interface Agent 2

Information Integration Conflict Resolution

Middle Agent 2

Info Agent 2 > Info Source 2

Answers

User u > Interface Agent i

Results

Solutions

Task Agent i > Task Agent t

Replies
Four parallel threads:
• Communicator
  • for conversing with other agents
• Planner
  • matches “sensory” input and “beliefs” to possible plan actions
• Scheduler
  • schedules “enabled” plans for execution
• Execution Monitor
  • executes scheduled plan
  • swaps-out plans for those with higher priorities
### MAS Infrastructure

<table>
<thead>
<tr>
<th>MAS Interoperation</th>
<th>Interoperation Modules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation Services</td>
<td>Capability to Agent Mapping</td>
</tr>
<tr>
<td>Interoperator Services</td>
<td>Middle Agent Components</td>
</tr>
<tr>
<td><strong>Capability to Agent Mapping</strong></td>
<td><strong>Name to Location Mapping</strong></td>
</tr>
<tr>
<td>Middle Agents</td>
<td>Agent Name Name Service</td>
</tr>
<tr>
<td><strong>Name to Location Mapping</strong></td>
<td><strong>Security</strong></td>
</tr>
<tr>
<td>Agent Name Name Service</td>
<td>Certificate Authority</td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td>Cryptographic Service</td>
</tr>
<tr>
<td></td>
<td>Performance Service Modules</td>
</tr>
<tr>
<td></td>
<td><strong>Multi-Agent Management Services</strong></td>
</tr>
<tr>
<td></td>
<td>MAS Monitoring</td>
</tr>
<tr>
<td></td>
<td>Reputation Services</td>
</tr>
<tr>
<td><strong>Multi-Agent Management Services</strong></td>
<td><strong>Performance Services</strong></td>
</tr>
<tr>
<td>Logging</td>
<td>Performance Service Modules</td>
</tr>
<tr>
<td>Activity Visualization</td>
<td>Management Services</td>
</tr>
<tr>
<td>Launching</td>
<td>Logging and Visualization Components</td>
</tr>
<tr>
<td></td>
<td><strong>ACL Infrastructure</strong></td>
</tr>
<tr>
<td></td>
<td>Parser, Private Ontology, Protocol Engine</td>
</tr>
<tr>
<td><strong>ACL Infrastructure</strong></td>
<td><strong>Communication Modules</strong></td>
</tr>
<tr>
<td>Public Ontology</td>
<td>Discovery</td>
</tr>
<tr>
<td>Protocol Servers</td>
<td>Message Transfer Modules</td>
</tr>
</tbody>
</table>

### Operating Environment

- Machines
- OS
- Network
- Multicast Transport Layer
- TCP/IP
- Wireless
- Infrared
- SSI
Outline

• Motivation
• Research Approach
• Taxonomy Findings
• Agent Development Process
• What’s Next
What’s Next

• Implement proof of concepts
• Monitoring agent that collects parameter settings during problem solving and stores them in a centralized location
• Implement resolution models
• Quantitative analysis of resolution agent use