OPENCOUNT

IMPROVED SUPPORT FOR MACHINE-ASSISTED BALLOT-LEVEL AUDITS

Eric Kim, Nicholas Carlini, Andrew Chang, George Yiu, Kai Wang†, David Wagner

University of California, Berkeley
†University of California, San Diego

How can we trust our election technology?

Answer: audit the election results.
Talk Overview

- Motivation
  - How can OpenCount help the audit process?
- Overview of OpenCount
- Experience using OpenCount in elections
What is OpenCount?

- OpenCount tabulates elections
- Generates ballot-level cast vote records

CVR 00001
President of the United States
Mitt Romney
Member, County Central Com.
Shawn Nelson
David John Shawver
Greg Sebourn
Steve Hwangbo
Motivation

- We want to perform a post-election audit
- Statistical, ballot-level audits are efficient
  - Typically only have to examine tens to hundreds of ballots (depends on margin)
  - More efficient than existing processes
    - CA: Each county hand-counts all ballots from 1% of precincts
Motivation (cont.)

- Ballot-level audits require knowledge of the **voting system’s interpretation** of each ballot

  - i.e., a Cast Vote Record (CVR) for each ballot
    - Electronic record of the cast votes
Motivation (cont.)

- Cast Vote Record (CVR)

CVR 00001
President of the United States
Mitt Romney
Member, County Central Com.
Shawn Nelson
David John Shawver
Greg Sebourn
Steve Hwangbo
Motivation (cont.)

- **Problem**: current deployed voting systems do not output CVRs for each ballot
  - They only output election totals
- **Consequence**: cannot perform ballot-level audits
Motivation (cont.)

- Can’t “upgrade” existing voting systems
  - Most vendors are focusing on next-gen systems
  - EAC certification process (U.S. Election Assistance Commission) would make upgrade too expensive
What is one to do?

If you can’t improve it, rebuild it!
OpenCount

- OpenCount tabulates elections
  - Input: Scanned ballot images
  - Output: Cast Vote Records, election totals.
- Built specifically with ballot-level audits in mind
- Open-source software (free!)
First Attempt: Blank Ballots

- Collect one **blank ballot** from each ballot style
- **Blank Ballot** = an unmarked ballot
With Blank Ballots... (1/6)
With Blank Ballots... (2/6)

Style A

Style B
With Blank Ballots... (3/6)
With Blank Ballots... (4/6)
With Blank Ballots… (5/6)

<table>
<thead>
<tr>
<th>Vote Both Sides</th>
<th>Vote Both Sides</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>President</strong></td>
<td><strong>President</strong></td>
</tr>
<tr>
<td>Vote for One</td>
<td>Vote for One</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>Donald Trump</td>
</tr>
<tr>
<td>Joe Biden</td>
<td>Joe Biden</td>
</tr>
</tbody>
</table>

**OFFICIAL BALLOT Democratic**

**Presidential Primary Election**

Orange County

June 05, 2012

Vote for One

United States Senator

Vote for One

SANDIEGIO HILTON

Democratic

**Presidential Primary Election**

Orange County

June 05, 2012

Vote for One

United States Representative

Vote for One

M. J. ROYCE

Republican

**United States Representative**

February 5, 2013

Vote for One

JIMMIE LOWE

Democratic

**United States Representative**

February 5, 2013

Vote for One

GREGG HOPWOOD

Republican

**United States Senator**

June 05, 2012

Vote for One

KARL BROWN

Democratic

**United States Senator**

June 05, 2012

Vote for One

SANDY ROYCE

Republican

**Member of the State Assembly**

March 5, 2013

Vote for One

JIM HAMMER

Republican

**Member of the State Assembly**

March 5, 2013

Vote for One

JEANINE S. CALVIN

Democratic

**NONPARTISAN OFFICES**

All offices, regardless of the party preference, are open for registration, or refuse to declare a party preference, may vote for any candidate for a non-partisan office. The party preference, if any, designated by a symbol in the column for the candidate, is determined by the candidate and is shown for the information of the voter only. It does not imply that the candidate is nominated or endorsed by the party or that the party approves of the candidate.

**State of California**

Judges of the Superior Court

Office No. 1

Vote for One

MONICA J. FLEMING

Judge of the Superior Court

Office No. 1

Vote for One

ROBERT J. JONES

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

SANDIEGIO HILTON

Judge of the Superior Court

Office No. 1

Vote for One

SANDY ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

GREGG HOPWOOD

Judge of the Superior Court

Office No. 1

Vote for One

KARL BROWN

Judge of the Superior Court

Office No. 1

Vote for One

JIMMIE LOWE

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

SANDIEGIO HILTON

Judge of the Superior Court

Office No. 1

Vote for One

SANDY ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

GREGG HOPWOOD

Judge of the Superior Court

Office No. 1

Vote for One

KARL BROWN

Judge of the Superior Court

Office No. 1

Vote for One

JIMMIE LOWE

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

SANDIEGIO HILTON

Judge of the Superior Court

Office No. 1

Vote for One

SANDY ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

GREGG HOPWOOD

Judge of the Superior Court

Office No. 1

Vote for One

KARL BROWN

Judge of the Superior Court

Office No. 1

Vote for One

JIMMIE LOWE

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

SANDIEGIO HILTON

Judge of the Superior Court

Office No. 1

Vote for One

SANDY ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

GREGG HOPWOOD

Judge of the Superior Court

Office No. 1

Vote for One

KARL BROWN

Judge of the Superior Court

Office No. 1

Vote for One

JIMMIE LOWE

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

SANDIEGIO HILTON

Judge of the Superior Court

Office No. 1

Vote for One

SANDY ROYCE

Judge of the Superior Court

Office No. 1

Vote for One

GREGG HOPWOOD

Judge of the Superior Court

Office No. 1

Vote for One

KARL BROWN

Judge of the Superior Court

Office No. 1

Vote for One

JIMMIE LOWE

Judge of the Superior Court

Office No. 1

Vote for One

M. J. ROYCE

Judge of the Superior Court

办公室 No. 1

投票给一个

MONICA J. FLEMING

法官

办公室 No. 1

投票给一个

ROBERT J. JONES

法官

办公室 No. 1

投票给一个

M. J. ROYCE

法官

办公
With Blank Ballots...

**Style A**

**Style B**
Previous Work

- EVT/WOTE 2012 (Bellevue, Washington)
  - First introduction of the OpenCount (2012) system
- **Required collecting all blank ballots**
Problem: Did not scale to large elections

Collecting blank ballots is a huge burden for election officials

- Blocked some counties from participating

Overall, too much required effort
Our Latest Work

- New approach: No blank ballots
How can we do this?
No Blank Ballots
No Blank Ballots

- How do we find:
  - Voting Targets?
  - Contests?
OpenCount Pipeline

- Overview of system
- Election experiences
  - California risk-limiting audit pilot program
Scan Ballots (1/6)

- Use any commercial, off-the-shelf scanner
Ballot Grouping (2/6)
Ballot Grouping (2/6)
Ballot Grouping (2/6)

~124,000 Ballots

~200 Styles
Ballot Grouping (2/6)
Ballot Grouping (2/6)
Implemented vendor-specific barcode decoders

- Diebold
- ES&S
- Hart
- Sequoia
Goal

- Specify location of contests and voting targets
- Perform data entry of contest text

Only need to annotate one ballot from each style
How to find voting targets automatically?
1.) Operator selects empty voting target
1.) Operator selects empty voting target
Search for empty voting target on ballots
Template Matching
  - Grid-search
  - Search for this:
Verify Matches
Problem: Voter marks interfere with template matching
Problem: Voter marks interfere with template matching

Idea: Voters vote differently. Can find missing targets on other ballots with the same style
Idea: Voters vote differently. Can find missing targets on other ballots with the same style.
Idea: Voters vote differently. Can find missing targets on other ballots with the same style.
**Union of detections from A + B**

**Idea:** Voters vote differently. Can find missing targets on other ballots with the same style
Contest text data entry
- Contest title, candidate names

Judge of the Superior Court (Office No. 1)
- Deborah J. Chuang
- Eugene Jizhak
Can’t rely completely on OCR
Manually labeling each contest takes forever
Number of distinct contests is small
  - A few hundred at most
Contests are duplicated on many ballot styles
  - “President of the US”
Should only have to label this contest once!
Want to detect contest duplicates

Simple idea: compare contest images

- Pixel-difference (L2 norm)
Judge of the Superior Court
Office No. 1
Vote for One
☐ DEBORAH J. CHUANG
Judge of the Superior Court
☐ EUGENE JIZHAK
General Practice Attorney
☐

Judge of the Superior Court
Office No. 1
Vote for One
☐ DEBORAH J. CHUANG
Judge of the Superior Court
☐ EUGENE JIZHAK
General Practice Attorney
☐

Judge of the Superior Court
Office No. 1
Vote for One
☐ DEBORAH J. CHUANG
Judge of the Superior Court
☐ EUGENE JIZHAK
General Practice Attorney
☐

Diff = 0.058
MATCH
Judge of the Superior Court
Office No. 1
Vote for One
☐ DEBORAH J. CHUANG
Judge of the Superior Court
☐ EUGENE JIZHAK
General Practice Attorney

STATE SENATOR
29th District
Vote for One
☐ GREG DIAMOND
Party Preference: Democratic
Workers' Rights Attorney
☐ ROBERT "BOB" HUFF
Party Preference: Republican
Lawmaker/Business Owner

Diff = 0.175
NO MATCH
Problem: contest visual appearance varies
- Word spacing, line wrapping, candidate re-ordering

Different Line Wrap

DISTRCT COURT OF APPEAL
Shall Judge Clay Roberts of the First District Court of Appeal be retained in office?

- YES
- NO
DID NOT MATCH
Our approach: utilize OCR + edit-distance

OCR

District Court of Appeals
Shall Judge Clay Roberts of the First District Court of Appeal be retained in office?

Yes No
Our approach: utilize OCR + edit-distance

OCR

District Court of Appeals
Shall Judge Dby Roberts of the First District Court of Appeal be retained in office?
Yes Mo

OCR

District Court of Appeals
Shall Judge Clay Roberts of the First District Court of Appeal be retained in office?
Yes No
Our approach: utilize OCR + edit-distance

OCR

District Court of Appeals
Shall Judge Dby Roberts of the First District Court of Appeal be retained in office?
Yes No

OCR

District Court of Appeals
Shall Iudg Clay Roberts of the First District Court of Appeal be retained in office?
Yes No

edit_dist(OCR(A), OCR(B)) = 20
Our approach: utilize OCR + edit-distance

OCR ( ) = District Court of Appeals
Sall Judge Dby Roberts of the First District Court of Appeal be retained in office?
Yes Mo

OCR ( ) = District Oourt of Appeal
Shall ludg Clay Roberts of the First Districf Courf of Appeal be refaimed in office?
Yes No

edit_dist(OCR(A), OCR(B)) = 20

Similarity: 82.8%

Match!
Ballot Interpretation (4/6)

- Goal
  - Determine if voting targets are “filled” or “empty”
Ballot Interpretation (4/6)

Separating Line

Sorted by Average Pixel Intensity
Ballot Interpretation (4/6)
Ballot Interpretation (4/6)
Ballot Interpretation (4/6)

Filled

Empty
Generate CVRs (5/6)

- Output CVRs

**CVR 0001**
President of the United States
Barack Obama
Proposition 19
Yes
City Council Representatives
UNDENVOTED

**CVR 0002**
President of the United States
Mitt Romney
Proposition 19
Yes
City Council Representatives
Jane Doe

...
Perform Audit (6/6)

- Finally, perform the audit!
- Done!
OpenCount has been used to support risk limiting pilot audits in several California counties
- Alameda, Madera, Merced, Napa, San Luis Obispo, Stanislaus, Ventura

OpenCount’s results matched all examined paper ballots perfectly
<table>
<thead>
<tr>
<th>County</th>
<th># Ballots</th>
<th># Ballot Styles</th>
<th>Total Time (2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanislaus</td>
<td>3,151</td>
<td>1</td>
<td>7m 18s</td>
</tr>
<tr>
<td>Merced</td>
<td>7,120</td>
<td>1</td>
<td>12m 31s</td>
</tr>
<tr>
<td>Ventura</td>
<td>17,301</td>
<td>1</td>
<td>23m 6s</td>
</tr>
<tr>
<td>Alameda</td>
<td>1,374</td>
<td>8</td>
<td>22m 1s</td>
</tr>
<tr>
<td>San Luis Obispo</td>
<td>10,689</td>
<td>27</td>
<td>30m 35s</td>
</tr>
<tr>
<td>Madera</td>
<td>3,757</td>
<td>1</td>
<td>6m 38s</td>
</tr>
<tr>
<td>Napa</td>
<td>6,809</td>
<td>11</td>
<td>1h 56m 9s</td>
</tr>
<tr>
<td>Yolo</td>
<td>35,532</td>
<td>623</td>
<td>3h 36m</td>
</tr>
</tbody>
</table>
### Election Experiences (cont.)

<table>
<thead>
<tr>
<th>County</th>
<th># Ballots</th>
<th># Ballot Styles</th>
<th>Total Time (2013)</th>
<th>Speedup (2012 / 2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanislaus</td>
<td>3,151</td>
<td>1</td>
<td>7m 18s</td>
<td>2.40x</td>
</tr>
<tr>
<td>Merced</td>
<td>7,120</td>
<td>1</td>
<td>12m 31s</td>
<td>2.04x</td>
</tr>
<tr>
<td>Ventura</td>
<td>17,301</td>
<td>1</td>
<td>23m 6s</td>
<td>2.52x</td>
</tr>
<tr>
<td>Alameda</td>
<td>1,374</td>
<td>8</td>
<td>22m 1s</td>
<td>1.29x</td>
</tr>
<tr>
<td>San Luis Obispo</td>
<td>10,689</td>
<td>27</td>
<td>30m 35s</td>
<td>2.78x</td>
</tr>
<tr>
<td>Madera</td>
<td>3,757</td>
<td>1</td>
<td>6m 38s</td>
<td>1.28x</td>
</tr>
<tr>
<td>Napa</td>
<td>6,809</td>
<td>11</td>
<td>1h 56m 9s</td>
<td>2.78x</td>
</tr>
<tr>
<td>Yolo</td>
<td>35,532</td>
<td>623</td>
<td>3h 36m</td>
<td>16.24x</td>
</tr>
</tbody>
</table>
### Election Experiences (cont.)

<table>
<thead>
<tr>
<th>County</th>
<th># Ballots</th>
<th># Ballot Styles</th>
<th>Total Time (2013)</th>
<th>Speedup (2012 / 2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanislaus</td>
<td>3,151</td>
<td>1</td>
<td>7m 18s</td>
<td>2.40x</td>
</tr>
<tr>
<td>Merced</td>
<td>7,120</td>
<td>1</td>
<td>12m 31s</td>
<td>2.04x</td>
</tr>
<tr>
<td>Ventura</td>
<td>17,301</td>
<td>1</td>
<td>23m 6s</td>
<td>2.52x</td>
</tr>
<tr>
<td>Alameda</td>
<td>1,374</td>
<td>8</td>
<td>22m 1s</td>
<td>1.29x</td>
</tr>
<tr>
<td>San Luis Obispo</td>
<td>10,689</td>
<td>27</td>
<td>30m 35s</td>
<td>2.78x</td>
</tr>
<tr>
<td>Madera</td>
<td>3,757</td>
<td>1</td>
<td>6m 38s</td>
<td>1.28x</td>
</tr>
<tr>
<td>Napa</td>
<td>6,809</td>
<td>11</td>
<td>1h 56m 9s</td>
<td>2.78x</td>
</tr>
<tr>
<td><strong>Yolo</strong></td>
<td><strong>35,532</strong></td>
<td><strong>623</strong></td>
<td><strong>3h 36m</strong></td>
<td><strong>16.24x</strong></td>
</tr>
</tbody>
</table>
## Election Experiences (cont.)

<table>
<thead>
<tr>
<th>County</th>
<th># Ballots</th>
<th># Ballot Styles</th>
<th>Total Time (2013)</th>
<th>Human Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marin</td>
<td>29,121</td>
<td>398</td>
<td>11h 53m</td>
<td>5h 45m</td>
</tr>
<tr>
<td>Santa Cruz</td>
<td>34,004</td>
<td>136</td>
<td>18h 50m</td>
<td>5h 27m</td>
</tr>
<tr>
<td>Leon</td>
<td>124,200</td>
<td>216</td>
<td>14h 2s</td>
<td>1h 53m</td>
</tr>
<tr>
<td>Orange</td>
<td>294,402</td>
<td>1,839</td>
<td>3d 22h 39s</td>
<td>1d 8h 25m</td>
</tr>
</tbody>
</table>

Previous version (2012) could not process elections of this size and complexity.

**Progress!**
Conclusion

- OpenCount facilitates efficient election auditing
- OpenCount is ready for election officials to use
- OpenCount has been used in California pilots
  - Audits made possible by OpenCount