Building an AS-Topology Model that Captures Route Diversity

Wolfgang Mühlbauer
Technische Universität München
wolfgang@net.in.tum.de

Anja Feldmann
Deutsche Telekom Laboratories

Olaf Maennel
University of Adelaide

Matthew Roughan
University of Adelaide

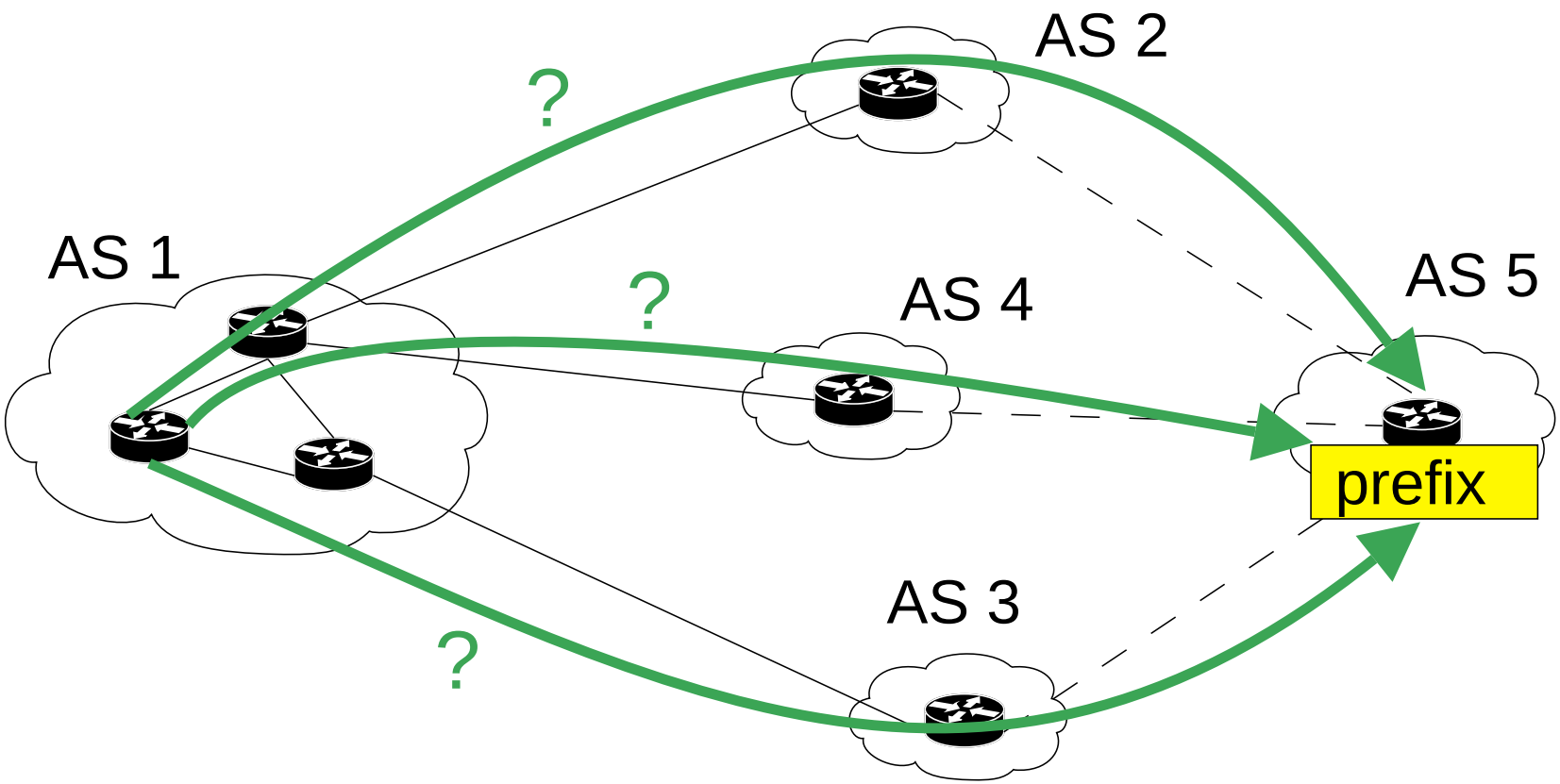
Steve Uhlig
Université Cath. de Louvain

- **Realistic** models of the Internet required
  - to evaluate protocols
  - to evaluate new architectures
  - ...
- **How to model** the Internet?
  - Which granularity/abstraction?
  - Which topology?
  - Which policies?
  - ...

Our contribution: Predict AS paths

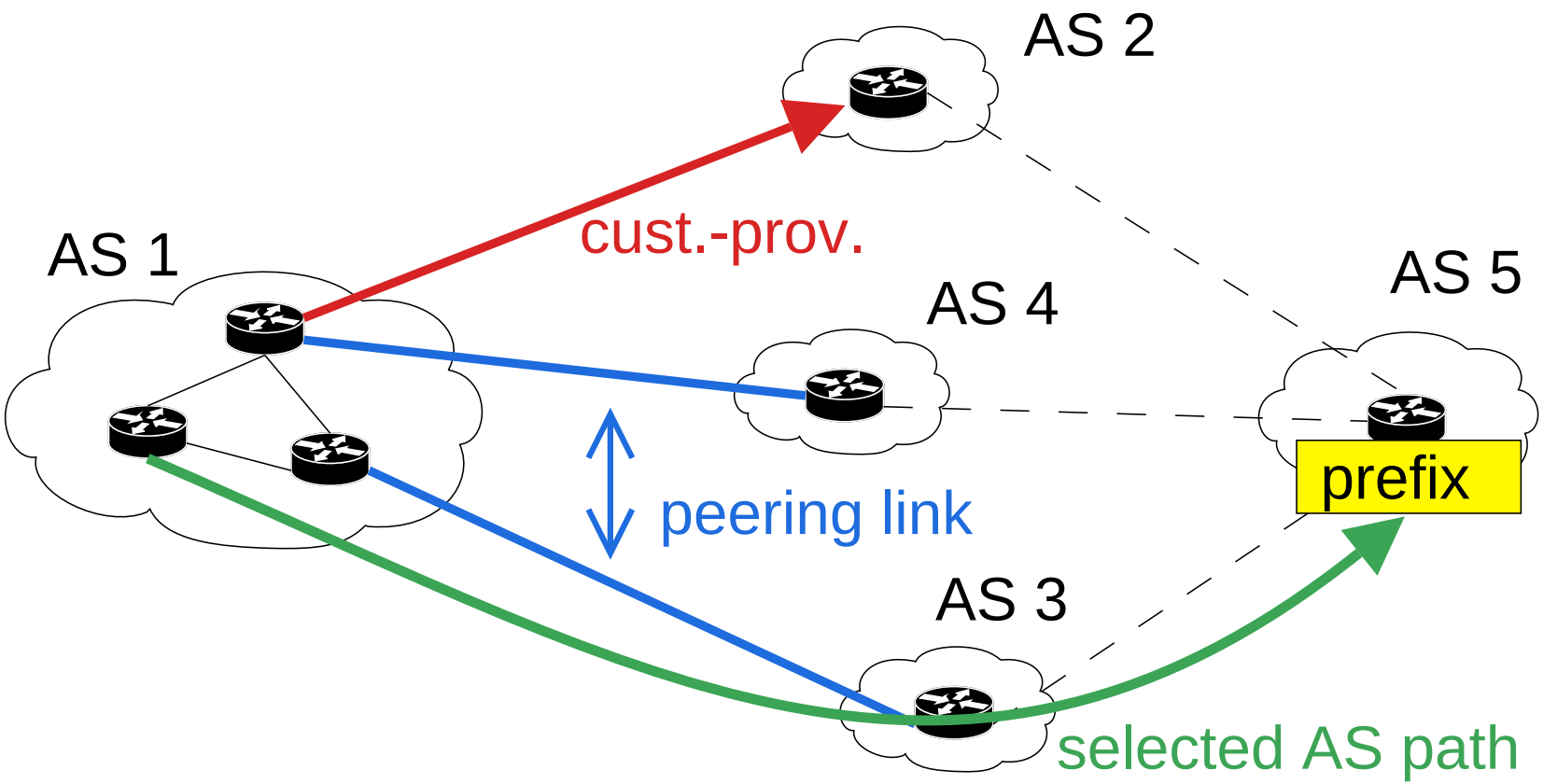
Goal: Predict AS Paths

- Difficulty: Route diversity at AS 1



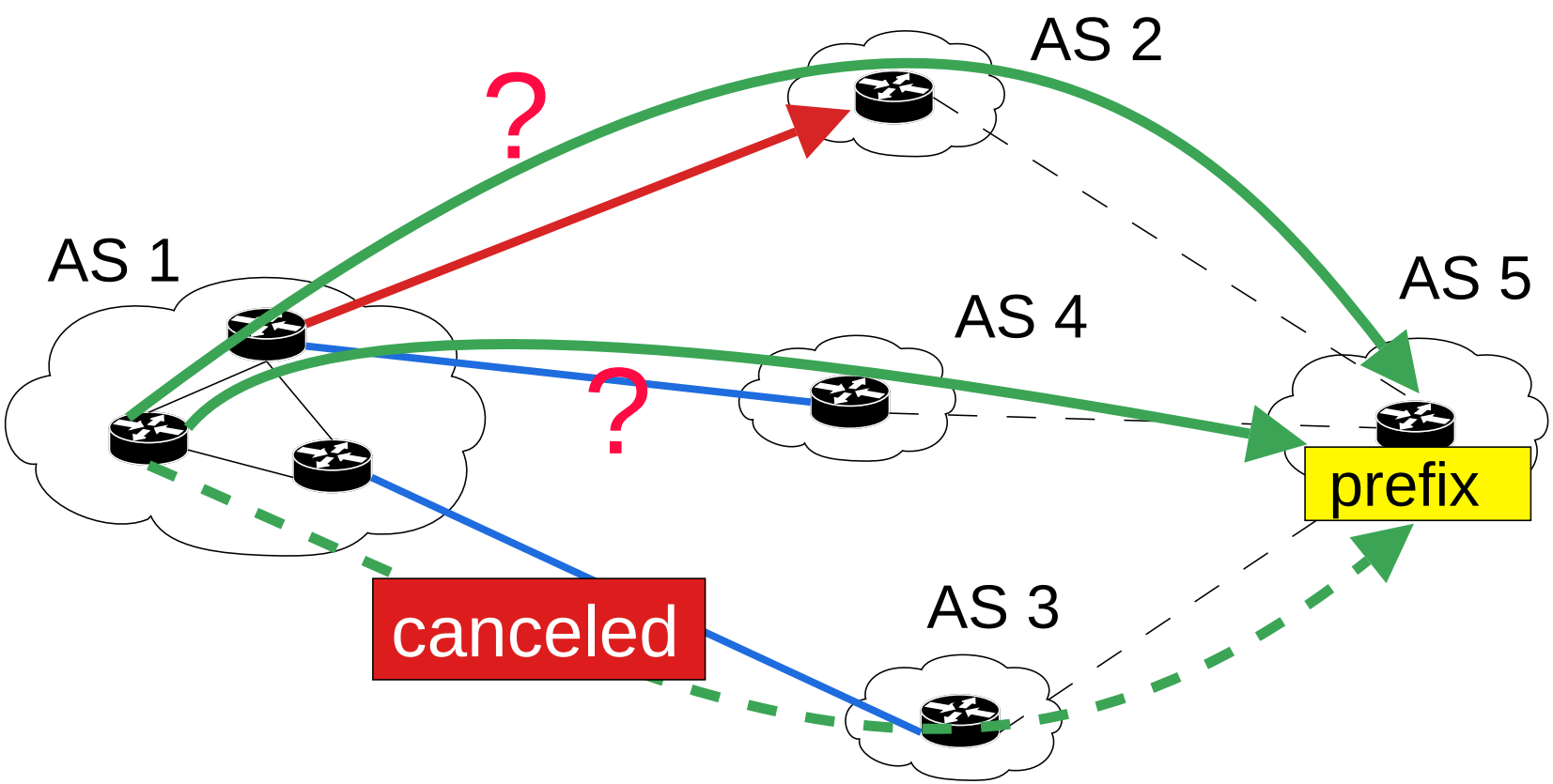
Goal: Predict AS Paths

- Difficulty: Policy routing



Long-term Vision: “What-If” Questions

- Example: Predict impact of policy changes



Building an AS-Topology Model

- **Prior models**:
  - Simplified policies/relationships: customer-provider/peering
  - Assumption: “AS $=$ one router”
- **Our approach**: Allow for
  - path diversity
  - agnostic routing policies

Outline

Data Sets

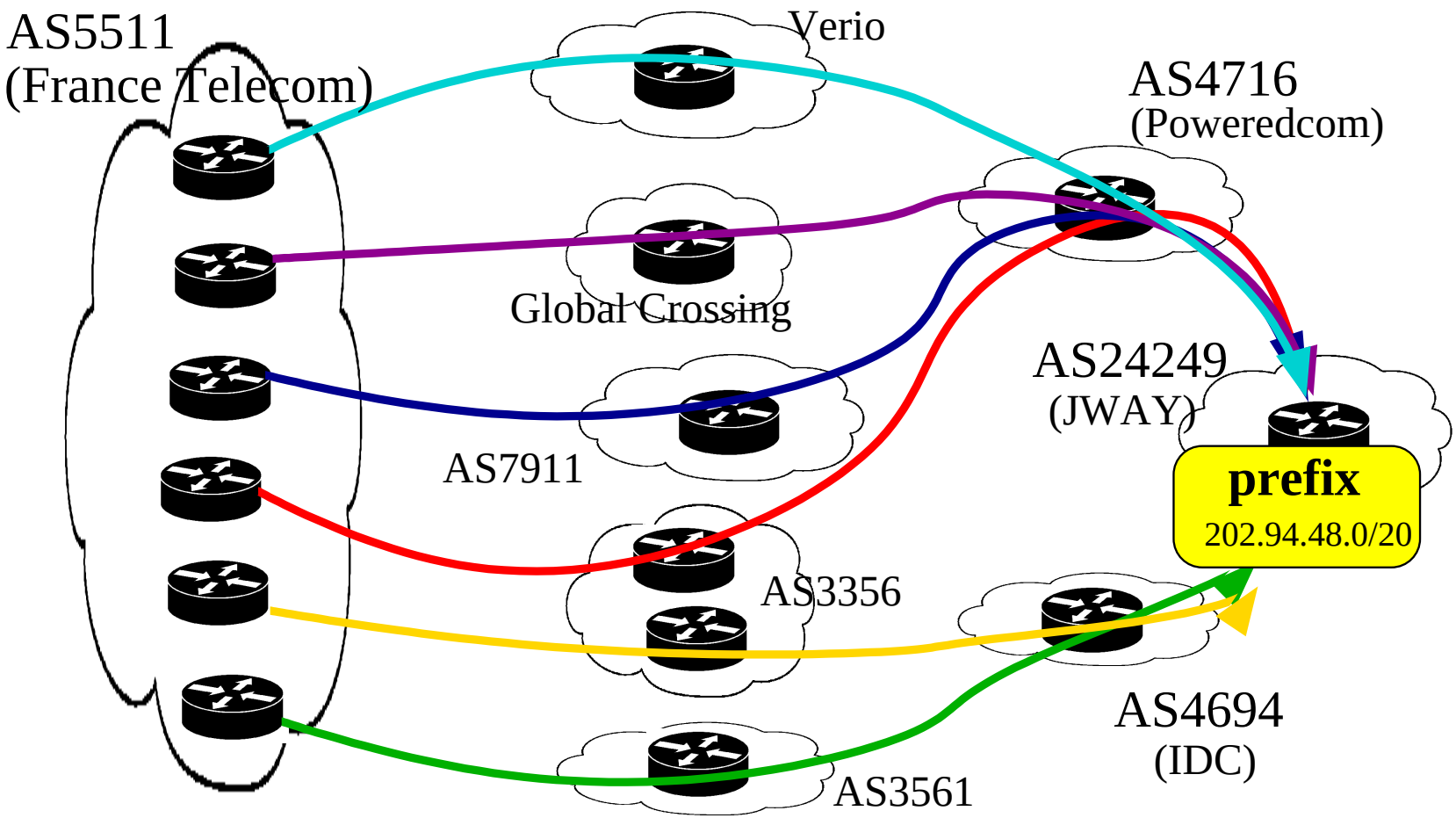
- BGP table dumps (Sun, November 13, 2005)
  - RIPE, RouteViews, ...
  - $>$ 1,300 observation points throughout the Internet
  - 4,730,222 distinct AS paths
- Infer AS graph
  - ASes: 21,178
  - AS-level edges: 58,903
- Focus on multi-homed ASes

Appropriate to Model ASes with Single Routers?

- **Simple experiment**:
  - Run simulation with C-BGP
    - Computes routing tables for each router
  - Check if obs. paths are selected in simulations for
    - No policies
    - Standard routing policies (customer-provider, peering)
- **Results**: AS-paths which agree
  - No policies: 23.5%
  - Standard routing policies: 12.5%
- **Possible Reasons**
  - Route diversity?
  - Routing policies?

Importance of Route Diversity - Example

- Observable AS paths for prefix `202.94.48.0/20`



Route Diversity Indicated by Observable AS Paths

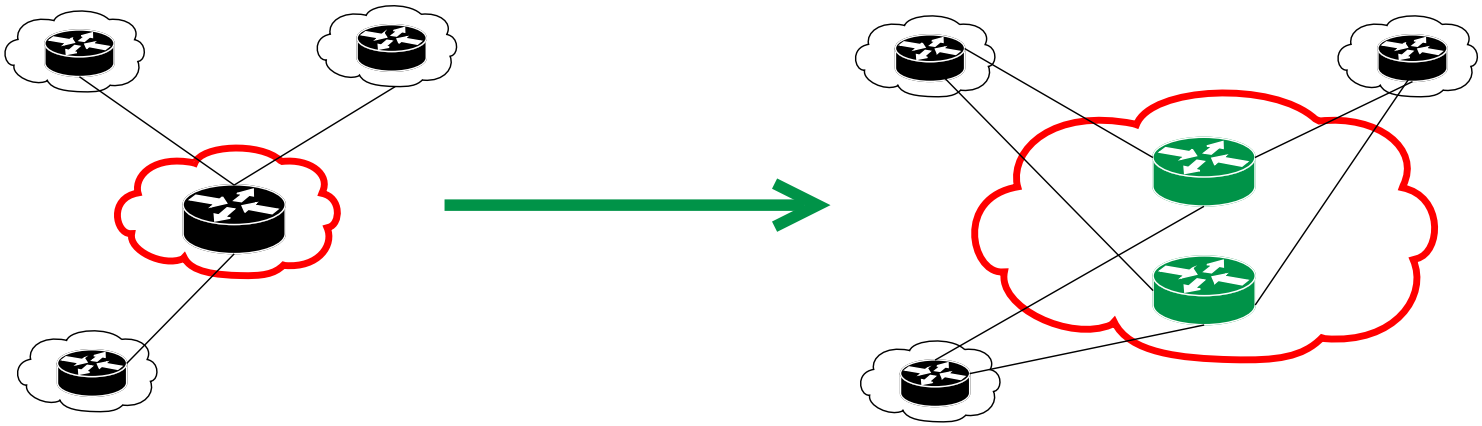
- 30% of AS-pairs: > 1 observed path
- Lower bound for number of routers inside an AS:
  Max. number of unique AS paths received towards **any** prefix
  - 50% of ASes: $>$ 2 routers
  - 10% of ASes: $>$ 5 routers
  - 2% of ASes: $>$ 10 routers

Approach (1)

- **Go beyond**
  - single router per ASes ($\rightarrow$ route diversity)
  - AS-relationship inference (customer-provider, peering)
- **Partition** data set
  - `Training Set`: used to build model
  - `Validation Set`: used to evaluate model
- Build model where **all** paths from `Training Set` are selected!

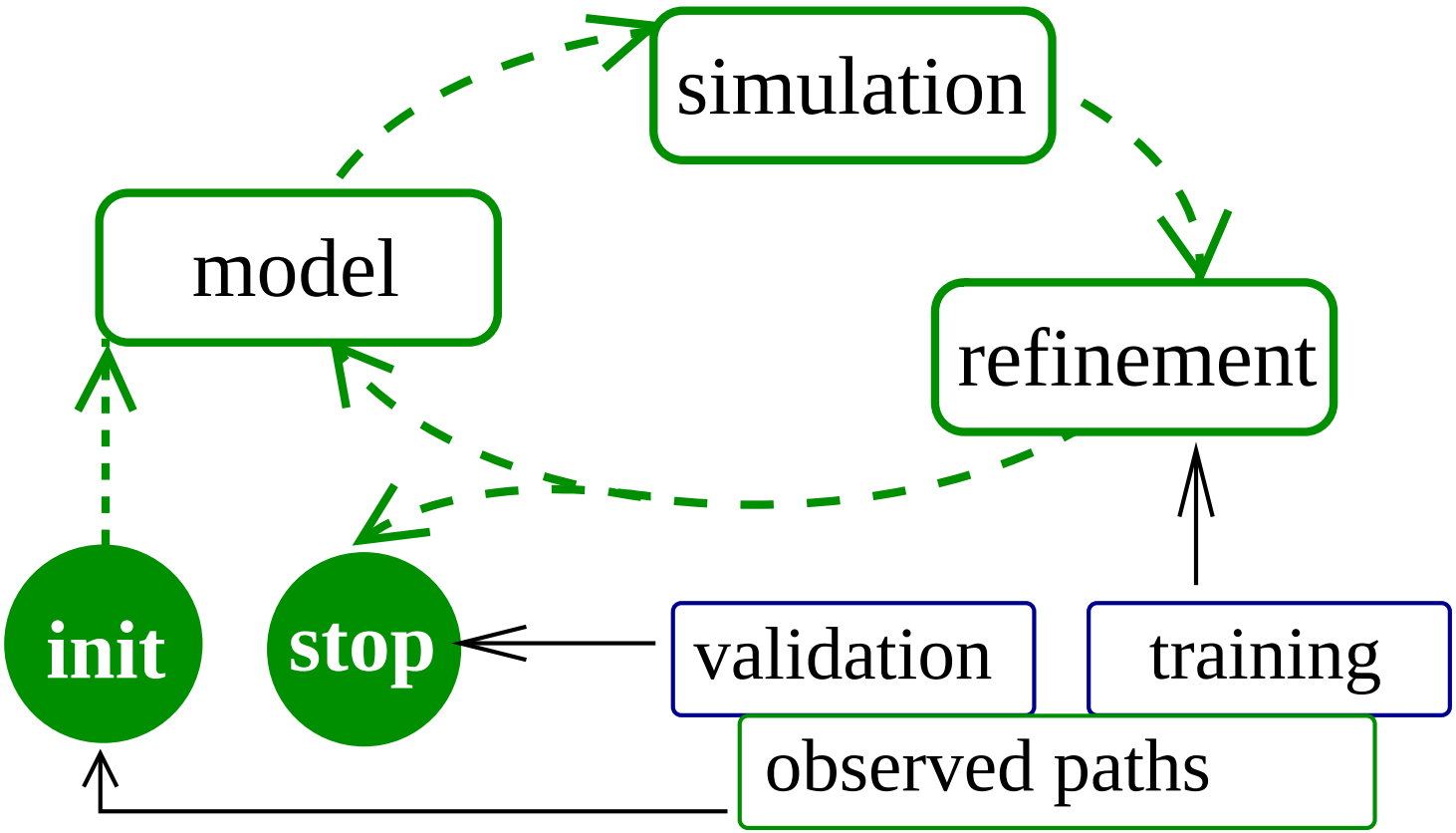
Approach (2)

- Capture Inter-domain connectivity: via **AS-graph**
- Manipulate route propagation: via **filters/policies**
- Capture relevant route diversity: via **quasi-router**
  - Group of routers in AS with same choice on routes



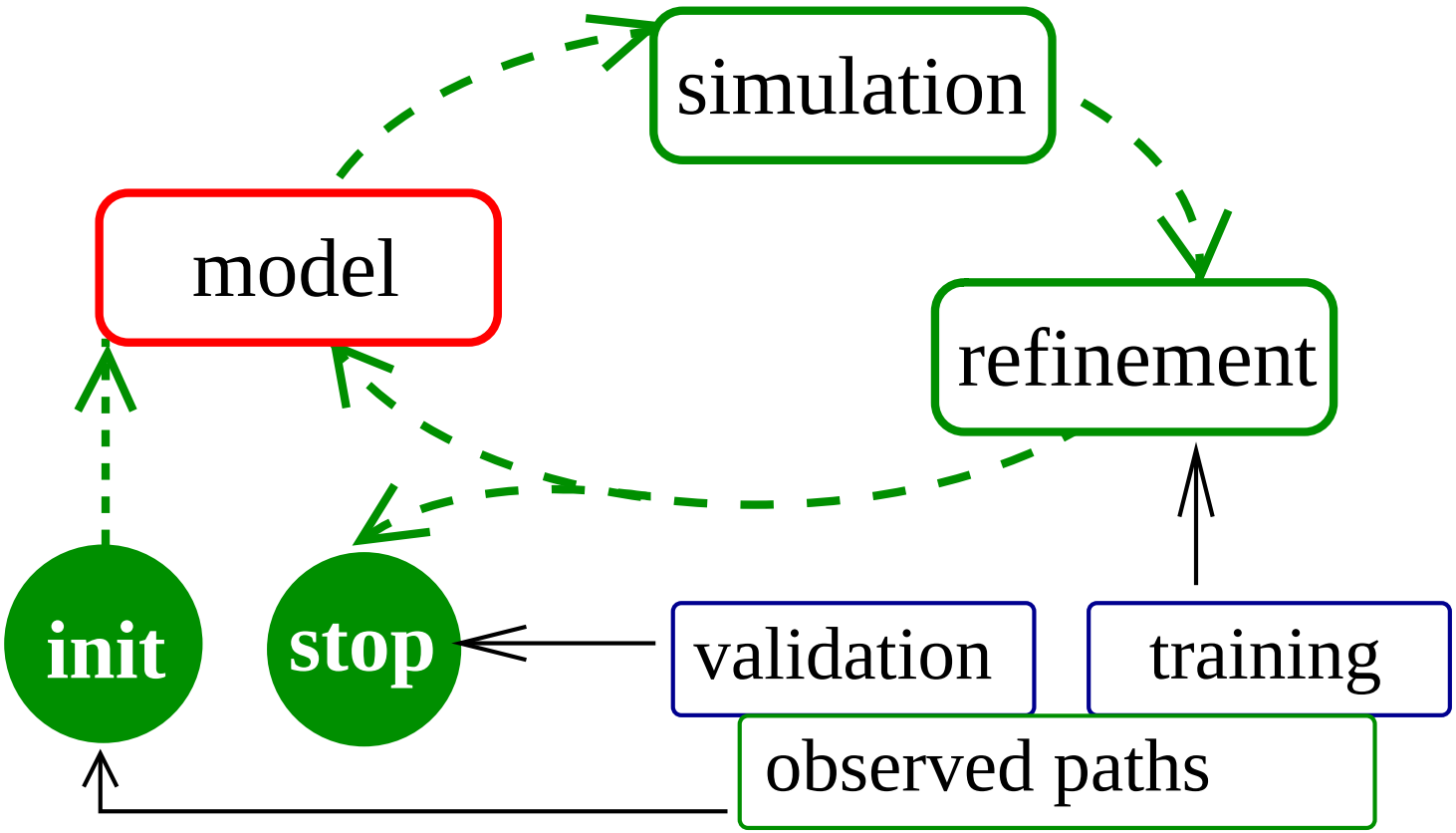
Methodology - Iterative Refinement

Goal: Train model to obtain observable paths



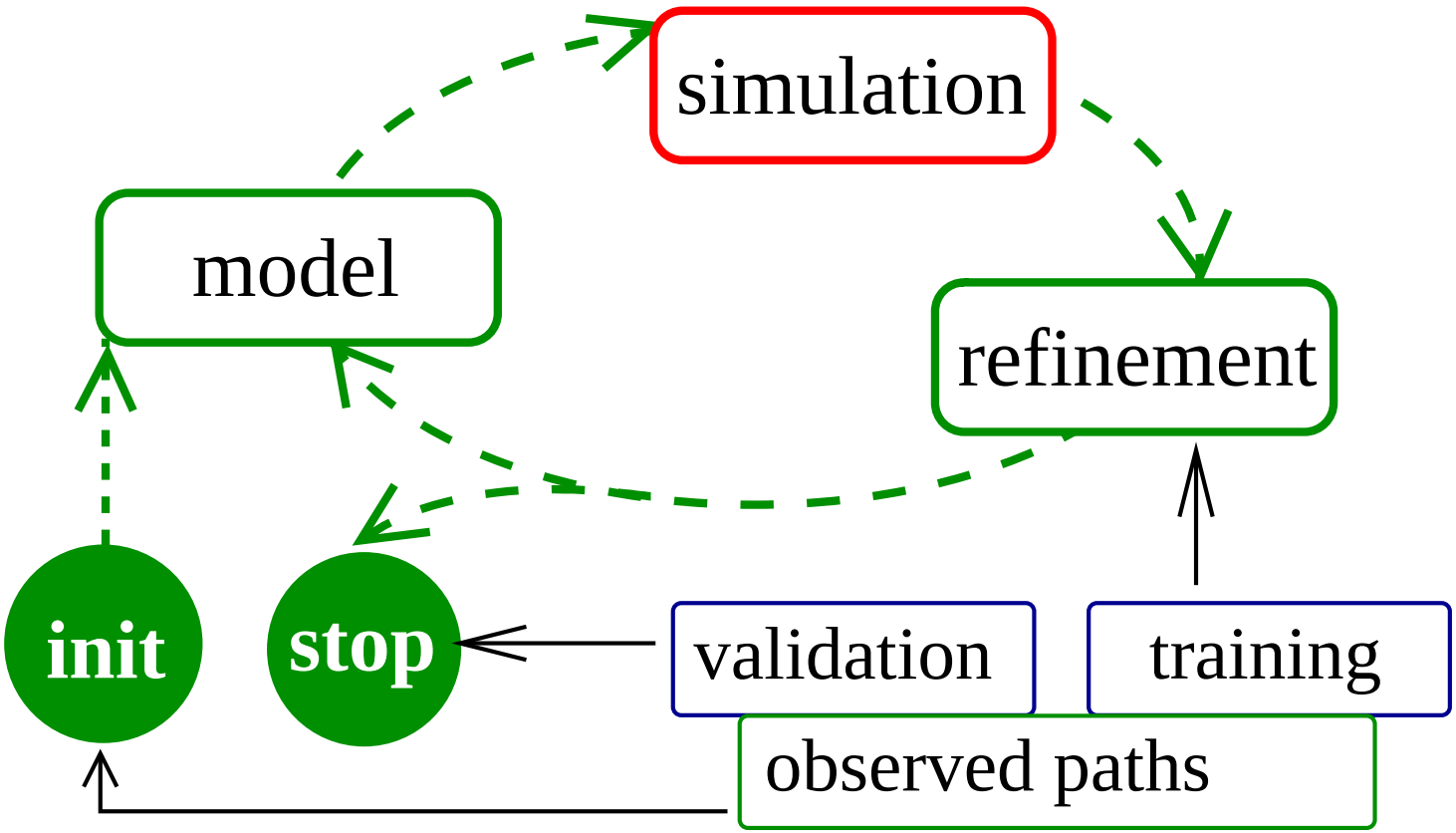
Methodology - Iterative Refinement

Goal: Train model to obtain observable paths



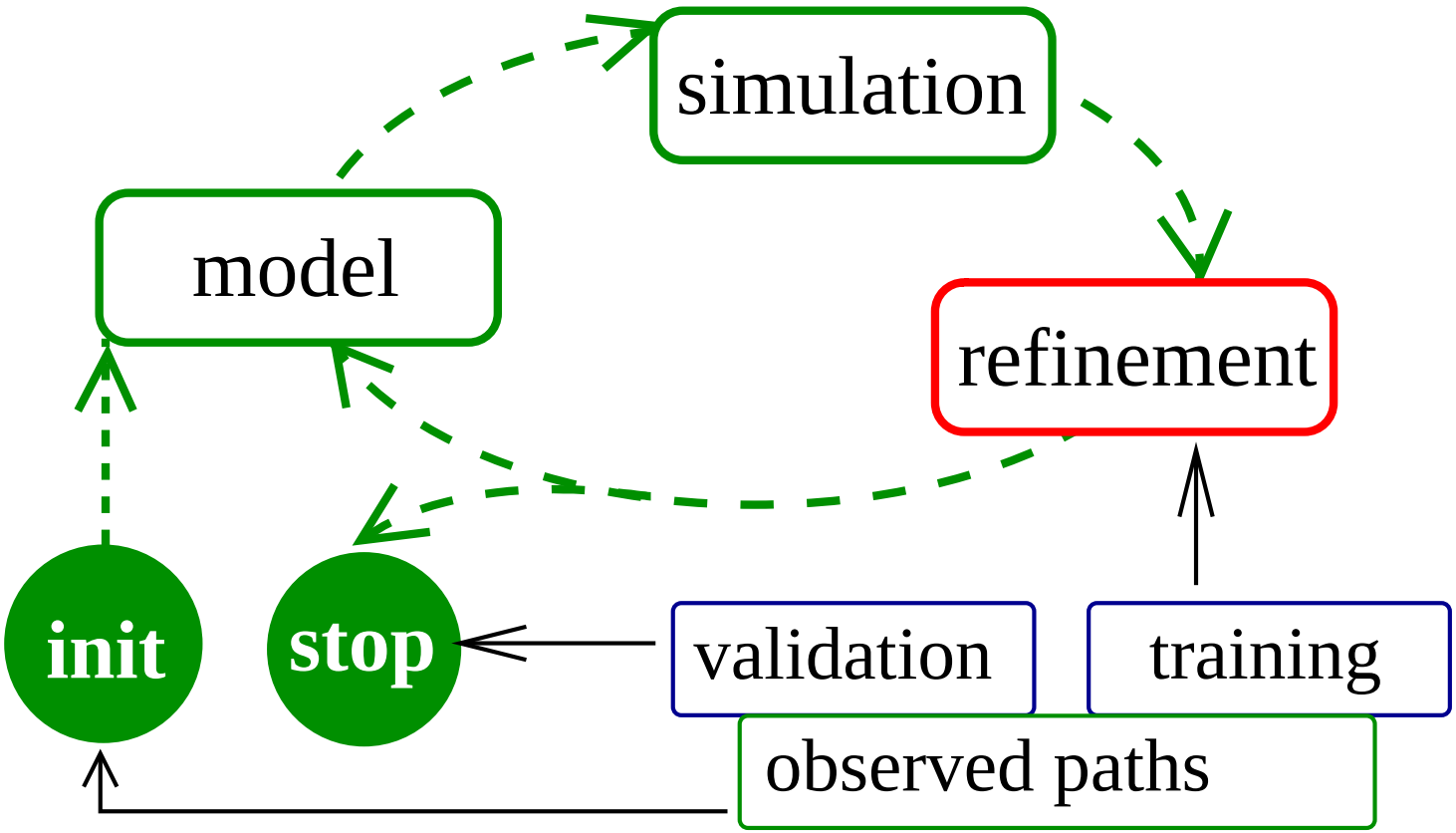
Methodology - Iterative Refinement

Goal: Train model to obtain observable paths



Methodology - Iterative Refinement

Goal: Train model to obtain observable paths



Metrics

- **Measure**
  - progress of iterative refinement
  - model performance for non-trained paths
- How many observable paths are
  - **selected** as best route by at least one *quasi-router*?
  - **learned** by at least one *quasi-router*?
    - But not necessarily selected as best route

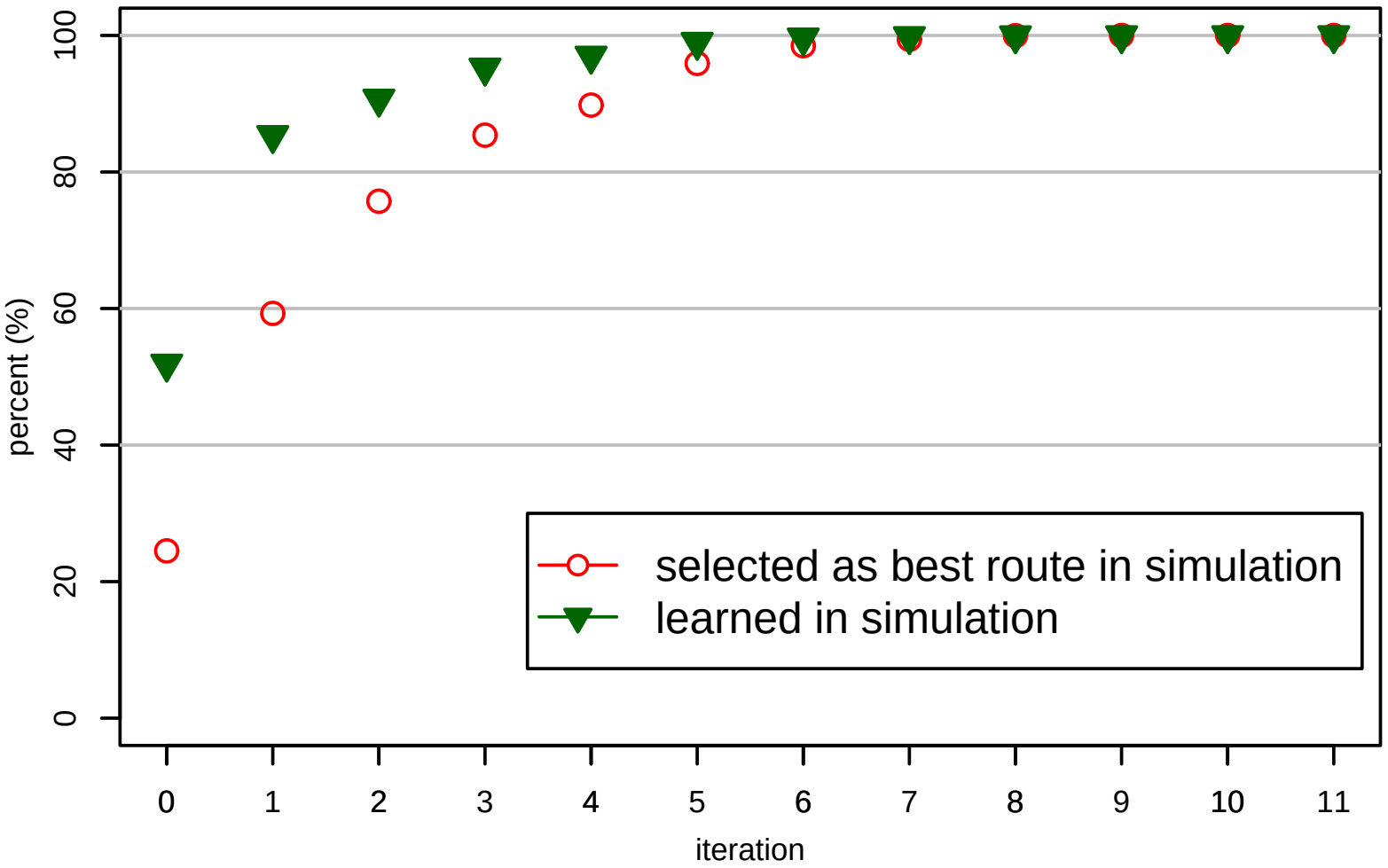
Learned Routes:

	Network	Next Hop
>	1.2.3.0/24	1.1.1.1
*>	1.2.3.0/24	2.2.2.2
>	5.6.7.0/24	8.8.8.8
*>	5.6.7.0/24	9.9.9.9

selected routes

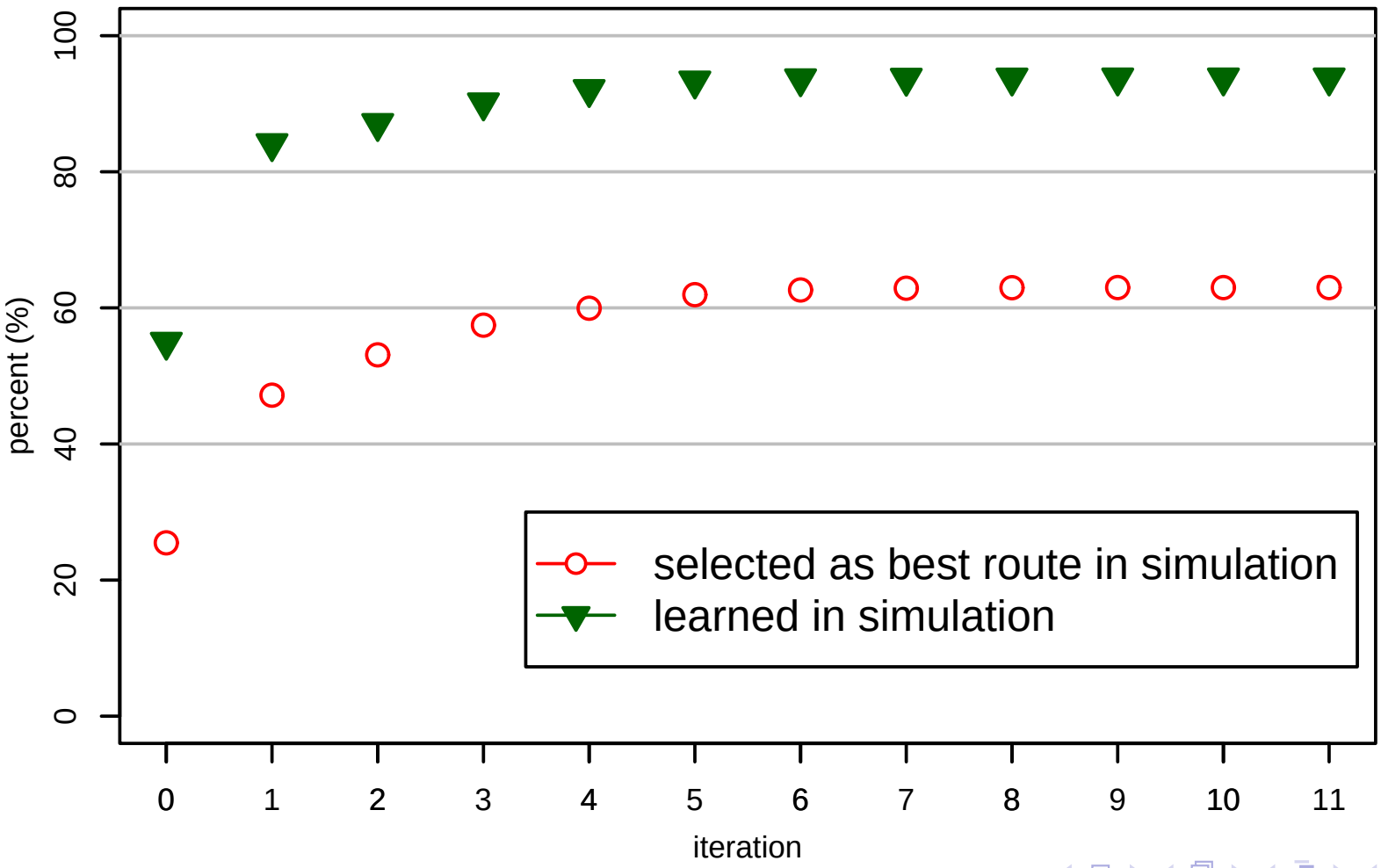
Experiment: Results for Training Set

- 11 iterations required
- Major improvements during first iterations



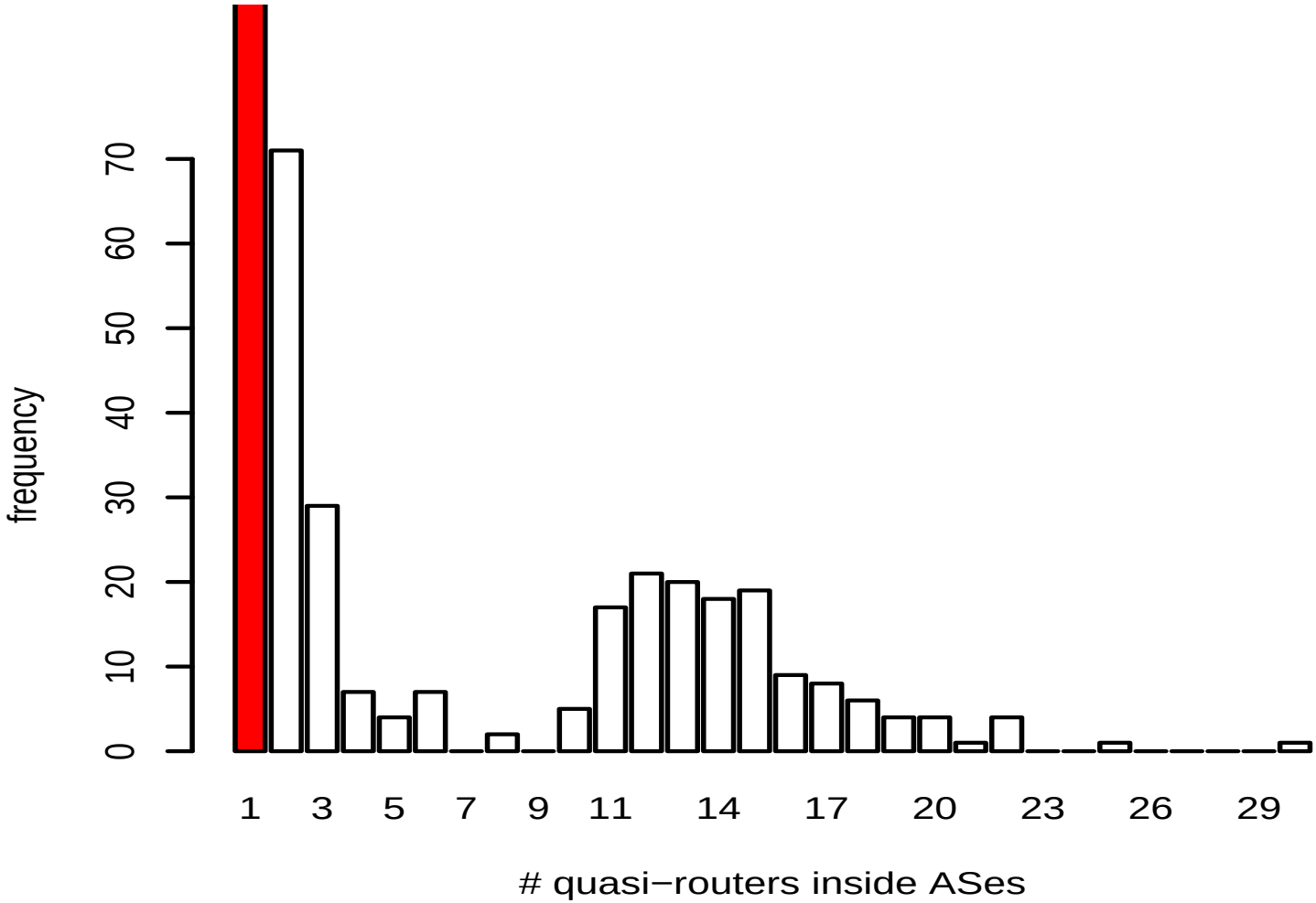
Experiment: Results for Validation Set

- 94% of observed paths are learned in the simulation
- Results seem better than previous work



How many quasi-routers?

- $\approx$ 98% of ASes: One *quasi-router*
- Well-known Tier1 ASes: Multiple *quasi-routers*



Summary

- **Insights**:
  - Single-router ASes not sufficient
  - Contractual AS-relationships seem too simplistic
- **Approach**: Model consistent with *all* observed paths
  - Introduce agnostic policies
  - Introduce route diversity
- **Predict** previously unobserved paths
  - $\approx 94\%$ of obs. paths learned
- **Long-term goal**: Answer what-if questions