APRICOT 2012
21February-2 March, 2012 - New Delhi, India

Extracting Benefit from Harm: Using Malware Pollution to Analyze the Impact of Political and Geophysical Events on the Internet

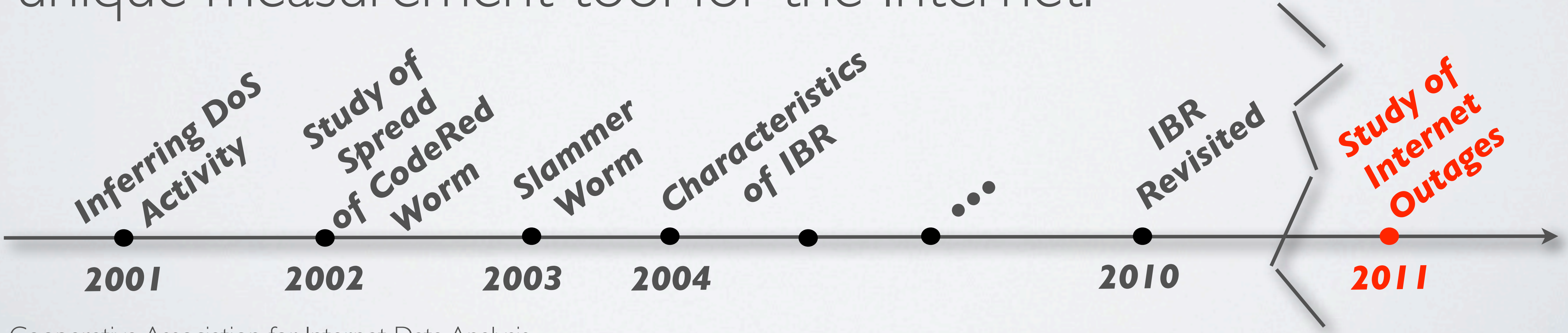
A. Dainotti, R. Amman, E. Aben, K. C. Claffy
kc@caida.org
CAIDA/UCSD

CONTEXT

Project goal & main message

- Analysis of **macroscopic Internet events** using multiple large-scale data sources

- Revival of Network Telescopes: ***Internet Background Radiation*** can be used as a unique measurement tool for the Internet!



Inferring DoS Activity
2001

Study of Spread of CodeRed Worm
2002

Slammer Worm
2003

Characteristics of IBR
2004

...

IBR Revisited
2010

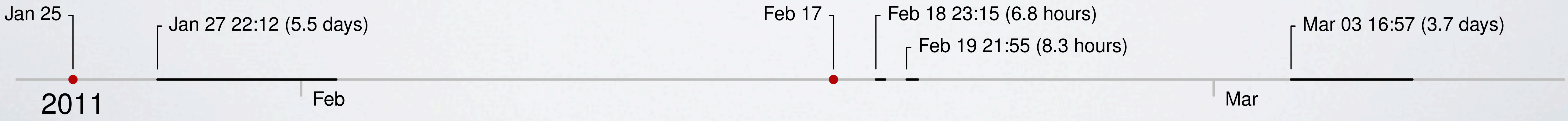
Study of Internet Outages
2011

THE EVENTS (1/2)

Internet Disruptions in North Africa

- Egypt
  - *January 25th, 2011*: protests start in the country
  - The government orders service providers to "shut down" the Internet
  - ***January 27th, around 22:34 UTC***: several sources report the withdrawal in the Internet's global routing table of almost all routes to Egyptian networks
  - The disruption lasts ***5.5 days***

- Libya
  - *February 17th, 2011*: protests start in the country
  - The government controls most of the country's communication infrastructure
  - ***February 18th (6.8 hrs), 19th (8.3 hrs), March 3rd (3.7 days):*** three different connectivity disruptions:



Jan 25

Jan 27 22:12 (5.5 days)

Feb 17

Feb 18 23:15 (6.8 hours)

Feb 19 21:55 (8.3 hours)

Mar 03 16:57 (3.7 days)

2011

Feb

Mar

NETWORK INFO

Prefixes, ASes, Filtering

- Egypt
  - **3165 *IPv4*** and 6 *IPv6* **prefixes** are delegated to Egypt by AfriNIC
  - They are managed by **51 Autonomous Systems**

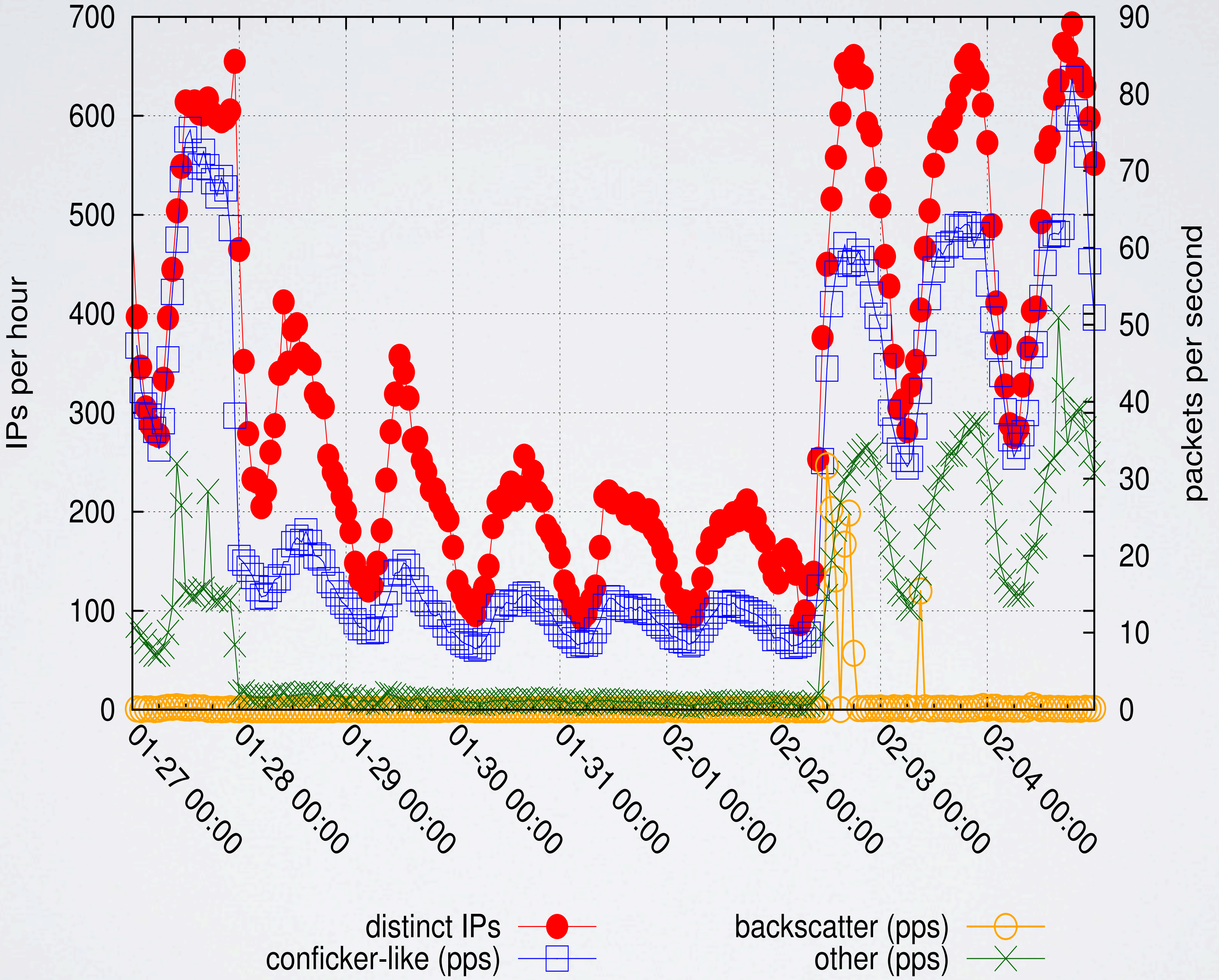
  - **Filtering** type: **BGP only**
  - Filtering dynamic: synchronized; progressive

- Libya
  - **13 *IPv4* prefixes**, no *IPv6* prefixes
  - **3 Autonomous Systems** operate in the country

  - **Filtering** type: mix of **BGP, packet filtering, satellite signal jamming**
  - Filtering dynamic: testing different techniques; somehow synchronized

EGYPT

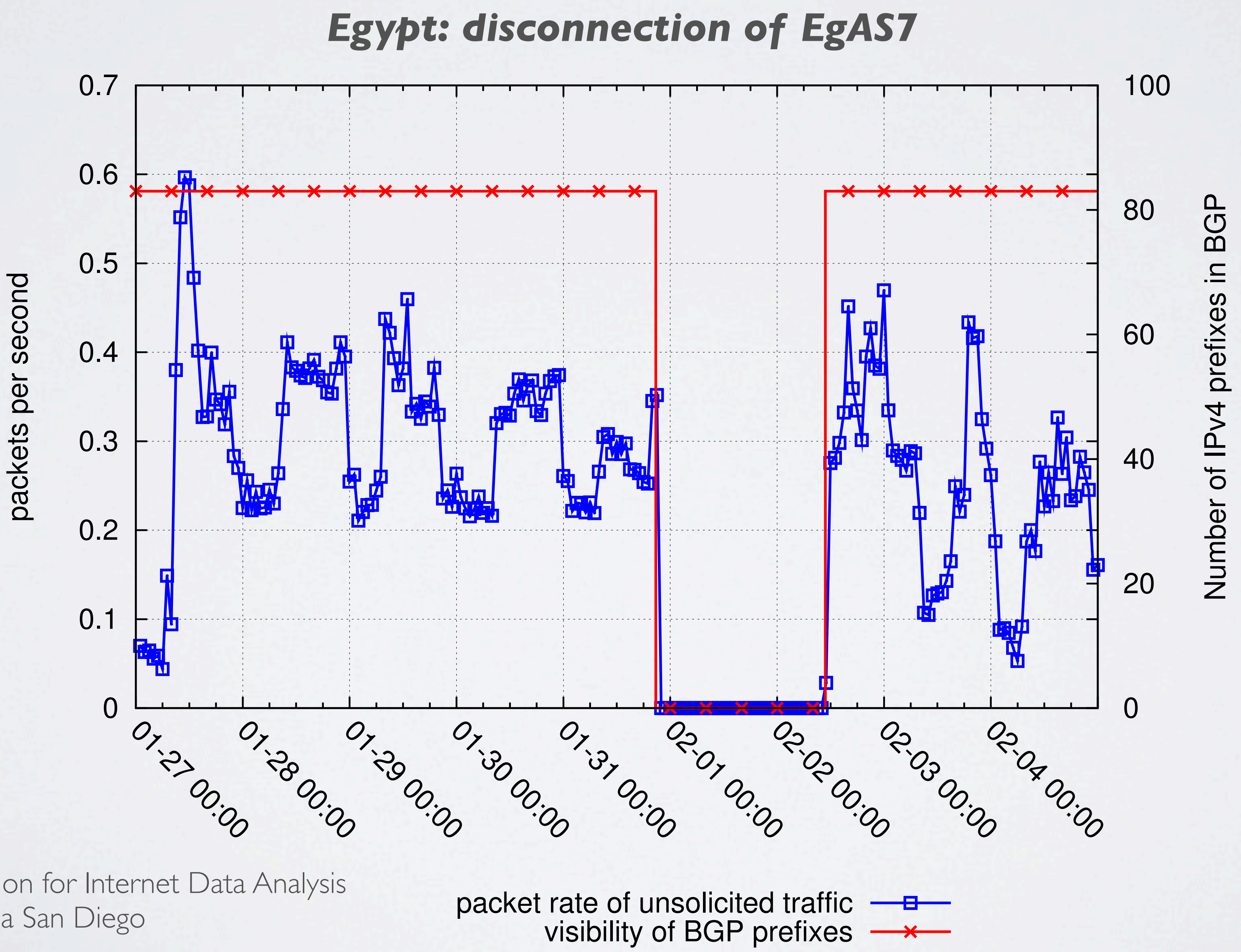
rate of distinct src IPs vs packet rate



TELESCOPE *vs* BGP

Consistency

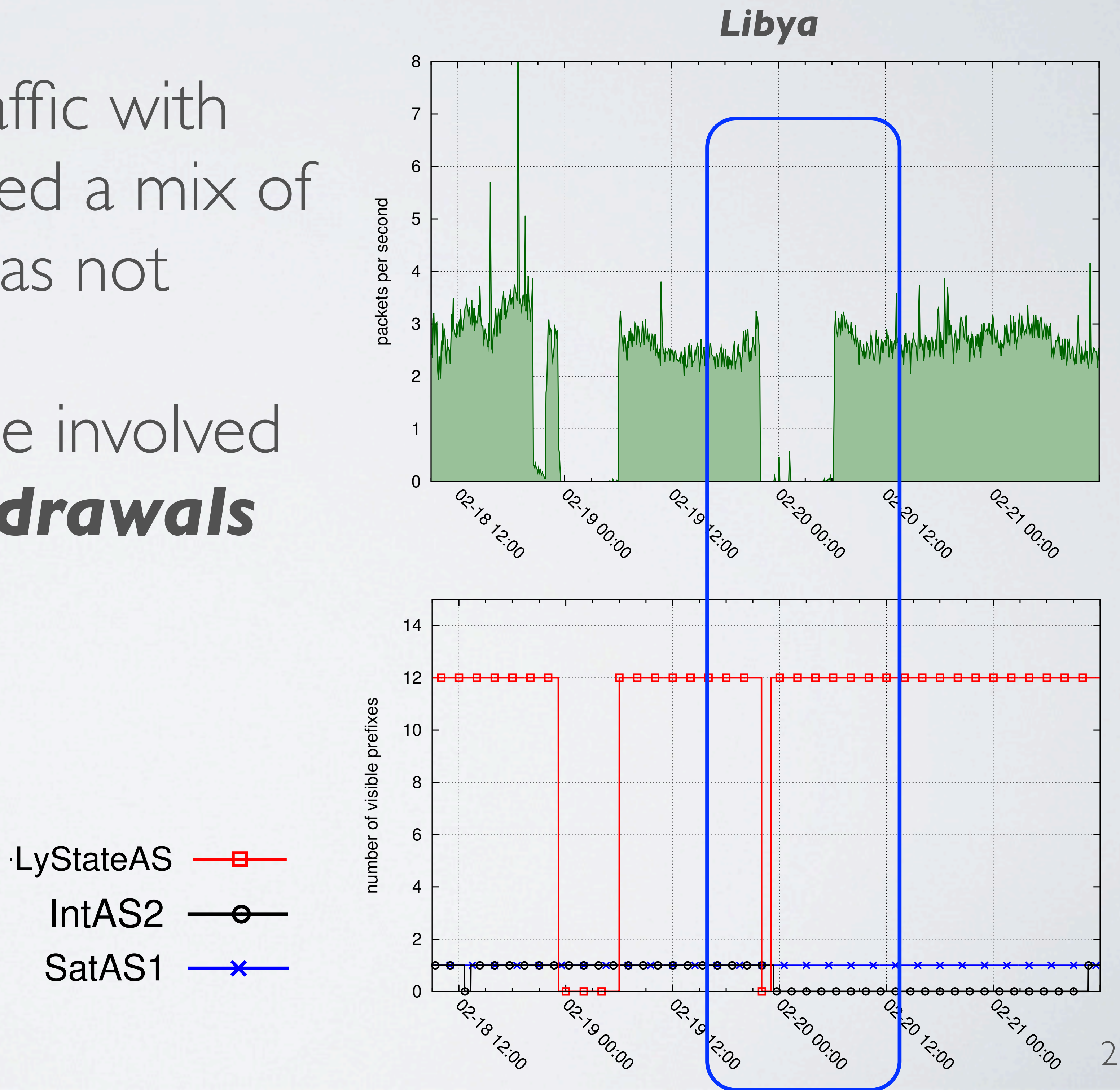
- The sample case of *EgAS7*shows the consistency between telescope traffic and BGP measurements



TELESCOPE vs BGP

Complementarity

- Contrasting telescope traffic with BGP measurements revealed a mix of blocking techniques that was not publicized by others
- The second Libyan outage involved overlapping of ***BGP withdrawals*** and ***packet filtering***



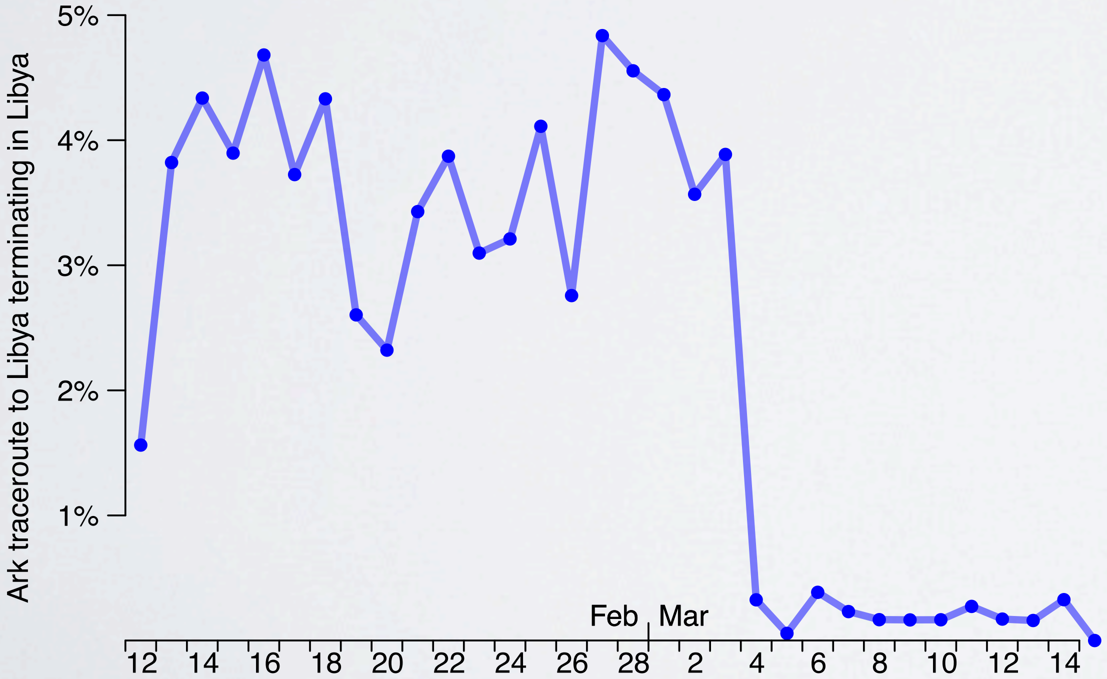
ARK

confirming telescope's findings

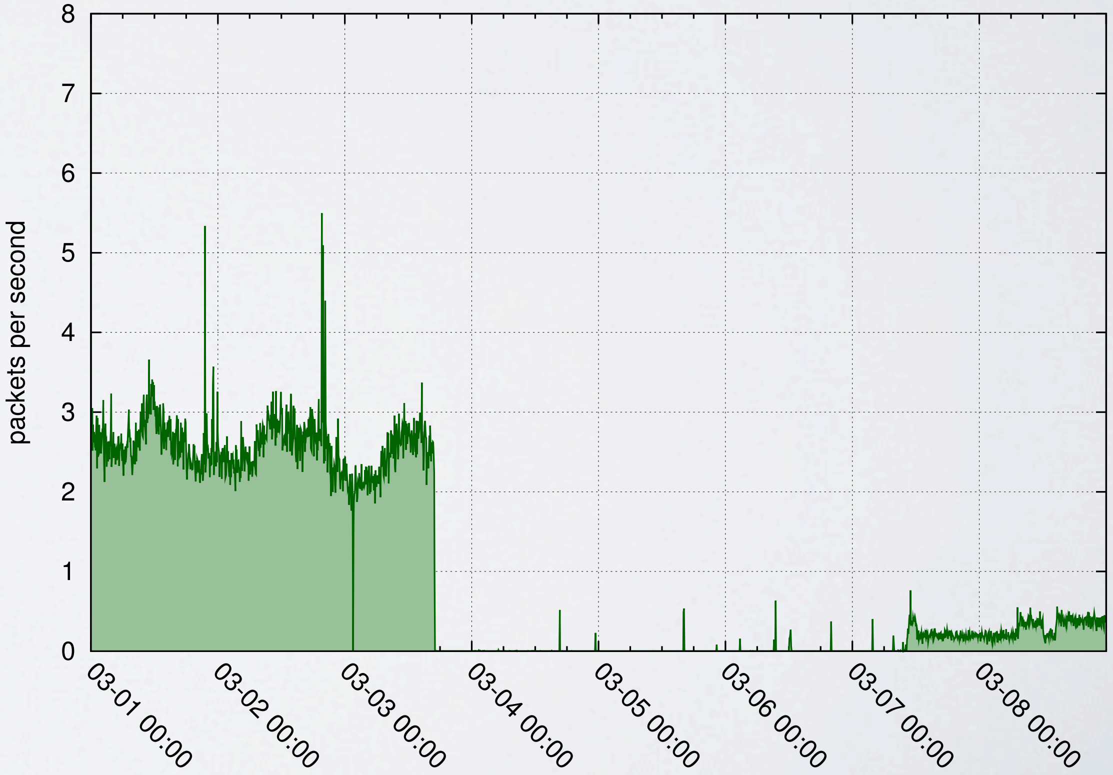
- Third Libyan outage: while BGP reachability was up, most of Libya was disconnected
  - ARK measurements confirmed the finding from the telescope
    1) disconnection
    2) identification of some reachable networks

  suggesting the use of packet filtering by the censors

Libya seen by ARK



Libya seen by the Telescope

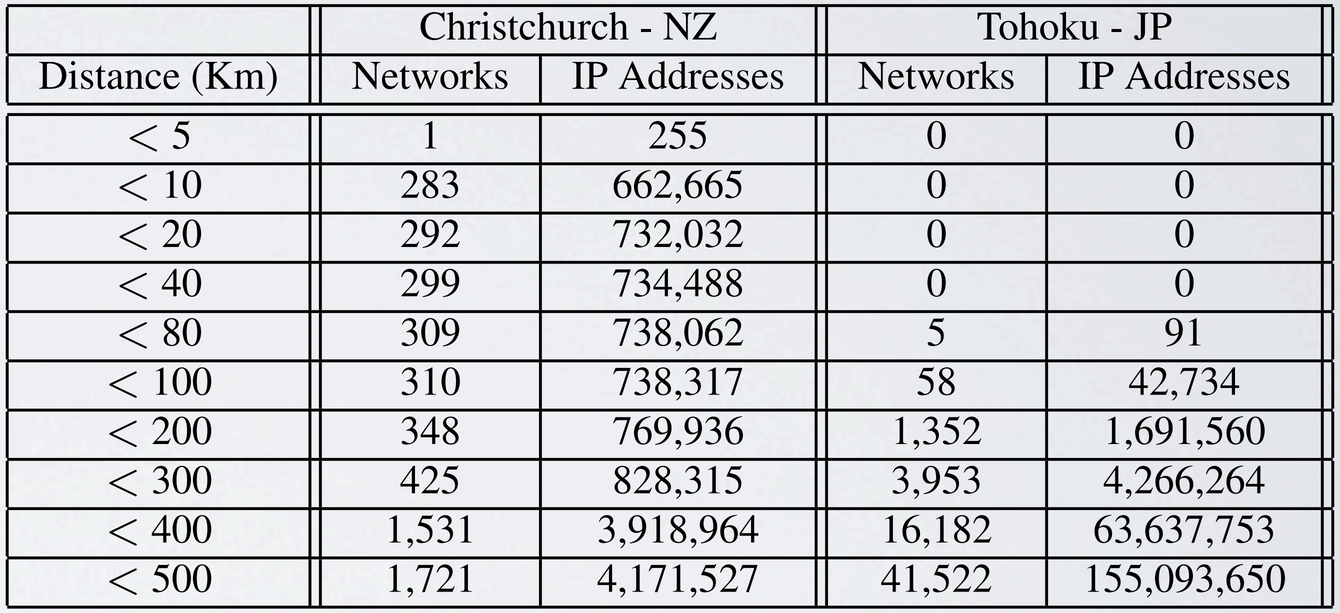


THE EVENTS (2/2)

Earthquakes

- Christchurch - NZ
  - *February 21st, 2011* 23:51:42 UTC
  - Local time 22nd, 12:51:42 PM
  - Magnitude: 6.1

- Tohoku - JP
  - *March 11th, 2011* 05:46:23 UTC
  - Local time 02:46:23 PM
  - Magnitude: 9.0



	Christchurch - NZ		Tohoku - JP	
Distance (Km)	Networks	IP Addresses	Networks	IP Addresses
< 5	1	255	0	0
< 10	283	662,665	0	0
< 20	292	732,032	0	0
< 40	299	734,488	0	0
< 80	309	738,062	5	91
< 100	310	738,317	58	42,734
< 200	348	769,936	1,352	1,691,560
< 300	425	828,315	3,953	4,266,264
< 400	1,531	3,918,964	16,182	63,637,753
< 500	1,721	4,171,527	41,522	155,093,650

We use MaxMind GeoLite City DB to compute distance from a given network to the epicenters

A SIMPLE METRIC

to evaluate impact and extension

- $I_{\Delta t_i}$ number of distinct source IP addresses seen by the telescope over the interval Δti,
- $\Delta t_1, ..., \Delta t_n$ 1-hour time slots ***following*** the event
- $\Delta t_{-1}, ..., \Delta t_{-n}$ 1-hour time slots ***preceding*** the event

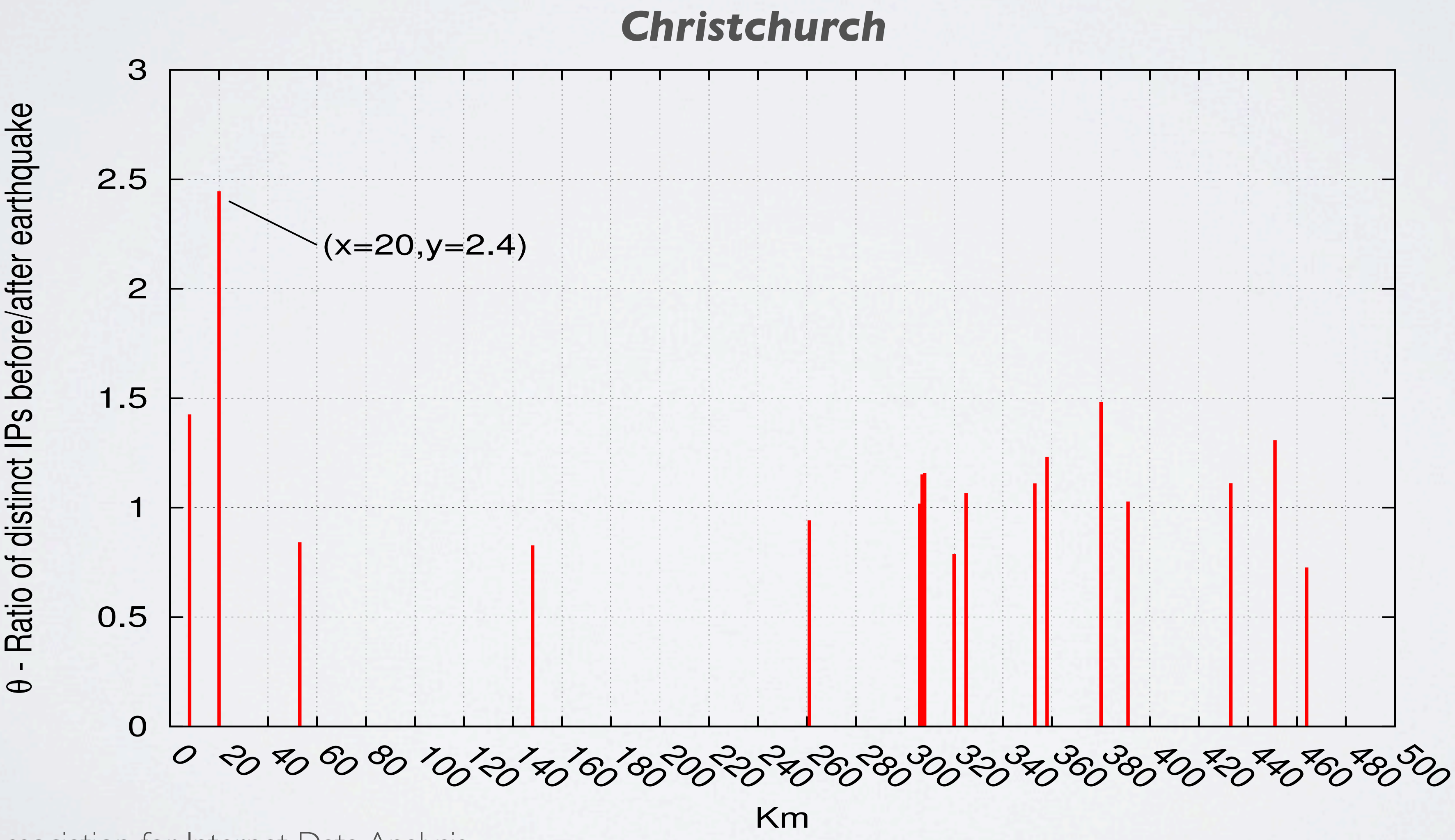
$$\theta = \frac{\displaystyle\sum_{i=-1}^{-24} I_{\Delta t_i}}{\displaystyle\sum_{j=1}^{24} I_{\Delta t_j}}$$

RADIUS OF IMPACT

rough estimate based on θ

- We compute θ for address ranges geolocated at different distances from the epicenter of the earthquake *(0 to 500km in bins of 1km each)*

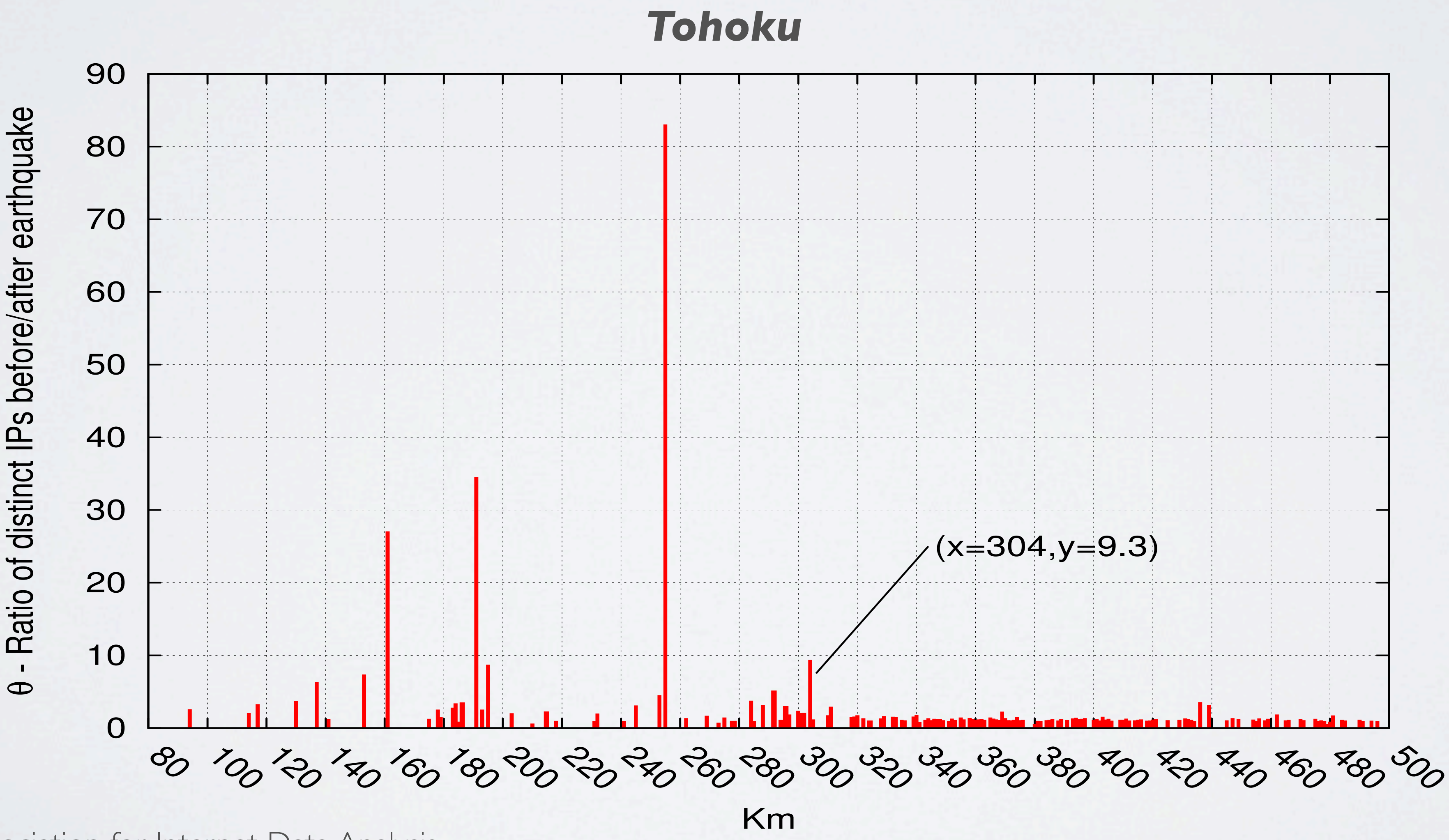
- θ around 1 indicates no substantial change in the number of unique IP addresses observed in IBR before and after the event.



RADIUS OF IMPACT

rough estimate based on θ

We call ρ_{max} the maximum distance at which we observe a value of θ significantly > 1

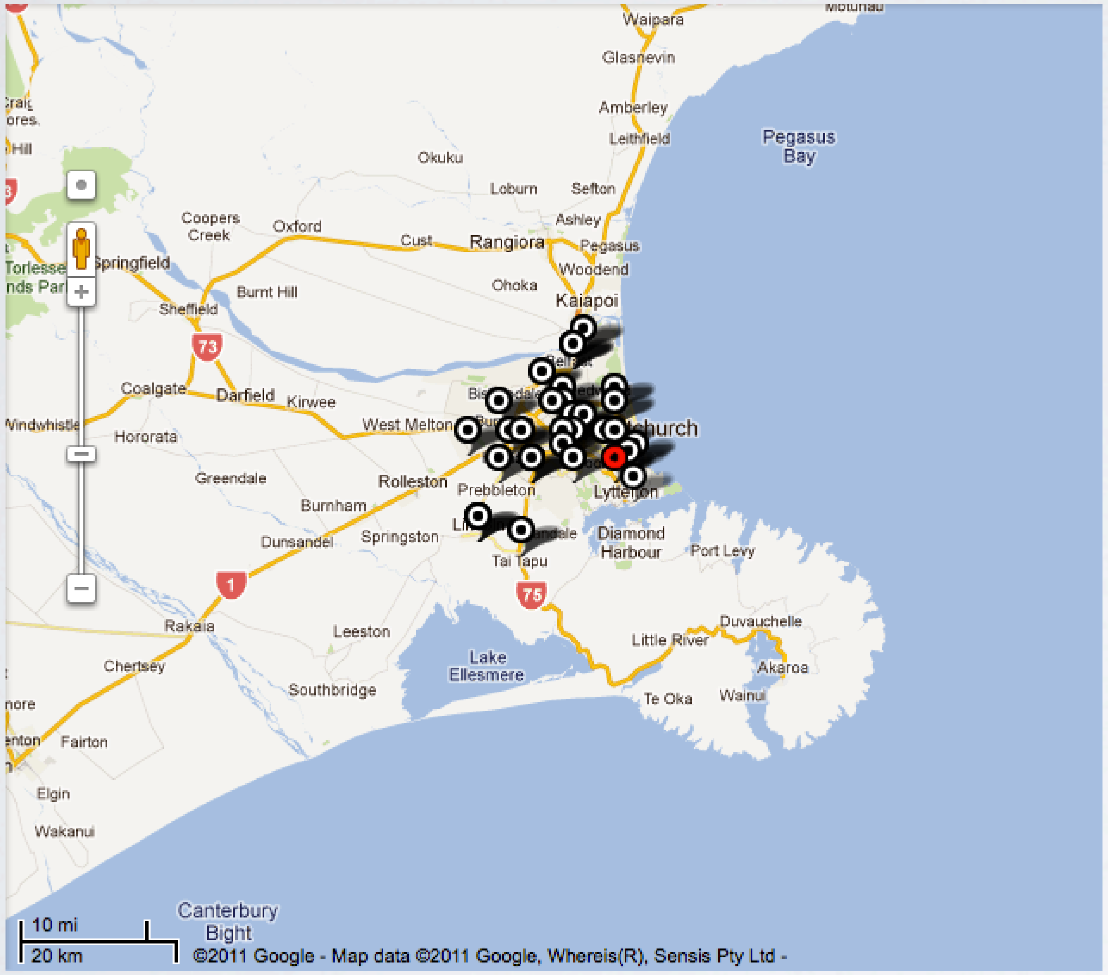


Cooperative Association for Internet Data Analysis
University of California San Diego

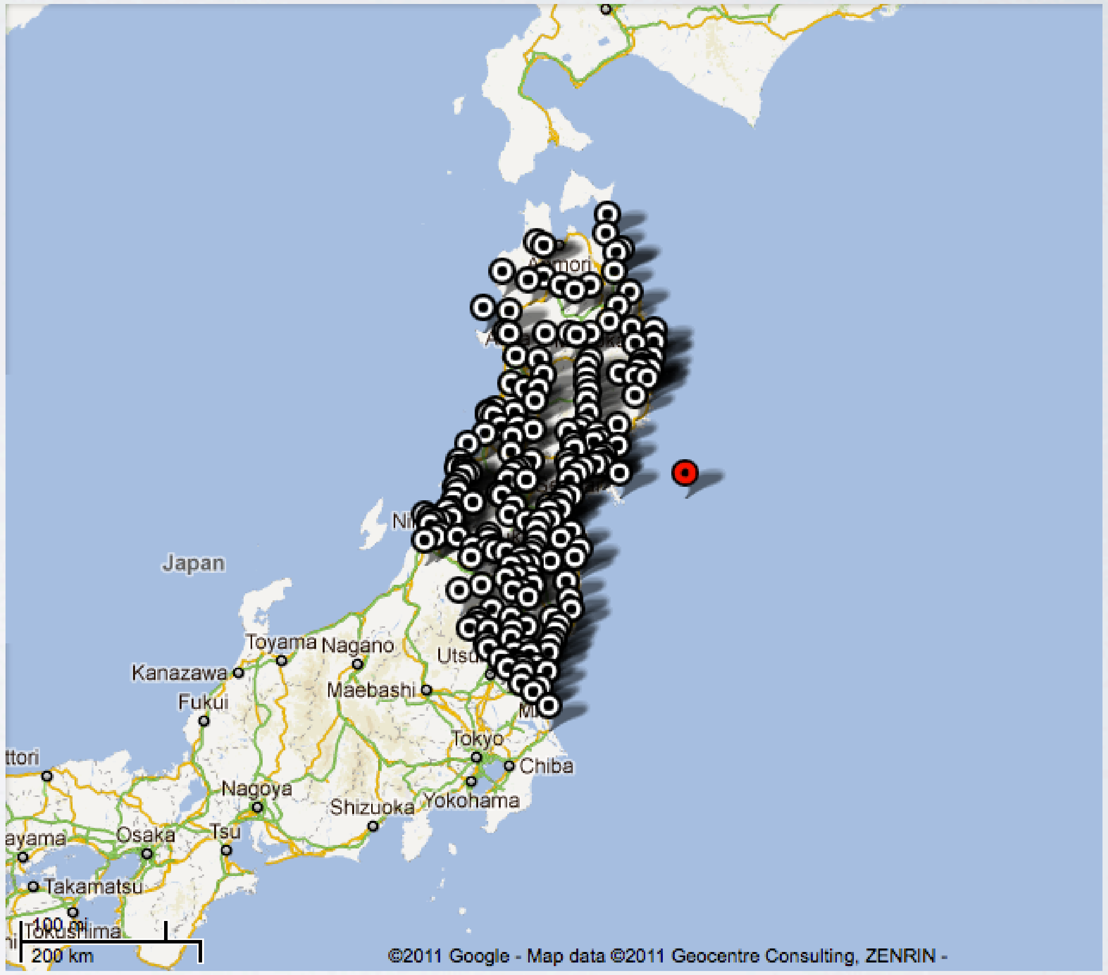
EXTENSION OF IMPACT

geo coordinates of most affected networks

Networks within each respective ρ_{max}



(a) Christchurch

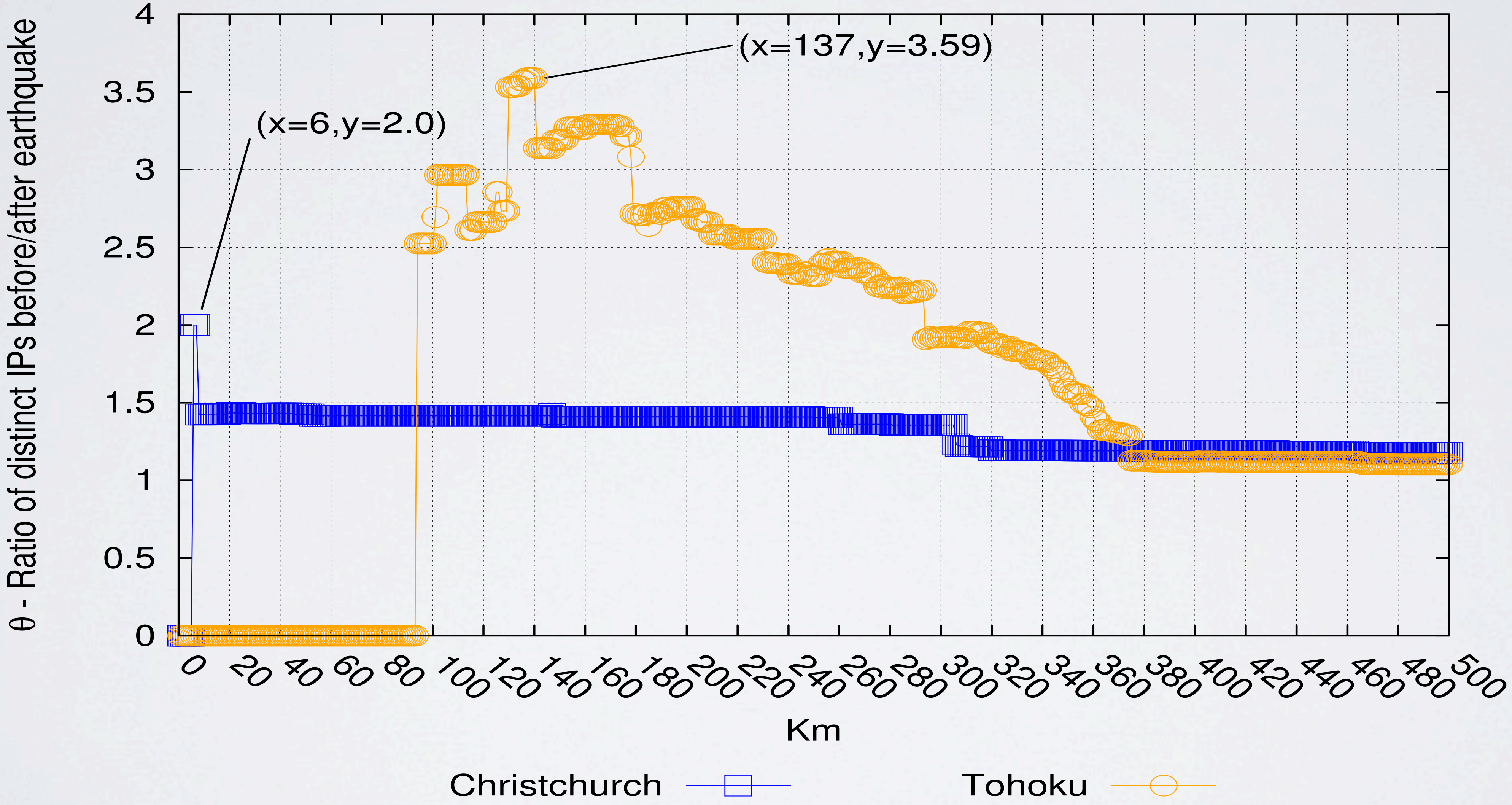


(b) Tohoku

“MAGNITUDE”

A measure of impact

- Varying the radius, we pick the highest value of θ calculated for *the whole set of* networks within the corresponding circle

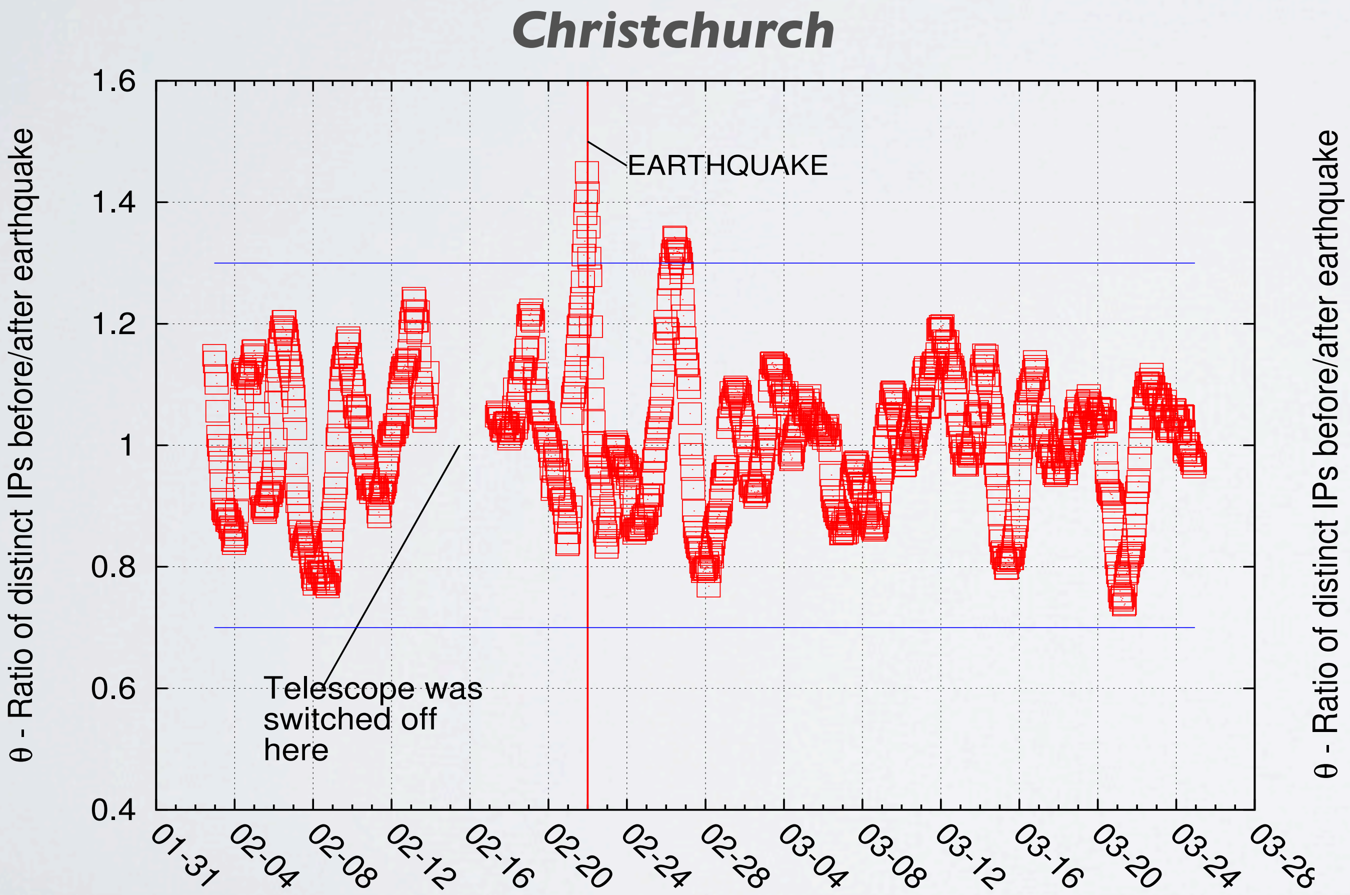


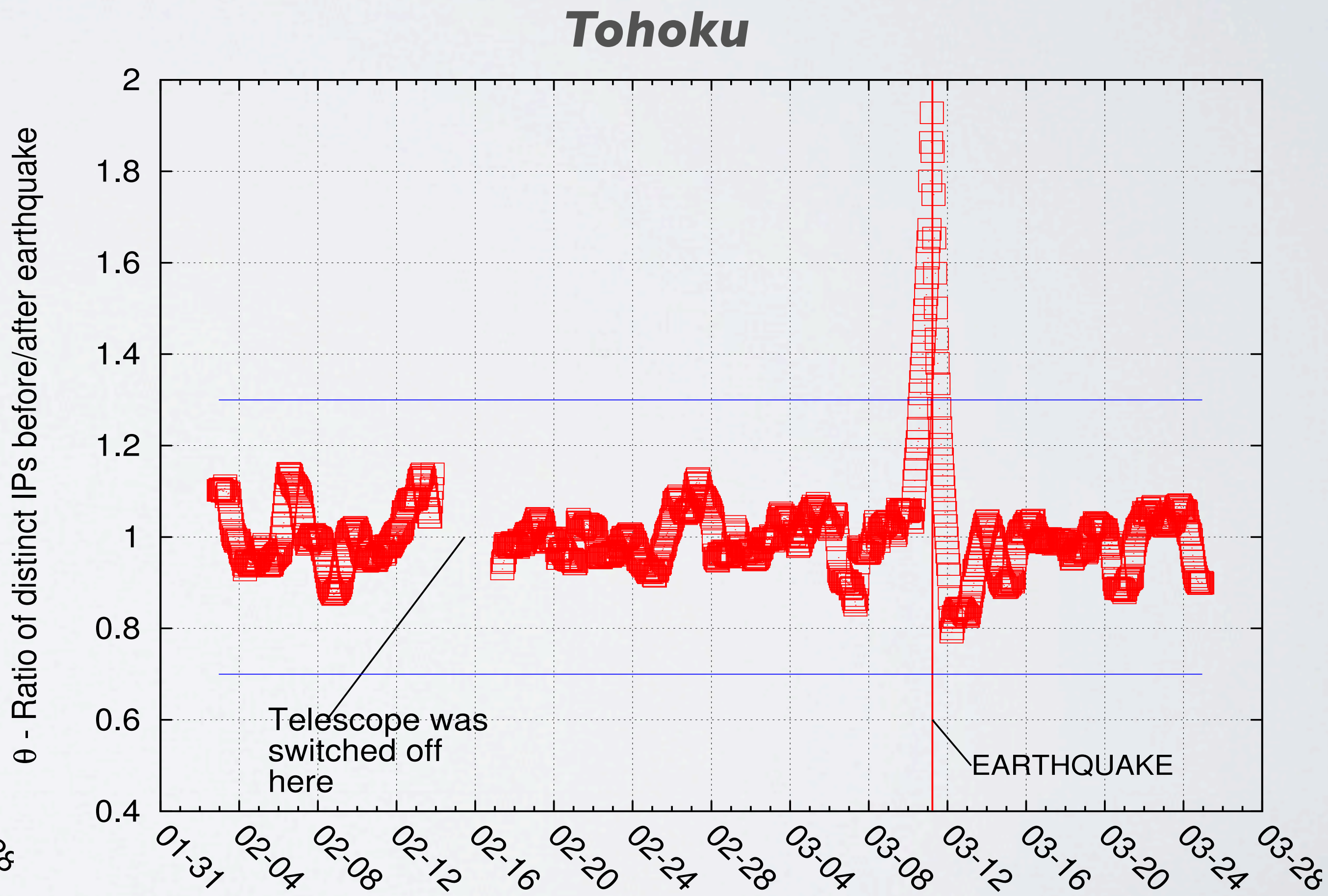
	Christchurch	Tohoku
Magnitude (θ_{max})	2 at $6km$	3.59 at $137km$
Radius (ρ_{max})	$20km$	$304km$

EVALUATING Θ

variations over a long time period

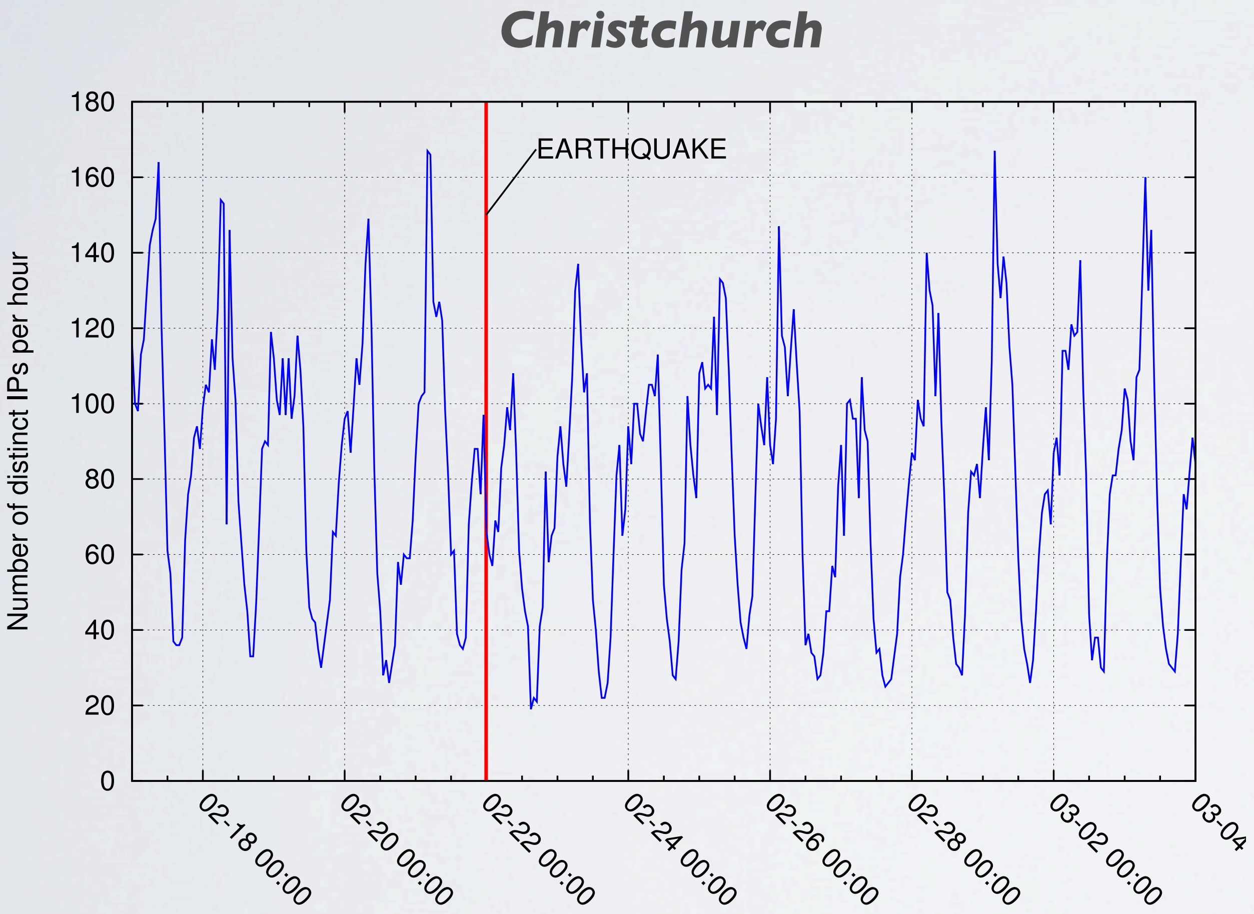
- 2 months period of observation
- θ normally stays within [0.7 - 1.3]

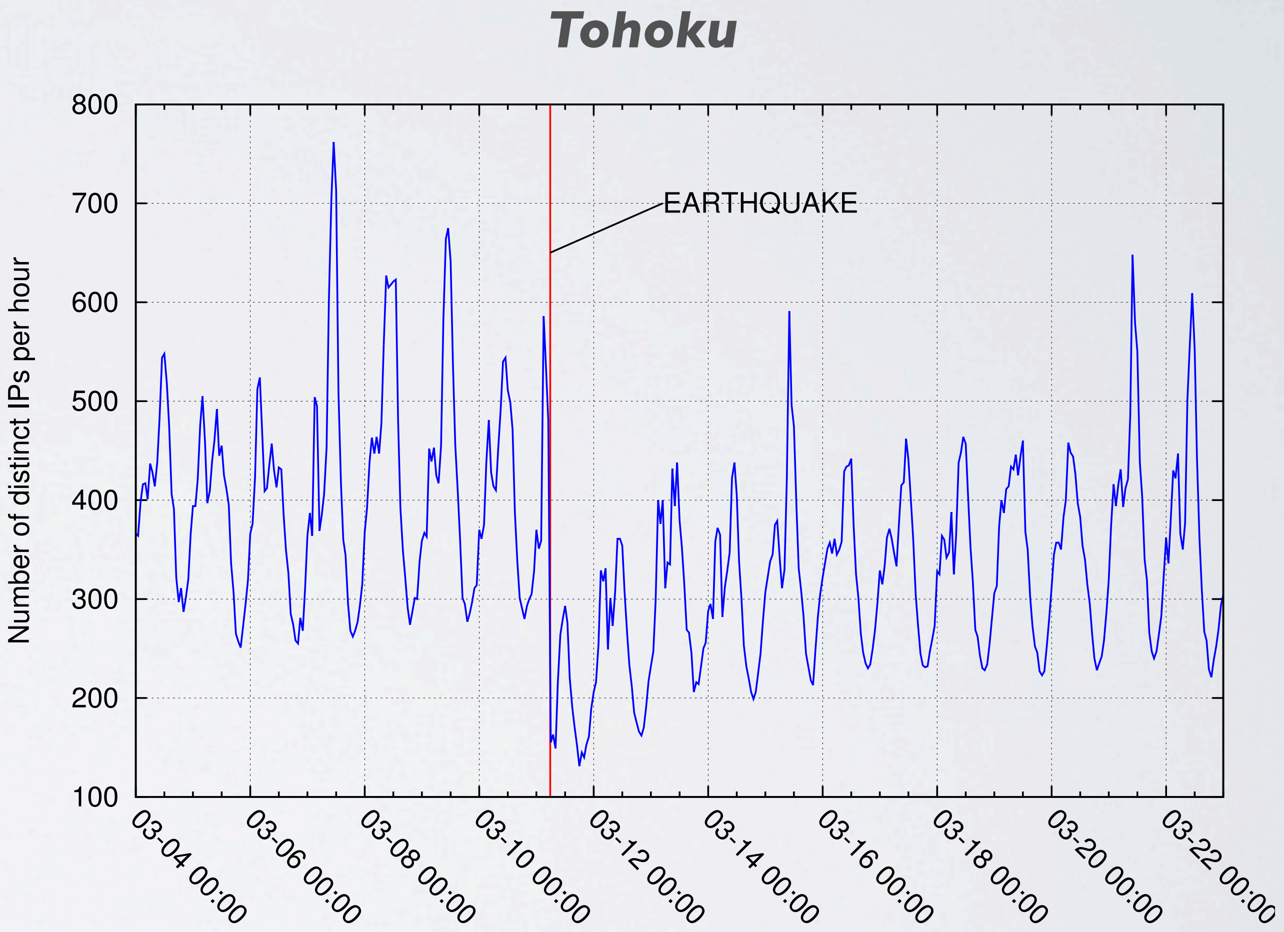




IP RATE IN TIME

reflects the dynamics of the event





CONCLUSION

ongoing work

- IBR is an effective source of data for the analysis of network outages caused by events of different typology

- Future work
  - Integrate and combine analysis of multiple data sources (BGP, IBR, active measurement, ...)
  - Analysis of AS/Link-level topology
  - Automated detection + triggered active measurements

THANKS

BKUP SLIDES

WHAT WE DID

Combined different measurement sources

- BGP
  - BGP updates from route collectors of ***RIPE-NCC RIS*** and ***RouteViews***
  - We combined information from both databases
  - Graphical Tools: ***REX***, ***BGPlay***, ***BGPviz***

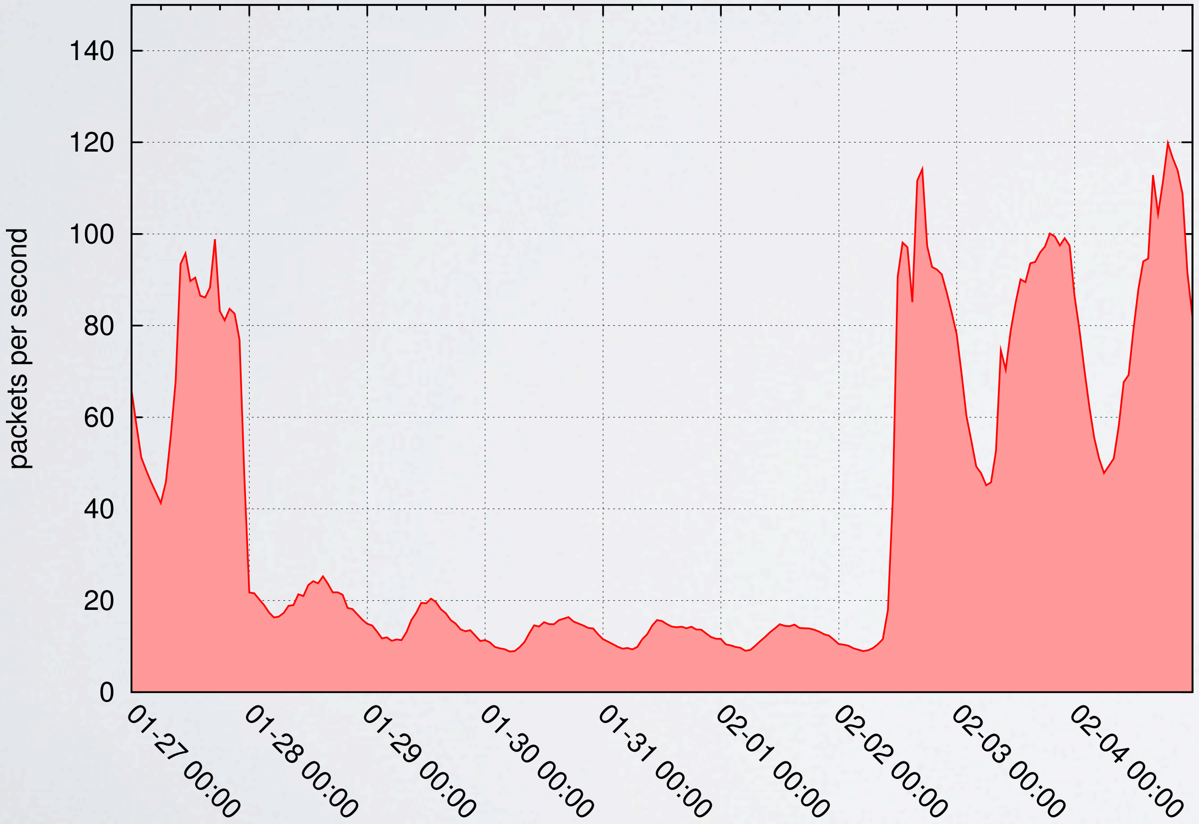
- Active Traceroute Probing
  - Archipelago Measurement Infrastructure (***ARK***)
  - We underutilized this data source..
- Internet Background Radiation (IBR)
  - Traffic reaching the ***UCSD Network Telescope***
  - Capable of revealing different kinds of blocking

UCSD TELESCOPE

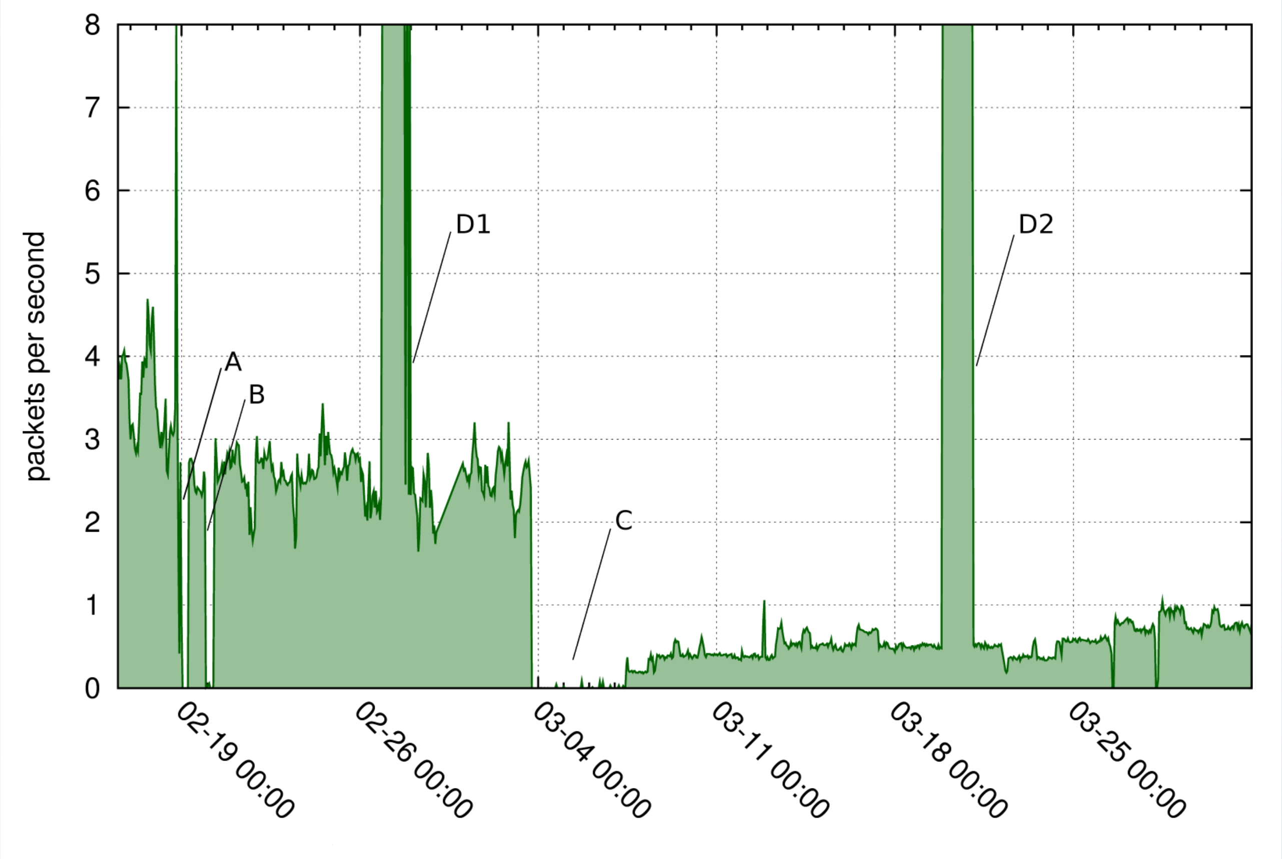
when malware helps..

- Unsolicited traffic, *a.k.a. Internet Background Radiation* - e.g. scanning from conficker-infected hosts - from the observed country reveals several aspects of these outages!

Egypt



Libya

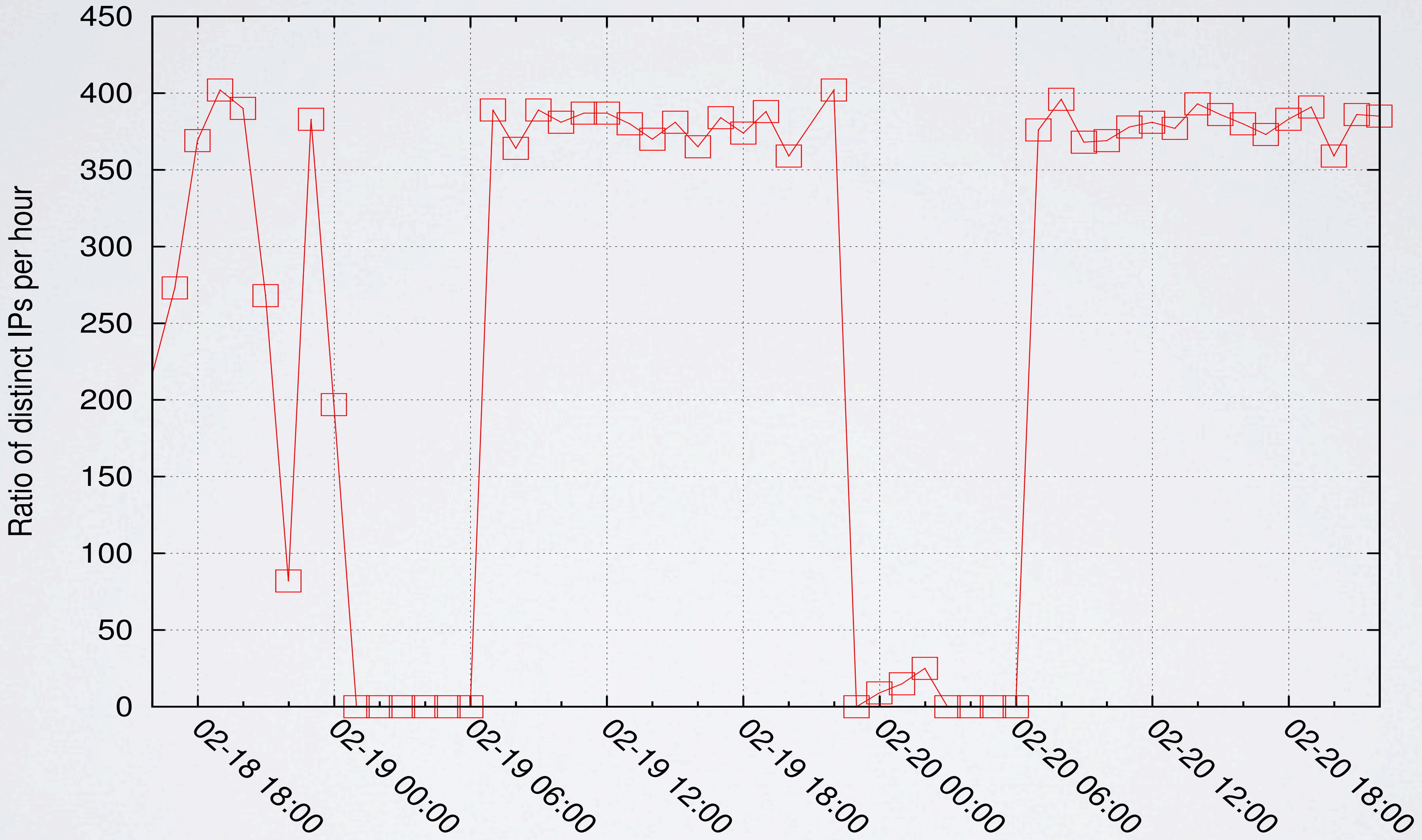


A,B,C: Outages
D1, D2: Denial of Service attacks

Cooperative Association for Internet Data Analysis
University of California San Diego

LIBYA

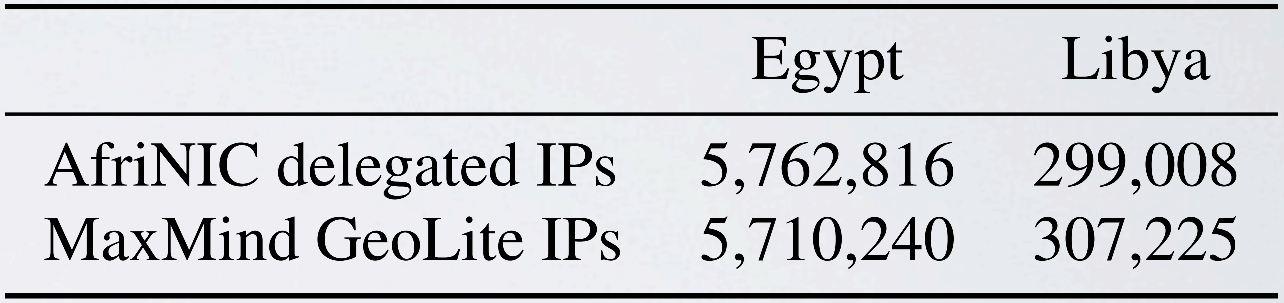
the first two outages



DATA SELECTION

Geolocation + announced prefixes

- IP ranges associated with the country of interest
  - Delegations from Regional Internet Registries (RIR)
  - Commercial geolocation database



	Egypt	Libya
AfriNIC delegated IPs	5,762,816	299,008
MaxMind GeoLite IPs	5,710,240	307,225

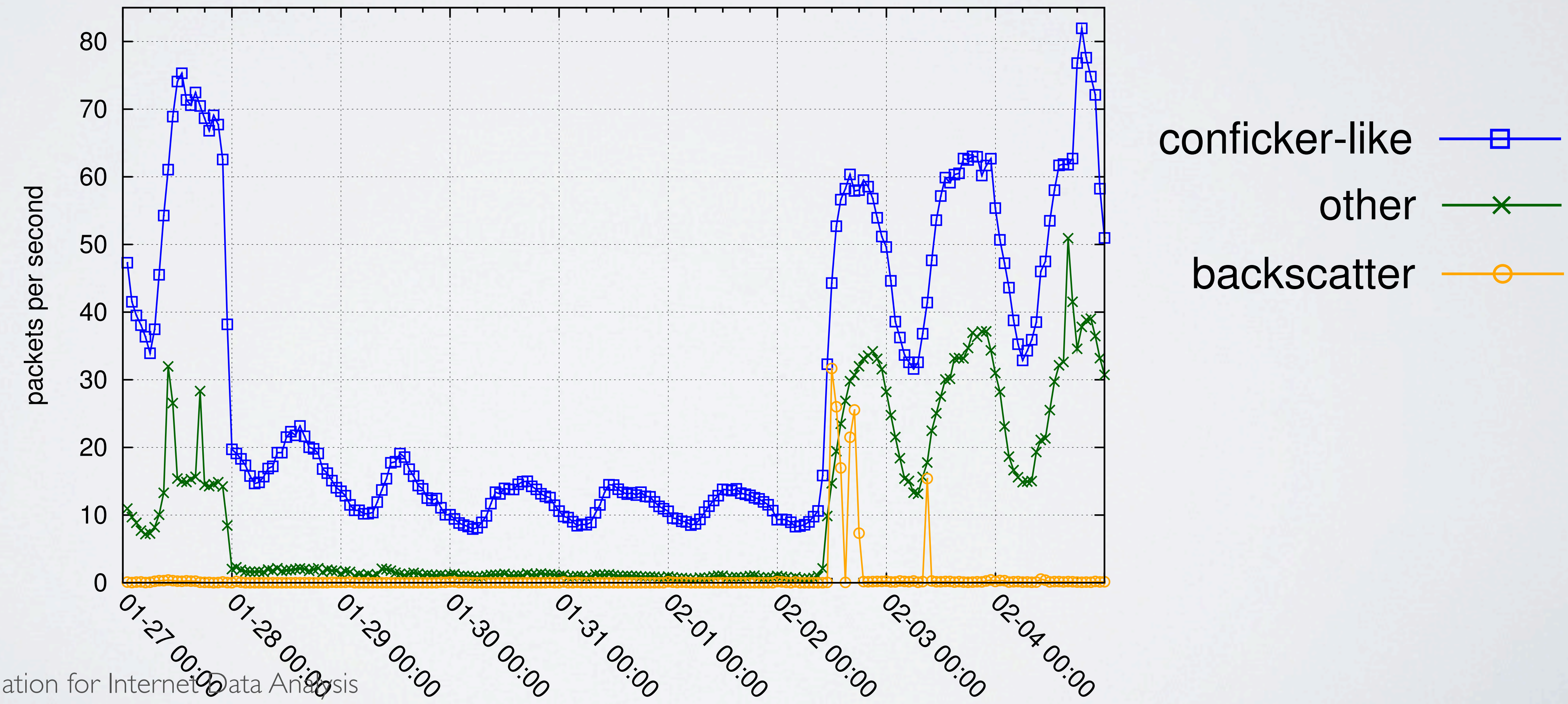
- Gather prefixes to be monitored by looking at BGP announcements. For each IP range:
  - Look up for an exactly matching BGP prefix
  - Find all the more specific (strict subset, longer) prefixes
  - Otherwise, retrieve the longest BGP prefix entirely containing it
- When referring to an AS, we actually refer to the IPs of that AS that are associated with the country of interest

UCSD TELESCOPE

need to dissect traffic

- We classified traffic to the telescope in
  - ***Conficker-like***
  - ***Backscatter*** (e.g. SYN-ACKs to randomly spoofed SYNs of DoS attacks)
  - ***Other***

Egypt: telescope traffic


packets per second
01-27 00:00
01-28 00:00
01-29 00:00
01-30 00:00
01-31 00:00
02-01 00:00
02-02 00:00
02-03 00:00
02-04 00:00
conficker-like
other
backscatter