A Tune-up for Tor:

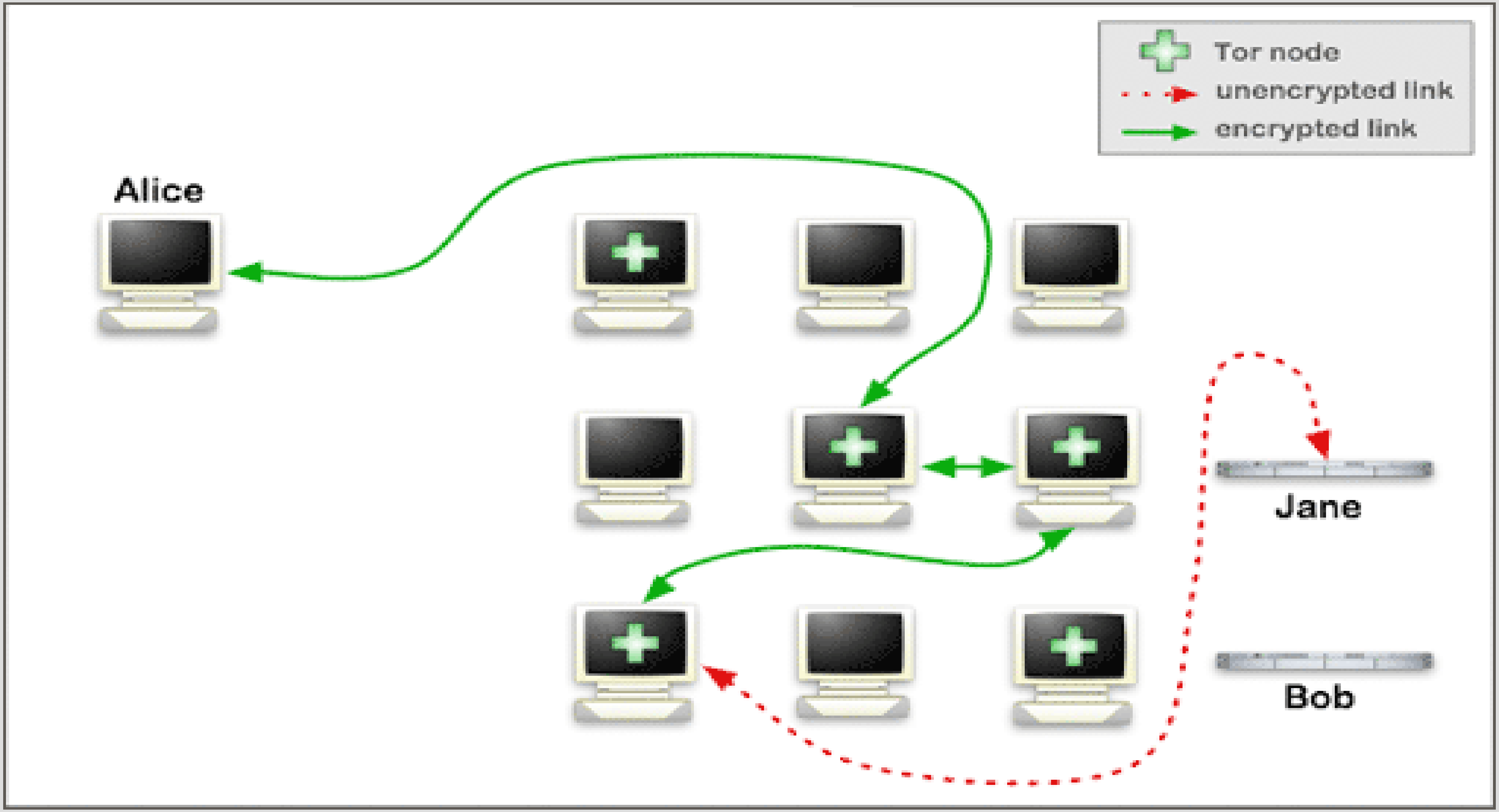
Improving Security and Performance in the Tor Network

Robin Snader
Department of Computer Science
rsnader2@cs.uiuc.edu

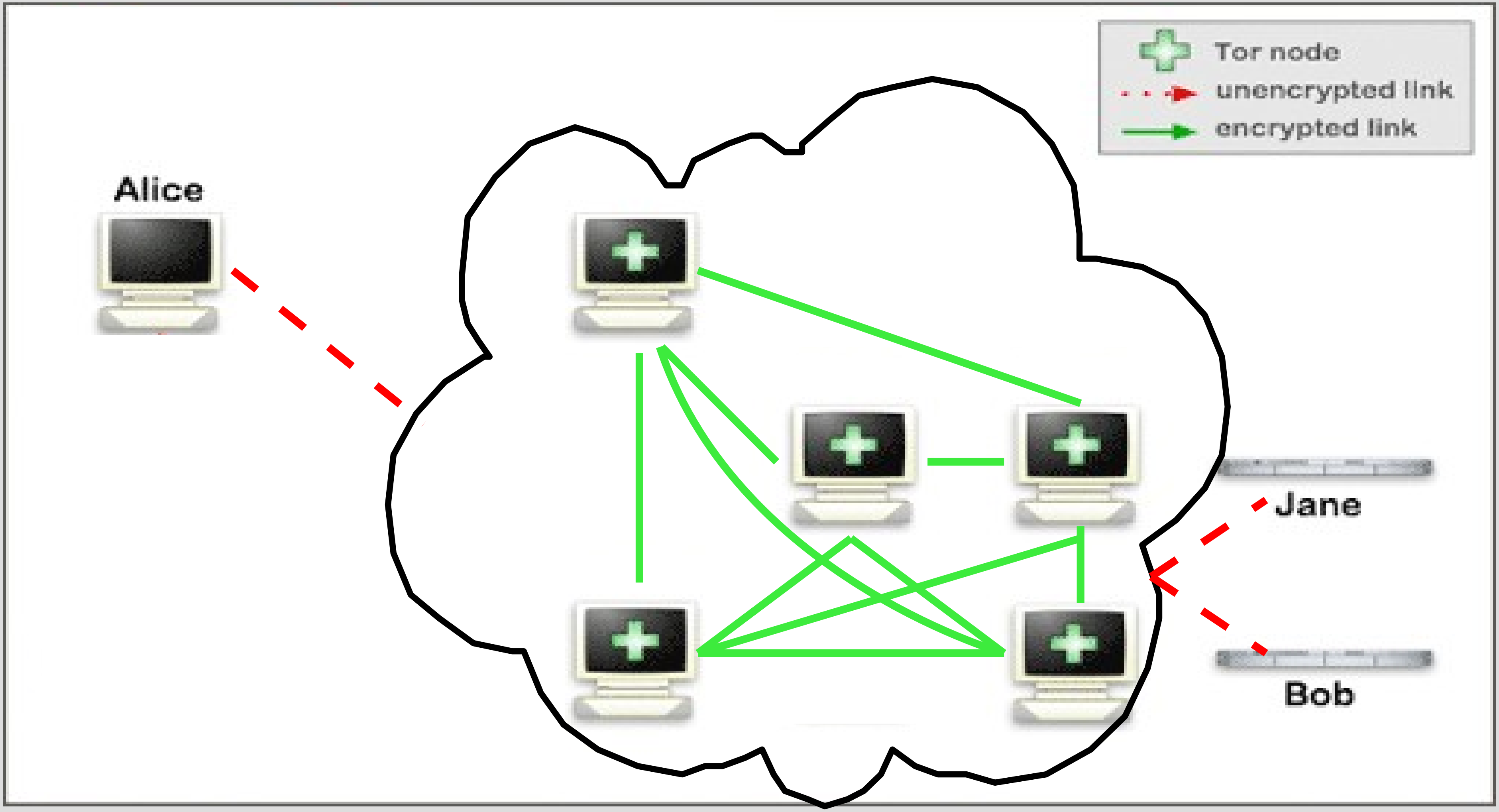
Nikita Borisov
Dept. of Electrical & Computer Engineering
nikita@uiuc.edu

University of Illinois at Urbana-Champaign

The Tor Network



Tor as an Overlay Network

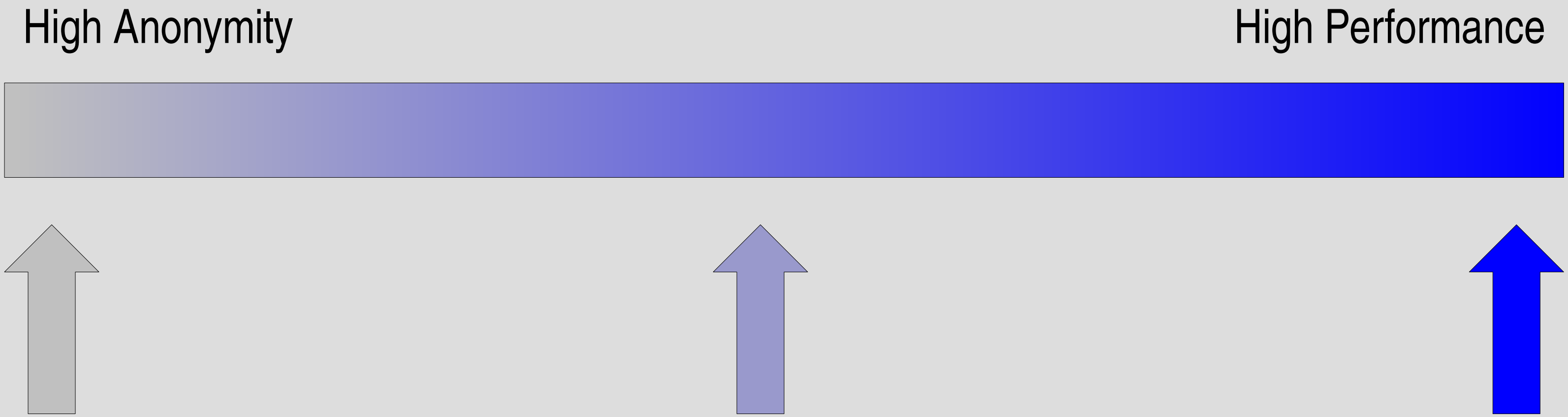


Overlay Network Considerations

- Overlay Networks
  - Link Evaluation
  - Efficient Route Selection
    - High Flow Bandwidth
    - High Aggregate Network Throughput
- Tor as an Overlay Network
  - Secure Link Evaluation
  - Secure, Anonymous and Efficient Route Selection

Anonymous vs. Efficient Route Selection

- Efficient Routes: prefer well-connected routers
- Anonymous Routes: choose routers uniformly at random



High Anonymity

High Performance

- Spies
- Dissidents
- Whistle blowers

- Casual Users
- Privacy Advocates

- P2P Users
- Cover Traffic Providers
- New Users

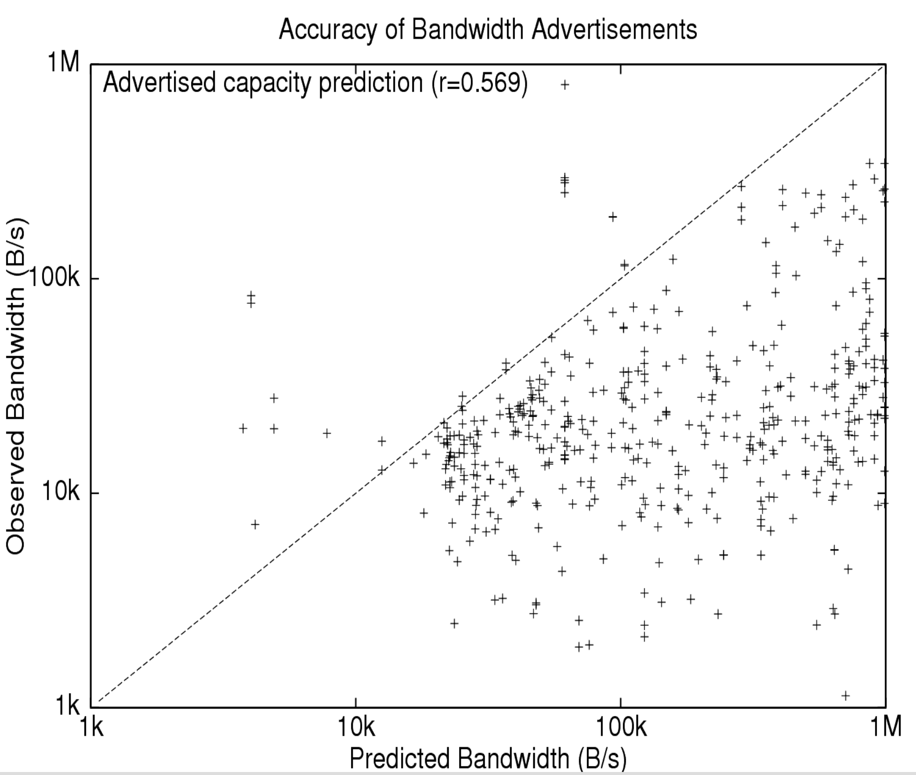
Evaluating Link Bandwidth (Current)

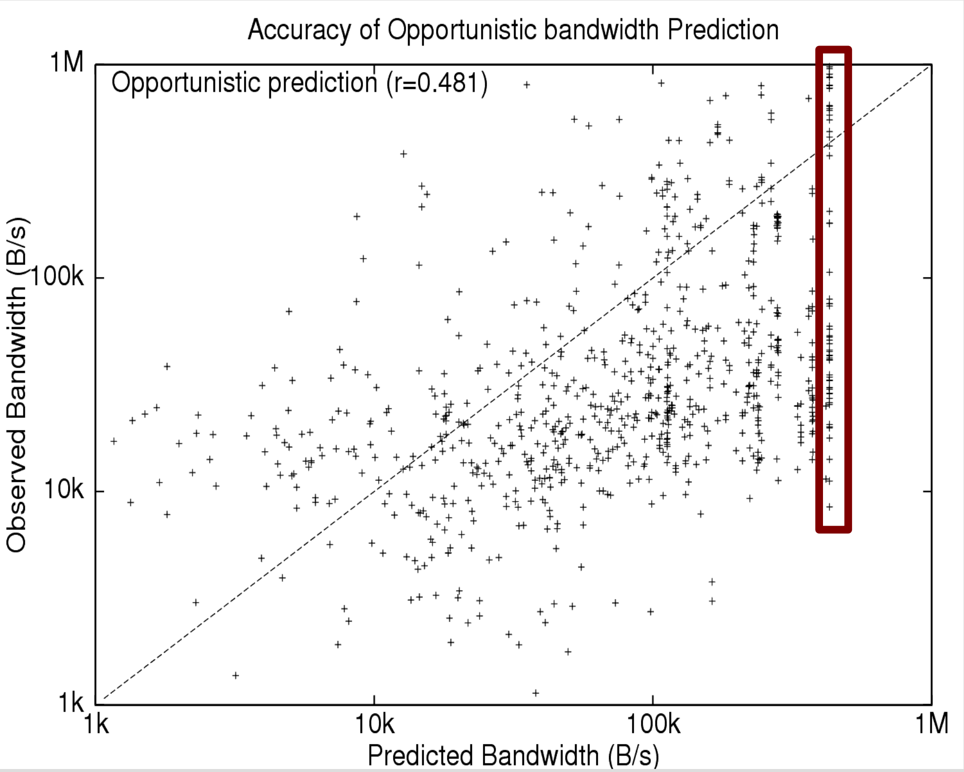
- Current Implementation
  - Each node estimates available bandwidth and reports it to directory server
  - Susceptible to manipulation by malicious nodes
    - Bauer, et al. “Low-resource routing attacks against anonymous systems” in WPES'07
  - Not sensitive to relative load
    - Static router popularity

Evaluating Link Bandwidth (Proposed)

- Proposed Method
  - Each node tracks the bandwidth to each of its peers
  - To estimate bandwidth, a node queries 5 of its peers and calculates the median values received
    - Nodes already query peers for lists of available nodes
  - Adjusts to relative load

Evaluation of Bandwidth Estimation



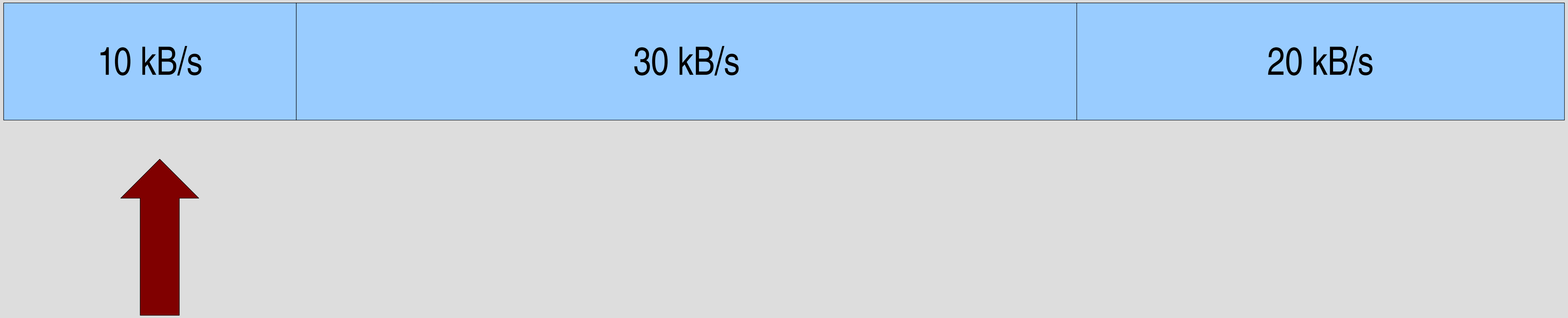


- Current Method Performance: r=0.57
  - Systematic overestimation
  - No malicious nodes

- Proposed Method Performance: r=0.48
  - Balanced prediction

Router Selection (Current)

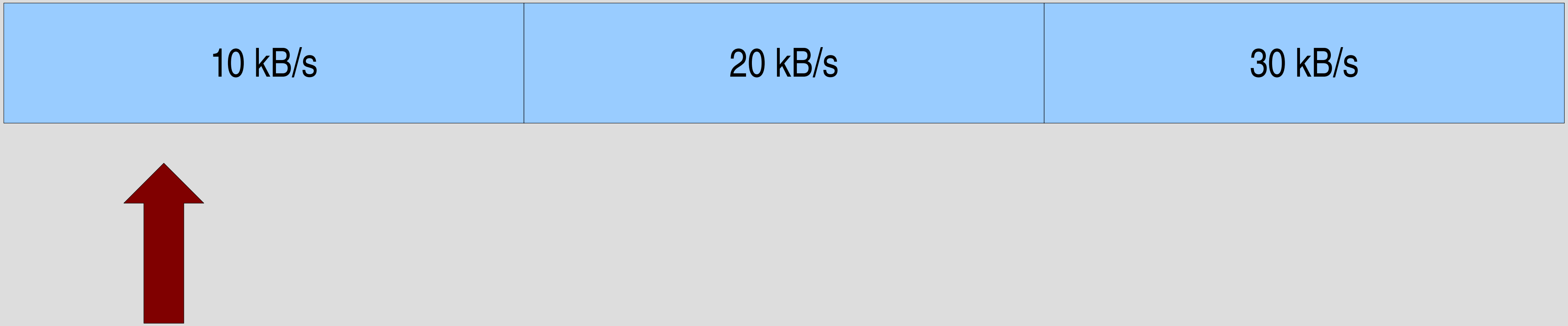
- Selection weighted by bandwidth



- Single Anonymity Level
- Bandwidth weight limited to 10 MB/s (was 1.5 MB/s)
  - Static tradeoff between underutilization and spoofing

Route Selection (Proposed)

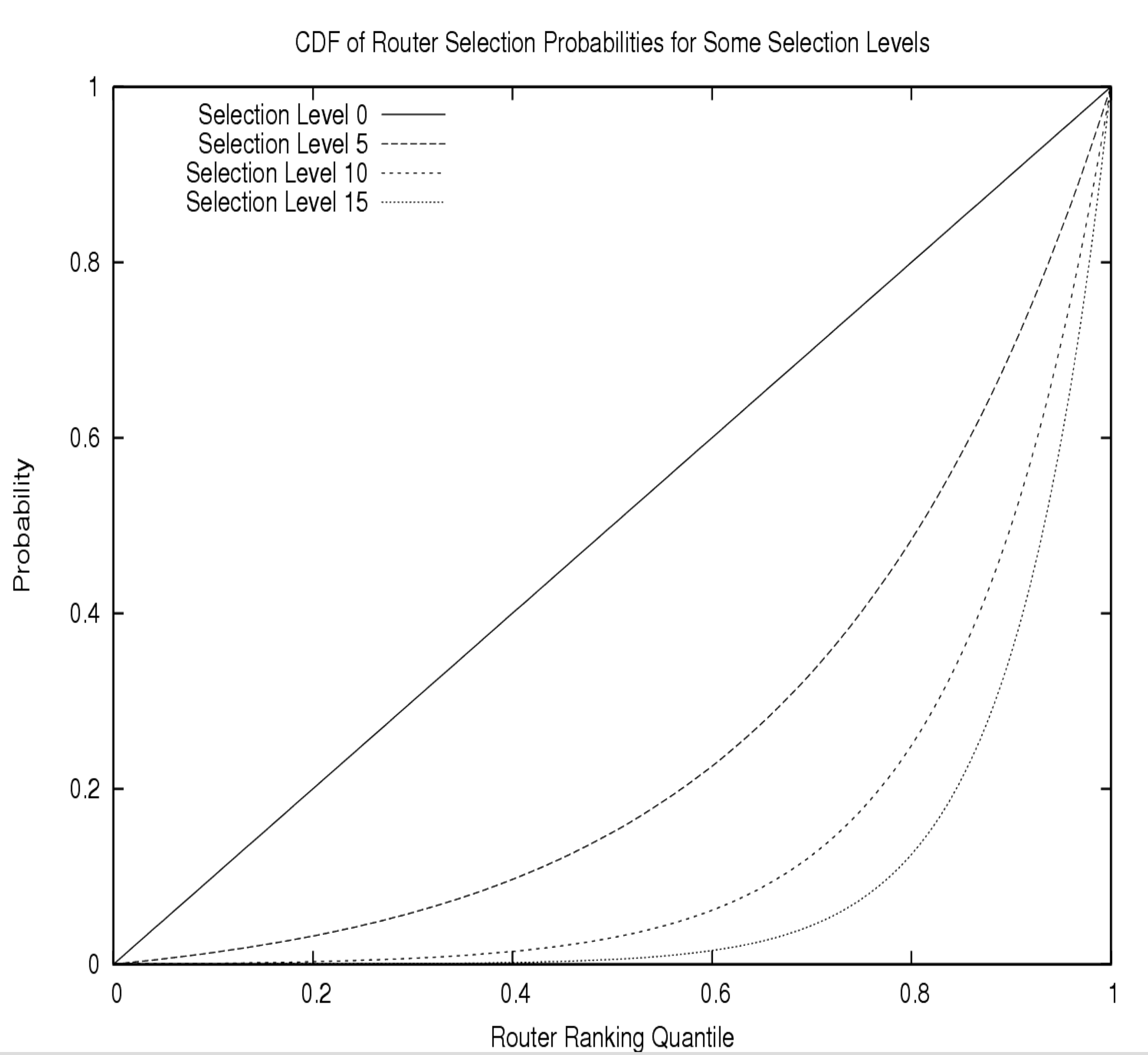
- Order routers by available bandwidth



10 kB/s	20 kB/s	30 kB/s

- Use non-uniform random variable to weight faster routers more heavily
- Parameterized RV => Parameterized Anonymity

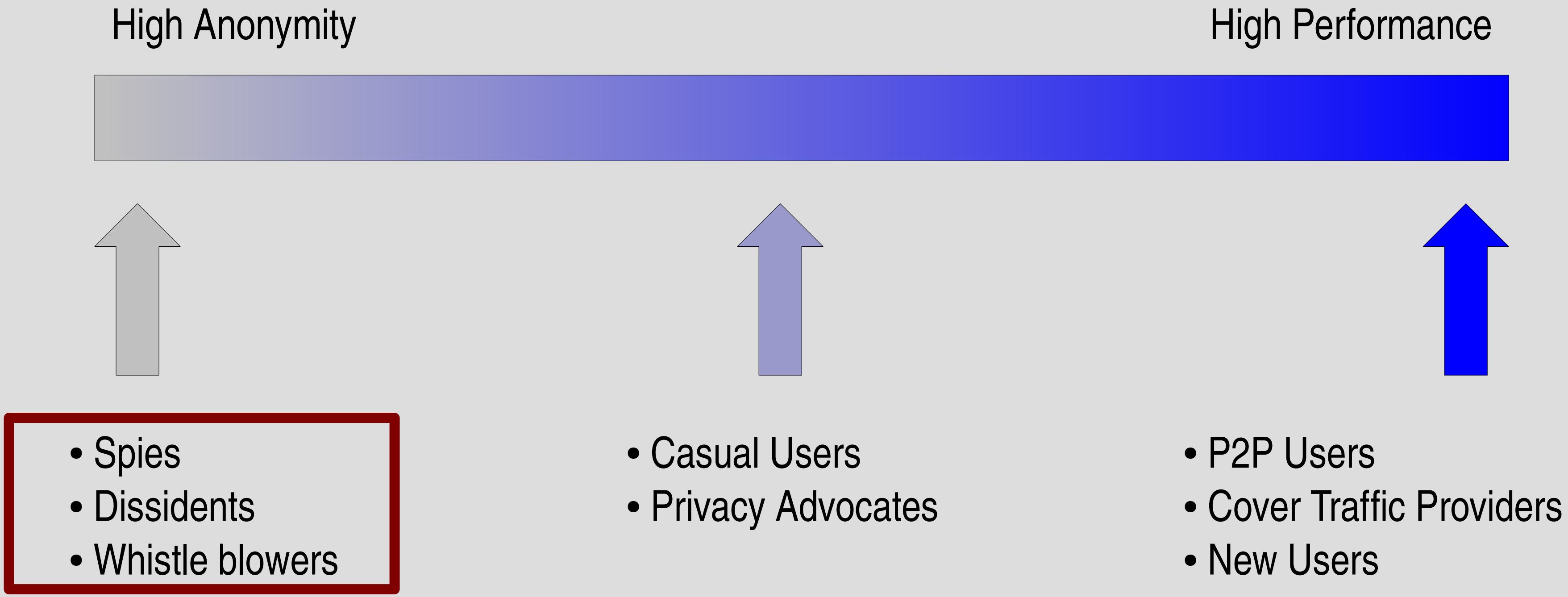
Route Selection (Proposed)



- Selection level 0 gives uniform selection
- Higher selection levels weight faster routers more heavily
- Weighted coin flip to choose known vs. unknown routers
  - Unknown routers always chosen uniformly at random

Evaluation of Router Selection

- Concern: how traceable is your selection level?
  - Attacker can focus on users more concerned with privacy



High Anonymity

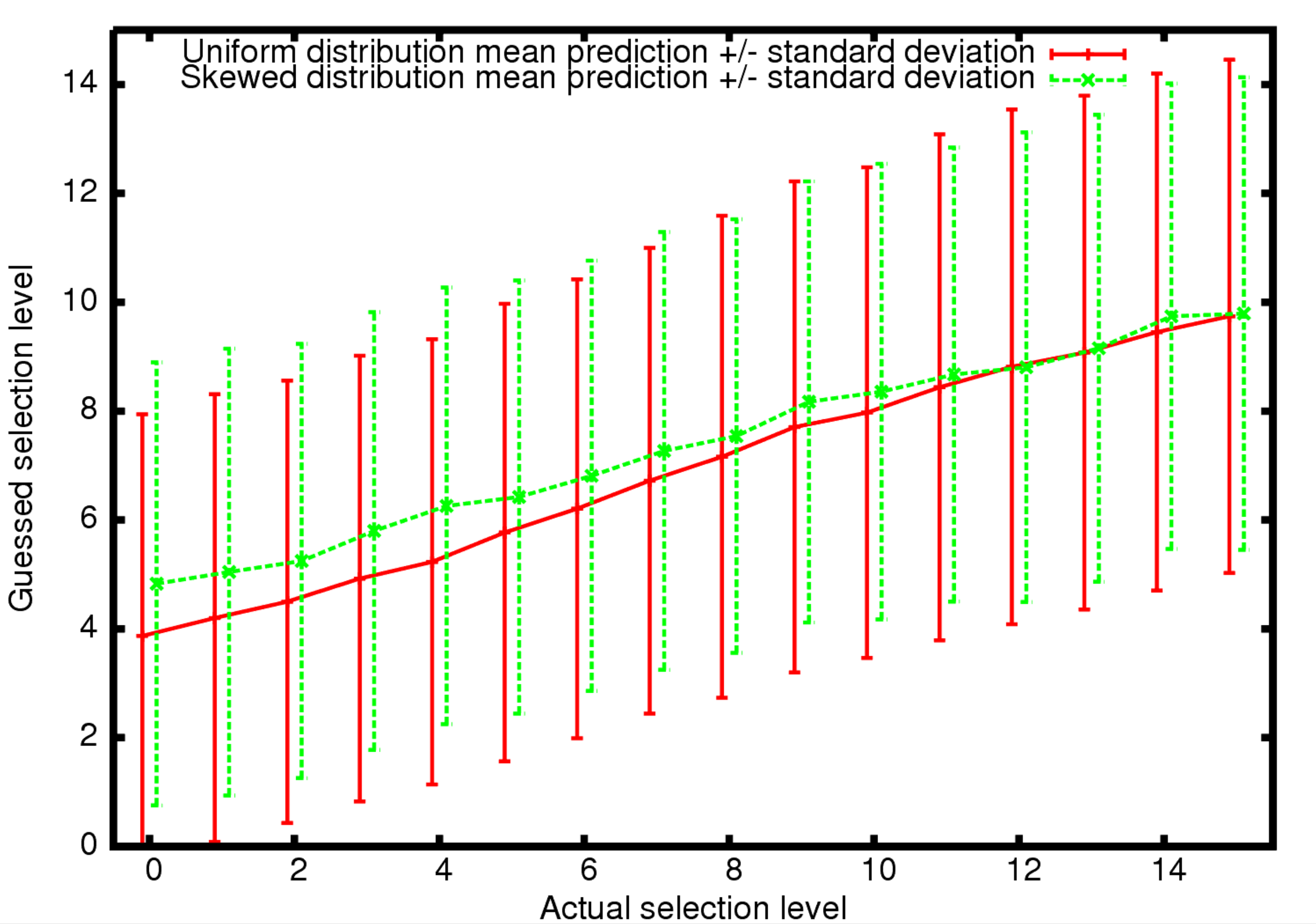
High Performance

- Spies
- Dissidents
- Whistle blowers

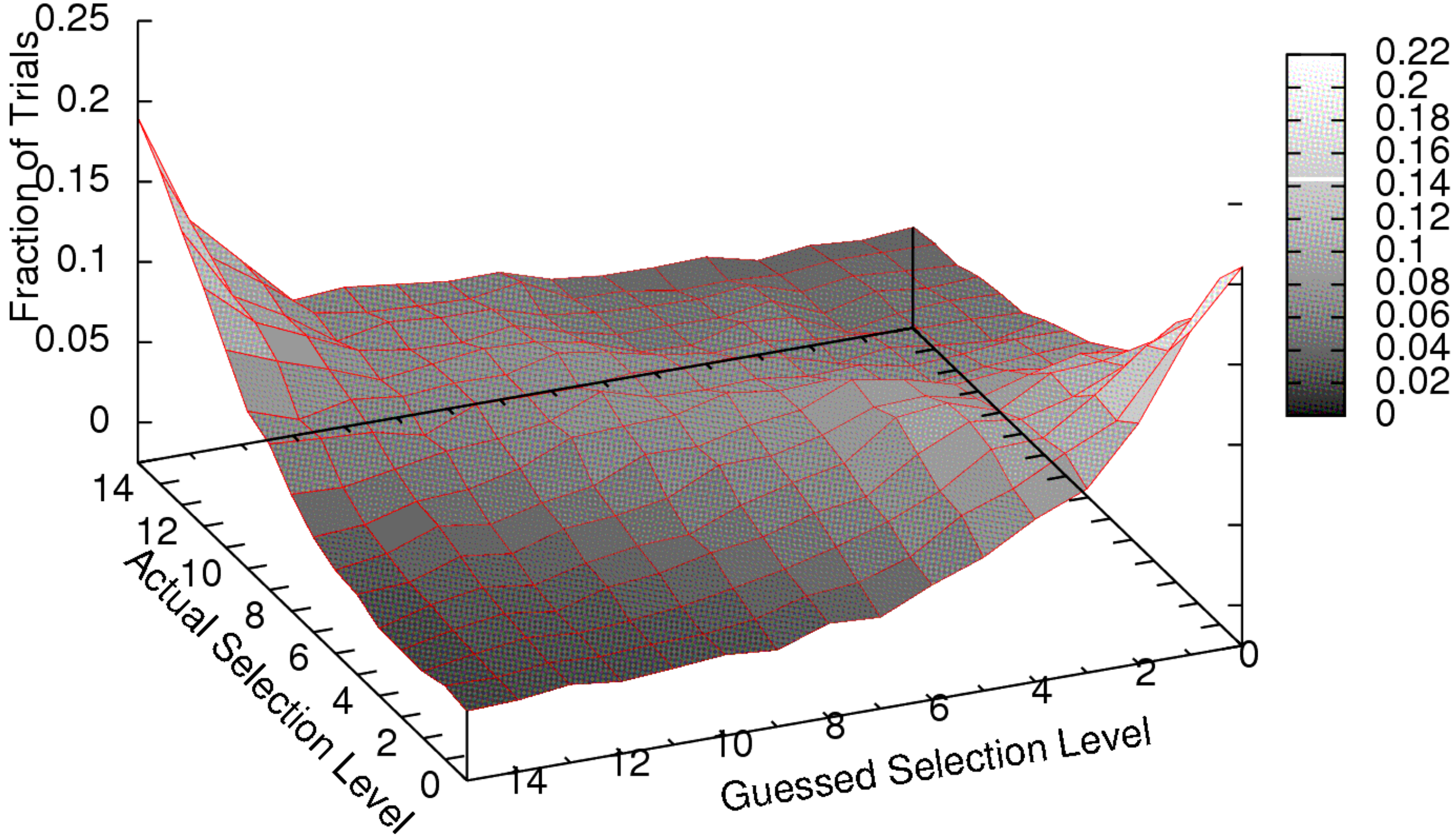
- Casual Users
- Privacy Advocates

- P2P Users
- Cover Traffic Providers
- New Users

Evaluation of Router Selection



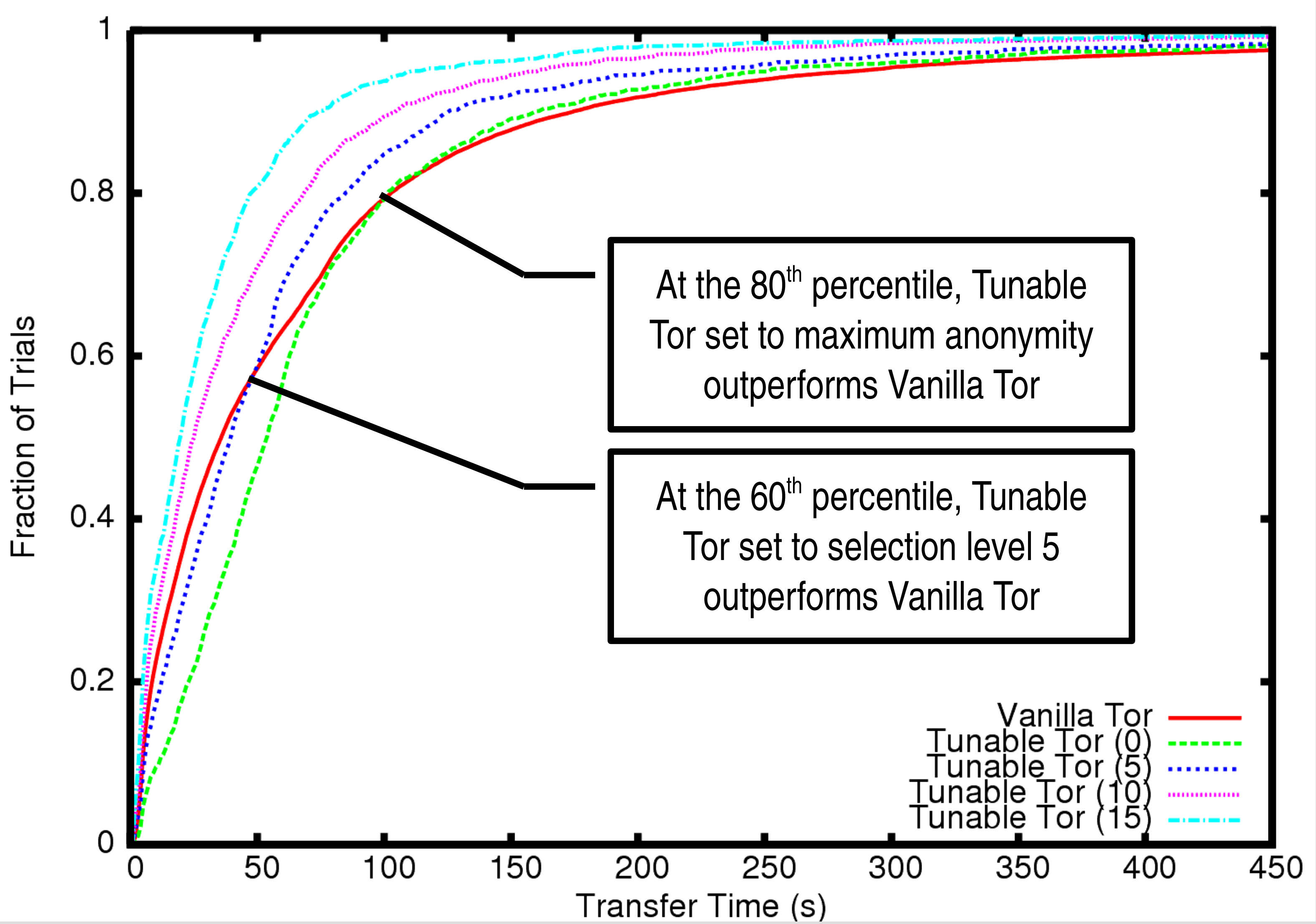
Evaluation of Router Selection



Tunable Tor: Combining both algorithms

- Evaluation setup:
  - Transfer 1 MB file
  - 40,000 trials for vanilla Tor over 4 weeks, various times of day
  - 20,000 trials for Tunable Tor over 6 weeks, various times of day
    - Selection level chosen uniformly at random
- Evaluate performance
  - Transfer time statistics
- Evaluate anonymity
  - Router selection equality
  - Effects of router compromise

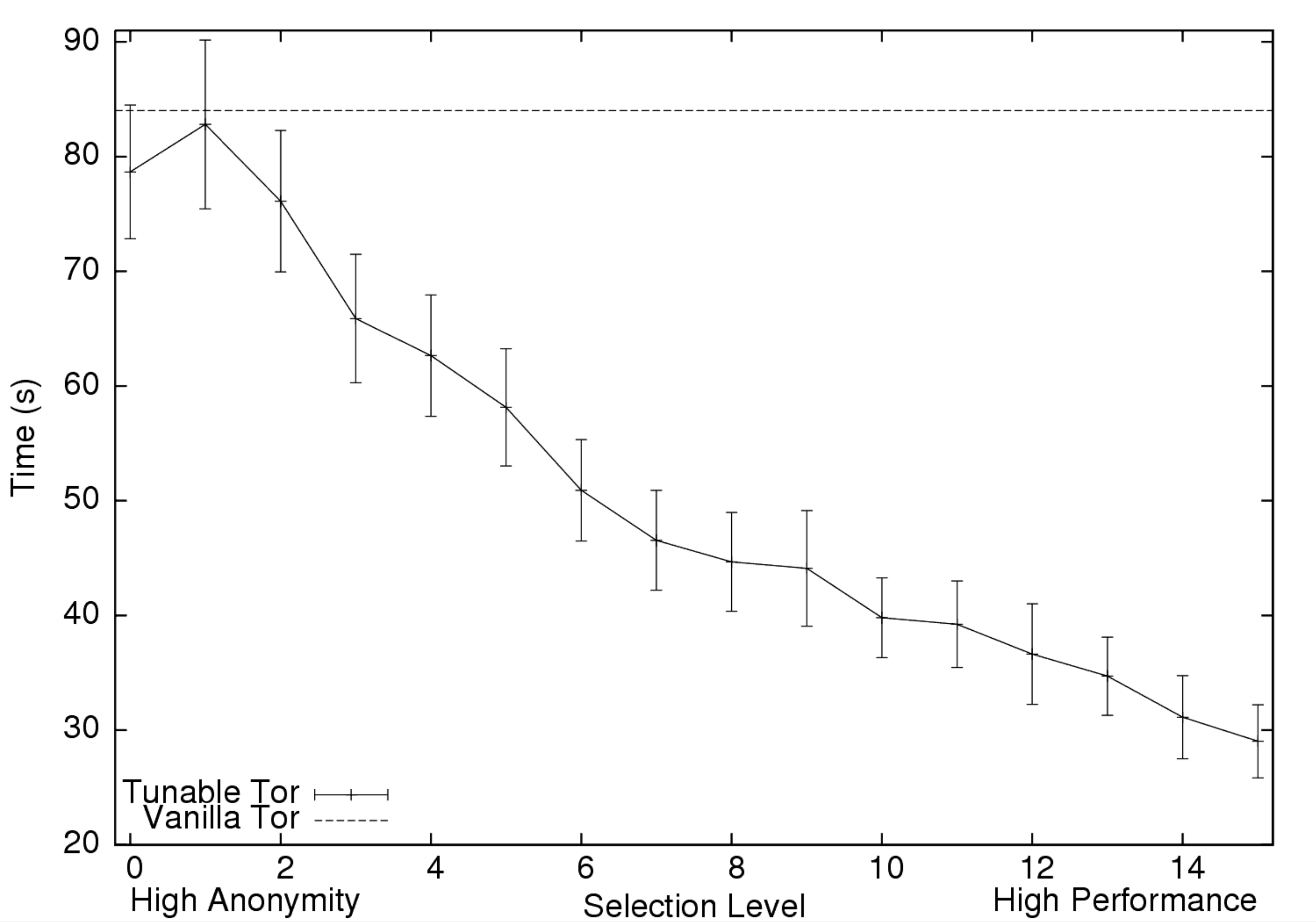
Whole System Evaluation (Performance)



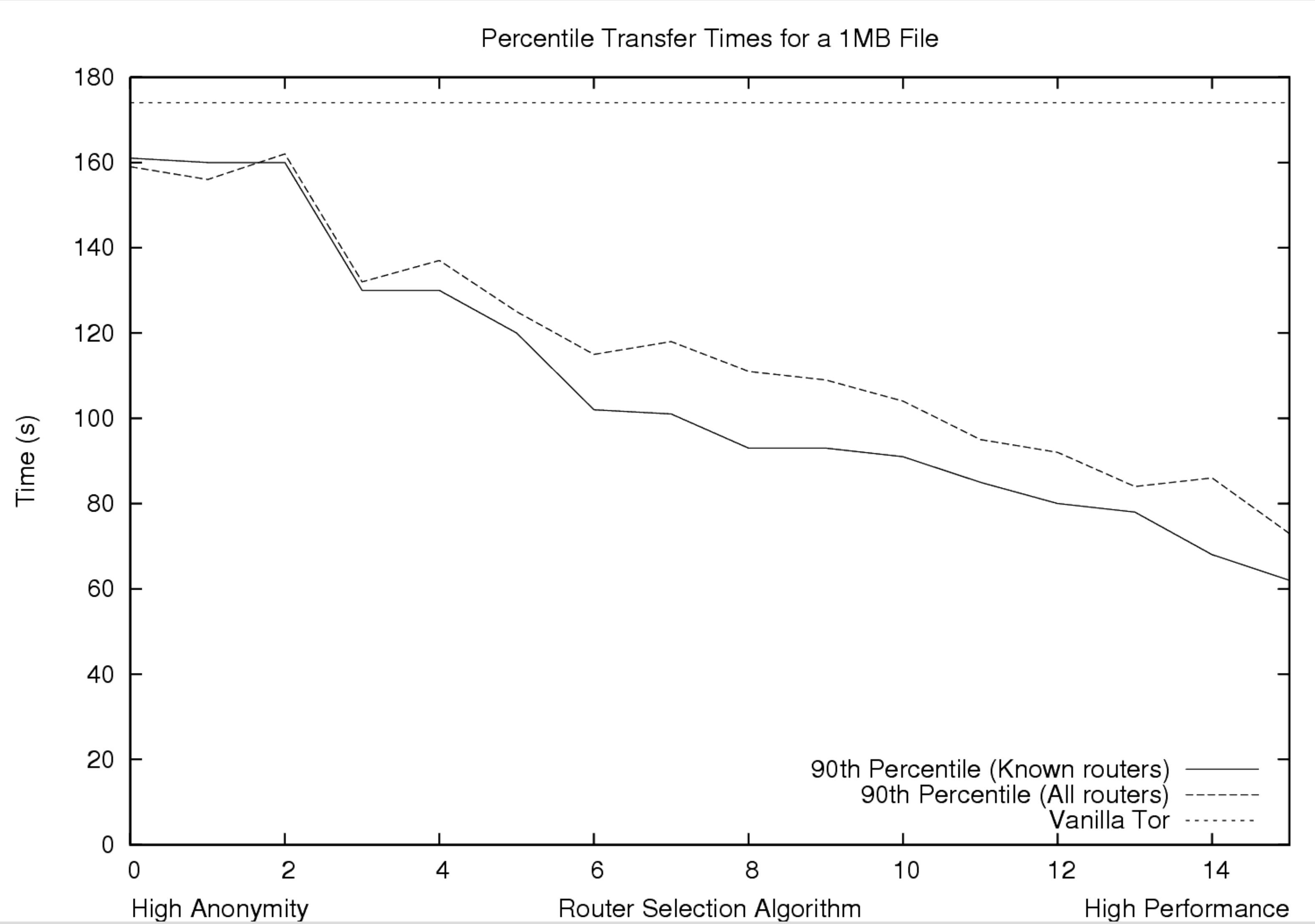
At the 80th percentile, Tunable Tor set to maximum anonymity outperforms Vanilla Tor

At the 60th percentile, Tunable Tor set to selection level 5 outperforms Vanilla Tor

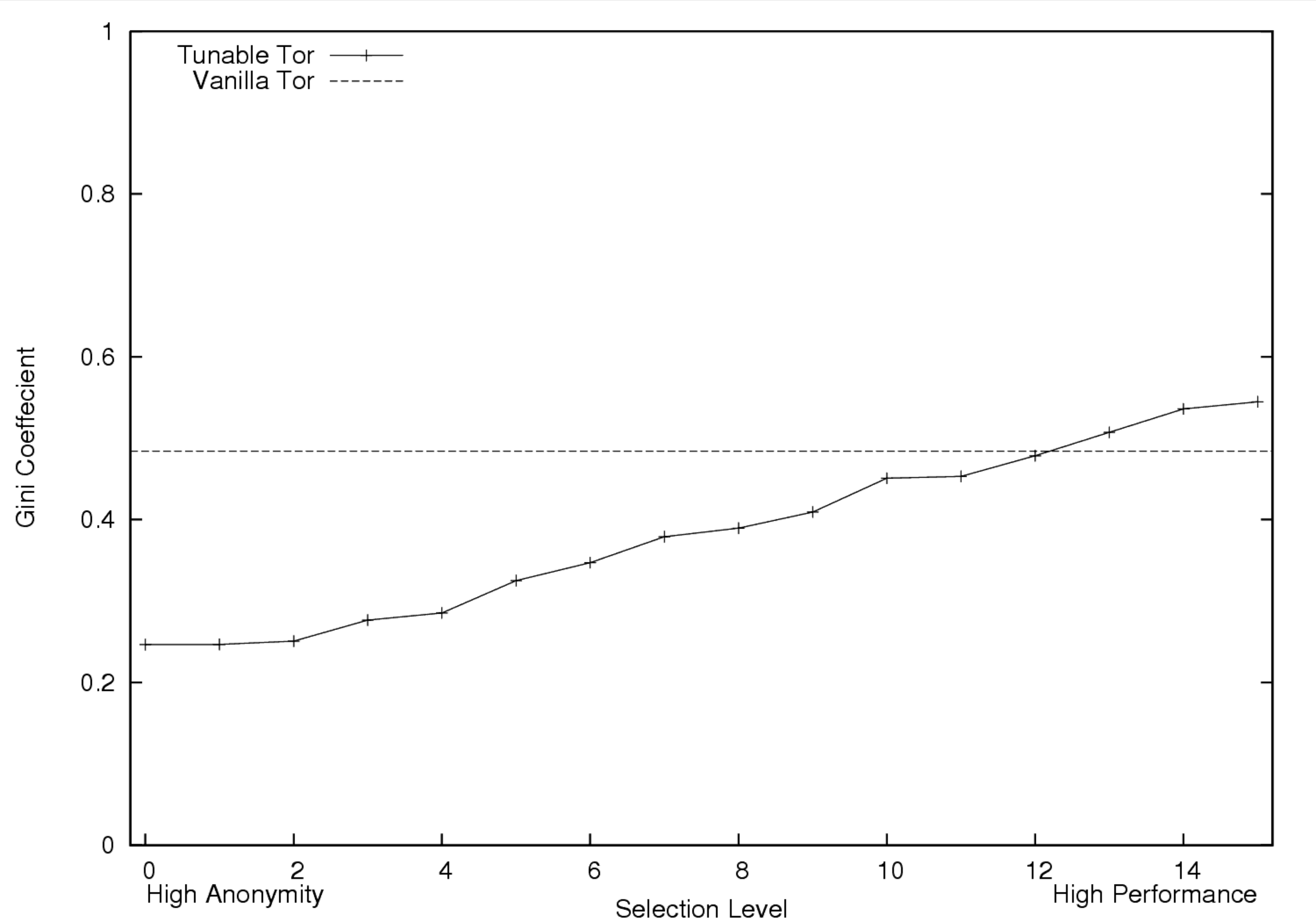
Whole System Evaluation (Performance)



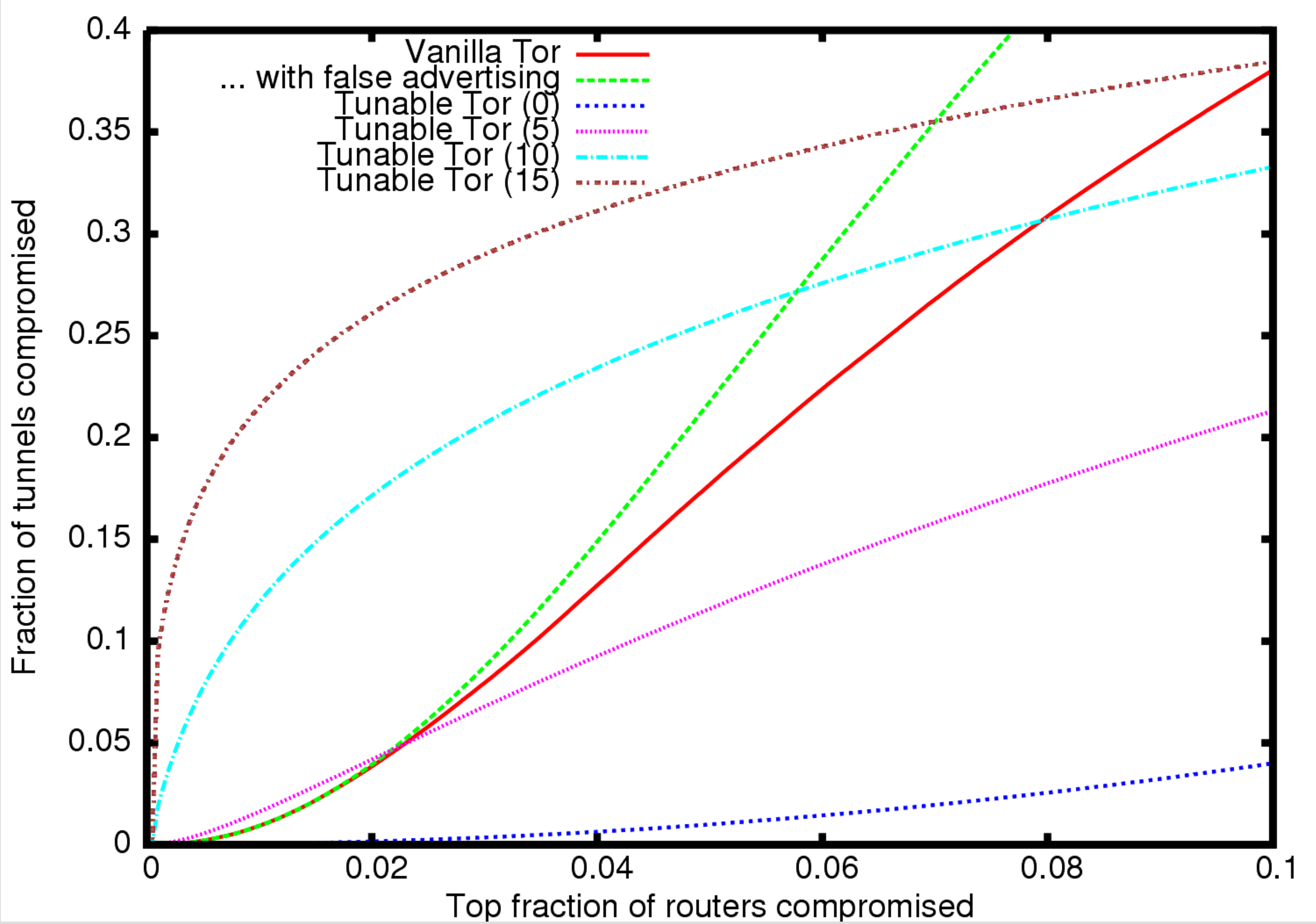
Whole System Evaluation (Performance)



Whole System Evaluation (Anonymity)



Whole System Evaluation (Anonymity)



Conclusions

- Tunable Tor provides:
  - Significantly more security
    - No reliance on self-reported information
    - Multiple, randomly selected, opportunistic router evaluations prevent targeted attacks
  - Tunability
    - 3x throughput improvement for the same anonymity
    - Dramatically more anonymity for the same performance
  - Much shorter “long tail”
  - But...

Current Work: Whole Network Simulation

- What happens when all nodes in the network are using these algorithms?
- Plan
  - Simulate 1000 nodes, 10,000 flows
  - Choose routes according to
    - Current Tor algorithm
    - All users using new algorithm
      - Everybody at a single selection level (for all levels)
      - Plausible mixes (20% level 0, 30% level 15, 5% each for the rest)
    - Transitional phase (some old, some new)

Current Work: Bandwidth Estimation Testing

- Can peer bandwidth measurements from low-bandwidth hosts be used?
- Plan:
  - Patch to monitor peer bandwidth periodically being distributed
  - Compare
    - Measured bandwidth
    - Measured bandwidth ranking

    from hosts with different available bandwidth