

Batch Reports for Audits - ElectionAudits and the Boulder 2008 Election

- Neal McBurnett
- NIST Common Data Formats
Workshop
- Oct 29 2009

Questions

- Ask questions any time
- Who has been involved with an audit?
- Who knows of a jurisdiction where ballots are reconciled by ballot style as well as by precinct?

Effective audits are critical

- Improve election processes
- Attain software independence and confidence in elections

Humboldt County 2008
Paper ballots, optical scan

197 ballots deleted without a trace

Discovered later by Humboldt County Election
Transparency Project audit

Kudos to Mitch Trachtenberg, Registrar of
Elections Carolyn Cernich, and colleagues

Goal:

software independence
(Rivest & Wack)

via auditable paper records,
good audits

Audit:

Compare system's reported results with the evidence

Take sample of detailed results by audit unit and compare to hand counts of those audit units

Auditing steps`

- Produce report by audit unit (precinct or batch)
- Reconcile number of ballots in each unit
- Recalculate sum of subtotals
- Randomly select audit units to audit
- Count audit units by hand
- Compare results
- Continue auditing until statistical evidence meets goal (limit risk of wrong outcome)

Report of Vote Counts by Audit Unit

Audit Unit	Susan	Nelson	Under	Over	Total
100	137	144	2	0	383
101	77	68	1	0	247
102	122	87	3	2	316
103	98	102	1	1	305
104	22	18	2	0	146
105	103	140	1	0	349
Total	559	559	10	3	1746

Must include undervotes and overvotes in the results for each audit unit. For multiple-winner contests, undervotes are ambiguous, so also need a count of ballots on which the contest appears

*Principles and Best Practices for
Post-Election Audits (2008)*

[*http://electionaudits.org/principles*](http://electionaudits.org/principles)

League of Women Voters
Election Audit Report (2009)

Check out typical Hart precinct report

PDF

No under/over

By precinct

Big numbers

Not suitable for auditing in many places -
vulnerable to many attacks!

Inefficient everywhere

From NIST's Call for Papers:

“Human-readable versus machine readable-human-readability matters for transparency and auditing”

Not true – auditing requires aggregation and summing subtotals to confirm totals.

The most efficient audits involve analyzing vote-count data for all audit units

Standard format helps with targeting units

Current reports typically only by precinct

Auditing by precinct is outmoded, inefficient
and problematic

Agreement on need for standardized
small-batch audit reporting from meeting
hosted by American Statistical Association,
October 2009

Batch reporting important:

Organization of the paper

Privacy of provisional ballots

Small batches more efficient to audit

Individual ballot interpretations needed for emerging methods that audit at the ballot level without breaching confidentiality

Organization of the paper:

Reports are by precinct

But often piles of paper aren't

- Mail in
- Early voting
- Provisional

Reporting by precinct before and after
provisional ballots are included can
compromise the privacy of the provisional
voters

Instead, report provisionals in batches, and
precincts only after all provisionals are in

Significance of the results driven by how many batches you audit, not how many ballots you count

Looking for incorrect vote counts

Don't care about total vote count for the sample

Typically, amount of work required to perform a vote-tabulation audit is roughly proportional to the batch sizes, with no reduction in statistical efficacy.

Getting batch data much more frustrating
than it should be

Boulder County, 2008

170,000 ballots, 65 races
13 contests audited, 40000 choices

Hart InterCivic system

Precinct reports

But only 17% cast in precinct

67% mail-in, 16% early voting

Solution: run cumulative report 525 times

Once for each batch

Subtract each report from the previous
report to get batch sub-totals

Like snapshots of election-night reporting
through the night

But we want ***lots*** of them....

Report of *Cumulative* Vote Counts by Audit Unit

Audit Unit	Susan	Nelson	Under	Over	Total
100	137	144	2	0	383
101	214	212	3	0	630
102	336	299	6	2	946
103	434	401	7	3	1251
104	456	419	9	3	1397
105	559	559	10	3	1746
Total	559	559	10	3	1746

Report of Vote Counts by Audit Unit

Audit Unit	Susan	Nelson	Under	Over	Total
100	137	144	2	0	383
101	77	68	1	0	247
102	122	87	3	2	316
103	98	102	1	1	305
104	22	18	2	0	146
105	103	140	1	0	349
Total	559	559	10	3	1746

No single Hart report had all the data we wanted in a reasonable format.

“Crystal Reports” XML schema reflected paper layout, not logical data

Didn't have ballot counts – dug those out of PDF reports

“CSV” reports didn't have columns that lined up

Sequoia reports are clean and complete

Good “text with headings” report (tab-separated)

Samples at

<http://bcn.boulder.co.us/~neal/electionaudits/formats/>

Reports also frequently take a long time to produce – painful when doing hundreds of times.

Audit results also need to be published in a standard format, suitable for aggregation across jurisdictions

Include hand counts and discrepancies for all units audited, whether they were targeted or randomly selected, information needed to validate the selections (e.g. algorithms used, random seeds used as input) , results of the audit

ElectionAudits software is open source,
BSD license

Supports the *Best Practices*

Automates many steps of the audit

Enter the data, publish with statistics
Roll 15 dice, publish all the selections

To do: automate discrepancies, escalation

Look at some real data from Boulder 2008

Features of ElectionAudits

- Imports standard election report files: Hart, Sequoia
- Produces auditable reports for the public.
- Protects voter anonymity by merging small audit units
- Doesn't require that paper ballots be sorted into piles by precinct
- Can produce batch reports from sequence of cumulative reports
- Facilitates risk-limiting audits
- Verifiably pseudo-random: dice + "Sum of Square Roots"

Biggest challenge: getting small-batch data
out of election systems

Summary

- Proper audits are critical
- Risk-limiting audits in Colorado law
- Small-batch reports help enormously
- Standard, machine-readable format is needed
- EML looks appropriate
- <http://launchpad.net/electionaudits>