# Internet Outage Detection Using Passive Analysis

USC Viterbi
School of Engineering
Information Sciences Institute

Asma Enayet
enayet@usc.edu

John Heidemann
johnh@isi.edu

## Introduction

Our need is to detect both short and long outages in IPv4 and IPv6. Outages are caused by natural disasters, political events, software and hardware issues, and human error and place a huge cost on today's e-commerce (an outage costs amazon $66k/minute).

We propose a new, principled approach to outage detection using passive data, to cover both IPv4 and IPv6, customizing parameters for each block to optimize the performance of our Bayesian inference model.

Contributions:

- We can detect **short-duration outages** as we control time precision
- We are the first one to report **outages in IPv6 address** space because we see data from clients that exists.
- We are the first to allow a **trade-off between spatial and temporal precision**

## Problem Statement

There are lots of short-duration outages, but prior works did not measure that

- Prior active detection systems cannot increase temporal precision without becoming intrusive resulting abuse complaints or discard.
- Prior passive detection systems detects short-duration outages, but at the cost of providing only much coarser, AS-level spatial precision.
- Our new passive approach can employ exact timestamps of observed data, allowing both fine spatial and temporal precision when possible.

Because the Internet is so diverse, outage detection systems need to be tuned to operate differently for differently-behaving regions.

- Prior passive systems are homogeneous, same parameters across all block providing only coarse spatial coverage or decreasing coverage.
- We exploit the ability to trade-off between spatial and temporal precision.
- We customize parameters to treat each blocks differently, allowing different regions to have different temporal and spatial precision.
- As a result we can get the coverage of more sparse blocks when we employ less temporal precision.

Of course there are outages in IPv6, but prior outage-detection systems have not been able to extend to IPv6, a growing part of the Internet today.

- Prior active monitoring systems cannot probe all unicast IPv6 address which requires centuries to scan,
- Our new approach extends coverage to IPv6, analyzing passive data, allowing the active addresses to come to us.

## Detection Algorithm

- **Goal**: detect short and long outages for the overall space
- **Input**: address blocks with unix timestamps of traffic arrival from B-Root
- **Output**: outages and availabilities with start time and duration.
- **Procedure**:
  - *P (a)* the rate at which traffic appears
  - Two extremes of *P(a)* is dense and sparse blocks
  - Bayesian inference to calculate the belief *B(a)* of the next time bin.
  - Finally, judge blocks as either down or up
  - Belief, $B(a)$ ranges from 0(DOWN) to 1(UP)

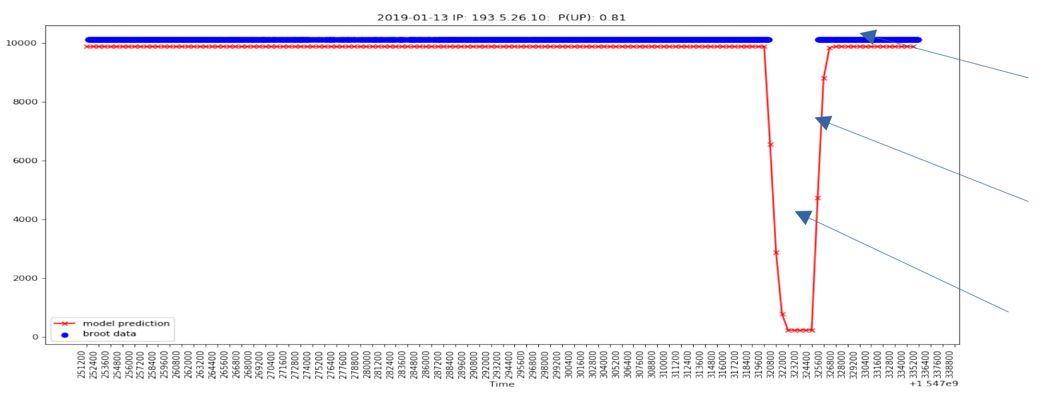
If there is a packet in the bin, then

$$\bar{B}(a) = \frac{P(a)B(a)}{P(a)B(a) + (1 - P(no|down))(1 - B(a))} \quad (1)$$

**else**

$$\bar{B}(a) = \frac{(1 - P(a))B(a)}{(1 - P(a))B(a) + (P(no|down))(1 - B(a))} \quad (2)$$

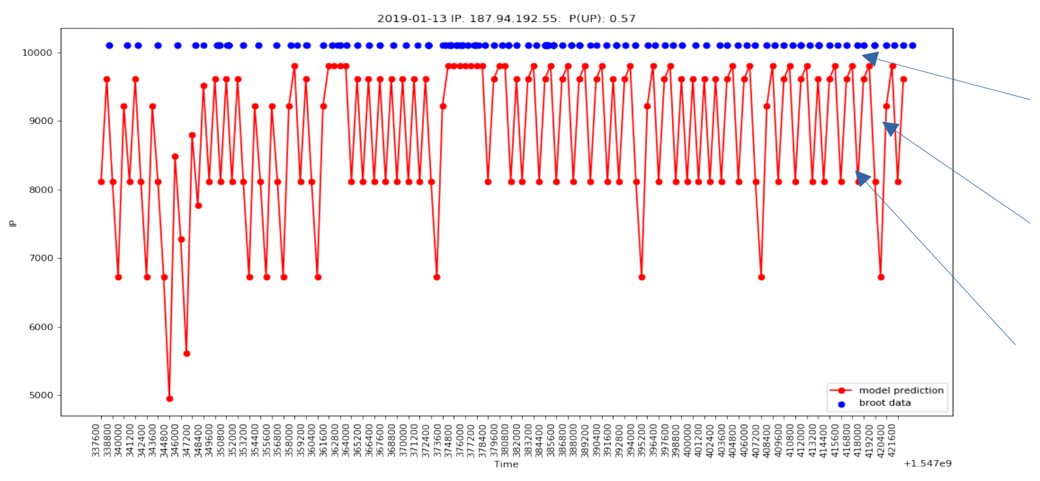
- To show that we can adapt and cover the whole range lets us show examples are two extreme cases of dense and sparse blocks



**Dense block**

Belief Changes more sharply

Can certainly say that DOWN



**Sparse block**

Belief Changes more frequently

Can not certainly say that DOWN

## Validation

- We use Trinocular and RIPE atlas data for comparison
- We analyze seven days of Internet outages observed from B-Root.

### Long outages : Comparing against active probing

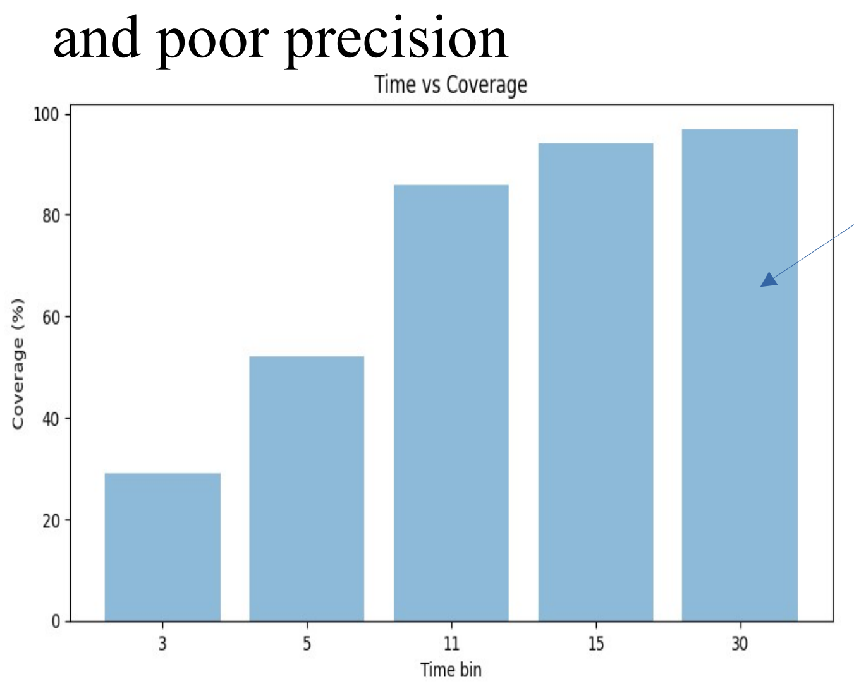
- We maximize the number of true outages detection.

| Observation | Ground truth (Trinocular) | | |
|---|---|---|---|
| (B-root) | availability (s) | outage (s) | |
| availability | TP = ta = 52525765695 | FP = fa = 2471178 | Precision 0.9999 |
| outage | FN = fo = 78163261 | TN = to = 13147965 | |
| | Recall 0.9985 | TNR 0.84178 | |

We have great precision and recall indicating our model have good accuracy

### Tradeoff between temporal and spatial precision

- We have good precision for the dense blocks and less precision for the sparse blocks
- We can either have good precision but less coverage or better coverage and poor precision



**Coverage** = percentage of observed /24 B-Root blocks with respect to total blocks.

Coverage increases with longer time bin

Observing longer duration there are more blocks that can have traffic.

The user can choose and get the best precision or coverage depending on the characteristics of data.

## Results

### Can we detect short-duration outages?

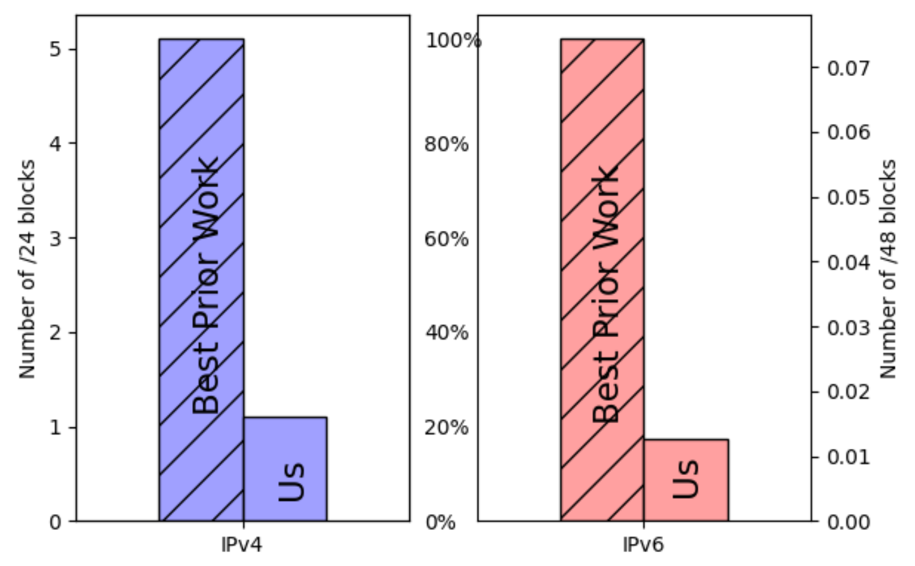
- Detects outages with short length unlike prior works.
- We study outages that are 5 minutes in length for both Broot
- and RIPE data on January 10, 2019.
- In RIPE data, there is status which defines if an address is available or out.

| Observation | Ground truth (RIPE) | | |
|---|---|---|---|
| (B-root) | availability (events) | outage (events) | |
| availability | 4445 | 105 | Precision 0.97692 |
| outage | 257 | 290 | |
| | Recall 0.9453 | TNR 0.7341 | |

We have great precision and recall for short outages.

- Evaluation of short outages is challenging: precision of ±180 s hides differences in uncertainty for short outages (300 s or less).
- We compare short outages by events (not time) to factor out imprecision in timing.

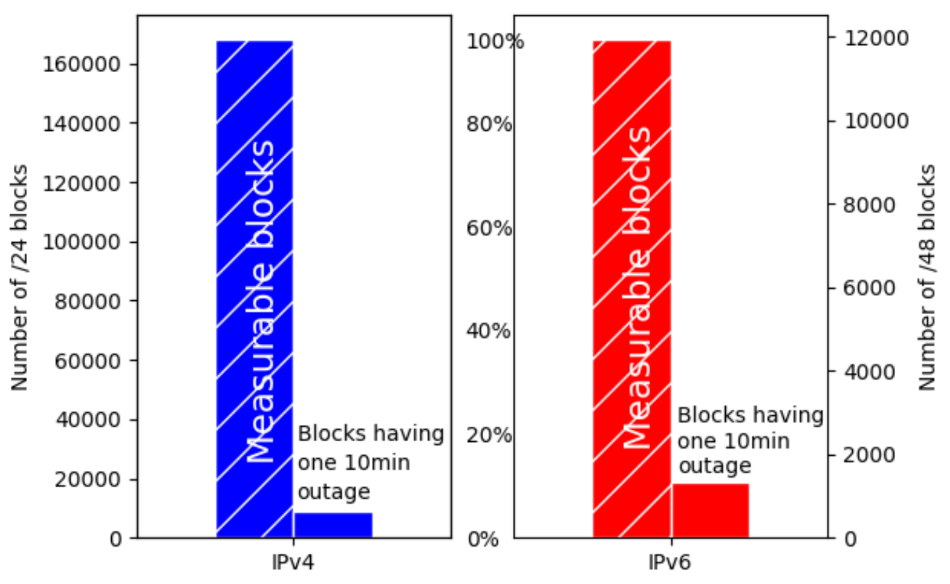
### Extending to IPv6



**Coverage report- IPv4 vs IPv6:**

- The fractions of best prior works' coverage are almost similar for both IPv4 (19.6% of Trinocular's) and IPv6 (17% of Gasser's)

**Outage report- IPv4 vs IPv6:**
Outage rate for IPv6 (12%) is greater than for IPv4 (5.5%)
—IPv6 reliability can improve.



**Our approach publish first reports on IPv6.**
**Our approach works on IPv6 is not as reliable as IPv4, and that our use of passive data can provide coverage that is a good fraction of current best IPv6 hitlists.**

## Conclusion

- We detect short outages which prior works could not detect before.
- Our method of IPv6 detection coverage is as consistent as IPv4 detection.
- We show that users can tradeoff between precision, coverage and correctness.
- We will continue to work on different types of passive data sets..

ACM IMC Poster 2022