Software Support for Software-Independent Auditing

Gabrielle A. Gianelli, Jennifer D. King,
Edward W. Felten, William P. Zeller

Center for Information Technology Policy
Department of Computer Science
Princeton University

EVT/WOTE ‘09
August 11, 2009
Goals of Post-Election Auditing

• Valid statistical guarantee
• Efficient
• Easy to use
• Instills confidence in election results for voters and officials
• Software independent
\[ n = \min \left\{ u \mid 1 - \prod_{k=0}^{u-1} \frac{N - B - k}{N - k} \geq c \right\} \]
Possible Solutions

- Efficiency
- Complexity
- Difficulty understanding
Possible Solutions

Ease of use ➔ Automation ➔ Software dependence
How to reconcile these tensions?

- One approach: eliminate computers
- Our approach: automate, but verify
Software Independence

Auditing System

Third Parties for Verification
Log Format

- User entered data
- Calculations
  - Pseudorandom numbers
  - Input
  - Precinct or ballot selections
Log

Attributes of log
- XML: can be easily parsed
- Stores all information necessary to recreate an audit, either by hand or with another machine

A log verifiable by a third party ensures software independence.
Our Solution

- Web application
- Python with Django web framework

Goal: auditing interface easy for non-expert users
Welcome, you are currently logged in.

Begin the audit process

Snapshot of previous audits:

<table>
<thead>
<tr>
<th>Audit</th>
<th>Precincts linked?</th>
<th>Dieroll complete?</th>
<th>Races Audited</th>
<th>View Results</th>
<th>Continue audit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>✓</td>
<td>29 / 29</td>
<td>Go</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>✓</td>
<td>29 / 29</td>
<td>Go</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>✗</td>
<td>0 / 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examine election data | Examine ballot data
Select races in election "November 2008 - Humboldt County":

<table>
<thead>
<tr>
<th>Audit race?</th>
<th>Race name</th>
<th>Algorithm</th>
<th>Percent of precincts:</th>
<th>% (0-100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑</td>
<td>President</td>
<td>Exact Percent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑</td>
<td>U.S. Representative</td>
<td>Exact Percent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Member of State Assembly</td>
<td>Constant Sample Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ferndale Mayor</td>
<td>Exact Percent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supported Algorithms

Precinct-based algorithms:
• Exact Percent
• Percent by Probability

Ballot-based algorithms:
• Constant Sample Size
• Varying Sample Size
Linking Precincts

- Assume two races A and B over the same set of precincts
- Goal: choose 2% of precincts for Race A and 3% of precincts for Race B
Unlinked Precincts

Set of all ballots, $S$

$S_A$  $S_B$
Linked Precincts

Set of all ballots, $S$

$S_A \& S_B \subseteq S_B$
Pseudorandom Number Generation

“1,2,1,4,4,…”

PRNG
Humboldt County Data

- Ballot images from Humboldt County (CA) Election Transparency Project (Nov 2008)
- Textual ballot representations from Mitch Trachtenberg’s Ballot Browser program
- 29 races; 145 precincts; 128,144 ballots
Process

- Loaded the data from individual ballots into our database
- Used the system to run a mock audit
- In order to simulate a manual recount, compared the ballot images against the data in our database
## Results

<table>
<thead>
<tr>
<th>Audit</th>
<th>Algorithm</th>
<th>Parameter</th>
<th>Precincts chosen</th>
<th>Ballots Chosen</th>
<th>Percent ballots chosen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exact Percent</td>
<td>1% of precincts</td>
<td>33</td>
<td>15,613</td>
<td>12%</td>
</tr>
<tr>
<td>2</td>
<td>Exact Percent (linking)</td>
<td>1% of precincts</td>
<td>15</td>
<td>5,768</td>
<td>4%</td>
</tr>
<tr>
<td>3</td>
<td>Constant Sample Size</td>
<td>99% confidence</td>
<td>N/A</td>
<td>3,006</td>
<td>2%</td>
</tr>
</tbody>
</table>
In closing...

Automation can
• Make post-election audits more efficient
• Expand the scope of complex auditing algorithms and reduce the number of ballots to be counted as long as the output can be independently verified.
Software Support for Software-Independent Auditing

Gabrielle A. Gianelli, Jennifer D. King, Edward W. Felten, William P. Zeller

Center for Information Technology Policy
Department of Computer Science
Princeton University

EVT/WOTE ‘09
August 11, 2009