Cheleby:
Subnet-Level Internet Topology

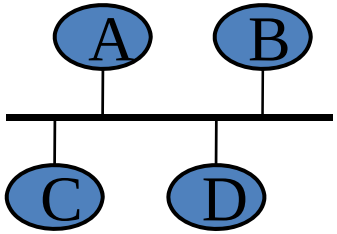
Mehmet Hadi Gunes

with Hakan Kardes and Mehmet B. Akgun

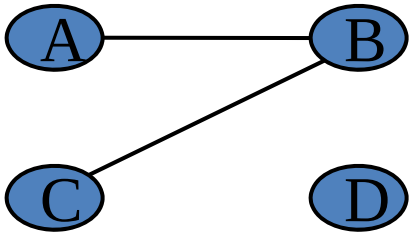
Department of Computer Science and Engineering

University of Nevada, Reno

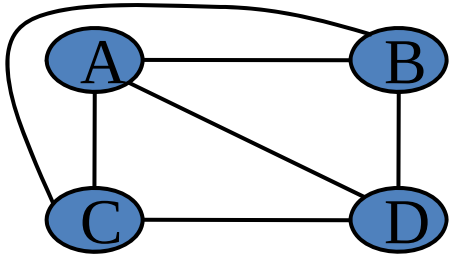
Subnet Resolution



genuine topology



observed topology

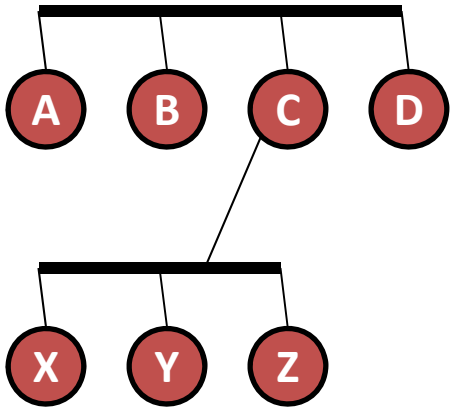


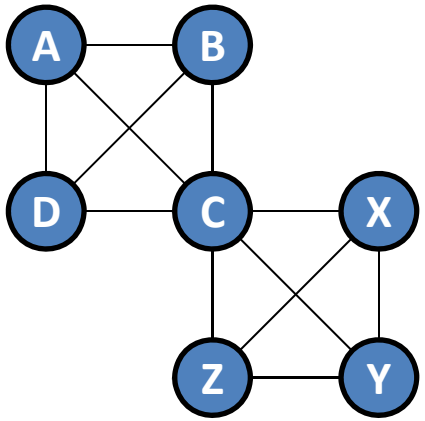
inferred topology

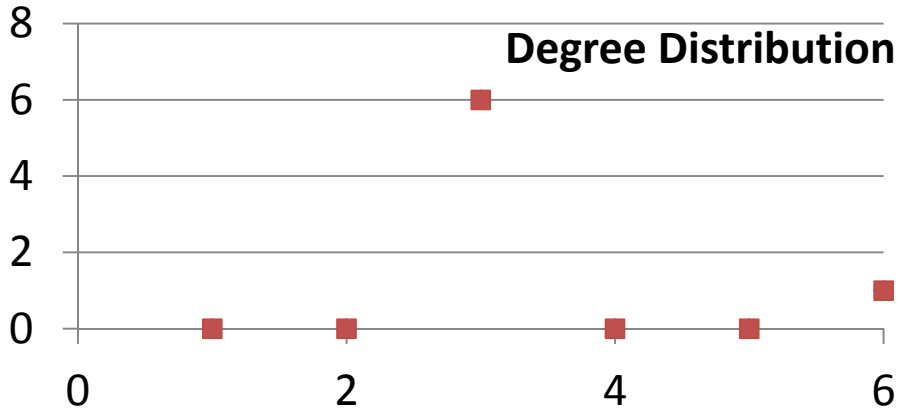
[Observed] Degree vs. [Actual] Interfaces

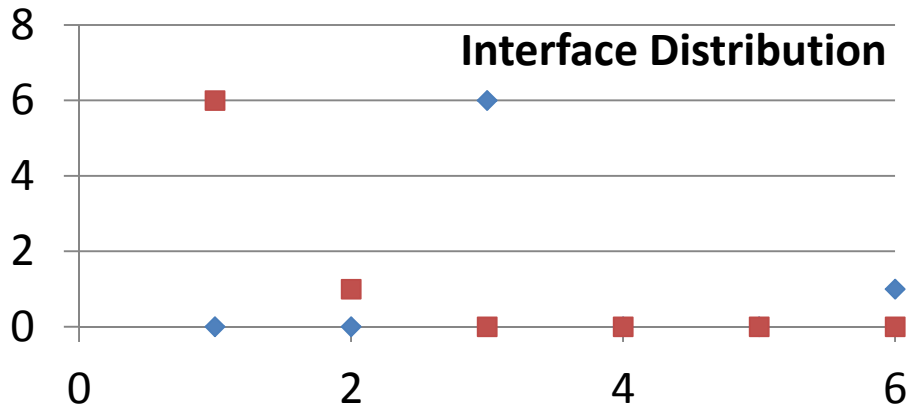
Degree: the number of one hop neighbors

Interface: the number of links the system is attached to



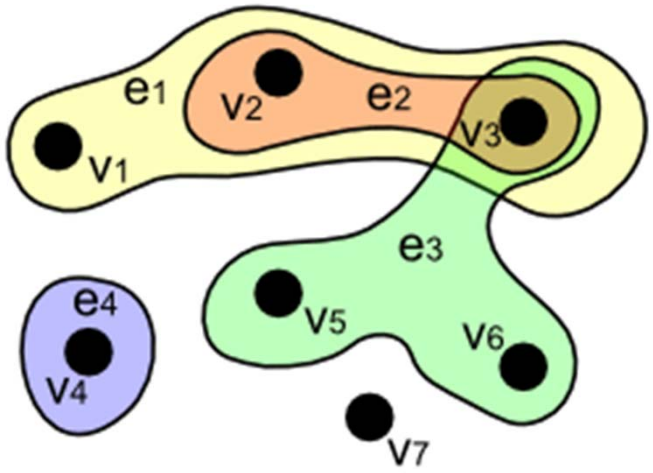


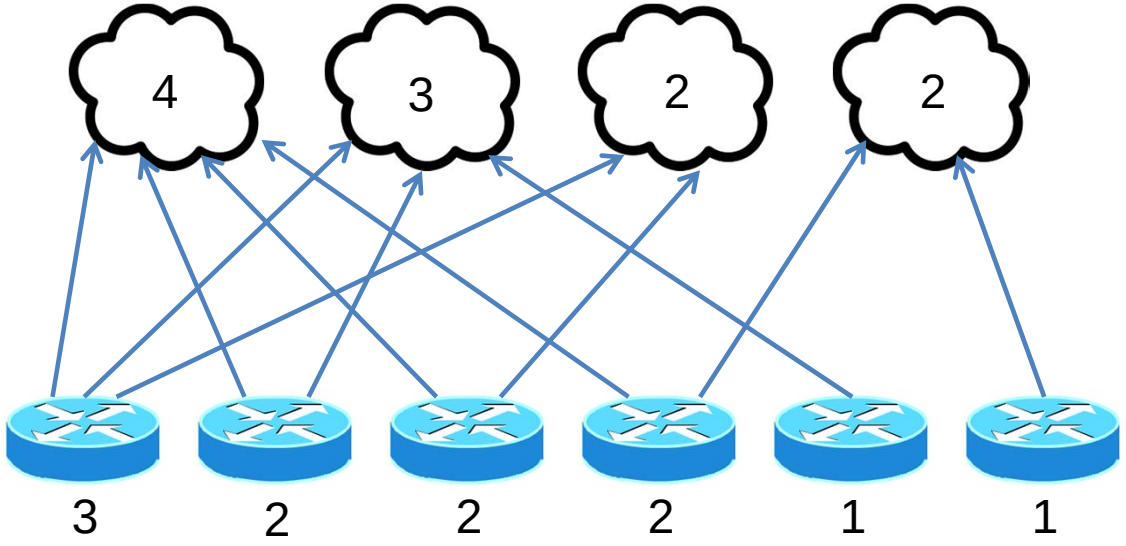




Hyper Graphs

- Networks modeled as graphs G=(V,E)
- Hyper graphs: H= (X,E) can accurately model *multi-access links*
  - also, bipartite (2-mode) graphs





Cheleby System Overview

PlanetLab
Vantage Points

Traces
- x - ✱ - L.2 - S.2 - y
- x - ✱ - A.1 - W.1 - ✱ - z
- y - S.1 - L.1 - ✱ - x
- y - S.1 – U.1 - ✱ - C.1 - ✱ - z
- z - ✱ - C.2 - ✱ - ✱ - x
- z - ✱ - C.2 - ✱ - U.2 - S.3 - y

Raw Data

Initial Pruner (IP)

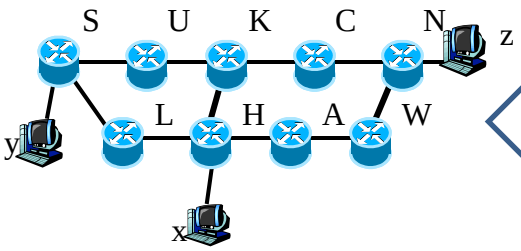
Structural Graph Indexer (SGI)

SubNet Inferrer (SNI)

Analytical IP Alias Resolver v2 (APARv2), iffinder

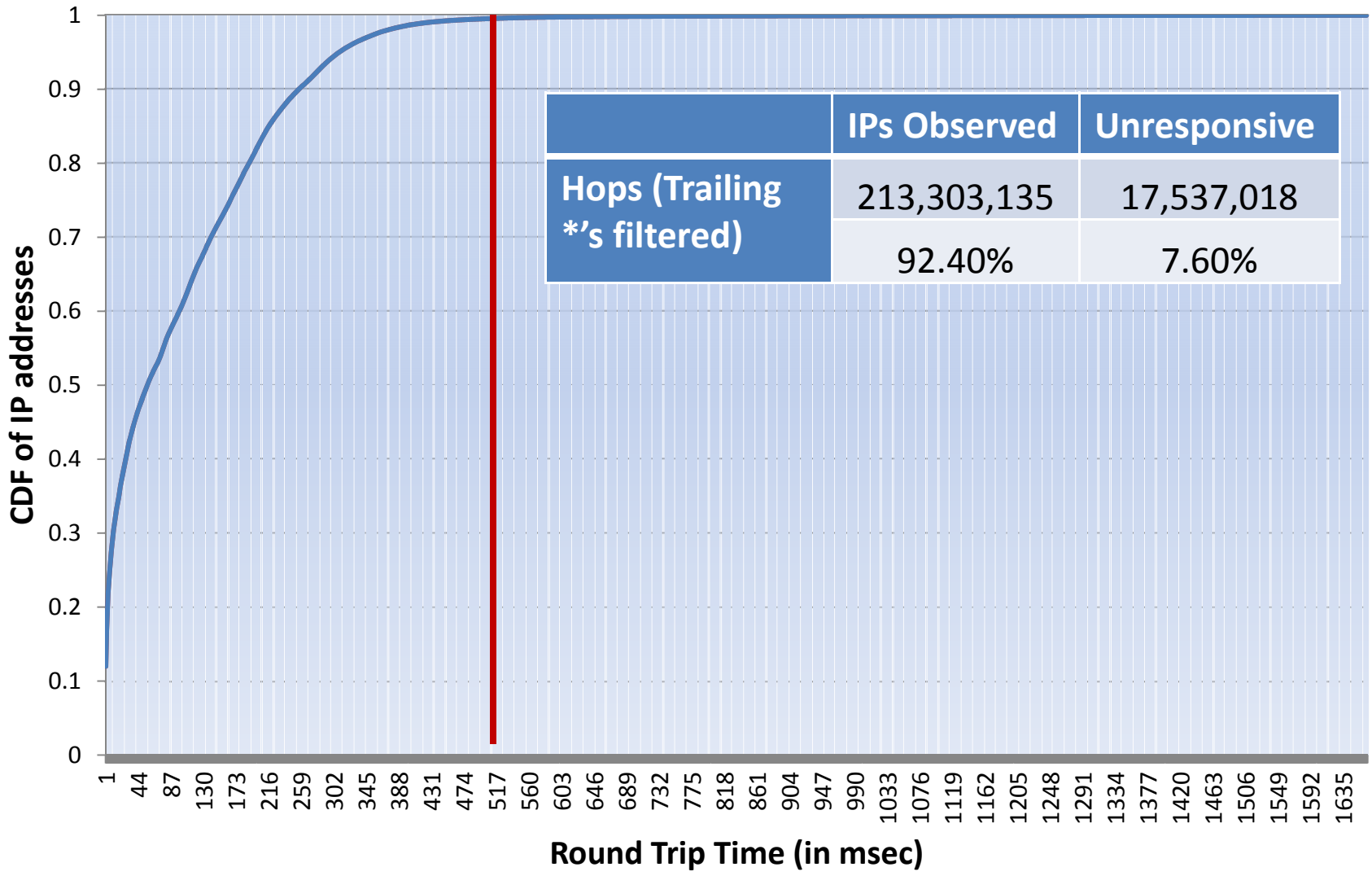
Graph Based Induction (GBI)

Network Topology



http://cheleby.cse.unr.edu

Round Trip Time Analysis



	IPs Observed	Unresponsive
Hops (Trailing *'s filtered)	213,303,135	17,537,018
	92.40%	7.60%

Unresponsive Routers

- Responsiveness to Direct Probes

Year / Data	Type	All	Router	End-System	.net	.com	.edu	.org	.gov
2008 / Router	ICMP	81.9 %	84.6 %	77.9 %	92.3 %	86.4 %	88.9 %	95.5 %	92.9 %
	TCP	67.3 %	70.4 %	62.8 %	76.7 %	72.6 %	83.2 %	77.3 %	83.0 %
	UDP	59.9 %	64.7%	50.3 %	63.5 %	61.7 %	57.3 %	64.4 %	62.8 %
2011 / WebSite	ICMP	80.4 %	-	80.4 %	84.9 %	86.7 %	53.2 %	83.6 %	37.2 %
	TCP	97.9 %	-	97.9 %	98.3 %	97.8 %	95.8 %	98.2 %	96.9 %
	UDP	46.7 %	-	46.7 %	47.6%	50.9%	21.0%	45.8%	14.4%

- Responsiveness to Indirect Probes

Type	Year	#Traces	Completed	#IPs	#Nodes	Unres.
ICMP	2008	306 K	93.1 %	313 K	1.0 M	68.7 %
	2011	537 K	88.8 %	770 K	2.0 M	66.5 %
TCP	2008	306 K	73.4 %	277 K	1.0 M	72.3 %
	2011	537 K	96.0 %	697 K	1.1 M	36.6 %
UDP	2008	306 K	45.0 %	210 K	1.5 M	86.0 %
	2011	537 K	64.4 %	201 K	1.5 M	86.6 %

Team Analysis

Teams	3	5	7	9	11
Time (min)	540	630	770	1,220	1,540
Traces	9.5M	15.9M	22.0M	28.7M	35.0M
Probes	151M	249M	347M	452M	552M
Total IPs	95.3M	157M	219M	285M	348M
Total *s	55.7M	92.4M	128M	167M	204M
Unique IPs	1.11M	1.18M	1.21M	1.24M	1.27M
IPs / all	79.3%	84.3%	86.3%	88.8%	90.7%
Per min IPs	2,057	1,874	1,571	1,020	825
Unique Edges	1.42M	1.76M	1.96M	2.13M	2.26M
Edges / all	46.1%	57.1%	63.6%	69.1%	73.1%
Per min Edges	2,636	2,794	2,550	1,747	1,465

Resolution results

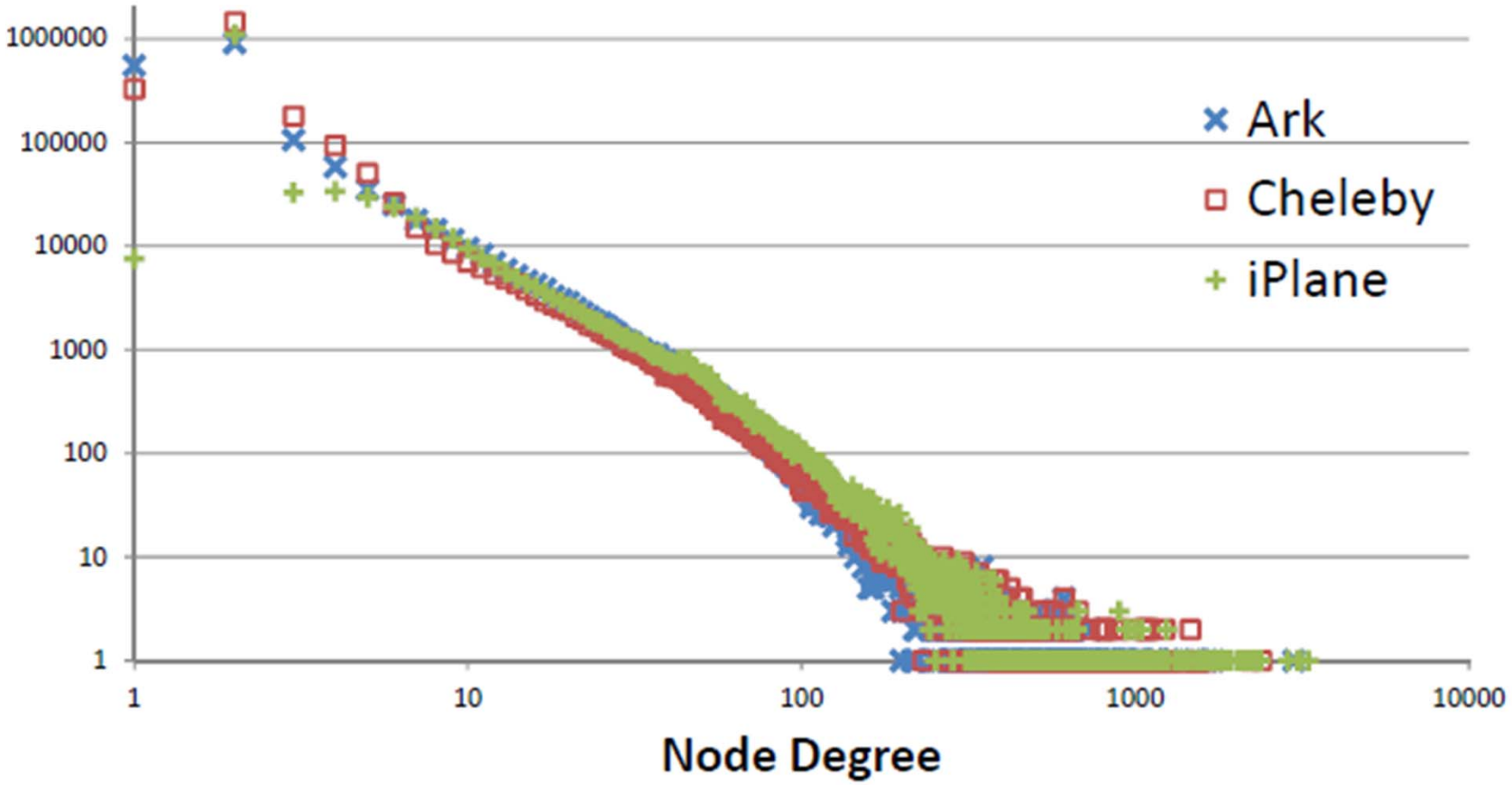
- Alias Resolution

Resolver	Alias Sets	Aliased IPs
APARv2	38,012	128,495
Ally (path traces)	32,860	65,720
Ally (common neighbor)	32,595	65,190
Ally (subnet)	25,436	50,872
Ally (combined)	55,027	110,054
Mercator	305	610
Combined	82,962	216,628

- Subnet Inference

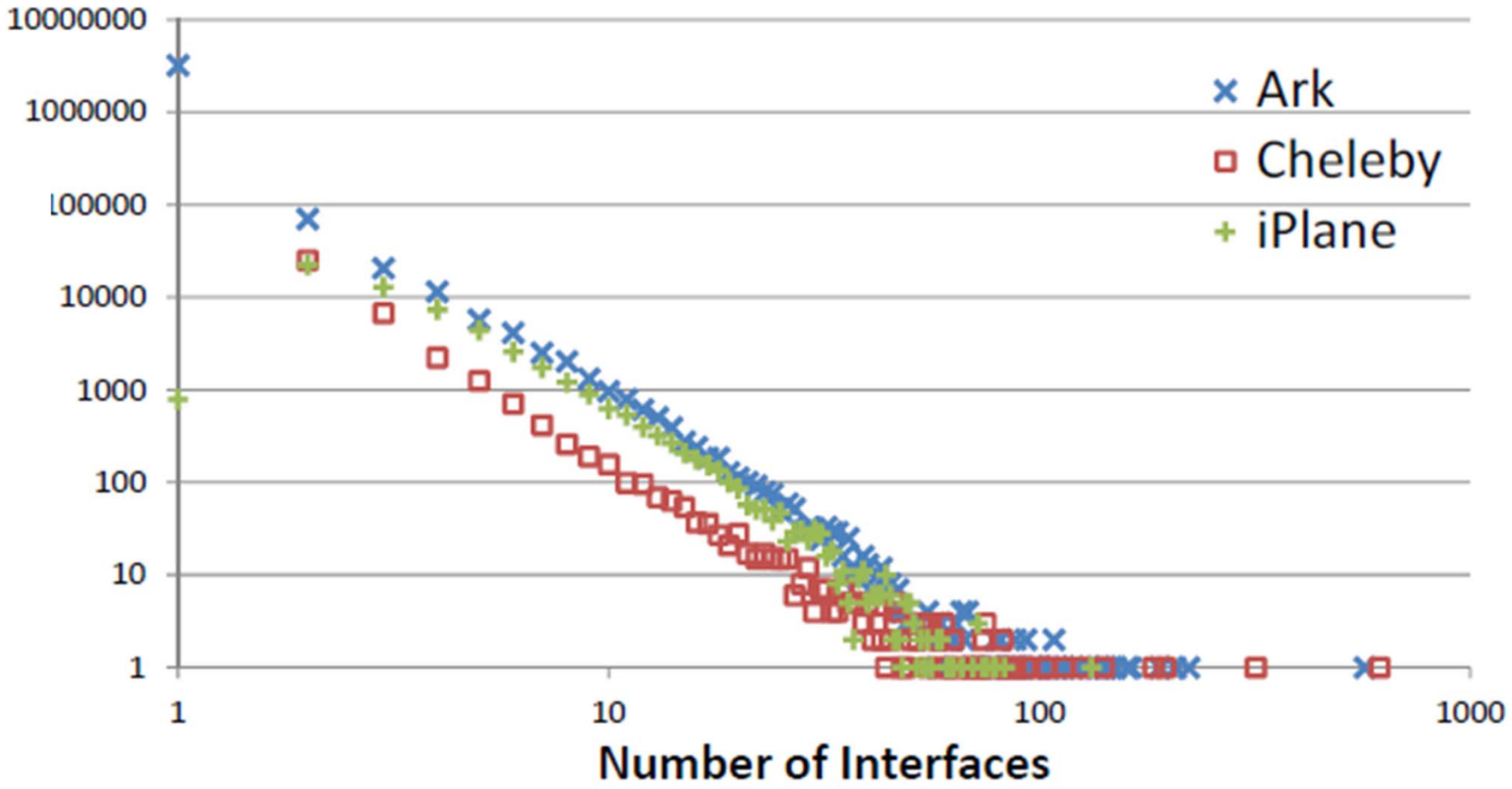
Subnet Size	/24	/25	/26	/27	/28	/29	/30	/31
Count	4	36	184	1,294	8,836	93,110	20,543	37,468
Completeness	26.3%	30.0%	28.3%	27.7%	28.0%	39.3%	100%	100%
All IPs	268	1,359	3,228	10,767	34,587	219,745	41,086	74,936

Degree Distribution



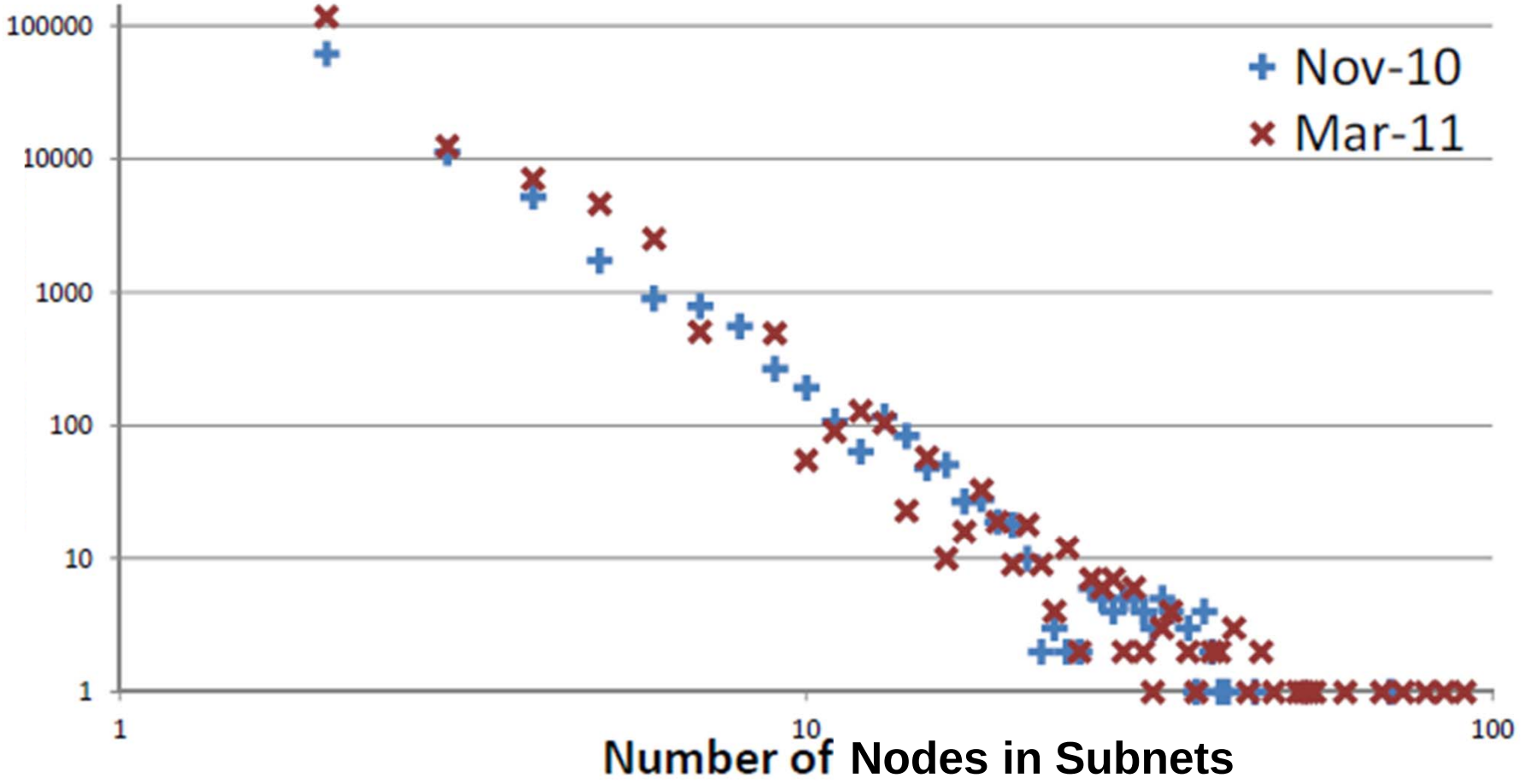
- Exponents : -2.17, -2.02, -1.92, respectively

Interface Distribution



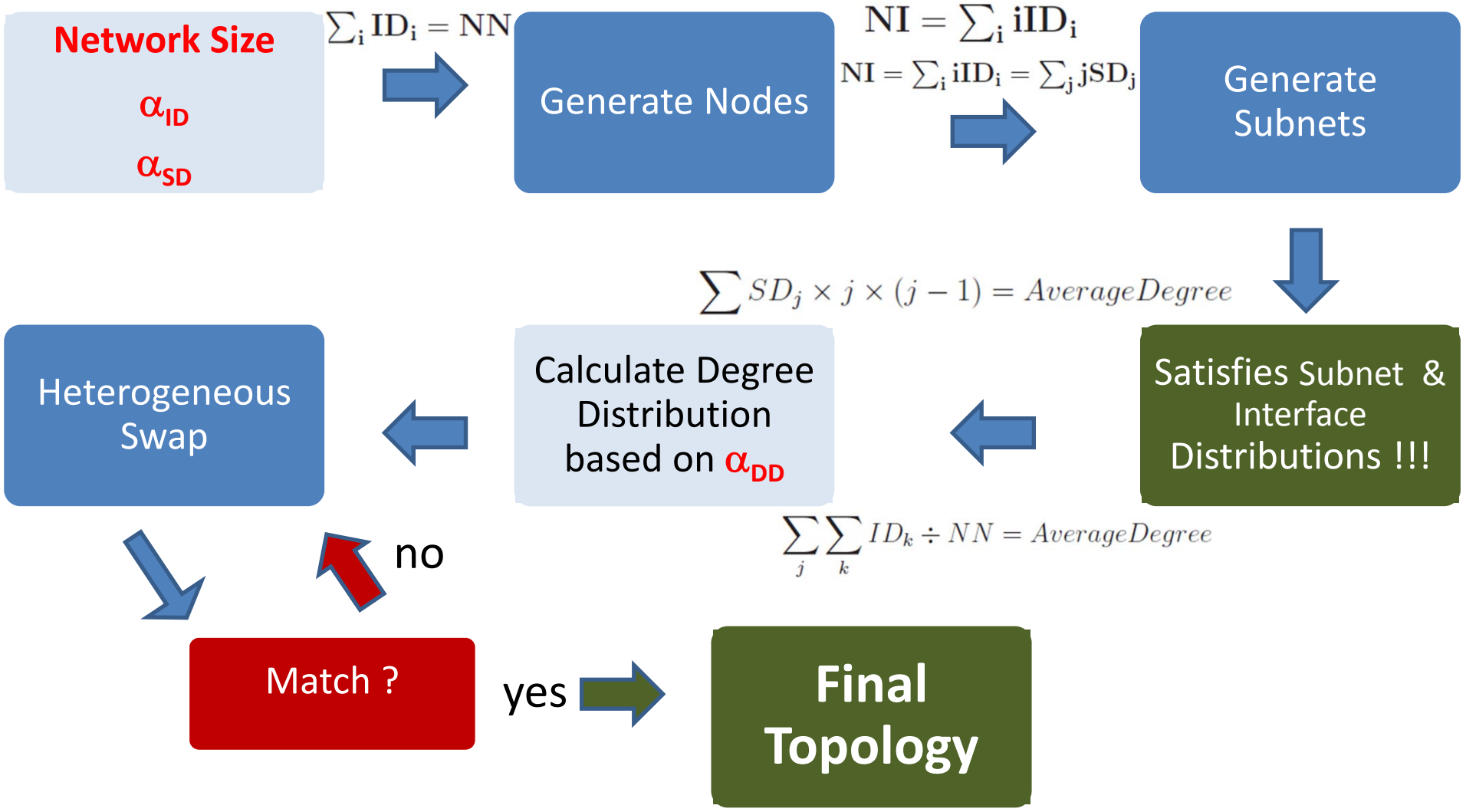
- Exponents : -2.71, -2.69, -2.74, respectively

Subnet Distribution



- Exponents : -3.42, 3.62, respectively

Synthetic Topology Generation



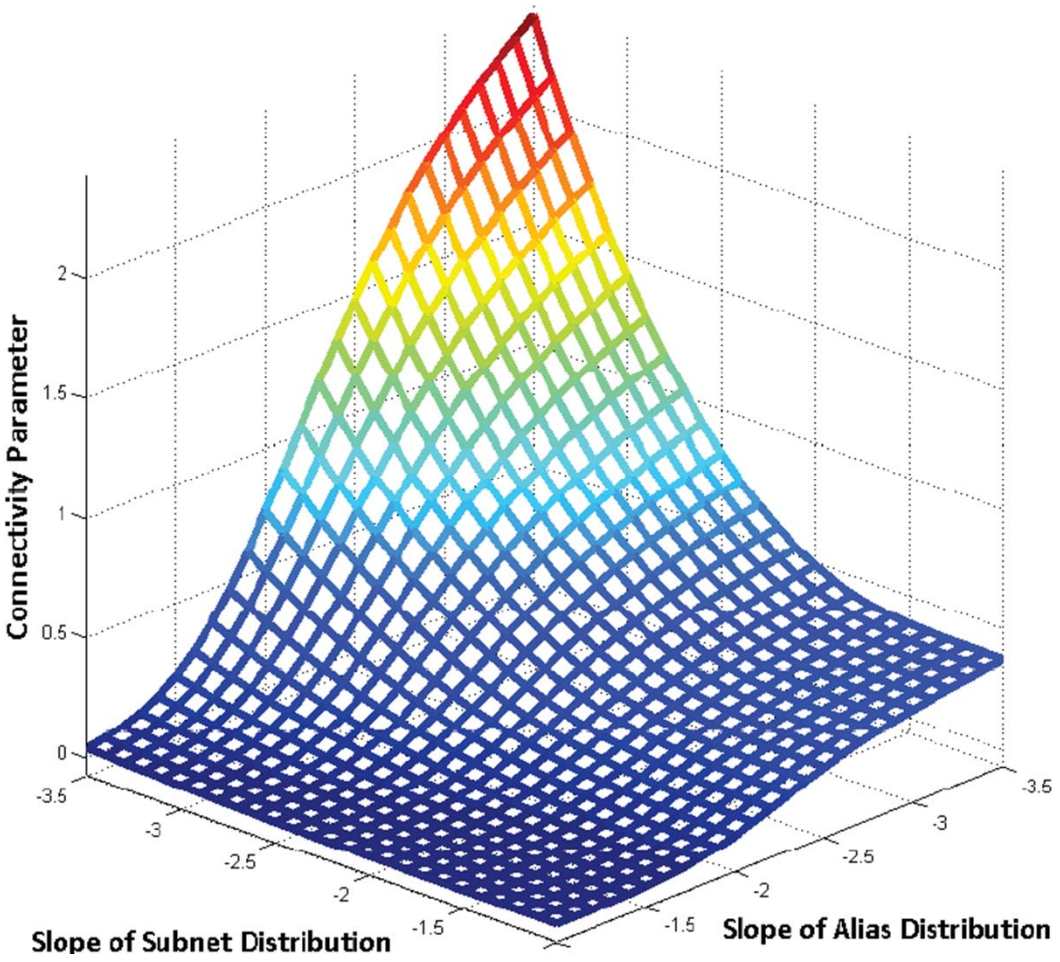
Connectivity Analysis

Relation between Interface Distribution and Number of Subnets

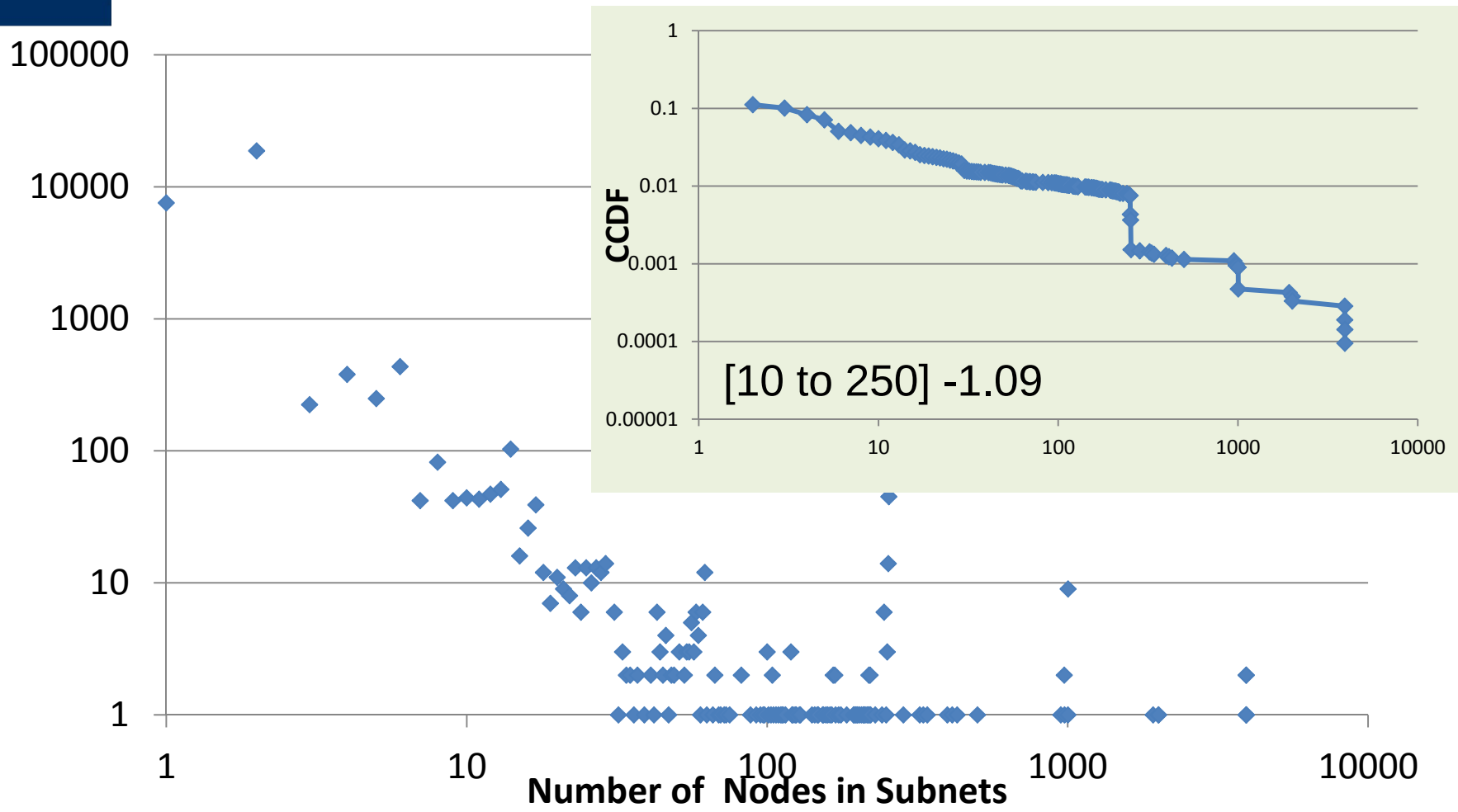
$$NS \leq 1 + \sum_{i=2} (i-1) \times ID_i.$$

- Single connected component is feasible only when
  - connectivity parameter <1

Feasible Region



Subnet Distribution: ExploreNET



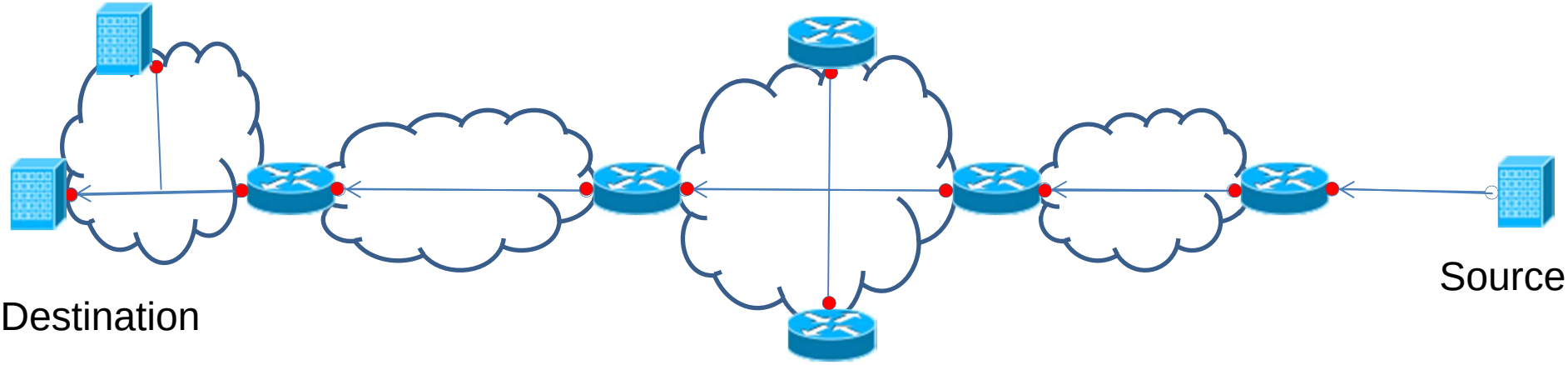
Estimating Network Layer Subnet Characteristics via Statistical Sampling, M. Engin Tozal and Kamil Sarac, IFIP/TC6 Networking, Prague, Czech Republic, May'12

TraceNET

Destination

Traceroute Path

Source



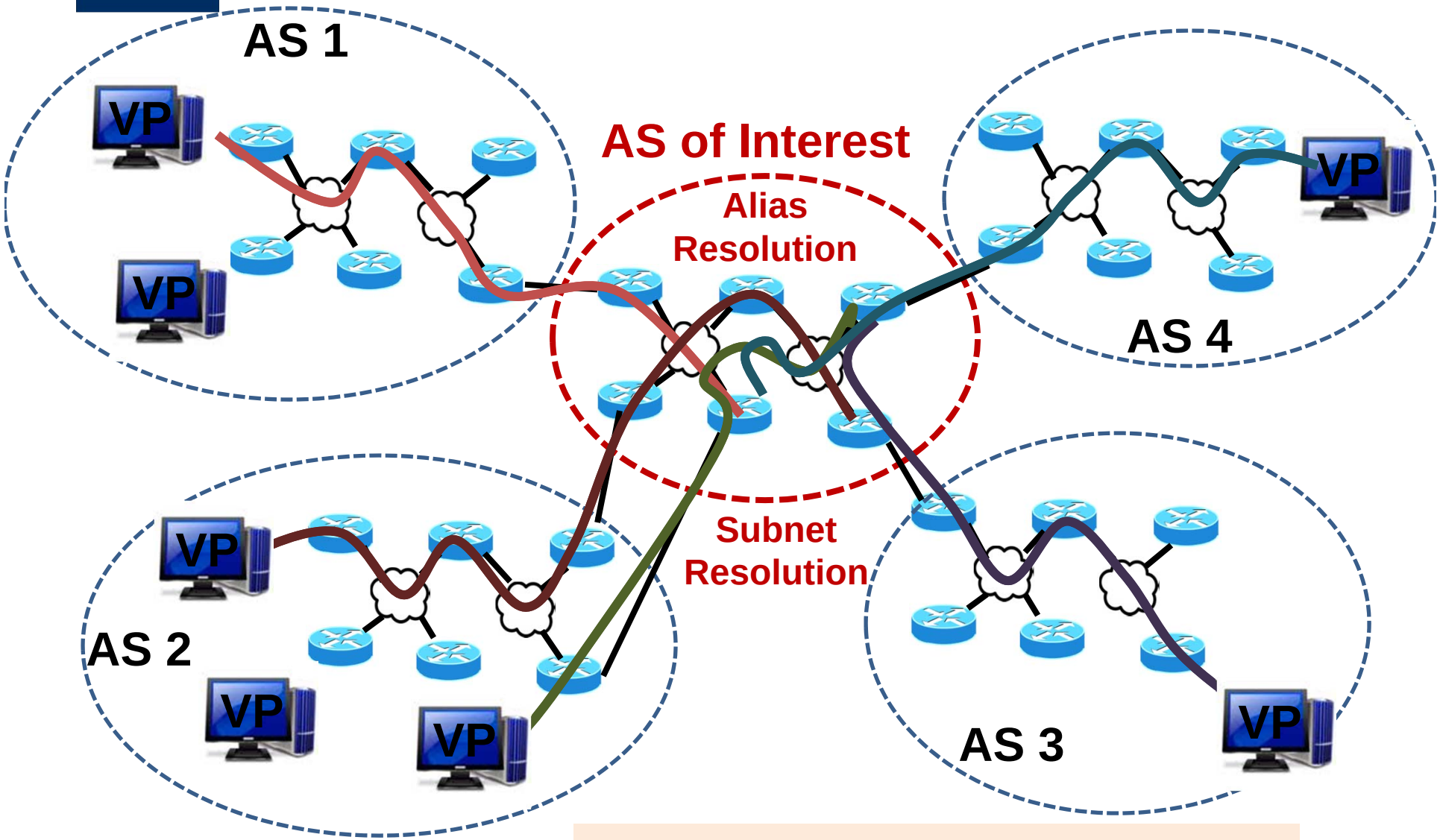
Destination

Source

TraceNET Path

TraceNET: An Internet Topology Data Collector, M. Engin Tozal and Kamil Sarac, ACM Internet Measurement Conference, Melbourne, Australia, November 2010

Work in Progress



Per Destination load balancers ?

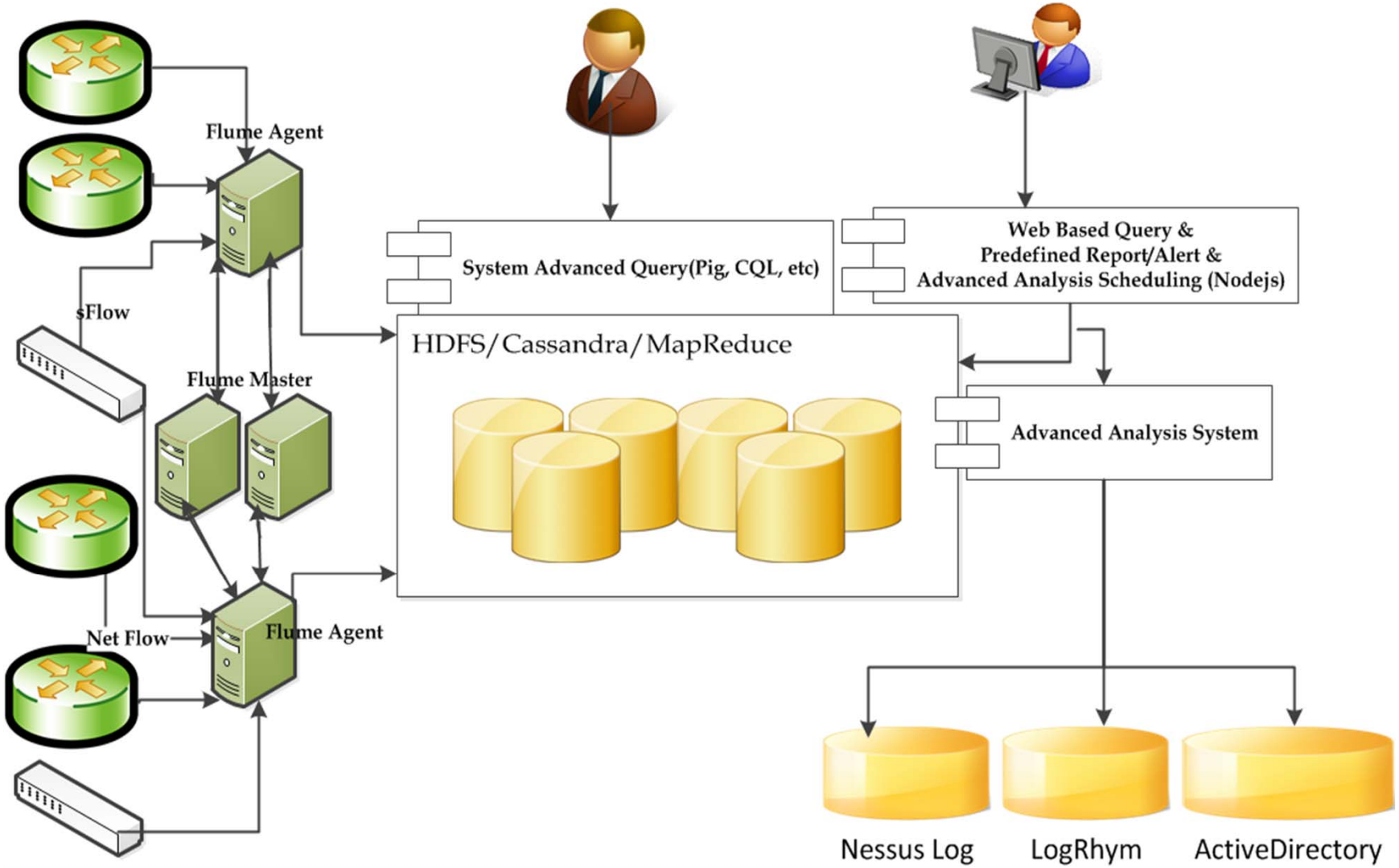
Network Traffic Analysis

with Bing Li, Jeff Springer, George Bebis

Design Goals

- Real time network query
  - near real time measurement and analysis
- Distributed system for
  - data collecting, storing, accessing, measuring and analyzing NetFlow
- Models of detection and classification based on profiling and behavior

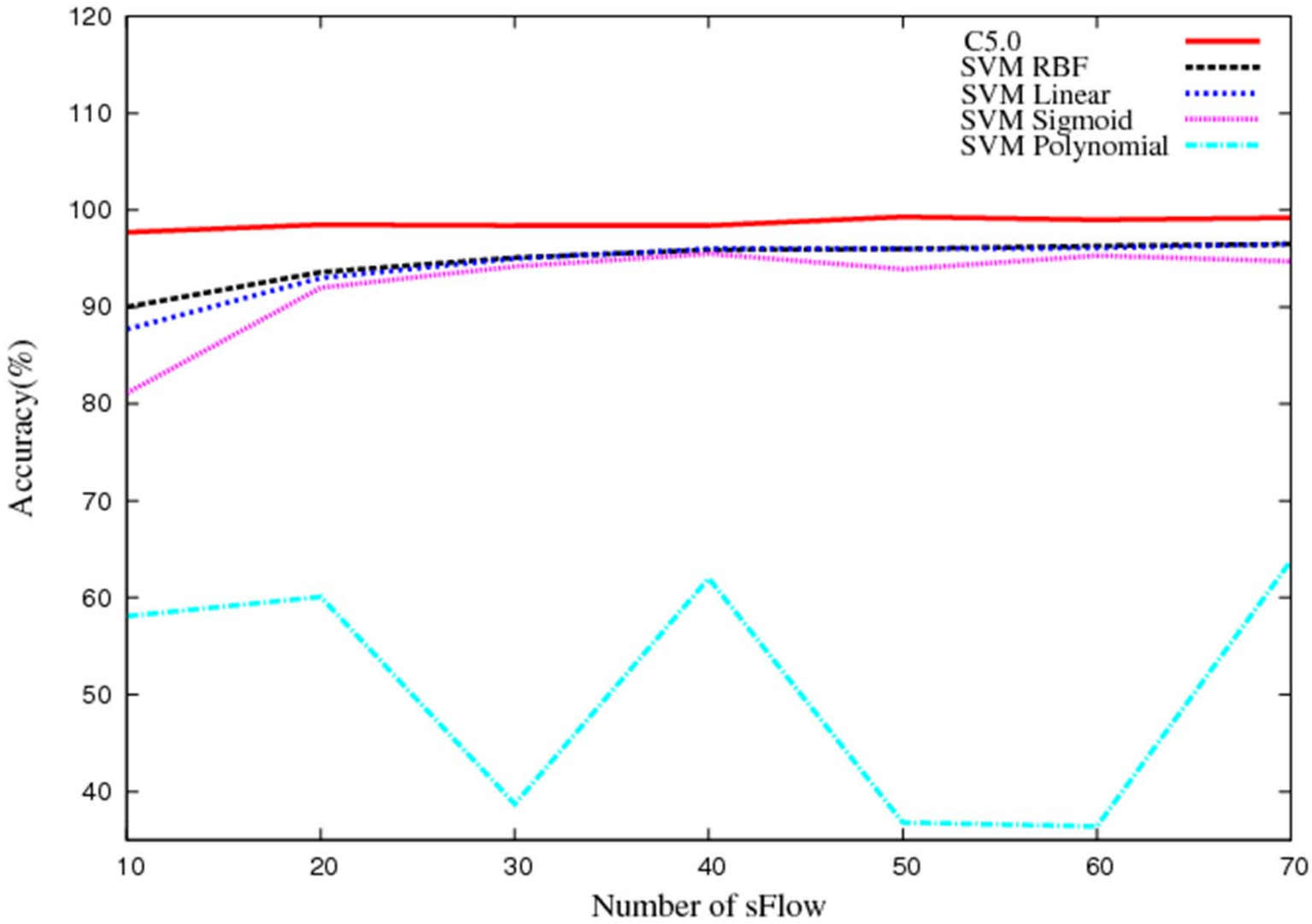
Design Components



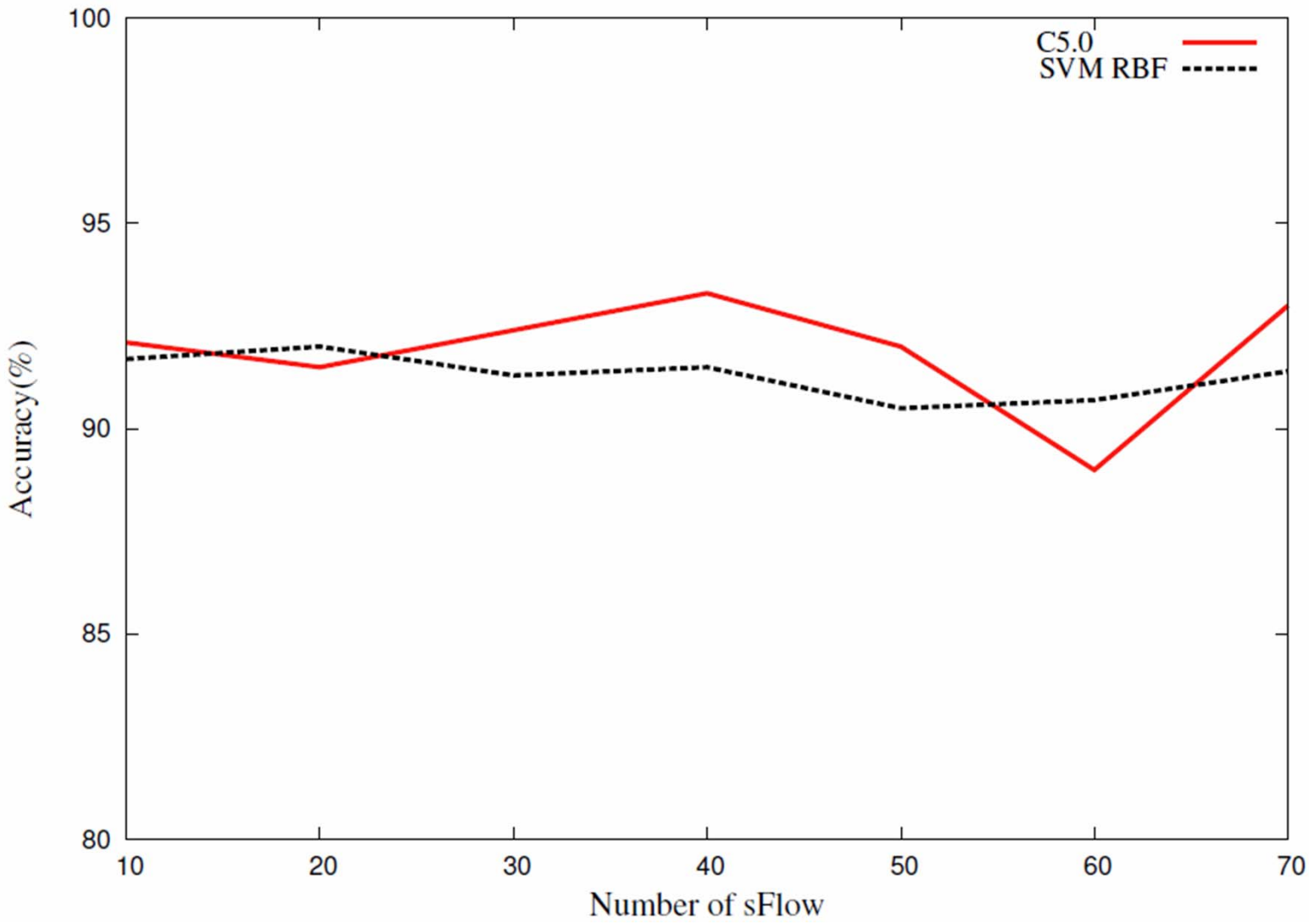
Demonstration

- Model Host Roles
- Algorithms:
  - On-line Support Vector Machine
  - Decision Tree
- Ground Truth:
  - Host Information in Active Directory and vulnerability scanner Nessus database

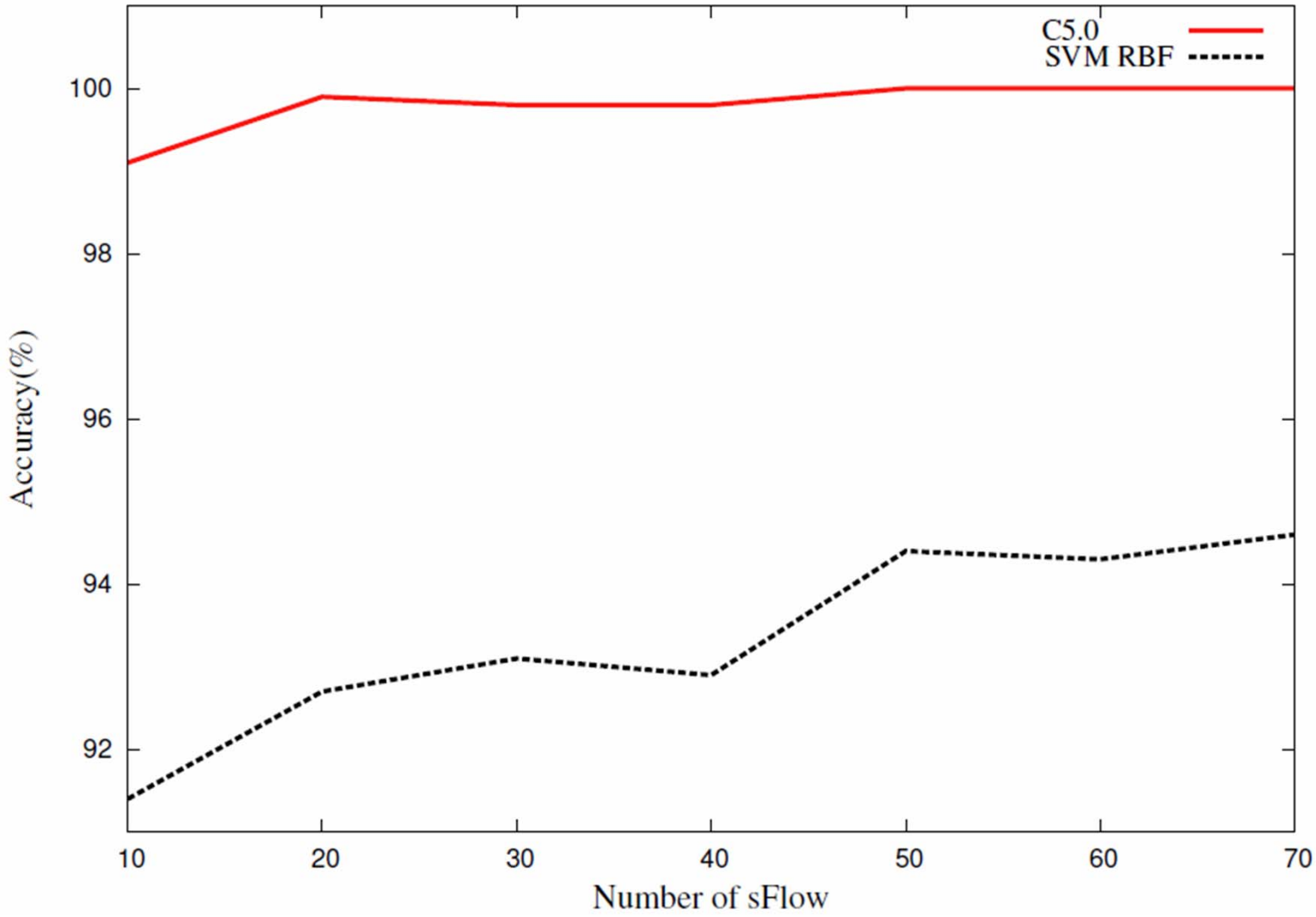
Client vs Server Classification



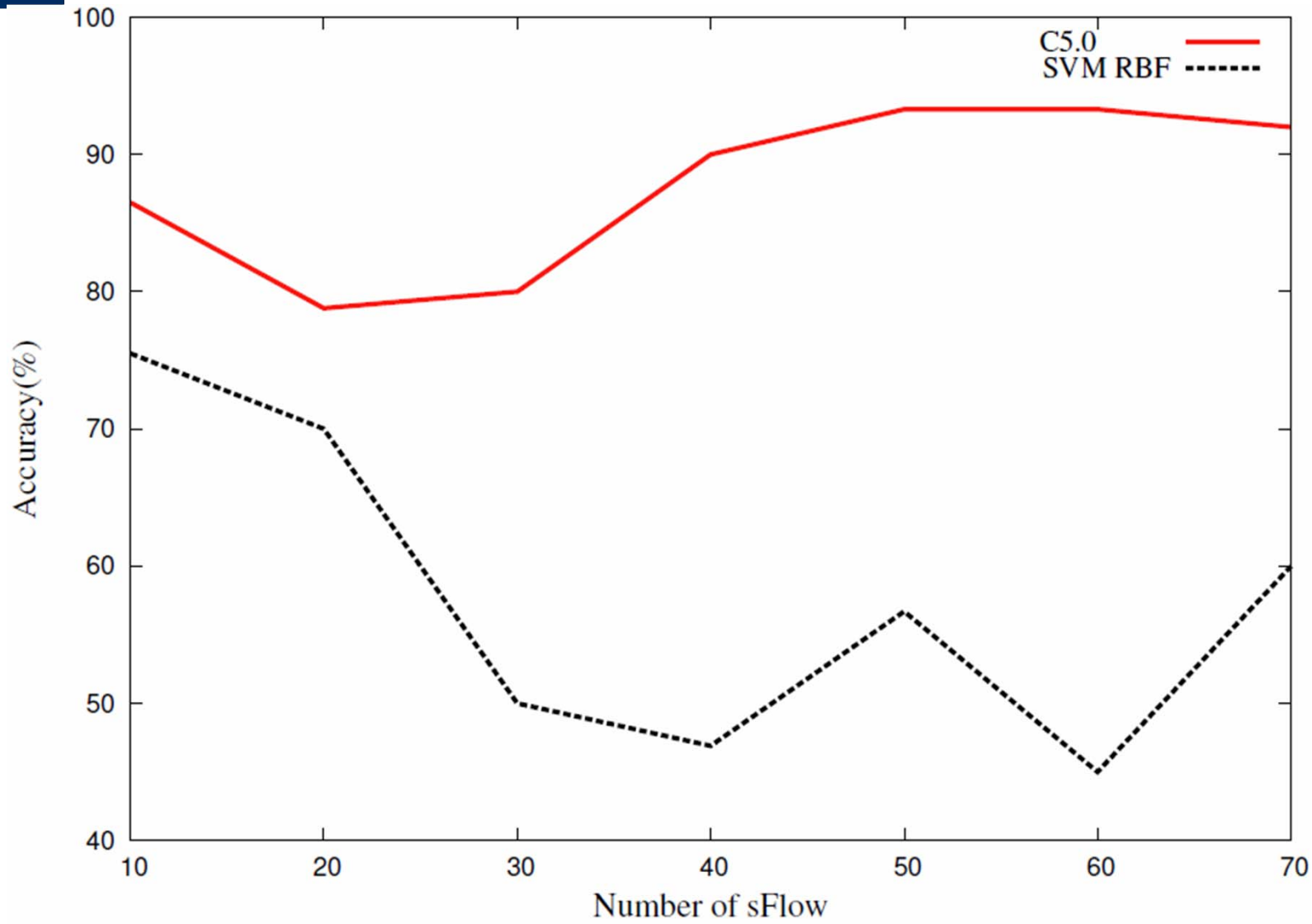
Personal System vs Public System



Web Server vs Email Server



Classifying Two Different Colleges

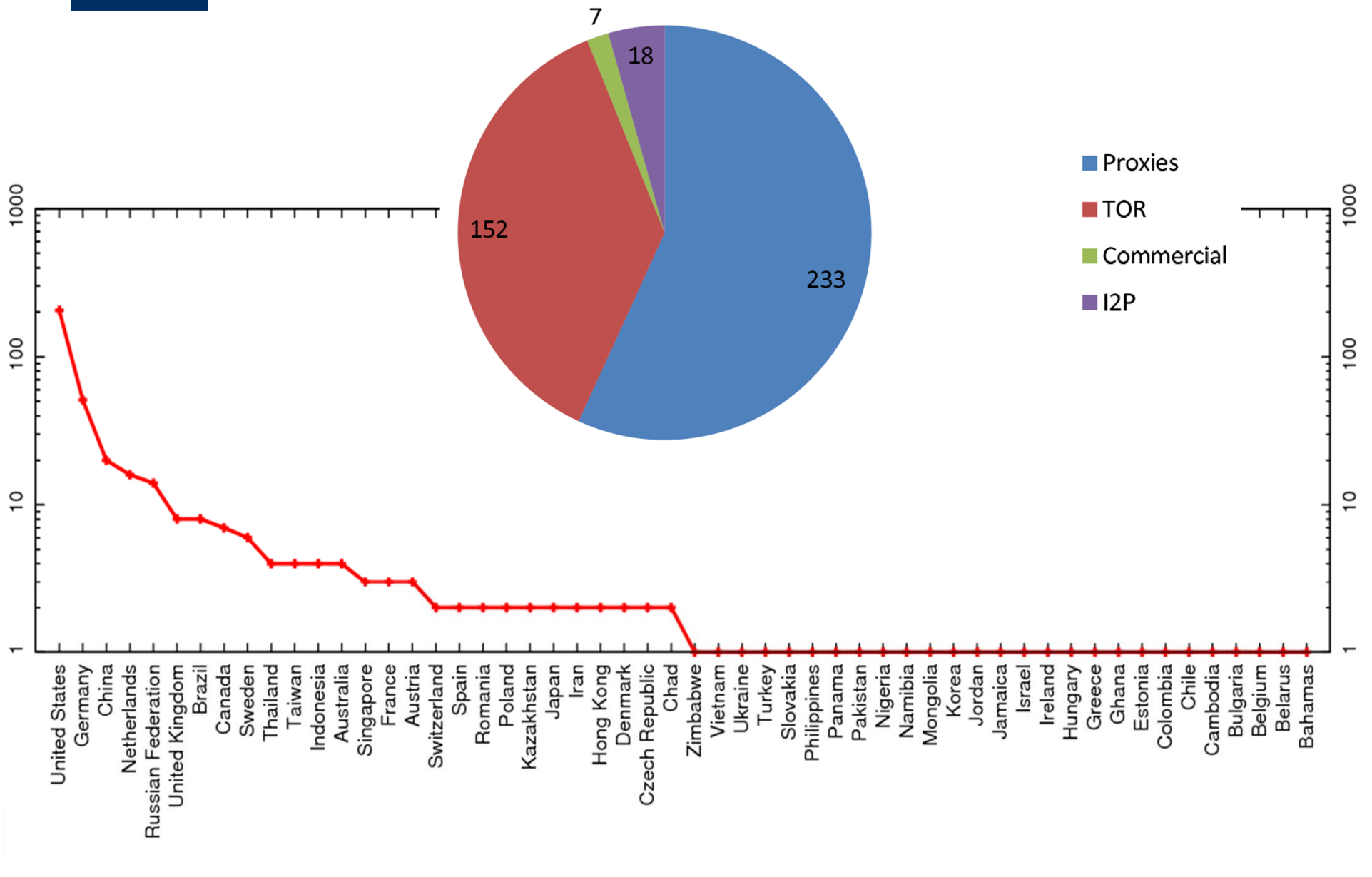


Anonymizer Usage

- Anonymity network usage via Pig scripting
  - 205 million packets
  - about 1.44TB data
- Analyzed Anonymity Networks

Network	Servers	Service
Tor	61,798	General
I2P	2,267	P2P
JAP	11	General
Remailers	15	Email
Proxies	7,246	General
Commercial	Anomymizer,Gotrusted	General

Anonymity Network Geolocation



Thanks