PROVIDENCE: an Efficient and Secure Ballot Polling Risk-Limiting Audit

Oliver Broadrick [1 5] Poorvi L. Vora[1] Filip Zagórski[34]

[1]Department of Computer Science, The George Washington University (odbroadrick@gmail.com)

[3]University of Wroclaw

[4]Votifica

[5]University of California, Los Angeles

August 12, 2023

Election security

What do we want?

- The right winner *and* strong evidence that they are the right winner

- Software independent
  - Configuration errors, bugs, hacking

Election security

An approach[1]:

verifiedvoting.org

- Compliance audits and tabulation audits

[1]Strongly supported by a report of the National Academy of Sciences and the Voluntary Voting Systems Guidelines, the closest we have to standards for election technology.

Our Contributions

- Rigorous tabulation audit (risk-limiting audit) PROVIDENCE, the most efficient and secure of its kind
- Open source implementation, included in Arlo, most popular audit software
- Pilot use of PROVIDENCE in the city of Providence, Rhode Island in 2022
- Comparison of PROVIDENCE with other ballot polling RLAs with new workload models

Background

Risk-Limiting Audits (RLAs)

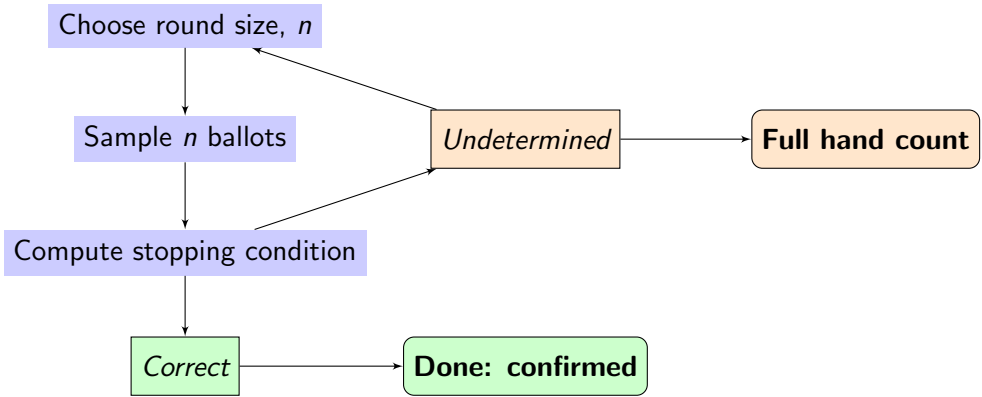
Assumption: successfully completed compliance audits

Risk-Limiting Audit (RLA) with risk limit α: A tabulation audit that, if the reported outcome is wrong, will detect and correct it with probability at least $1 - \alpha$.

Small α is good.

The **stopping probability** for a given sample is the probability that–given the reported outcome is correct–the audit confirms the result. Large stopping probabilities are good.

Ballot Polling RLAs



Existing ballot polling RLAs

BRAVO

- Theory: the most efficient RLA (requires the smallest expected number of ballots) when ballots are sampled one at a time (ballot-by-ballot).
- Practice: in real audits, decisions are taken after many ballots are drawn (round-by-round).

MINERVA

- Recent RLA designed for round-by-round use.
- In a first round chosen to give a typical 0.90 stopping probability, MINERVA requires
  - 50-80% as many ballots as BRAVO.
- Proven to be risk-limiting if all round sizes are predetermined, before the audit begins.

Problems We Will Address

1. Predetermined round sizes give inflexible audits
   - May be more efficient to choose future round sizes as a function of previous samples
2. Existing workload measures don't capture the cost of a round
   - We are unaware of any RLAs that have ever actually drawn a single ballot at a time

Our work

The Adversary in an RLA

Definition (α-RLA)

An audit $\mathcal{A}$ is an α-RLA if for samples $X \in \mathcal{X}$

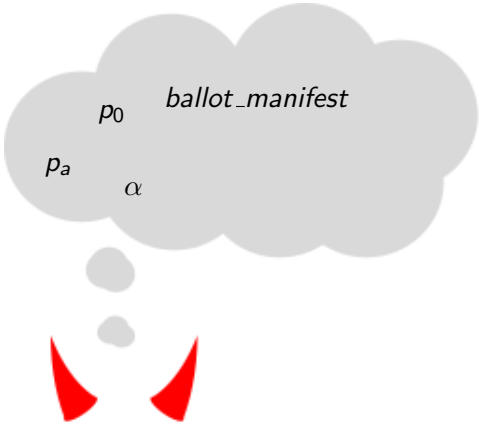
$$\Pr[\mathcal{A}(X) = \textit{Correct} \mid H_0] \leq \alpha,$$

where H_0 corresponds to the incorrectly reported outcome *closest* to the reported outcome (i.e. a tie)

Adversarial goal: to increase the risk above α

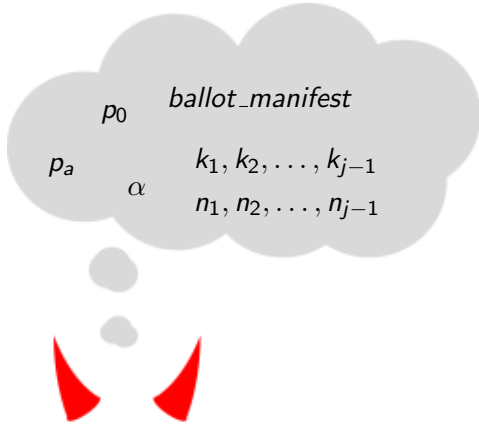
$$\Pr[\mathcal{A}(X) = \text{Correct} \mid H_0] > \alpha$$

Weakly round-choosing[2]



BRAVO secure
MINERVA secure

Strongly round-choosing

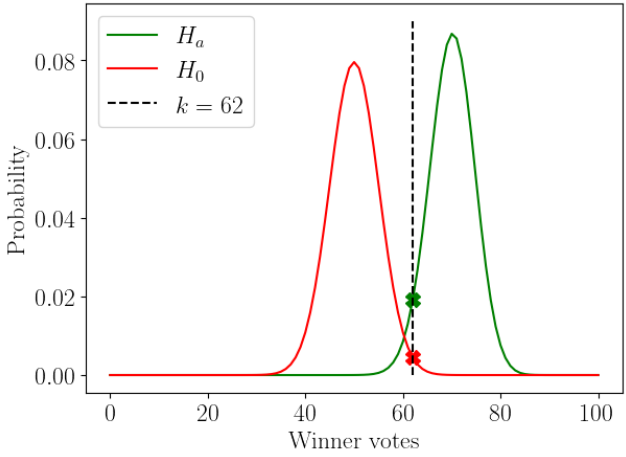


BRAVO secure
MINERVA not known secure

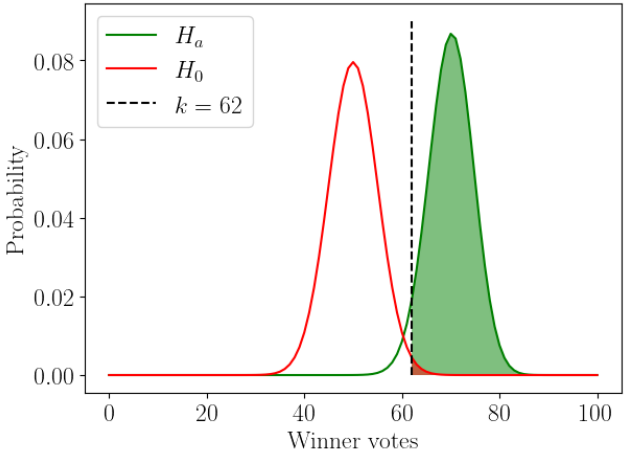
[2] Adversary names thanks to our anonymous *USENIX Security* shepherd.

Bravo

$$P(k_j \mid H)$$

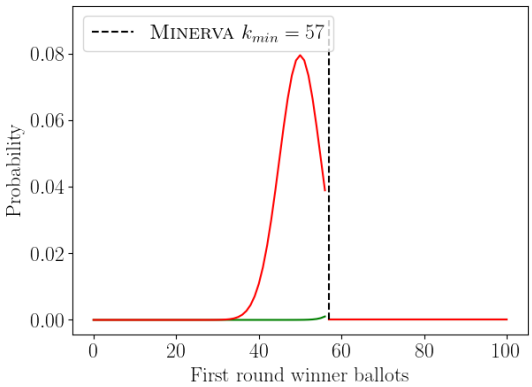


$P(k_j \mid H)$

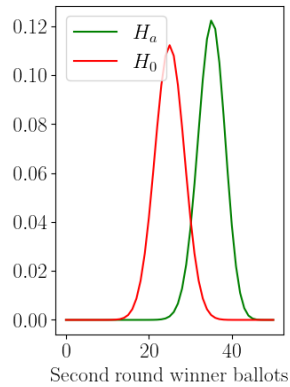

H_a
H_0
$k = 62$
Probability
0.00
0.02
0.04
0.06
0.08
0
20
40
60
80
100
Winner votes

How MINERVA proceeds

$$P(k_2 \wedge K_1 \leq k_{min,1} \,|\, H, n_1) =$$



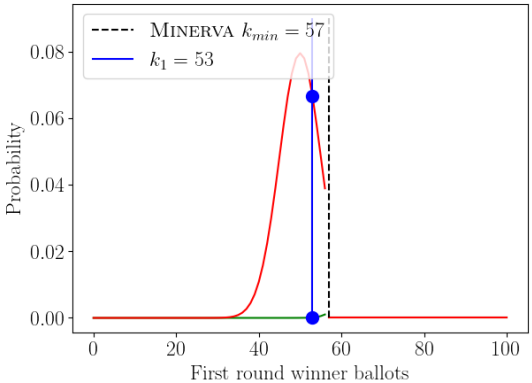
$\circledast$

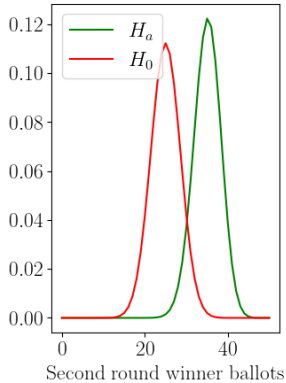


Implicitly assumes that n_2 is the same for all k_1

How PROVIDENCE proceeds

$$P(k_2 \wedge k_1 | H) =$$





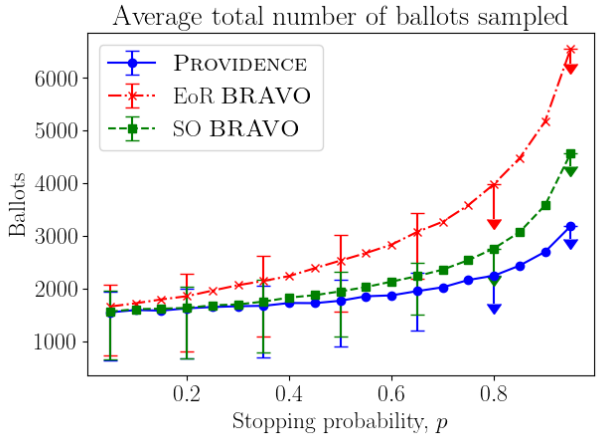
PROVIDENCE Properties

- Risk-Limiting in the presence of a strongly-round choosing adversary[3]
- Efficiency comparable to MINERVA, shown through simulations

[3] Arkady Yerukhimovich points out that random seeds should be freshly generated at the start of each round so that adversaries do not know which ballots will be drawn in a round before they choose the round size.

2016 Presidential contest in VA

Margin ≈ 0.053



Workload

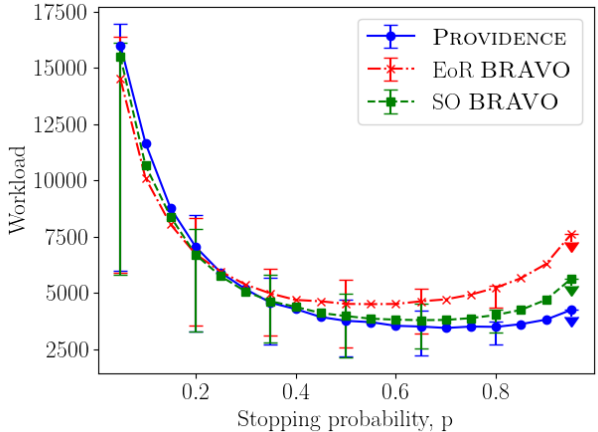
With a round cost:

$$W(E_b, E_r) = E_b c_b + E_r c_r$$

E_b: expected number of ballots
E_r: expected number of rounds
c_b: per ballot cost
c_r: per round cost

Workload

$W(E_b, E_r) = E_b c_b + E_r c_r$ with $c_b = 1$ and $c_r = 1000$



Conclusion

- PROVIDENCE: efficient and flexible
- Introduction of workload models accounting for the cost of a round
- Other round-size considerations (misleading samples and per-precinct cost)
- Piloted in the city of Providence, Rhode Island
- Implemented in Arlo, most commonly used RLA software

Thank you