Security Through Entertainment: Using a Memory Game for Secure Device Pairing

Alexander Gallego, Nitesh Saxena, and Jonathan Voris
agalle01@students.poly.edu, nsaxena@poly.edu, jvoris@isis.poly.edu
Polytechnic Institute of New York University
Department of Computer Science and Engineering
The Problem: Device Pairing

- How can a secure channel be initialized between two wireless devices when they have:
  - No prior association or shared secret
  - No common source of trust
- Examples of device pairing scenarios:
  - Pairing a Bluetooth cell phone with a headset
  - Connecting a WLAN laptop with an access point
• **Out of band** channels can be used between devices
  - Created using human perceptible forms of output
    - Audio
    - Visual
    - Tactile
  - Physically authenticable by device owners
• But introduce a **usability burden** for users
  - Pairing tasks may be tedious and time consuming
  - Users may make **errors** while pairing
  - Process is limited by **human factors**
    - Must be slow enough for a human to perceive
    - Security is dependant on an individual’s attention span
The “Tom Sawyer” Effect

• “The Adventures of Tom Sawyer” by Mark Twain
  ▪ Tom is forced to paint a fence
  ▪ He pretends to enjoy it rather than resent it
  ▪ Friends insist on helping!
• Apply this principle to security tasks
  ▪ Reframe device pairing as a game
  ▪ Introduce entertainment value
• Competitiveness between individuals causes them to perform better at certain tasks
  ▪ Such as random number generation and recognition - Halprin and Naor [SOUPS ‘09]
To leverage the Tom Sawyer Effect, designed a game for device pairing
- Inspired by Hasbro’s Simon
- Performed a usability study with 20 participants
- Conclusion: Incentivizing users with entertainment is a promising way to address device pairing and other security issues
Coordination and Bargaining over the Gaussian Interference Channel

Xi Liu

Polytechnic Institute of NYU
Advisor: Prof. Elza Erkip
**Problem:** Performance of communication systems is often limited by interference. Interference management methods require cooperation in transmission strategies among users. In practice, **selfish** users may not have an incentive to cooperate.

The following problems are treated using tools from game theory and information theory:

- What are the incentive conditions under which users have incentives to cooperate?
- How to determine an operating point for two selfish users if they do cooperate?

**Main Result:** A two-phase mechanism for coordination is proposed.

- Phase 1: allow users to negotiate for the use of a simple Han-Kobayashi (H-K) type scheme.
- Phase 2: employ the Nash bargaining solution (NBS) from game theory to select a fair operating point.

The H-K NBS is shown to not only provide a fair operating point, but also maintains a good overall performance.
Motivations ...

1. **Cache Updates;**
   - Update the result caching periodically by issuing batches of queries.

2. **Data Mining;**
   - Various internal quality testing by issuing a lot of not-interactive queries;
   - In machine learning researchers usually issue queries to mine people’s attitude;
   - Individual group wants to pull data from a search system, as in the so called “deep web”.

3. **Query by Outside Parties;**
   - Yahoo! BOSS (Build your Own Search Service)
   - In May 2009, Yahoo! reported that about 30 million queries per day were issued through BOSS.
Ideas ...

**Improvement for Disk (better caching technique):**

(1) Query Reordering:
(2) Clairvoyance:
   Since we know all the queries, we can use clairvoyant algorithms for caching lists. It is known that for list caching this is the optimal.

**Improvement for CPU:**
Modern search engines store the index in main memory. By caching the sub queries we can improve the query processing performance.
Results ...

Disk improvement

Figure 1: Caching performance using LRU and Clairvoyant, with different query stream reordering, on our 1.16 million query stream. The x-axis is the size of the cache in Millions, and the y-axis is the size of data needs to be transferred from disk in TB.

CPU improvement

Figure 8: Performance gains with caching pairs (2-term sub-queries) using our greedy algorithm, with different cache size with respect to the inverted index size. Zero means no sub-queries are cached and re-used, 110% means the cache size is even larger than the inverted index size. We average the numbers over our 1.1 million queries.

Figure 9: Performance gain with different numbers of queries (in thousands). The benefit is showed as the sum of lists lengths.
Improved Index Compression Techniques for Versioned Document Collections

Jinru He, Junyuan Zeng, Torsten Suel, Polytechnic Institute Of NYU

Two-level Indexes by Altingovde et al 2008
- Top level index the union of all versions
- Lower level using bit vectors.

Monkey

10, 30, 34...

0 1 1 0 0 1 0 0 1 0

- The length of each bit vector is the number of versions in the document.
- For bit vector of term t, if t appears in ith version, the ith position of bitmap is set to 1 otherwise, it is set to 0

Background – Inverted Index
An inverted index consists of inverted lists. One inverted list per distinct term in collection. Each inverted list is a sequence of postings. Inverted lists are sorted by DocID and compressed

Issues in Versioned Document
- Index representation and compression
- Index traversal techniques
- Support for temporal range queries
- Aggregate query processing
- (e.g. stable top-k by Leong et al 2010)

<table>
<thead>
<tr>
<th></th>
<th>Wikipedia</th>
<th>Ireland Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td># documents</td>
<td>2,401,799</td>
<td>1,059,670</td>
</tr>
<tr>
<td>Versions/Doc</td>
<td>35.54</td>
<td>15.78</td>
</tr>
<tr>
<td>Total Postings</td>
<td>5.82 Billion</td>
<td>3.42 Billion</td>
</tr>
</tbody>
</table>
Improved Index Compression Techniques for Versioned Document Collections

Our Solutions
Bitwise Approach:
1. Model Approach:
   Given the information of known versions in a document, we derive models to predict what it is like in the next versions.
2. Featured Approach:
   Exploit features, machine learning tools to do prediction
Practical Approach:
   - Transforming DIFF and MSA into two-level methods
   - Bit vector reordering based on amount of change
   - Hybrid of DIFF and MSA

Query Processing
   - Bit vectors corresponding to the result docIDs are fetched
   - AND the bit vector
   - First level index is small, many bit vectors can be skipped to speed up the second level query processing!

Table 3. Query processing on Wikipedia on all methods
   Actually, we can do it even faster than previous work!

<table>
<thead>
<tr>
<th></th>
<th>Wiki</th>
<th>Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous Result</td>
<td>140</td>
<td>304</td>
</tr>
<tr>
<td>Model upper bound</td>
<td>92</td>
<td>254</td>
</tr>
<tr>
<td>Model lower bound</td>
<td>88</td>
<td>249</td>
</tr>
<tr>
<td>Feature-Based</td>
<td>77.5</td>
<td>209</td>
</tr>
</tbody>
</table>

Our combinatorial approach:
   Wikipedia: up to 50% times smaller.
   Ireland: up to 30% smaller

<table>
<thead>
<tr>
<th></th>
<th>Wiki</th>
<th>Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>docID</td>
<td>freq</td>
</tr>
<tr>
<td>HUFF</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2R-DIFF-IPC</td>
<td>214</td>
<td>86</td>
</tr>
<tr>
<td>2R-DIFF-PFD</td>
<td>242</td>
<td>98</td>
</tr>
<tr>
<td>2R-MSA-IPC</td>
<td>197</td>
<td>41</td>
</tr>
<tr>
<td>2R-MSA-PFD</td>
<td>212</td>
<td>51</td>
</tr>
<tr>
<td>2R-HUFF</td>
<td>210</td>
<td>64</td>
</tr>
<tr>
<td>2R-HUFF combined</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Our practical approach:
   Wikipedia: up to 30% times smaller.
   Ireland: up to 5% smaller

Table 2. Compression Size in MB for Wikipedia and Ireland Dataset by applying different proposed methods
“Hence if too much salt is used for food, the pulse hardens ...”

---- An Ancient Chinese Medicine Book, 2698–2598 BC.
Closed-Loop Control System

- Kidneys
- osmoreceptor (sensor)
- Extracellular Fluid (state)
- Hypothalamus (actuator)

- optimal controller (LQR)
- open loop model

Perfectly matches the experimental data !!!
Allocation of Distributed Generations (DG) in a Power Network

Students: Po-Chen Chen, Qincheng Zhu
Advisers: Profs. Z. P. Jiang, F. De Leon, D. Czarkowski
Control and Telecommunications Research Laboratory, NYU Poly

Project Objectives

I. Develop possible scenarios to distribute DGs into a network.
II. Consider both sizing and placement of DGs.
III. Study the impact of DGs to the utility network.
IV. Key Tools: Markov Chain Monte Carlo method and Gibbs sampling algorithm.
Batch Query Processing for Web Search Engines

Shuai Ding
Joint work with
Josh Attenberg, Ricardo Baeza-Yates, Torsten Suel
Power Network of Our Study

Figure: All available locations (loads) and transformers with network protectors
EMTP Results and Their Analysis

<table>
<thead>
<tr>
<th>Voltage of Loads in p.u.</th>
<th>NETWORK PROTECTOR OPEN</th>
<th>DG PROTECTION OPEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak value over ±5% limited valuation</td>
<td>Max. Value</td>
<td>Min. VALUE</td>
</tr>
<tr>
<td>20% - Light</td>
<td>0</td>
<td>1.041</td>
</tr>
<tr>
<td>30% - Light</td>
<td>0</td>
<td>1.020</td>
</tr>
<tr>
<td>40% - Light</td>
<td>0</td>
<td>1.021</td>
</tr>
</tbody>
</table>

Figure : EMTP results
A Framework for Femtocells to Access Both Licensed & Unlicensed Bands

Feilu Liu and Elza Erkip

• Cellular operators are offloading data traffic from licensed spectrum to unlicensed bands
  o E.g., AT&T owns 23,000+ WiFi hotspots
• Femtocells are being deployed to increase cellular capacity
• Femtocells can be a perfect platform to access both licensed & unlicensed bands, and offload traffic from licensed to unlicensed bands in a smart way.
  o Interference level in licensed bands
  o Channel usage in unlicensed bands
• We propose that femtocells opportunistically access unlicensed bands on top of licensed cellular bands.
Bayesian-inference Based Recommendation in Online Social Networks

Xiwang Yang, Yang Guo*, and Yong Liu

* Technicolor (Thomson)
Bayesian Recommendation in OSN

- CF-based approach
  - people who (dis)agreed in the past tend to (dis)agree again in the future

- Exploiting Structure of OSN
  - measure similarity between direct friends by conditional rating probability distributions
  - predict local rating from ratings of direct and indirect friends
  - address Rating Sparsity, good coverage for all users
    - set prior distr. based on user input
  - address Cold Start, good recommendation for new users

- Evaluation
  - beat MoleTrust and TidalTrust by 10% in MAE on Epinions dataset.

- Distributed Protocol Design, Prototype in Facebook
Optimal Distributed Voltage Control in Smart Grid

Student: Li Yu, Mingyu Li
Advisor: Dariusz Czarkowski
Smart Grid with Smart Controllers and Distributed Energy Sources
Control Algorithm

\[ S = \begin{bmatrix}
S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} & S_{17} & S_{18} & \ldots & S_{1n} \\
S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} & S_{27} & S_{28} & \ldots & S_{2n} \\
S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} & S_{37} & S_{38} & \ldots & S_{3n} \\
S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46} & S_{47} & S_{48} & \ldots & S_{4n} \\
S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56} & S_{57} & S_{58} & \ldots & S_{5n} \\
S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66} & S_{67} & S_{68} & \ldots & S_{6n} \\
S_{71} & S_{72} & S_{73} & S_{74} & S_{75} & S_{76} & S_{77} & S_{78} & \ldots & S_{7n} \\
S_{81} & S_{82} & S_{83} & S_{84} & S_{85} & S_{86} & S_{87} & S_{88} & \ldots & S_{8n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\
S_{m1} & S_{m2} & S_{m3} & S_{m4} & S_{m5} & S_{m6} & S_{m7} & S_{m8} & \ldots & S_{mn}
\end{bmatrix} \Rightarrow \tilde{S} = \begin{bmatrix}
S_{a1} & S_{c2} & \ldots & \ldots \\
S_{b1} & S_{d2} & \ldots & \ldots \\
S_{e3} & S_{f4} & S_{g5} & \ldots \\
S_{h3} & S_{i4} & S_{j5} & \ldots \\
S_{k3} & S_{l5} & \ldots & \ldots \\
S_{m6} & S_{p7} & \ldots & \ldots \\
S_{n6} & S_{q7} & \ldots & \ldots \\
S_{o6} & S_{r7} & S_{s8} & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ldots & \ldots \\
S_{xn} & \ldots & \ldots & \ldots & \ldots & \ldots
\end{bmatrix} \]
NYU POLY WiMAX 4G PROJECT

ABHIJEET MATE
HECTOR MORENO
SANJAY GOYAL
KILLOL MUNISHI
DEEPIKA LAMBA

PROFESSOR THANASIS KORAKIS

POLYTECHNIC INSTITUTE OF NYU
A WiMAX 4G network have been installed at NYU:poly
The efforts are funded by a GENI grant
The network is based on a NEC Base Station with partial of the MAC functionality implemented in open source software
It supports applications for mobile devices over high speed connectivity
• Testing of the network has been done using WiMAX clients such as laptops, AWB dongles and HTC EVO phones
• Measurements show good coverage in the greater area of Downtown Brooklyn
• Today we will show a demo on real time conference call that involves mobility

The coverage range of the Poly WiMAX network
A Demo of ParkNet System using WiMAX GENI Network of NYU Poly

ABHIJEET MATE
TONG JIN
SANJIT KAUL
BIN ZAN
HECTOR MORENO
SANJAY GOYAL
KILLOL MUNSHI
DEEPIKA LAMBA

PROFESSOR MARCO GRUTESER
RUTGERS UNIVERSITY

PROFESSOR THANASIS KORAKIS
POLYTECHNIC INSTITUTE OF NYU
- Demonstration of the synergy of the WiMAX GENI network at NYU Poly with the ParkNet mobile system
- A joint GENI demo between NYU:Poly, Rutgers, NICTA
- 8 cars equipped with ultrasound sensor and GPS receivers drove around Brooklyn, collecting information about available parking spots.
- The data is forwarded through the 4G WiMAX network of NYU:Poly to a central server which builds a real-time map of parking availability through a real-time application.
- Different routes were covered by the vehicles, the data was successfully presented in real-time
- Recorder videos of real-time traces on a Brooklyn map will be presented in this demo
Real-time map of parking availability in downtown Brooklyn.
Sip Filter Hardware Implementation

Jianping Quan
Nallammai Sekar
Ramesh Karri

11/08/10
Session Initiation Protocol (SIP)

- Application-layer signaling protocol
- Text-based protocol like HTTP
- Runs on static port 5060
- Used to create and terminate sessions and to negotiate media stream parameters
  - Type of media
  - Transport protocol used
  - Port numbers for media transfer
    - Defined using session description protocol (SDP)
    - SDP is transported along with the SIP packet
Motivation

• New security vulnerabilities at the application layer
  – SIP over UDP – spoof SIP requests
    • Registration/Call Hijacking
    • Session teardown
    • Request and Response flooding
  – SIP packet format - Plain text
    • Easy to modify and intercept
• Layer-3 filters are not sufficient
  – Arbor Networks, Cisco (formerly Riverhead Networks), TopLayer, Mazu networks, Juniper etc
• Key differentiator: Application layer (Layer 7) filtering
  – Prototype demonstration: Columbia-Verizon
  – Carrier class implementation: NYU-Poly Verizon and Cisco
Basic Data Flow

- CAM1 and CAM2 are both 32-bit by 4k;
- RAM1 is about 26 bytes by 4k; Time stamp RAM is 35-bit by 4k; Current state RAM is 3-bit by 4k.
• Thanks
Video Conferencing with Adaptive FEC and Video Rate Allocation

Yen-Fu Ou, Minyi Yang, Yao Wang
Polytechnic Institute of NYU
Motivation

• ARQ induces unacceptable delay for video conferencing
• Application layer FEC is necessary to recover lost packets
• Both the total data rate ($R_T$) and the allocation between video rate ($R_V$) and FEC redundancy ($\alpha$) should be adjusted based on network feedback
• Establish a testbed on which our adaptation algorithm can be validated over real network impairments.
Flow Diagram of Proposed Scheme

Source raw data (repeated)

Quality Metrics (PSNR, VQMTQ)

Frequently Updated

Video playback

Video Encoder

FEC and Video Coder Parameter Adaptation

Parameter adjustment

Feedback Information

Video Decoder

FEC Decoder

Internet with Network Impairment Emulator Embedded

Performance Monitoring
FEC and Video Rate Allocation

• Total data rate ($R_T$) is adjusted according to the estimated network throughput based on the delay (RTT) and packet loss (P)
• Given $R_T$, allocation between video and FEC is adapted based on the raw and residual packet loss rate
• Given $R_V$, we choose the frame rate at which to code the video, based on prior subjective quality test results.

• For More Details, please see our poster and demonstration !!
Sender

test_v3_sender

Packet loss rate [%] 0.00%
Successful src pkt 0.00%
FEC redundancy [%] 25.00%
FEC_K:12 FEC_N:16

CNC bit rate [k bps] 400
Actual bit rate [k bps] 413
Total bit rate [k bps] 738

Estimate BW [kbps] 0.00
Receiver
Perceptual Video Quality Assessment on a Mobile Platform Considering both Spatial and Quantization

Yuanyi Xue, Yen-Fu Ou, Zhan Ma, Yao Wang
Polytechnic Institute of NYU

• Motivation
  – Video can be coded at different frame size \(s\), frame rate \(t\) and quantization stepsize \(q\) and one should choose \(s,t,q\) to maximize the perceptual quality for a given rate constraint;
  – Our prior work investigated the impact of \(t\) and \(q\) on perceptual quality and rate;
  – This work examines the impact of \(s\) and \(q\) on perceptual quality.

• Test Protocol
  – 7 sequences, 5 spatial resolutions, 4 QPs.
  – Single stimulus.
  – Developed a subjective test interface using a smartphone platform. (TI Zoom 2 with Android)
Perceptual Video Quality Assessment on a Mobile Platform Considering both Spatial and Quantization

Yuanyi Xue, Yen-Fu Ou, Zhan Ma, Yao Wang
Polytechnic Institute of NYU

• Results

  – MOS vs. Spatial Resolution follows a falling exponential decay

• More Results – See poster
An Efficient Querying Framework For Network Monitoring and Forensics

Paul Giura

PhD Student
Polytechnic Institute of NYU