Trustworthy Decentralized Publication, Search and Retrieval in Heterogeneous Networks

Ph.D. Dissertation Defense by Isaí Michel Lombera
1 Introduction
2 Design of iTrust
3 iTrust over HTTP
4 iTrust over SMS
5 iTrust over Wi-Fi Direct
6 Conclusions and Future Work
1 Introduction
2 Design of iTrust
3 iTrust over HTTP
4 iTrust over SMS
5 iTrust over Wi-Fi Direct
6 Conclusions and Future Work
The convergence of research and innovation.
The convergence of research and innovation.
The convergence of research and innovation.
Server-based

P2P-network
GOVERNMENT CENSORSHIP

PROTECTING YOU FROM REALITY

Dislike
Imagine a World Without Free Knowledge

For over a decade, we have spent millions of hours building the largest encyclopedia in human history. Right now, the U.S. Congress is considering legislation that could fatally damage the free and open Internet. For 24 hours, to raise awareness, we are blacking out Wikipedia. Learn more.

Contact your representatives.

Your ZIP code: 

Look up
iTrust
iTrust Contributions

1. Trustworthy information sharing
2. HTTP for desktop, laptop searching
3. SMS extension to generalized data
4. Wi-Fi Direct enhancements
iTrust Heterogeneous Network
1 Introduction

2 Design of iTrust

3 iTrust over HTTP

4 iTrust over SMS

5 iTrust over Wi-Fi Direct

6 Conclusions and Future Work
Distribution of Metadata

Source of Information
Distribution of a Request

Source of Information

Request Encounters Metadata

Requester of Information
Retrieval of Information

Source of Information

Request Matched

Requester of Information

Isaí Michel Lombera / Ph.D. Dissertation Defense / 20 May 2013
Probability of a Match

\[ p = 1 - \left( \frac{n - mx}{n} \frac{n - 1 - mx}{n - 1} \ldots \frac{n - r + 1 - mx}{n - r + 1} \right) \]

- **p**: probability of an encounter occurring (match or hit)
- **n**: number of participating nodes
- **m**: number of nodes to which metadata is distributed
- **r**: number of nodes to which requests are distributed
- **x**: proportion of operational nodes
iTrust Performance Analysis

![Graph showing probability of a match vs. number of nodes for metadata and request distribution. The x-axis represents the number of nodes ranging from 10 to 80, while the y-axis represents the probability of a match ranging from 0 to 1. The graph includes four lines, each representing a different proportion of operational nodes: 1.0, 0.8, 0.6, and 0.4. The lines are colored blue, red, green, and cyan, respectively.]
1 Introduction
2 Design of iTrust
3 iTrust over HTTP
4 iTrust over SMS
5 iTrust over Wi-Fi Direct
6 Conclusions and Future Work
iTrust over HTTP

metadata functions
metadata xml engine
register metadata list
apply xml
publish xml list

helper functions
nodes wrapper
keywords wrapper
resource wrapper
tag keyword resource
search functions
globals / navigation

tika / lucene / dictionary

public interface

tools
user settings
statistics
metadata inbox
leave membership
delete nodes
search
query
inbox

apache
PHP
cURL
SQLite
session
log
PECL http
iTrust, node: test.isaim.com:81

Tools

Database tools
- Recreate database (and delete old database) Warning: delete all resources first!
- Create new keyword(s)
- Create new node(s)

Resource tools
- Upload a resource
- Tag keywords to resource
- Delete all resources Warning: this will delete all uploaded resources and related database entries

Metadata tools
### Insert resource

**Select a file below to upload to this node**

<table>
<thead>
<tr>
<th>Resource file:</th>
<th><code>./home/michel/apple.full.jpg</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Website address:</td>
<td></td>
</tr>
<tr>
<td>Keywords you want to save to (separate with space):</td>
<td><code>fruit, organic</code></td>
</tr>
<tr>
<td>Days you want to keep this resource (default 7 days):</td>
<td></td>
</tr>
<tr>
<td>Index keywords from content or metadata:</td>
<td>Contents OMetadata</td>
</tr>
</tbody>
</table>

[Insert button]
\[
p = 1 - \left( \frac{n-m}{n} \right) \left( \frac{n-m-1}{n-1} \right) \cdots \left( \frac{n-m-r+1}{n-r+1} \right)
\]

\[n = \text{membership nodes}\]
\[m = \text{metadata nodes}\]
\[r = \text{requested nodes}\]
\( n = \text{membership nodes} \)
\( m = \text{metadata nodes} \)
\( r = \text{requested nodes} \)
\( x = \text{operational nodes} \)

\[
p = 1 - \left( \frac{n-mx}{n} \right) \left( \frac{n-mx-1}{n-1} \right) \cdots \left( \frac{n-mx-r+1}{n-r+1} \right)
\]
1 Introduction
2 Design of iTrust
3 iTrust over HTTP
4 iTrust over SMS
5 iTrust over Wi-Fi Direct
6 Conclusions and Future Work
iTrust with SMS
iTrust with SMS App

![iTrustHwTest](image1)

Enter iTrust node telephone number

1760689

Query

Tahrir Square

Send SMS

Status: waiting to send.

![iTrustHwTest](image2)

Enter iTrust node telephone number

1760689

Query

Send SMS

Status: waiting to send.

Hit from iTrust node @ +1760689: meet near talaat harb street @ 2330.
Probability of a Match

\[ p = 1 - \left( \frac{n - mx}{n} \cdot \frac{n - 1 - mx}{n - 1} \cdots \frac{n - r + 1 - mx}{n - r + 1} \right) \]

\( p \)  probability of an encounter occurring (match or hit)

\( n \)  number of participating nodes

\( m \)  number of nodes to which metadata are distributed

\( r \)  number of nodes to which requests are distributed

\( x \)  proportion of operational nodes
Emulation vs. Analysis

250 Node Network with 100% of the Nodes Operational

250 Node Network with 60% of the Nodes Operational

Probability of a Match

Number of Nodes for Distribution of Metadata or Requests
Number of Messages for a Match

\[ p(k) = \frac{(mx \frac{mx-1}{k-1} \cdots mx-k+1)(n-mx \frac{n-mx-1}{r-k-1} \cdots n-mx-r+k+1)}{(\frac{n}{r} \frac{n-1}{r-1} \cdots \frac{n-r+1}{1})}. \]

\[ y = 2 + r + \sum_{k=1}^{\min\{mx,r\}} kp(k). \]

\( p(k) \) probability of \( k \) matches
\( k \) number of matches (\( k \) reporting matches)
\( n \) number of participating nodes
\( m \) number of nodes to which metadata is distributed
\( r \) number of nodes to which requests are distributed
\( x \) proportion of operational nodes
\( y \) mean number of messages for a match
Emulation vs. Analysis

250 Node Network with 100% of the Nodes Operational

250 Node Network with 68% of the Nodes Operational
iTrust over SMS nodes

user

iTrust over SMS

mobile network

app

signal parser

SMS receiver

SMS transmitter

node core

DB adapter

SMSC
iTrust over SMS Protocol

<table>
<thead>
<tr>
<th>Message</th>
<th>Identifier</th>
<th>Parameter 1</th>
<th>Parameter 2</th>
<th>Parameter 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEND_QUERY</td>
<td>itq</td>
<td>&lt;caller_number&gt;</td>
<td>&lt;query_id&gt;</td>
<td>&lt;query&gt;</td>
</tr>
<tr>
<td>NOTIFY_MATCH</td>
<td>itr</td>
<td>&lt;source_number&gt;</td>
<td>&lt;query_id&gt;</td>
<td>&lt;resource_id&gt;</td>
</tr>
<tr>
<td>REQUESTRESOURCE</td>
<td>itq</td>
<td>now</td>
<td>&lt;query_id&gt;</td>
<td>&lt;resource_id&gt;</td>
</tr>
<tr>
<td>SENDRESOURCE</td>
<td>itr</td>
<td>data</td>
<td>&lt;query_id&gt;</td>
<td>&lt;data&gt;</td>
</tr>
<tr>
<td>NOTIFY_METADATA</td>
<td>itm</td>
<td>&lt;source_number&gt;</td>
<td>&lt;expiry_date&gt;</td>
<td></td>
</tr>
<tr>
<td>REQUEST_METADATA</td>
<td>itm</td>
<td>pull</td>
<td>&lt;query_id&gt;</td>
<td>unused</td>
</tr>
<tr>
<td>SEND_METADATA</td>
<td>itm</td>
<td>push</td>
<td>&lt;query_id&gt;</td>
<td>&lt;data&gt;</td>
</tr>
</tbody>
</table>
Distributing Metadata

<table>
<thead>
<tr>
<th>Message</th>
<th>Identifier</th>
<th>Parameter 1</th>
<th>Parameter 2</th>
<th>Parameter 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTIFY_METADATA</td>
<td>itm</td>
<td>&lt;source_number&gt;</td>
<td>&lt;expiry_date&gt;</td>
<td></td>
</tr>
<tr>
<td>REQUEST_METADATA</td>
<td>itm</td>
<td>pull</td>
<td>unused</td>
<td></td>
</tr>
<tr>
<td>SEND_METADATA</td>
<td>itm</td>
<td>push</td>
<td>&lt;data&gt;</td>
<td></td>
</tr>
</tbody>
</table>
# Searching / Retrieving Information

<table>
<thead>
<tr>
<th>Message</th>
<th>Identifier</th>
<th>Parameter 1</th>
<th>Parameter 2</th>
<th>Parameter 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEND_QUERY</td>
<td>itq</td>
<td>&lt;caller_number&gt;</td>
<td>&lt;query_id&gt;</td>
<td>&lt;query&gt;</td>
</tr>
<tr>
<td>NOTIFY_MATCH</td>
<td>itr</td>
<td>&lt;source_number&gt;</td>
<td>&lt;query_id&gt;</td>
<td>&lt;resource_id&gt;</td>
</tr>
<tr>
<td>REQUEST_RESOURCE</td>
<td>itq</td>
<td>now</td>
<td>&lt;query_id&gt;</td>
<td>&lt;resource_id&gt;</td>
</tr>
<tr>
<td>SENDRESOURCE</td>
<td>itr</td>
<td>data</td>
<td>&lt;query_id&gt;</td>
<td>&lt;data&gt;</td>
</tr>
</tbody>
</table>

Source of Information

Requester of Information

SEND_QUERY
1: itq@15551234567@r4nd0m1d@tahrir square

NOTIFY_MATCH
2: itr@15550011223@r4nd0m1d@456

REQUEST_RESOURCE
3: itq@now@r4nd0m1d@456

SEND_RESOURCE
4: itr@r4nd0m1d@meet near talaat harb street
Probability of a Match

\[ P(k \geq 1) = 1 - \frac{\binom{n-mx}{r} \binom{n-mx-1}{r-1} \cdots \binom{n-mx-r+1}{1}}{\binom{n}{r} \binom{n-1}{r-1} \cdots \binom{n-r+1}{1}} \]

\textbf{n} \quad \text{Number of participating nodes (membership size).}

\textbf{m} \quad \text{Number of participating nodes to which metadata are distributed.}

\textbf{r} \quad \text{Number of participating nodes to which requests are distributed.}

\textbf{x} \quad \text{Proportion of the } n \text{ participating nodes that are available.}

\textbf{k} \quad \text{Number of participating nodes that report matches to a requesting node.}
\[ P(k \geq 1) > 0.9817 \quad \text{if} \quad m = r = 2\sqrt{n} \]
\[ m = r = 2 \times \sqrt{n} \]
Android Emulator (x16)
Probability of a Match

\[ n = 16 \]
\[ m = r = 8 \]

<table>
<thead>
<tr>
<th>Matches</th>
<th>Analysis</th>
<th>Observed</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000155</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>1</td>
<td>0.004974</td>
<td>0.011000</td>
<td>0.011000</td>
</tr>
<tr>
<td>2</td>
<td>0.060926</td>
<td>0.094143</td>
<td>0.105143</td>
</tr>
<tr>
<td>3</td>
<td>0.243705</td>
<td>0.315714</td>
<td>0.420857</td>
</tr>
<tr>
<td>4</td>
<td>0.380790</td>
<td>0.366796</td>
<td>0.787653</td>
</tr>
<tr>
<td>5</td>
<td>0.243705</td>
<td>0.177327</td>
<td>0.964980</td>
</tr>
<tr>
<td>6</td>
<td>0.060926</td>
<td>0.031020</td>
<td>0.996000</td>
</tr>
<tr>
<td>7</td>
<td>0.004974</td>
<td>0.004000</td>
<td>1.000000</td>
</tr>
<tr>
<td>8</td>
<td>0.000000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
</tbody>
</table>
iTrust active searches

- JPEGs of 25 December 2011
- JPEGs of last christmas
- Mobile number of John Smith
- Moonlight

New Search
New iTrust search

Search for

Search now

Mobile number of John Smith
26 Jun 2012

No documents found yet
iTrust nodes

17601234567
15551234567
15559876543

iTrust preferences

General settings

Clear memory
Clear all query, node, keywords, etc.

Reset settings
Reset all options on their default values.

Show protocol
Show the raw protocol on SMS app.

Save retrieve documents
Save retrieved documents to this device.

SMS

Limit SMSes
Limit the number of SMS messages sent from this device.

SMS limit
The maximum number of SMS messages that can be sent.
iTrust preferences

Save Retrieve documents
Save retrieved documents to this device.

SMS

Limit SMSes
Limit the number of SMS messages sent from this device.

SMS limit
The maximum number of SMS messages that can be sent.

Documents

Share contacts
Share your contacts information with others.

Share documents
Share documents with others.

Document location
Location of shareable documents on this device.
1  Introduction
2  Design of iTrust
3  iTrust over HTTP
4  iTrust over SMS
5  iTrust over Wi-Fi Direct
6  Conclusions and Future Work
iTrust over Wi-Fi Direct

Diagram:
- User
- iTrust over WiFi Direct
- OS
- App
  - Signal parser
  - Node core
  - Wi-Fi P2P service
  - DB adapter
- Inbox thread
- Outbox thread
- Wi-Fi P2P broadcast receiver
- Android/Linux
iTrust over Wi-Fi Direct Network Stacks
Wi-Fi Direct on Android
iTrust over Wi-Fi Direct Peer Management

1: NEW_PEER
2: PEER_LIST

A

1: NOTIFY_METADATA
2: REQUEST_METADATA
3: SEND_METADATA

B

1: SEND_QUERY
2: NOTIFY_MATCH
3: REQUEST_RESOURCE
4: SEND_RESOURCE

C
Data Transfer

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Small files</th>
<th>Medium files</th>
<th>Large files</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum file size (bytes)</td>
<td>188236</td>
<td>3783308</td>
<td>144715776</td>
</tr>
<tr>
<td>Maximum file size (bytes)</td>
<td>955165</td>
<td>5755928</td>
<td>162666496</td>
</tr>
<tr>
<td>Mean file size (bytes)</td>
<td>453962</td>
<td>5013612</td>
<td>151327949</td>
</tr>
<tr>
<td>Mean transfer time (s)</td>
<td>0.39</td>
<td>9.74</td>
<td>78.82</td>
</tr>
<tr>
<td>Mean transfer rate (kbytes/s)</td>
<td>2493</td>
<td>620</td>
<td>1940</td>
</tr>
</tbody>
</table>
1 Introduction
2 Design of iTrust
3 iTrust over HTTP
4 iTrust over SMS
5 iTrust over Wi-Fi Direct
6 Conclusions and Future Work
iTrust Contributions

1. Trustworthy information sharing

2. HTTP for desktop, laptop searching

3. SMS extension to generalized data

4. Wi-Fi Direct enhancements
iTrust over HTTP

iTrust over HTTP nodes

apache PHP
cURL
SQLite
session
log
PECL http

metadata functions
metadata xml engine
register metadata list
apply xml
publish xml list

helper functions
nodes wrapper
keywords wrapper
resource wrapper
tag keyword resource
search functions
globals / navigation

tika / lucene / dictionary

public interface
tools
user settings
statistics
metadata inbox
leave membership
delete nodes
search
query
inbox

144 Node Network with 100% of the Nodes Operational

Analysis
Simulation

Tools
Database tools
- creates database (and delete old database) Warning: delete all resources first!
- create_xml_conversion

iTrust, node: test.isaim.com:81
iTrust over SMS

![Diagram of iTrust over SMS nodes and HTTP nodes]
iTrust over Wi-Fi Direct

Diagram showing the architecture of iTrust over Wi-Fi Direct, with nodes connected through Wi-Fi Direct and Android/Linux interfaces.
Related Work

1. Centralized and decentralized search
2. Structured and unstructured networks
3. Trust, reputation and malicious nodes
4. Mobile search over cell / ad-hoc networks
Future Work

- **Distributed Ranking**: design / implementation
- **Distributed Reputation**: design / implementation
- **Open source release**: HTTP, SMS, Wi-Fi Direct
- **Release of iTrust**: Google Play store
Further Information
http://itrust.ece.ucsb.edu

Advisors: L. E. Moser & P. M. Melliar-Smith

Isaí imichel@ece.ucsb.edu

Research supported by NSF: CNS 10-16193