A distributed infrastructure supporting personalized services for the Mobile Web

Claudia Canali
Michele Colajanni
Riccardo Lancellotti

University of Modena and Reggio Emilia

Philip S. Yu
IBM T.J. Watson Research center
The Mobile Web scenario

Mobile clients
Personalized services

Web-based services tailored to
Client device
User preferences
User context

User preferences stored in a user profile
User profile contains sensitive information

We should guarantee the privacy of the user profile
Performance vs. Privacy

Need for high performance
- Computationally expensive Web-based services
- Use of highly distributed infrastructure

Need for privacy
- High security levels for the nodes of the infrastructure
- Centralized, controlled infrastructure
- Using a highly distributed infrastructure may be unfeasible (every node should be secure)

Fully centralized and fully distributed infrastructures are not viable
Intermediate infrastructure

Core node
Powerful (cluster)
High security level

Edge nodes
Highly distributed
Security level are relaxed

Observation:
Different mobile services
Different privacy requirements

Request dispatching
Privacy awareness
Do not forget performance
Request dispatching process

Web resource
- Multiple Web resource components

Privacy
- Each component has a different privacy requirement
  - Strong = *must* be processed on core node
  - Light = *should* be processed on core node
  - None = can be processed anywhere

Performance
- Considers CPU utilization
- Load sharing: avoids overload
The request dispatching process
Algorithm idea:

Use of a threshold

If utilization is below the threshold assign components according to privacy requirements

If utilization of a node is beyond the threshold, less components are assigned to that node

For performance reasons we can relax privacy requirements

Dispatching mismatch

Should minimize dispatching mismatch while preserving adequate performance
Other algorithms

Performance-oriented
- Does not consider privacy while dispatching requests

Privacy-oriented
- Does not consider performance
Experimental testbed

Personalized Web portal
- RSS feed aggregation
- Banner insertion
- Content adaptation

Different workload mixes
- Components with Light privacy (PL-components)
  - PL-components = 10%, 40%, 70%

System
- 16 servers
- 1 core node, 4 edge nodes

WAN emulation
- Based on netem packet scheduler
- Network delay + BW limitation
Experimental results: Performance

PL = 10%
PL = 40%
PL = 70%
Experimental results: Privacy

Dispatching mismatch for the algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>PL=10%</th>
<th>PL=40%</th>
<th>PL=70%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance and Privacy</td>
<td>2.5%</td>
<td>4.6%</td>
<td>14.7%</td>
</tr>
<tr>
<td>Performance -oriented</td>
<td>11.5%</td>
<td>25.2%</td>
<td>43.8%</td>
</tr>
<tr>
<td>Privacy-oriented</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

The performance and privacy algorithm preserves privacy better than the performance oriented alternative.
Summary of experiments

**Performance-oriented**
High mismatch

**Privacy-oriented**
Poor performance

**Performance and Privacy**
Combines the goals of performance and privacy

![Bar chart showing response time and mismatch for Priv, Perf, and P&P with PL=70%]
Impact of WAN effects

Distributed infrastructure

Impact of Network delay

Performance degradation

Same impact on every algorithm

PL=40%
Conclusion

Performance vs. Privacy trade-off

Performance and privacy can be successfully combined through

Intermediate infrastructure based on heterogeneous nodes
Performance and Privacy-aware request dispatching algorithms

Experimental results:
Better performance than privacy-oriented algorithms
Better privacy than performance-oriented algorithms
Future research directions

More complex performance and security models
- Multiple privacy classes
- Multiple security levels

More sophisticated algorithms
- Load balancing instead of Load sharing
- Algorithms that consider the computational cost of each Web service
A distributed infrastructure supporting personalized services for the Mobile Web

Claudia Canali
Michele Colajanni
Riccardo Lancellotti

University of Modena and Reggio Emilia

Philip S. Yu
IBM T.J. Watson Research center