

# Range-Based Set Reconciliation

---

Aljoscha Meyer

# Set Reconciliation

- set union over a network
- between (exactly) two machines
- unstructured data
- no shared state or history

## Trivial Reconciliation



Sam Gwilym

# Model and Analysis

- Alfie and Betty talk over a network
- reliable communication, rounds of unit length, unlimited bandwidth
- probabilistic solutions

# Model and Analysis

- Alfie and Betty talk over a network
- reliable communication, rounds of unit length, unlimited bandwidth
- probabilistic solutions
- $n$ : size of the union
- $n_{\triangle}$ : size of the symmetric difference

# Model and Analysis

- roundtrips
- communicated bytes
- computation time per reconciliation session
- computation space per reconciliation session
- computation time per item
- computation space per item



- my interest: p2p systems -  $\bar{c}$  low resources (cannot load sets into memory, cannot iterate over full set), peers are untrusted
- traditional: obtain estimate of the size of symdif, then reconcile in  $O(1)$  communication rounds
- cpi: compute loooong time ( $O(n)$ ), needing space of  $O(\text{symdif})$  (tends to crash)
- iblt: same
- hence relax to logarithmic number of rounds
- divide-and-conquer -  $\bar{c}$  explain rbsr
- complexity analysis
- arbitrary recursion anchor (protocols)
- arbitrary partition techniques in principle
- two issues, start with the subtle one
- deadlock
- solution
- analysis and model