

On the Constancy of Internet Path Properties, Opportunistic Measurement: Extracting Insight from Spurious Traffic

Madhav Suresh

Northwestern University, Department of Computer Science

January 18, 2013

Motivation

- "Measurements are most valuable when they are a useful guide to the future."

Motivation

- "Measurements are most valuable when they are a useful guide to the future."
 - This happens when there is constancy

Motivation

- "Measurements are most valuable when they are a useful guide to the future."
 - This happens when there is constancy
- Three types of Constancy

Motivation

- "Measurements are most valuable when they are a useful guide to the future."
 - This happens when there is constancy
- Three types of Constancy
 - Mathematical

Motivation

- "Measurements are most valuable when they are a useful guide to the future."
 - This happens when there is constancy
- Three types of Constancy
 - Mathematical
 - Operational

Motivation

- "Measurements are most valuable when they are a useful guide to the future."
 - This happens when there is constancy
- Three types of Constancy
 - Mathematical
 - Operational
 - Predictive

Mathematical Constancy

- Can be described with a single time-invariant mathematical model

Mathematical Constancy

- Can be described with a single time-invariant mathematical model
- Finding appropriate model is the challenge and key

Mathematical Constancy

- Can be described with a single time-invariant mathematical model
- Finding appropriate model is the challenge and key
- However, lack of mathematical constancy does not imply that applications care

Mathematical Constancy

- Can be described with a single time-invariant mathematical model
- Finding appropriate model is the challenge and key
- However, lack of mathematical constancy does not imply that applications care

Operational Constancy!!!!!!!!!!

- Whether or not applications care about the changes in the dataset

Operational Constancy!!!!!!!!!!

- Whether or not applications care about the changes in the dataset
- "Quantities of interest remain within bounds considered operationally equivalent"

Operational Constancy!!!!!!!!!!

- Whether or not applications care about the changes in the dataset
- "Quantities of interest remain within bounds considered operationally equivalent"
- The degree to which the network remains in a particular operating regime

Predictive Constancy

- Past measurements allow one to reasonably predict future characteristics

Predictive Constancy

- Past measurements allow one to reasonably predict future characteristics
- Reflects the degree to which changes in path properties can be tracked

Measurement Methodology

- NIMI measurement infrastructure

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80
 - \mathcal{W}_2 - Winter 2000-2001, 49 Hosts 73

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80
 - \mathcal{W}_2 - Winter 2000-2001, 49 Hosts 73
- University and research networks

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80
 - \mathcal{W}_2 - Winter 2000-2001, 49 Hosts 73
- University and research networks
- “we might plausibly argue that our observations could apply fairly well to the better connected commercial Internet of the not-too-distant future, if not today”

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80
 - \mathcal{W}_2 - Winter 2000-2001, 49 Hosts 73
- University and research networks
- “we might plausibly argue that our observations could apply fairly well to the better connected commercial Internet of the not-too-distant future, if not today”
 - lol

Measurement Methodology

- NIMI measurement infrastructure
- Two sets of Data, $\mathcal{W}_1, \mathcal{W}_2$
 - \mathcal{W}_1 - Winter 1999-2000, 31 Hosts 80
 - \mathcal{W}_2 - Winter 2000-2001, 49 Hosts 73
- University and research networks
- “we might plausibly argue that our observations could apply fairly well to the better connected commercial Internet of the not-too-distant future, if not today”
 - lol
- **The point of the paper was not to provide definitive results, but to provide a framework that could be used in the future**

Loss Constancy

- “Loss process are better thought of as spikes”

Loss Constancy

- “Loss process are better thought of as spikes”
- Loss episode rates are mathematically non-constant

Loss Constancy

- “Loss process are better thought of as spikes”
- Loss episode rates are mathematically non-constant
- Operational constancy coincides with mathematical constancy

Loss Constancy

- “Loss process are better thought of as spikes”
- Loss episode rates are mathematically non-constant
- Operational constancy coincides with mathematical constancy
- Predictive Constancy performs well independent of other constancy

Loss Constancy

- “Loss process are better thought of as spikes”
- Loss episode rates are mathematically non-constant
- Operational constancy coincides with mathematical constancy
- Predictive Constancy performs well independent of other constancy
- All the fancy schmancy math is introduced here.

Delay Constancy

- Zing poisson packet streams on NIMI hosts

Delay Constancy

- Zing poisson packet streams on NIMI hosts
- Spikes vs. non spikes

Delay Constancy

- Zing poisson packet streams on NIMI hosts
- Spikes vs. non spikes
- Mathematically less steady than loss

Delay Constancy

- Zing poisson packet streams on NIMI hosts
- Spikes vs. non spikes
- Mathematically less steady than loss
- Operationally not steady

Delay Constancy

- Zing poisson packet streams on NIMI hosts
- Spikes vs. non spikes
- Mathematically less steady than loss
- Operationally not steady
- Predictively steady

Delay Constancy

- Zing poisson packet streams on NIMI hosts
- Spikes vs. non spikes
- Mathematically less steady than loss
- Operationally not steady
- Predictively steady
- Is this surprising?

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period
 - Much different these days for commercial broadband users (speedboost, fios, andy)

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period
 - Much different these days for commercial broadband users (speedboost, fios, andy)
 - Also these measurements were taken on well provisioned networks

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period
 - Much different these days for commercial broadband users (speedboost, fios, andy)
 - Also these measurements were taken on well provisioned networks
- No simple relationship between operational and mathematical constancy

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period
 - Much different these days for commercial broadband users (speedboost, fios, andy)
 - Also these measurements were taken on well provisioned networks
- No simple relationship between operational and mathematical constancy
- Given parameters, predictive constancy is good.

Throughput Constancy

- Throughput constancy didn't change minute to minute, only over long time period
 - Much different these days for commercial broadband users (speedboost, fios, andy)
 - Also these measurements were taken on well provisioned networks
- No simple relationship between operational and mathematical constancy
- Given parameters, predictive constancy is good.
 - This makes sense if bandwidth is generally continuous, is this the case?

Conclusions

- They introduced notions of constancy.

Conclusions

- They introduced notions of constancy.
- helps to interpret and contextualize measurement dataset

Conclusions

- They introduced notions of constancy.
- helps to interpret and contextualize measurement dataset
- paper cited 432 times, seemingly relevant

Conclusions

- They introduced notions of constancy.
- helps to interpret and contextualize measurement dataset
- paper cited 432 times, seemingly relevant
- Their dataset was weak, but their methods for interpreting are what's important

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard
- Why not leverage existing traffic sources?

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard
- Why not leverage existing traffic sources?
 - worm probes

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard
- Why not leverage existing traffic sources?
 - worm probes
 - DDoS backscatter

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard
- Why not leverage existing traffic sources?
 - worm probes
 - DDoS backscatter
 - botnet scans

Opportunistic Measurement: Extracting Insight from Spurious Traffic

- Measuring on the edge is really hard
- Why not leverage existing traffic sources?
 - worm probes
 - DDoS backscatter
 - botnet scans
 - spam floods

Some Results

- Things they were able to infer:

Some Results

- Things they were able to infer:
 - Access Link bandwidth

Some Results

- Things they were able to infer:
 - Access Link bandwidth
 - number of attached disks

Some Results

- Things they were able to infer:
 - Access Link bandwidth
 - number of attached disks
 - uptime