USING THE ECLIPSE PLUGIN DEVELOPMENT ENVIRONMENT (PDE) TO IMPROVE THE USABILITY OF BOA: A DOMAIN-SPECIFIC LANGUAGE AND INFRASTRUCTURE FOR ULTRA-LARGE-SCALE SOFTWARE REPOSITORY DATA MINING.

SAMBHAV P. SRIRAMA
AGENDA

• Status Quo
• What is Boa?
• Barriers
• The Solution
• My Specific Work
STATUS QUO

• Online repositories
• Collaborative spaces
• Revision-control systems
WHAT IS BOA?

• Domain-specific language
  • Query-language for non-experts

• Data mining infrastructure
  • Contains software projects
BARRIERS

• Three main barriers to entering this field:
  • Requires expert knowledge of data mining
  • Requires an infrastructure to data mine (which can be expensive to construct)
  • Can take days to execute due to inefficiencies
THE SOLUTION

• Boa offers:
  • Infrastructure
    • Efficiently executes jobs
    • Scalable
  • Web-based UI and compiler
  • Easily learnable query language
    • Designed with higher level abstractions
BOA WEB APP

Web-based UI and compiler

• Lacks ability to save programs and resume later
• Vulnerable to browser crashes

Requires network connection which may not always be available
THE ECLIPSE PLUGIN FOR BOA

Authentication

• Boa credentials can be loaded into Eclipse persistent storage
• Credentials are encrypted for security with built in Eclipse master key

Boa Programs

• Java HTTP frameworks/facilities
• Boa Web API
THE ECLIPSE PLUGIN FOR BOA

Job management: 3 views

• All Jobs
• Job Details
• Job Output

Additional Features:

• Source code
• File and Project
FUTURE WORK

• Boa program example templates
• Local error reports for submitted jobs
• Improved syntax checking and content assist
• Output modeling and visualization (in progress)
• Formal Eclipse facilities for bug reporting
QUESTIONS?