Usage Behavior Mining Based Recommendation Engine

Principal Investigator: Dr. Jaideep Srivastava
Students: Kalyan Beemanapalli, Amit Bose
Intel Collaborators: Dr. Sigal Louchheim

Publications: Incorporating Concept Hierarchies into Usage Mining

Research Sponsored by Intel IT Research
Problem Statement

- Improving user satisfaction and reducing call center escalation from a self-help web site, by means of providing recommendations based on mining web usage data for behavioral patterns.

- In addition to technology, the problem has other dimensions such as privacy of usage behavior, user acceptability of automated recommendations, organizational acceptance of such techniques, etc.
Example Concept Hierarchy

Student Services
(I = 0)

- Advising (I = ...)
  - Returning after absence (I = ...)
    - ...

- Registration (I = 2.17891)
  - Pre-registration (I = 5.29699)
    - ...

- How-to Guides (I = 4.5362)
  - Credit Requirements (I = 4.9578)
    - Grading Options (I = ...)
      - ...

- Career Services (I = ...)

13-creditpolicy.htm

Figure 4: Annotated concept hierarchy for student-services example
Algorithm Overview

Web Logs

Session Identification

Sessions

Session Similarity

Graph Partitioning

Session Clusters

Clickstream Trees

Recommendations

Ranking of Recommendations

Website

Hierarchy Generation

Concept Hierarchy
System Architecture

- Web Client
  - HTTP Request
    - Circuit Web-server (IIS)
      - ASP.NET DLL
        - ISAPI Filter
          - Circuit Content Management Server (MCMS)
            - URL modification
              - Placeholder Content
                - Recommendation in HTML Format
                  - RS Placeholder Control
                    - Service Contract
                      - Request Parameters (XML document)
                        - Recommendations (XML document)
                          - Recommendation Service Provider
                            - Recommendation Generator
                              - Online
                                - Usage Analyzer
                                  - Recommendation Store
                                    - Offline
                                      - Usage Store
                                        - Common Usage Format
                                          (Contract)
                                            - Log Converters
                                              - W3C Extended
                                                Converter
                                                  - IIS Log
                                                    Converter
                                                      - NCSA
                                                        Common
                                                          Converter
                                                            - Apache
                                                              Log
                                                                Converter
                                                                  - Telephone
                                                                    Call Record
                                                                      Converter
                                                                        - W3C
                                                                          Log
                                                                            - IIS Log
                                                                              - NCSA
                                                                                Log
                                                                                  - Apache
                                                                                    Log
                                                                                      - Phone
                                                                                       Record
Status and Future Plans

**STATUS**

- Concept Hierarchy
  - Page Similarity Based on Concept Hierarchy

**FUTURE PLANS**

- Site Connectivity
  - Site Topology
    - Page Similarity Based on Site Topology
  - Semantic Hierarchy
    - Page Similarity Based on Semantic Hierarchy

Clusters of User Sessions

Web Logs

Online Phase of the Recommendation Engine