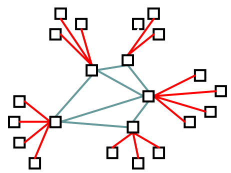
Random Annotated Graphs

Xenofontas Dimitropoulos
Dmitri Krioukov
George Riley

Georgia Tech
CAIDA

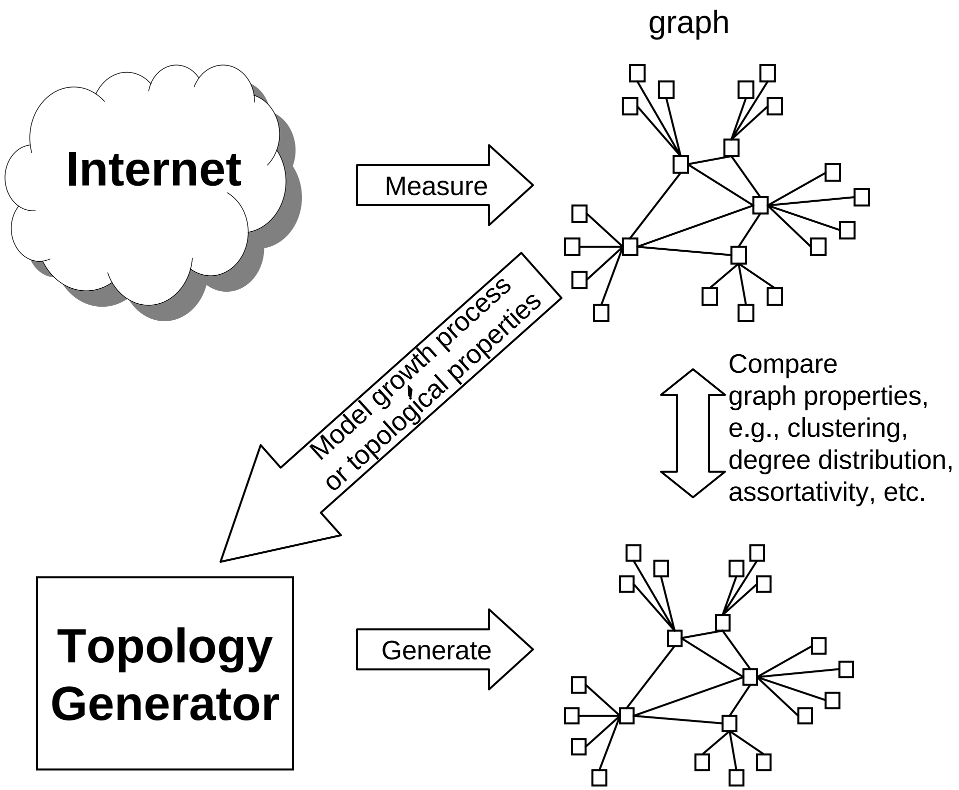


Outline

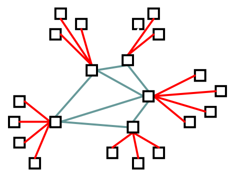
- Background and annotated graphs.
- Framework for generating random annotated graphs.
- Evaluation results.
- Conclusions and future work/directions.

Topology Generation Background

- Goal: Generate synthetic network topologies for protocols evaluation.
- Use graphs to represent measured topologies.
- Model growth process or topological properties, e.g., degree distribution.
- Generate synthetic graphs and compare.

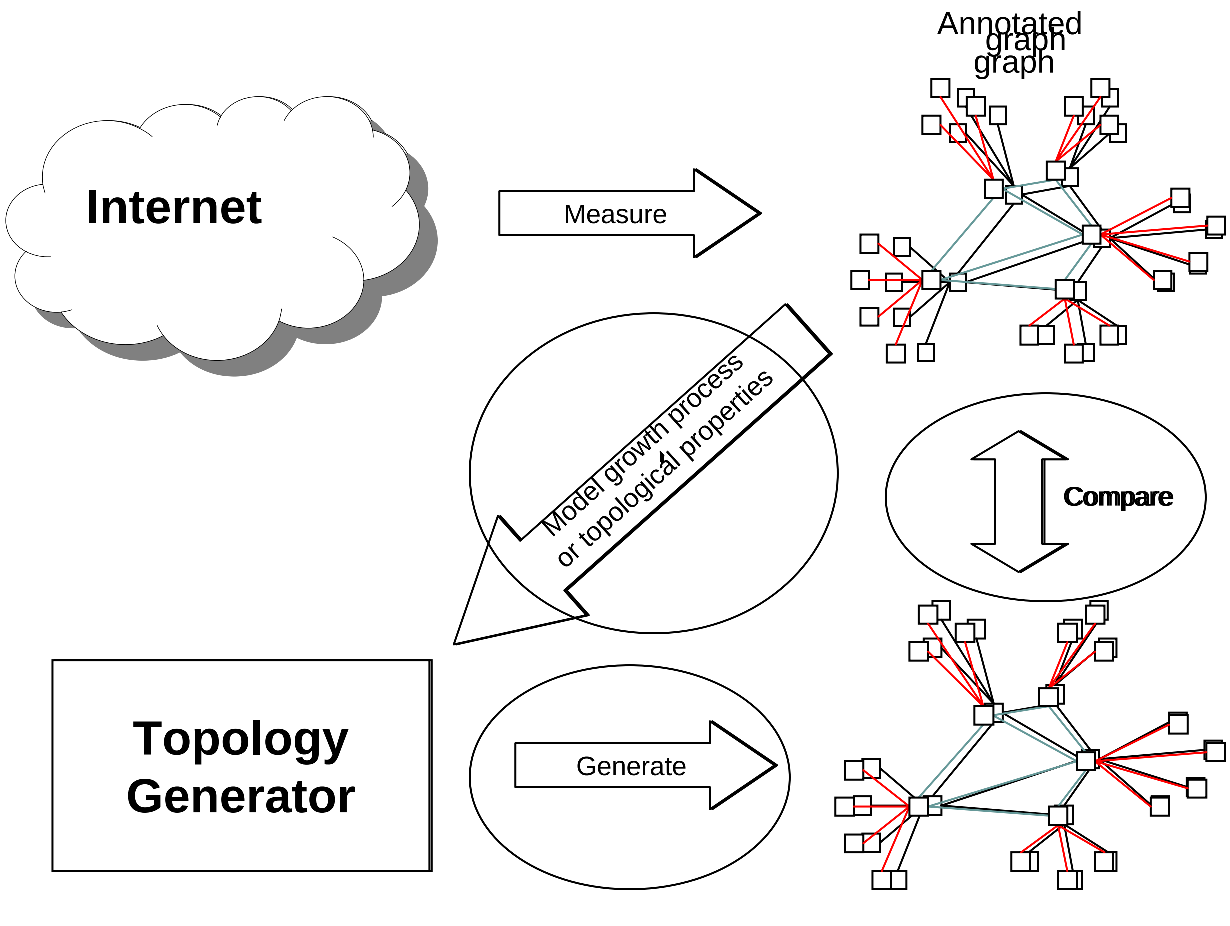


Annotated Graphs

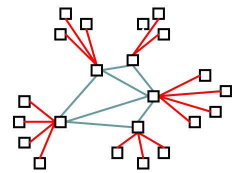


- An *annotated graph* is a graph in which:
  - each link has a single annotation from a finite set of link annotations.
  - links can be directed (asymmetric annotations) or undirected (symmetric annotations).
- Annotated graphs can represent useful network information, e.g.:
  - in router-level topologies, link annotations can represent capacities or latencies.
  - in AS-level topologies, link annotations can represent AS relationships, e.g., peer to peer (p2p), customer to provider (c2p), etc.
- Annotated graphs capture more information than plain (un)directed graphs.

Topology Generation Diagram

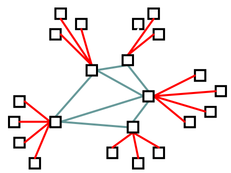


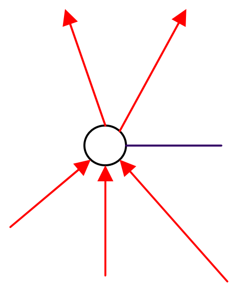
Outline



- Background and annotated graphs.
- Framework for generating random annotated graphs.
  - Annotation-aware topological properties.
  - Reproduce annotation-aware properties in synthetic graphs.
- Evaluation results.
- Conclusions and future work/directions.

Topological Properties

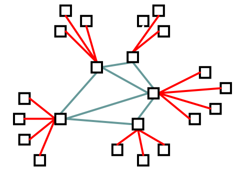




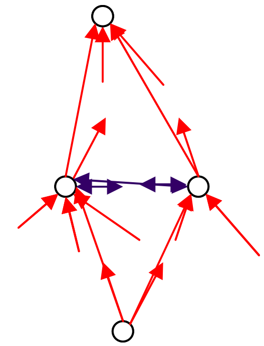
customer-degree: 3
provider-degree: 2
peer-degree: 1

- Annotation-degrees:
  - customer-degree (d_{p2c}): number of customers of a node.
  - provider-degree (d_{c2p}): number of providers of a node.
  - peer-degree (d_{p2p}): number of peers of a node.
- Joint Annotation-Degree Distribution (JADD): joint distribution of annotation-degrees.
- Joint Degree Distribution (JDD) of p2p (c2p) links: joint distribution of total degrees of connected nodes with p2p (c2p) edges.

Reproduce JADD

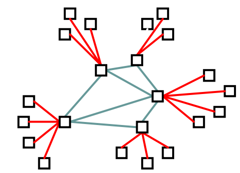


- Generate N random triplets of p2c-, p2p-, and c2p-degrees (d^i_{p2c}, d^i_{p2p}, d^i_{c2p}), $0 \leq i < N$.
- For each degree triplet introduce a node with d^i_{p2c} p2c-, d^i_{c2p} c2p-, and d^i_{p2p} p2p-stubs.
- Perform one random matching between c2p and p2c stubs and one between p2p stubs.

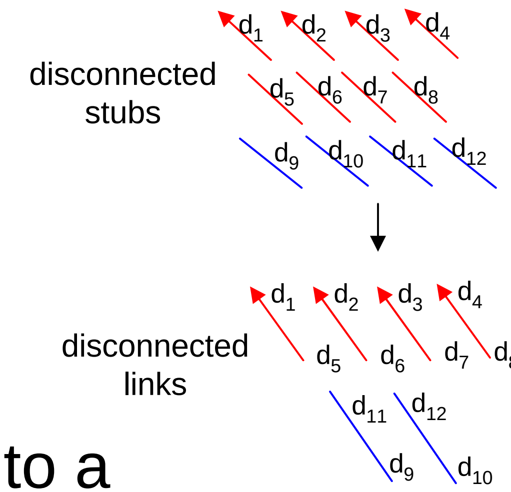


- Extract largest connected component and remove self-loops and multi-edges to get final graph.

Reproduce JADD and JDD



- Generate random triplets of p2c, c2p-, and p2p-degrees $(d^i_{p2c}, d^i_{p2p}, d^i_{c2p})$, $0 \le i < N$, and introduce p2c, c2p, and p2p stubs labeled w/ total degrees.

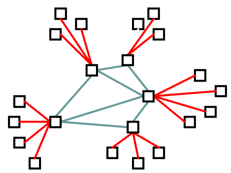


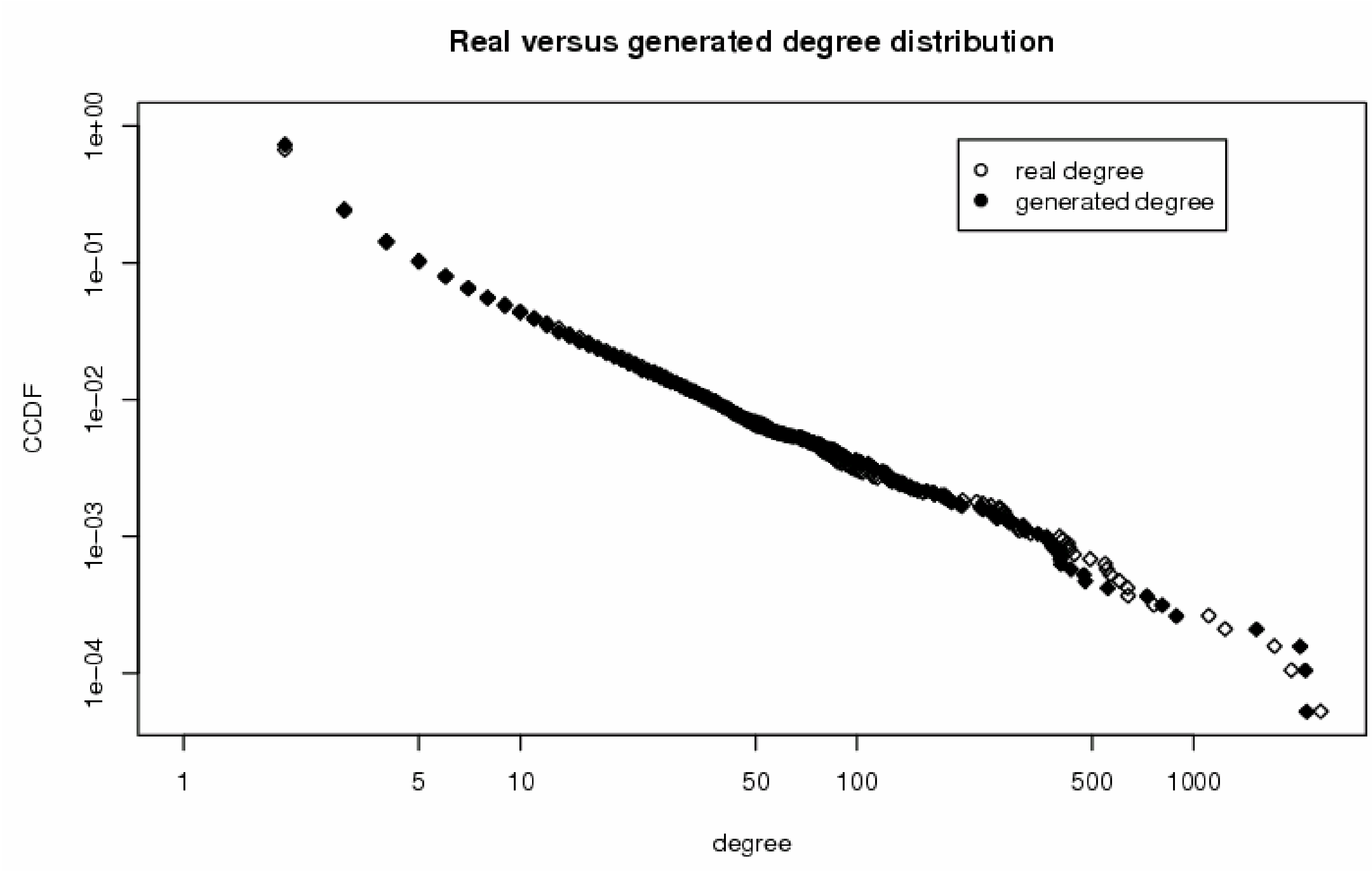
- For each of the three stub-types create a sequence of degrees.
- Join c2p (p2p) and p2c (p2p) sequences into a sequence of degree pairs that reflects c2p (p2p) JDD. Each degree pair reflects a disconnected link.
- For each triplet of total degree d, randomly select d_{p2c} p2c edge-ends, d_{p2p} p2p edge-ends and d_{c2p} c2p edge-ends from the set of edge-ends labeled with d and construct a node.

Outline

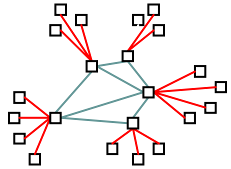
- Background and annotated graphs.
- Framework for generating random annotated graphs.
  - Define a set of annotation-aware topological properties.
  - Reproduce these properties in synthetic annotated graphs.
- Evaluation results.
- Conclusions and future work/directions.

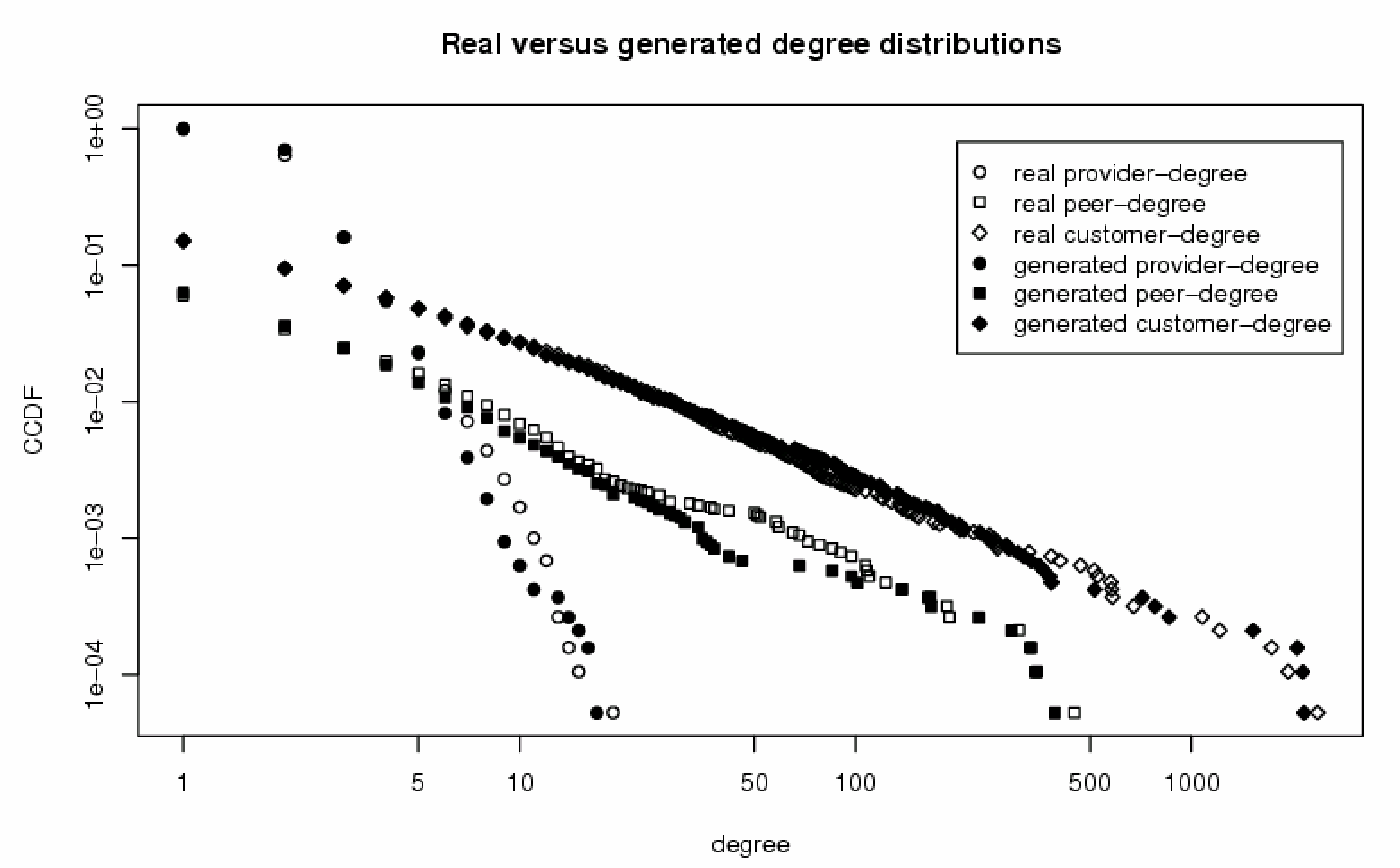
Degree distribution





Annotation-degree distributions

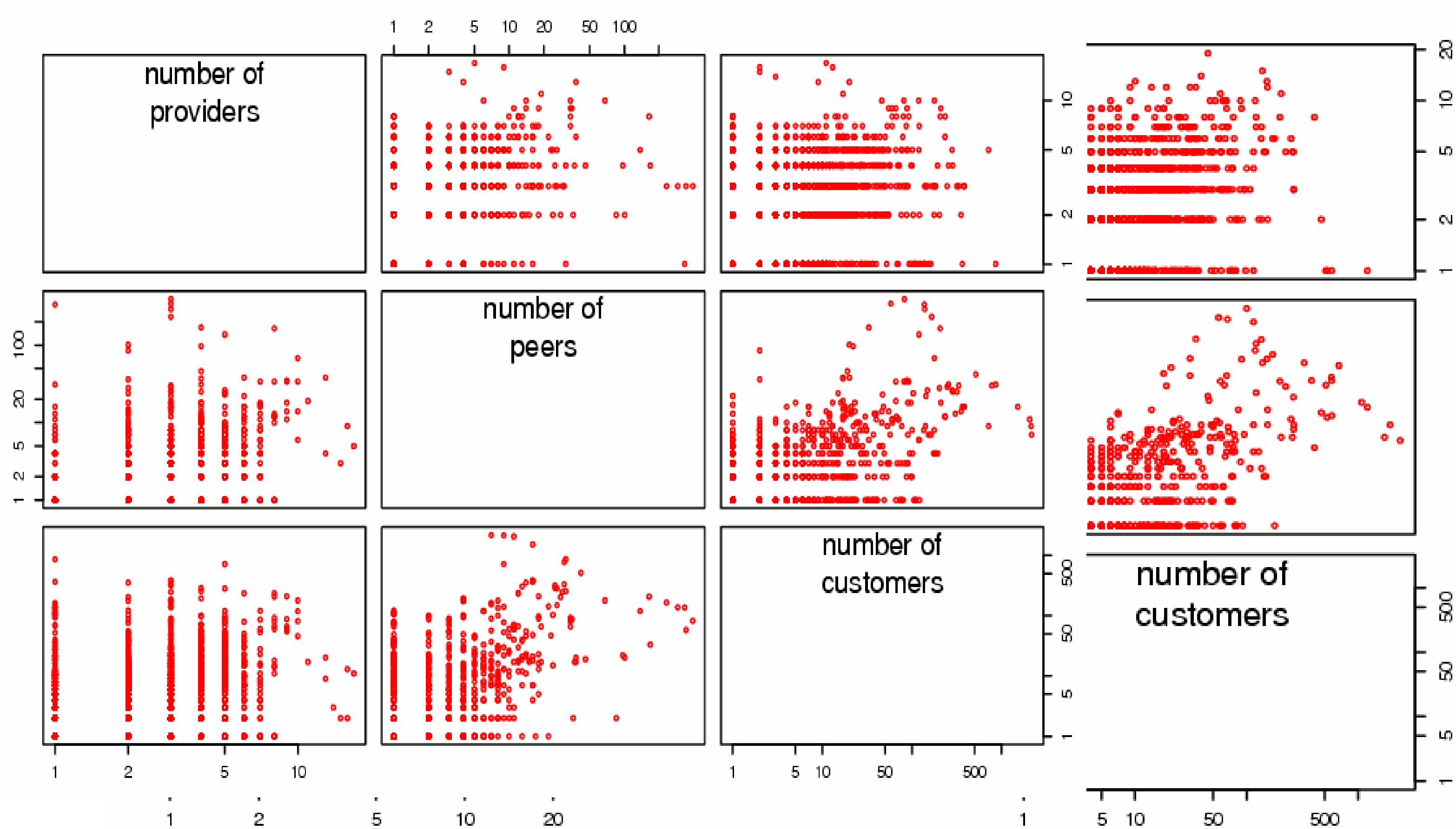




JADD

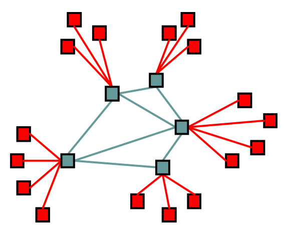
Matrix scatterplot for measured topology

Matrix scatterplot for synthetic topology



Conclusions and Future Work

- **Proposed using annotated graphs to model network topologies.**
- **Described framework to generate synthetic annotated graphs.**
- **Implemented our framework for generating synthetic AS topologies with synthetic c2p and p2p annotations.**
- **Outlined evaluation results.**
- **Work in progress:**
  - **Richer evaluation and comparison with other generators.**
- **Future directions:**
  - **Public release of generator.**
  - **Generalize to introduce node annotations, which can represent router models, AS types, etc.**



Questions and publications

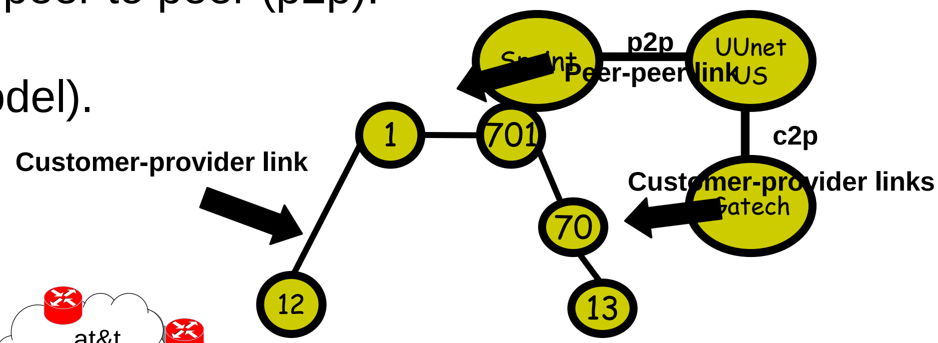
- **Towards a Topology Generator Modeling AS Relationships**
  Xenofontas Dimitropoulos; George Riley; Dima Krioukov; Ravi Sundaram
  IEEE ICNP (extended abstract), 2005.
- **Modeling Autonomous System Relationships**
  Xenofontas Dimitropoulos; George Riley
  To appear in 20th Principles of Advanced and Distributed Simulation (PADS), 2006

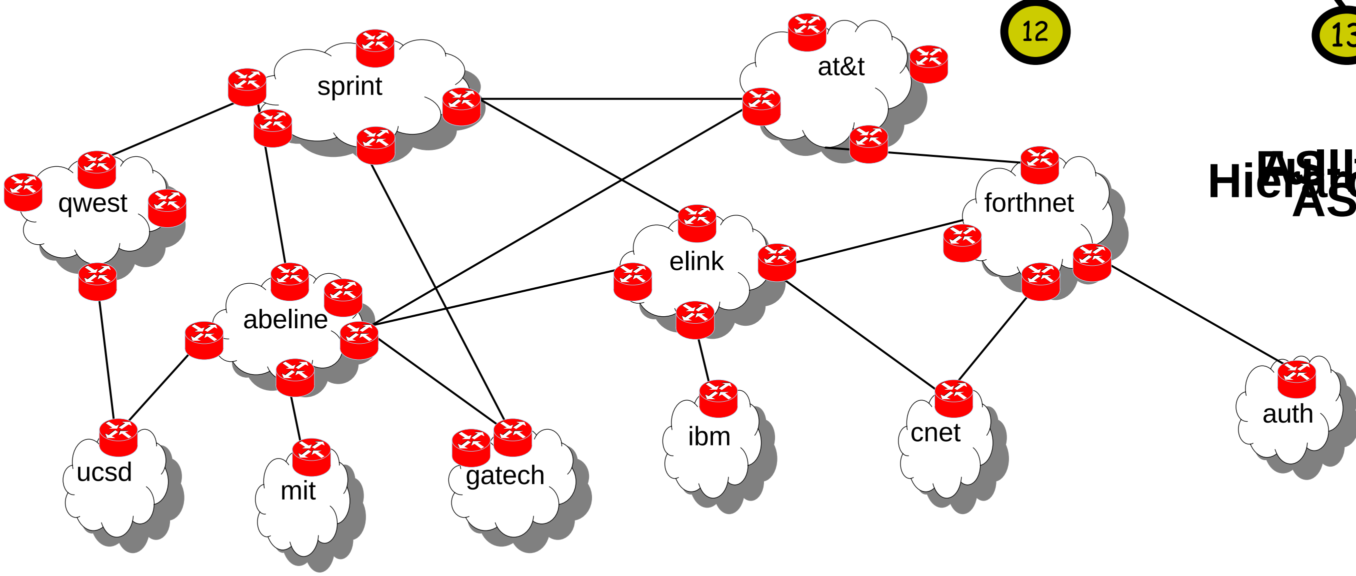
- **Inferring AS Relationships: Dead End or Lively Beginning?**
  Xenofontas Dimitropoulos; Dima Krioukov; Bradley Huffaker; kc claffy; George Riley
  4th Workshop on Efficient and Experimental Algorithms (WEA), 2005.
- **AS Relationships: Inference and Validation**
  Xenofontas Dimitropoulos; Dima Krioukov; Marina Fomenkova; Bradley Huffaker; Young Hyun; kc claffy; George Riley
  Under submission.
  http://www.caida.org/publications/papers/2006/as_relationships_inference/
- **AS Relationships Repository**
  http://as-rank.caida.org/data/

Extra Slides

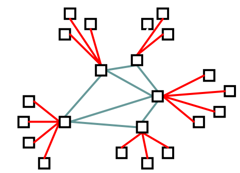
AS relationships

- AS-level topology of the Internet, i.e., interconnections between ASs.
- AS relationships:
  - are customer to provider (c2p) or peer to peer (p2p).
  - reflect business agreements.
  - determine routing (valley free model).



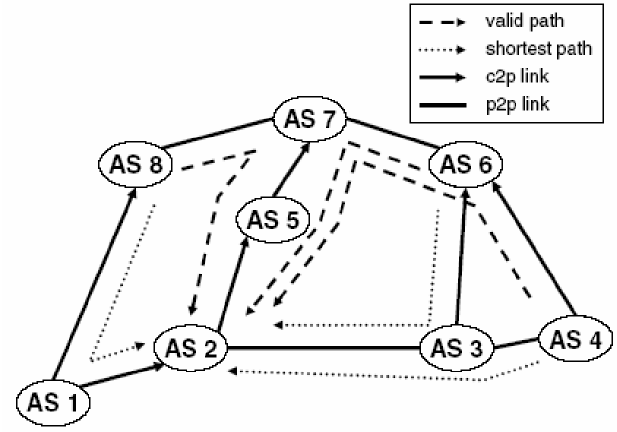


Hierarchical AS paths
AS-level topology
AS relationships



Simulation Examples

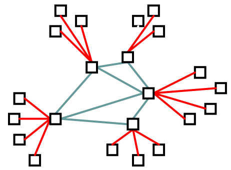
- Present topology generators do not model AS relationships.
- Simulation artifacts:
  - AS paths are shorter than in reality.
  - Number of alternative AS paths available to an AS is larger than in reality.
  - The traffic load on ASs and AS links is lower than in reality.



AS number	1	2	3	4	5	6	7	8
AS relationships enabled	12	9	10	8	8	7	9	6
AS relationships disabled	12	13	16	15	13	15	15	13

Number of paths available to each AS

Computing JADD

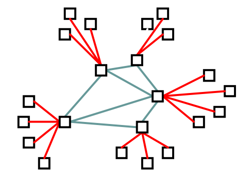


- Collect AS topology.
- Infer c2p and p2p relationships.
- Fit JADD:
  - Fit annotation-degree distributions using splines.
  - Model correlations using historical copula data.

How to produce random pairs from an empirical bivariate distribution?

- Use splines to fit marginal distributions.
- Generate N random number from each of the two marginal models.
- Joint the two degree sequence into a sequence of degree pairs so that these degree pairs respect a given JDD.

How to join two degree sequences?



- Input: two degree sequences, p_i and q_i, of length N.
- Output: one sequence of degree pairs so that these pairs respect a historical JDD.

- Algorithm:
  - Randomly select N degree pairs (x_i,y_i) from historical data.
  - Map each degree pair $(x_i,y_i) \rightarrow (R(x_i),R(y_i)) \rightarrow (p_i,q_i)$

degree: 6