Searching the World Wide Web in Low Connectivity Communities

Libby Levison
Massachusetts Institute of Technology
Web search

- Alternative information sources exist
- Web search replaces other info sources
- Existing telecomms infrastructure

Optimized for fast delivery
Search: low-connectivity communities

- Few sources of information
- Web could provide information access
- But - little infrastructure:
  - Telephone
  - Electric
Costs

High-bandwidth search in low-connectivity settings

- Cost of infrastructure
- ISP access fees
- Telephone fees

Is the goal still fast delivery?
Proposal: email-based search

- User searches on local machine
- If information not found, email query to Server on Internet.
- Server performs search, returns results
- Results downloaded from ISP
- User reads results off-line
Talk Outline

• Introduction
• TEK Implementation
• Rationale
• TEK for the Global Community
• Discussion
TEK Design Goals

• Reduce dependence on telecomms
  - Low & intermittent connectivity
  - Low-bandwidth
• Reduce number of Internet searches
• Similar to existing search engines
• User friendly
Decouple search from Internet access:

1. TEK Client: user interface
2. Email protocol: transfers info via email
3. TEK Server: finds, returns results
Demo
TEK Protocol

- Email *not* reliable: messages get lost
- TEK Protocol tracks all messages
TEK Protocol

- Runs on both TEK Client and TEK Server
- Request-reply model
- If no reply, automatic retransmission
- Other functions:
  - Client Registration
  - Administration messages
• Receives information request
  - General search, page fetch
• Finds & returns information
• Provides front page navigation tool
TEK Server

Goals: 1. Reduce bandwidth
   2. Send best content

- Remove duplicate pages
- Remove extraneous code
- Select authority pages over hubs
TEK Server: Track Client state

- Record all client search requests
- Record all URLs sent to client
- Checks all candidate URLs, does not send previously sent URLs
TEK: Summary

- Low-connectivity
- Low-bandwidth
- User friendly
- Similar to existing search engines
- Manageable amount of information
Talk Outline

• Introduction
• TEK Implementation
• Rationale
• TEK for the Global Community
• Discussion
Rationale I: Decreased Cost

• Connection time is shorter
• Email-only accounts less expensive
• Call ISP at will:
  Calls cheaper, lines less-noisy, off peak
• Local web cache prevents some searches
## Email and Internet rates

<table>
<thead>
<tr>
<th>Country</th>
<th>Email only (month)</th>
<th>Internet (month)</th>
<th>Extra/Discounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armenia</td>
<td>$8.50</td>
<td>$42</td>
<td>Night discount: 50%</td>
</tr>
<tr>
<td>Arminico</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malawi</td>
<td>$15</td>
<td>$30</td>
<td></td>
</tr>
<tr>
<td>E &amp; O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>$11</td>
<td>$15</td>
<td>Extra hours: $1.32 peak</td>
</tr>
<tr>
<td>LankaNet</td>
<td></td>
<td></td>
<td>$0.88 off peak</td>
</tr>
</tbody>
</table>
Email as communication medium:

- Never need continuous path from client to server
Rationale II: Improved Reliability

Email as communication medium:
- Never need continuous path from client to server
- Reduce dependence on global network
Rationale II: Improved Reliability

Email as communication medium:
• Never need continuous path from client to server
• Reduce dependence on global network
• Store and forward
Rationale II: Improved Reliability

Email as communication medium:
- Never need continuous path from client to server
- Reduce dependence on global network
- Store and forward
- Asynchronous
Rationale III: Improved Convenience

• View results offline: quick, reliable
  - Build local URL library
  - More people can use computer
• Send email during off peak hours
• Filtered results returned by TEK Server
• Manageable amount of information
Talk Outline

• Introduction
• TEK Implementation
• Rationale
• TEK for the Global Community
• Discussion
TEK for the Global Community

• Extends Web search to email-only users
• Extend still farther:
  collect outgoing emails on diskette,
  take diskette to Internet
Applications: DakNet

Fletcher & Hasson, 2002
Percentage population online

2001 data

Africa: 0.7%
Asia: 5%
Latin America: 6%
Europe: 14%
US & Canada: 64%
Appropriate Information Technology

- “Technology that fits”
- Need to understand:
  - What technology is possible
  - What technology, infrastructure exist
  - Cultural context
- Numerous technical challenges, if we can find them
Talk Outline

• Introduction
• TEK Implementation
• Rationale
• TEK for the Global Community
• Discussion
Open Issues

- TEK: Web search only
  Access must be bi-directional
- Different cultural information needs
- Summarization
- Send graphics on request
Summary

• TEK: email-based Search engine
  - More affordable
  - More reliable
  - More convenient
• Extends access to Web search
• TEK: Appropriate Info Technology
• Alpha deployment this summer
TEK Participants

- William Thies
- Janelle Prevost
- Tazeen Mahtab
- Genevieve Cuevas
- Saad Shakhshir
- Bihn Vo
- Alexandro Artola
- Yuliya Litvak
- Sheldon Chan
- Sid Henderson
- Mark Halsey
- Damon Berry
- Libby Levison
- Saman Amarasinghe