Internet Outage Detection using Passive Analysis (Poster Abstract and Poster)

Asma Enayet
University of Southern California
Information Science Institute and CS Dept.
Los Angeles, California, USA

John Heidemann
University of Southern California
Information Science Institute and CS Dept.
Los Angeles, California, USA

Outages from natural disasters, political events, software or hardware issues, and human error [2] place a huge cost on e-commerce ($66k/minute at Amazon [1]).

While several existing systems detect Internet outages [4–8], these systems often too inflexible, fixed parameters across the whole internet with CUSUM-like change detection. We instead propose a system using passive data, to cover both IPv4 and IPv6, customizing parameters for each block to optimize the performance of our Bayesian inference model.

Our poster describes our three contributions: First, we show how customizing parameters allows us often to detect outages that are at both fine timescales (5 minutes) and fine spatial resolutions (/24 IPv4 and /48 IPv6 blocks). Our second contribution is to show that, by tuning parameters different for different blocks, we can scale back temporal precision to cover more challenging blocks. Finally, we show our approach extends to IPv6 and provide the first reports of IPv6 outages.

Approach summary: Our approach uses passive traffic observations from network-wide services. In our evaluation we consider traffic from `B-root`'s DNS service, but in principal we could use data from a large website (like Wikipedia, Google, or Amazon) or other infrastructure (like NTP). We build a model of historical traffic from each source to the service, then detect interruptions that violate model history, using with Bayesian inference to detect outages. Some strong sources directly support detection, but when possible, we correlate multiple signals from the same region to corroborate results. When sources are weak, we can aggregate data over larger time periods or spatial scales to gain confidence.

Currently we set parameters per-block based on history. Future work will consider seasonal and diurnal effects.

Detecting short outages: Prior work either detects 11-minute outages at fine spatial scales (typically /24 blocks, [6, 7]) or 5-minute outages, but for coarse spatial scales (entire ASes, [4, 8]), or very fast reaction but with a very large amount of input data (seconds, but requiring all TCP flows [5]). In each case, prior systems only improve temporal resolution by increasing active traffic, passive spatial scale or input data. Our new approach interprets passive data and can employ exact timestamps of observed data, allowing both fine spatial and temporal precision in many cases.

Although our approach can apply to many traffic sources, we quantify its effectiveness for both long-duration and short-duration outages, by evaluate `B-root` [9] as a passive data source. We test against Trinocular active outage detection [6] and observations from RIPE Atlas (inspired by Chocolatine [4]) as ground truth. Since `B-root` coverage is limited, we compare only /24 IPv4 blocks that overlap between our observations from `B-root` and Trinocular's observations.

Observation (`B-root`)	Ground truth (Trinocular) 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	

Table 1: Confusion matrix for long-duration outages (in seconds)

Observation (`B-root`)	Ground truth (Trinocular) availability (s)	outage (s)	
availability	TP = ta = 7644527262	FP = fa = 77152	Precision 0.99
outage	FN = fo = 387011	TN = to = 2233042	
	Recall 0.99	TNR 0.96	

Table 2: Confusion matrix for long-duration outages on dense blocks (in seconds)

We evaluate our accuracy in the confusion matrix in Table 1. We define a false outage (*fo*) as a prediction of down when it's really up in Trinocular, with analogous definitions of false availability (*fa*), true availability (*ta*), and true outages (*to*). High precision means the outages that we report are true, and strong recall means we correct estimate duration. TNR suggests that we often find shorter outages than Trinocular. This difference may be due to Trinocular's precision (±330 s), while using exact timestamps of data allows us to be more precise and often shorter. To avoid uncertainty and show our model works for finding the maximum number of outages we also test on only dense data having high frequency of traffics in Table 2. This shows that we have very good precision and recall for very dense blocks and TNR shows that we can detect 96% of the outages. Evaluation of short outages is challenging: precision of ±180 s hides differences in uncertainty for short outages (300 s or less). The poster will compare short outages by events (not time) to factor out imprecision in timing.

In the Table 3 we report outage events that are 5 minutes in length for both `B-root` and RIPE data. Our results show that we have great precision (0.9979) and recall (0.9479) indicating our model has good accuracy in long-duration outages (11 minutes or more). Similarly, we have great precision (0.9769) and recall (0.9453) for short-duration outages (5 minutes or more). Our measurements show that on 2019-01-10, around 5% of total blocks that have 5 minute outages that were not seen in prior work. These short outages add up—when we add the outages from 5 to 11 minutes that were previously omitted to observations, we see that total outage duration increases by 20%.

Optimizing across a diverse Internet: Because the Internet is so diverse, outage detection systems need to be tuned to operate differently for differently-behaving regions. We describe the first passive system that optimizes parameters to each block to provide fine spatial and temporal precision when possible, but falling back on coarser temporal precision when necessary. By contrast, prior

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

Table 3: Confusion matrix for short-duration outages (events)

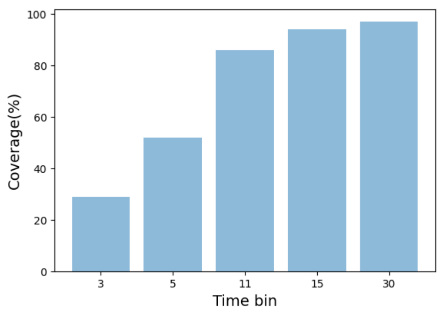
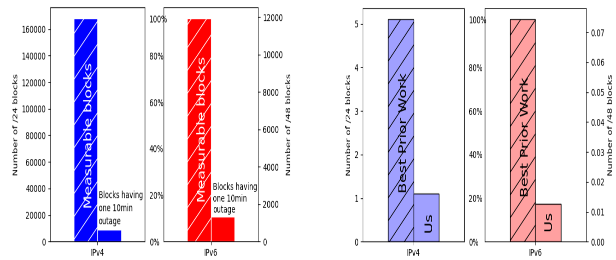


Figure 1: Trading temporal and spatial precision.



(a) Outage **(b) Coverage**

Figure 2: IPv6 coverage is similar to IPv4

passive systems are often homogeneous, using the same parameters across all block and therefore providing only coarse spatial coverage (at the country or AS level [4], 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.

In Figure 1 we evaluate this trade-off, showing that we have good precision for the dense blocks and less precision for the sparse blocks. We can observe almost 90% of B-root blocks in one day including sparse blocks if we take longer time bin which means less temporal precision. Therefore, the user can choose and get the best precision or coverage depending on the characteristics of the data.

IPv6: IPv6 is a growing part of the Internet today, and of course it has outages, but prior outage-detection systems have not extended to IPv6. Prior active monitoring systems cannot possibly probe all unicast IPv6 address, since 2^{128} addresses requires centuries to scan, and privacy-preserving addressing makes most client addresses ephemeral. Our new approach extends coverage to IPv6, by analyzing passive data, allowing the active addresses to come to us.

We evaluate our IPv6 coverage based on one representative day of passive data from B-root, comparing results in IPv4 and IPv6. In Figure 2a we see 11,918 /48 IPv6 blocks that are measurable (they have enough data to provide a reliable outage signal), and we see at least one 10 minute outages in 1338 (12% of measurable blocks). By comparison, the same system sees 167,851 /24 IPv4 blocks that are measurable, and 8689 /24 IPv4 blocks have one 10 minute outage (5.5% of measurable blocks). The absolute number of IPv4 outages is larger than IPv6 because there are far more measurable IPv4 blocks. However, the outage *rate* (the percentage measurable blocks with outages) for IPv6 seems somewhat greater than for IPv4, suggesting IPv6 reliability can improve.

Our coverage in IPv6 is surprisingly large: Figure 2b compares coverage relative to best prior system. For IPv6, our approach with B-root sees about 12,765 IPv6 /48 blocks, 17% of the 74,373 /48 blocks in the Gasser IPv6 hitlist [3]. This coverage is similar to what we see in IPv4, where the 1M /24s blocks our system with B-root is about 20% of the 5.1M in Trinocular. In both cases, B-root coverage is limited (it sees only recursive resolvers), but it seems about the same fraction of IPv6 as IPv4. We expect to add additional passive sources to increase IPv6 coverage.

Although our work is still in progress, our early results suggest a significant advance in the ability to observe both long and short outages of the network edge for both IPv4 and IPv6. We show that users can tradeoff between spatial and temporal precision depending on the type of data. We show our IPv6 provides coverage similar to our IPv4 coverage, and provide the first report IPv6 outages.

This work was supported by NSF grant CNS-2007106 (EIEIO).

REFERENCES

[1] Amazon down 15 minutes, loses over $66,000 per minute. In *https://smallbiztrends.com/2013/08/amazon-down-custom-error-page.html*, [Accessed on 17-Feb-2022].
[2] G. Aceto, A. Botta, P. Marchetta, V. Persico, and A. Pescapé. A comprehensive survey on internet outages. *J. Netw. Comput. Appl.*, 2018.
[3] O. Gasser, Q. Scheitle, P. Foremski, Q. Lone, M. Korczyński, S. D. Strowes, L. Hendriks, and G. Carle. Clusters in the expanse: understanding and unbiasing IPv6 hitlists. In *Proc. of the IMC 2018*, pages 364–378, 2018.
[4] A. Guillot, R. Fontugne, P. Winter, P. Merindol, A. King, A. Dainotti, and C. Pelsser. Chocolatine: Outage detection for internet background radiation. In *2019 TMA Conference*, pages 1–8. IEEE, 2019.
[5] T. Holterbach, E. C. Molero, M. Apostolaki, A. Dainotti, S. Vissicchio, and L. Vanbever. Blink: Fast connectivity recovery entirely in the data plane. In *16th USENIX Symposium on NSDI 19*, pages 161–176, 2019.
[6] L. Quan, J. Heidemann, and Y. Pradkin. Trinocular: Understanding Internet reliability through adaptive probing. In *Proc. ACM SIGCOMM Conference*, pages 255–266, Hong Kong, China, Aug. 2013. ACM.
[7] P. Richter, R. Padmanabhan, N. Spring, A. Berger, and D. Clark. Advancing the art of internet edge outage detection. In *Proc. of IMC*, pages 350–363, 2018.
[8] A. Shah, R. Fontugne, E. Aben, C. Pelsser, and R. Bush. Disco: Fast, good, and cheap outage detection. In *2017 TMA Conference*. IEEE, 2017.
[9] USC. B-Root DNS service. https://b.root-servers.org/.

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)

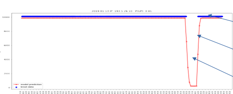
If there is a packet in the bin, then

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

else

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

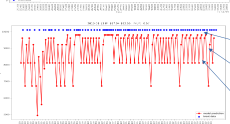
- To show that we can adapt and cover the whole range lets us show examples are two extreme cases of dense and sparse blocks



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

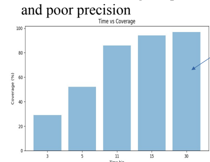
- We maximize the number of true outages detection.

Observation (B-root)	**Ground truth (Trinocular)** 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?

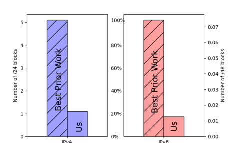
- Detects outages with short length unlike prior works.
- We study outages that are 5 minutes in length for both Broot
- and RIPE data on January 10, 2019.
- In RIPE data, there is status which defines if an address is available or out.

Observation (B-root)	**Ground truth (RIPE)** 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.

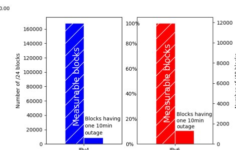
Extending to IPv6



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..