Empower. Partner. Lead.

Ohio Supercomputer Center

Modeling of Multi-resolution Active Network Measurement Time-series

Prasad Calyam, Ph.D.
pcalyam@osc.edu

Ananth Devulapalli
ananth@osc.edu

Third IEEE Workshop on Network Measurements

October 14th 2008

Topics of Discussion

- Background
- Measurement Data Sets
- Time Series Analysis Methodology
- Results Discussion
- Conclusion and Future Work

Background

- Internet ubiquity is driving common applications to be network-dependent
  - Office (e.g. Videoconferencing), Home (e.g. IPTV), Research (e.g. Grid)
- ISPs monitor end-to-end *Network Quality of Service* (QoS) for supporting existing and emerging applications
  - Network QoS metrics: bandwidth, delay, jitter, loss
  - Active measurement tools: Ping, Traceroute, Iperf, Pathchar, …
    - Inject test packets into the network to measure performance
- Collected active measurements are useful in network control and management functions
  - E.g., Path switching or Bandwidth on-demand – based on network performance anomaly detection and network weather forecasting

Challenges in using Active Measurements

- High variability in measurements
  - Variations manifest as short spikes, burst spikes, plateaus
  - Causes: user patterns, network fault events, cross-traffic congestion
- Missing data points or gaps are not uncommon
  - Compound the measurement time-series analysis
  - Causes: network equipment outages, measurement probe outages
- Measurements need to be modeled at multi-resolution timescales
  - Forecasting period is comparable to sampling period
    - E.g., Long-term forecasting for bandwidth upgrades
  - Troubleshooting bottlenecks at timescales of network events
    - E.g., Anomaly detection for problems with plateaus, and periodic bursts

Our goals

- Address the challenges and requirements in modeling multi-resolution active network measurements
  - Analyze measurements collected using our ActiveMon framework that is being used to monitor our state-wide network viz., OSCnet
  - Develop analysis techniques in ActiveMon to improve prediction accuracy and lower anomaly detection false-alarms
- Use Auto-Regressive Integrated Moving Average (ARIMA) class of models for analyzing the active network measurements
  - Many recent works have suggested suitability for modeling network performance variability
    - Zhou et al., combined ARIMA models with non-linear time-series models to improve prediction accuracy
    - Shu et al., showed seasonal ARIMA models can predict performance of wireless network links
  - We evaluate impact of multi-resolution timescales due to absence and presence of network events on ARIMA model parameters

Topics of Discussion

- Background
- Measurement Data Sets
- Time Series Analysis Methodology
- Results Discussion
- Conclusion and Future Work

ActiveMon Measurements

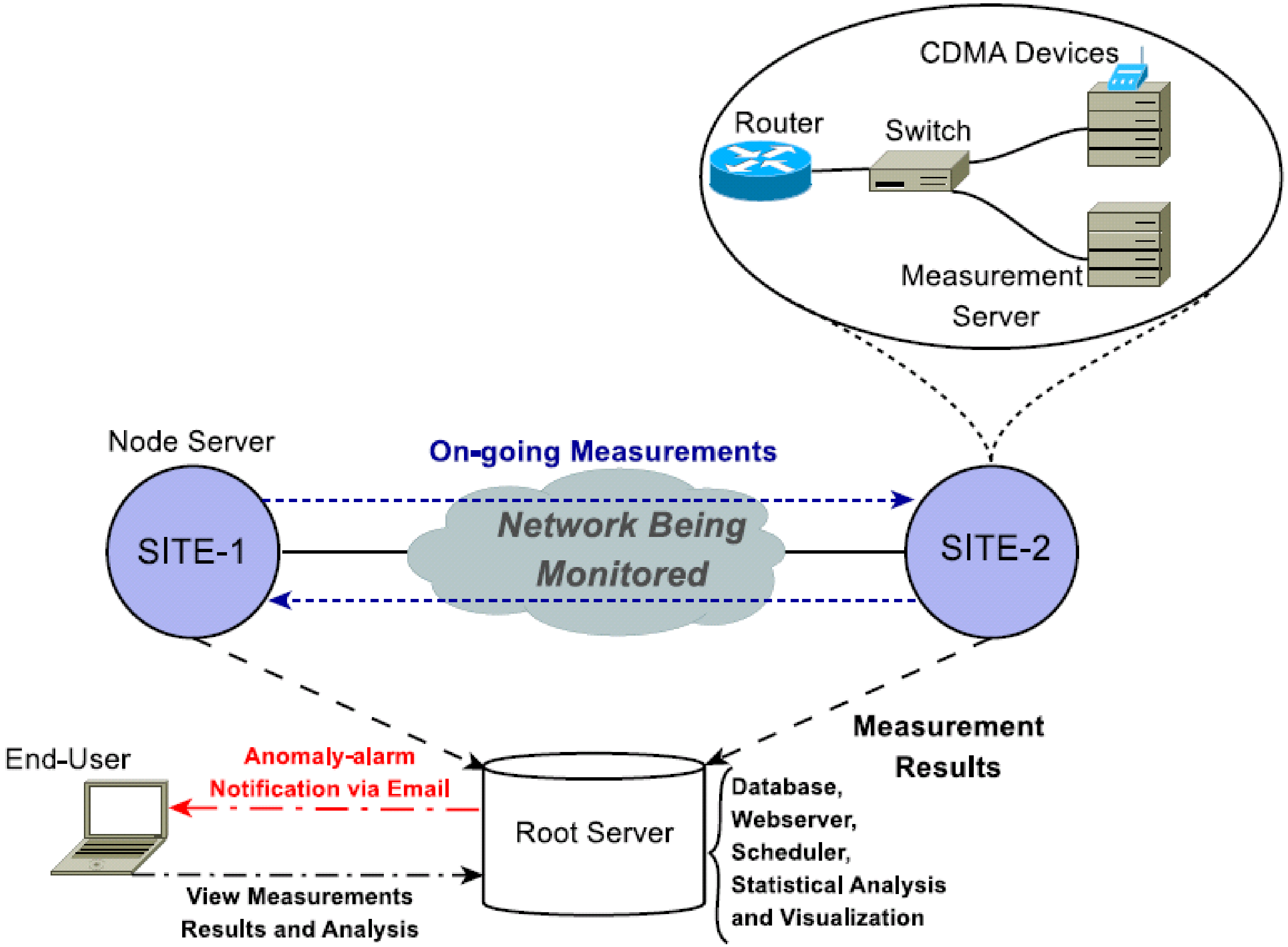
- We collected a large data set of active measurements for over 6 months on three hierarchically different Internet backbone paths
  - Campus path on The Ohio State Uni. (OSU) campus backbone
  - Regional path between OSU and Uni. of Cincinnati (UC) on OSCnet
  - National path between OSU and North Carolina State Uni. (NCSU)
- Used in earlier studies
  - How active measurements correlate to network events?

P. Calyam, D. Krymskiy, M. Sridharan, P. Schopis, "TBI: End-to-End Network Performance Measurement Testbed for Empirical-bottleneck Detection", *IEEE TRIDENTCOM*, 2005.

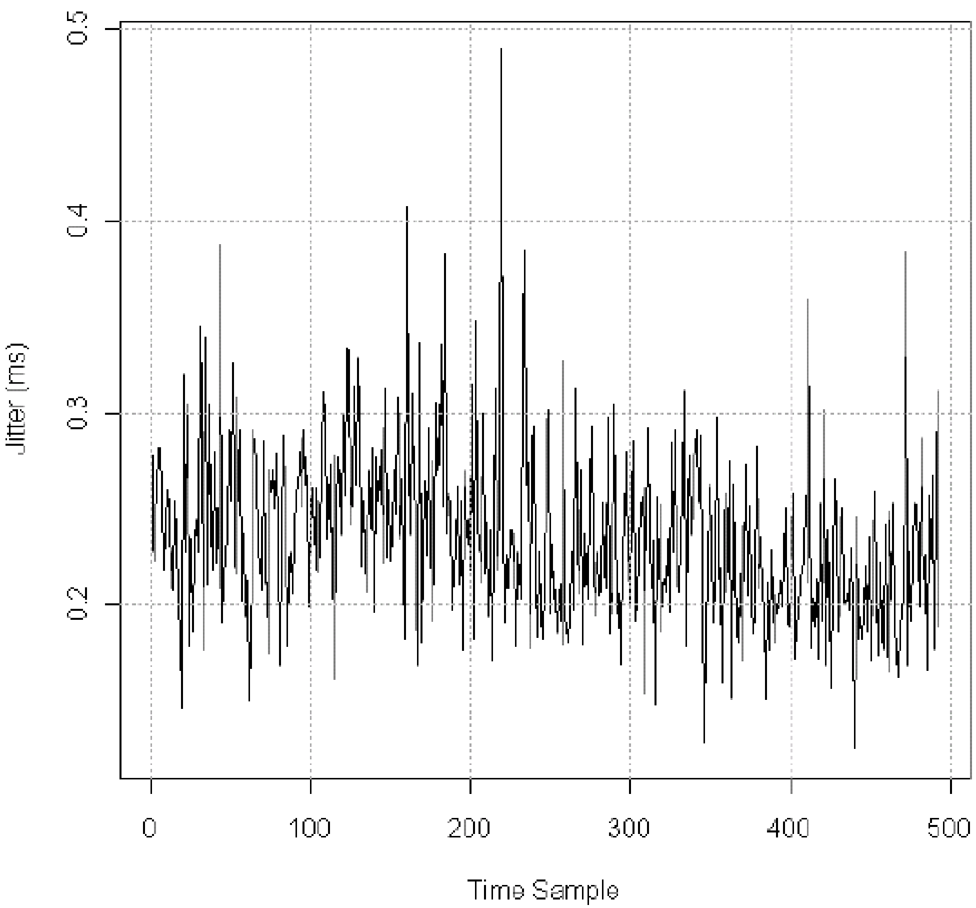
  - How long-term trends of active measurements compare on hierarchical network paths?

P. Calyam, D. Krymskiy, M. Sridharan, P. Schopis, "Active and Passive Measurements on Campus, Regional and National Network Backbone Paths", *IEEE ICCCN*, 2005.

OSC ActiveMon Setup

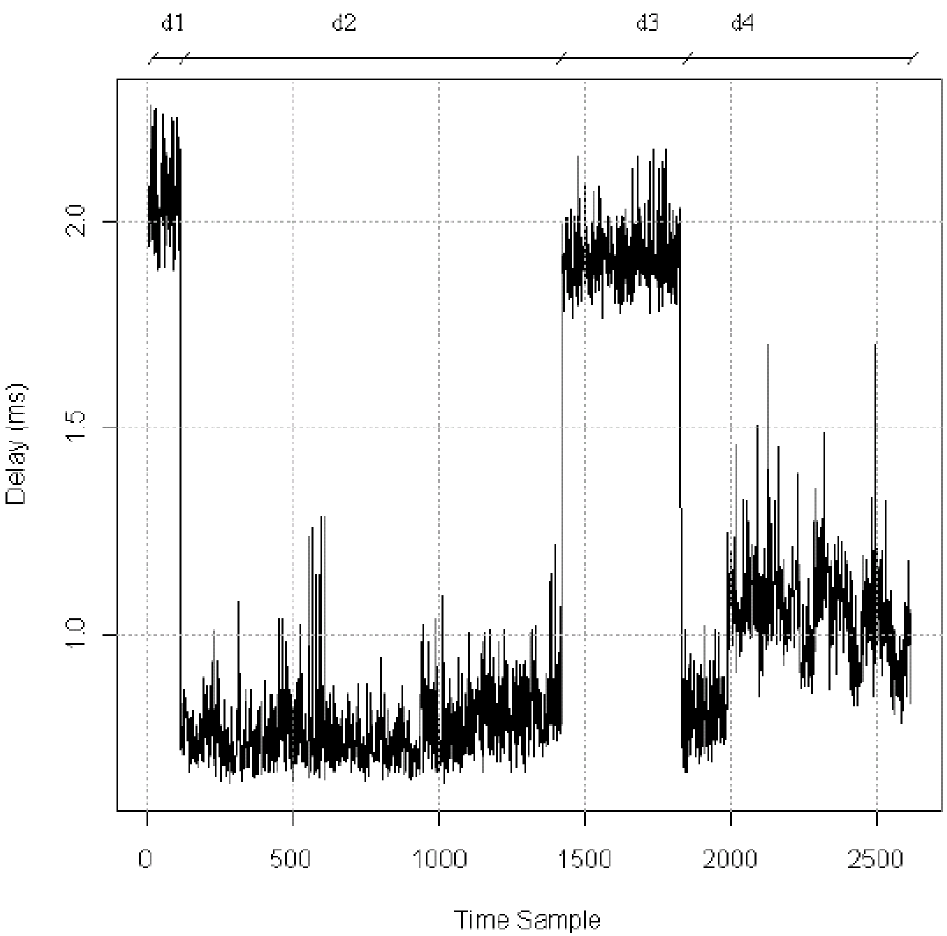


“Routine” Jitter Measurement Data Set



- Collected between OSU and UC border routers
- Iperf tool measurements over a two-month period
- Iperf probing comprised of UDP traffic at 768 Kbps
- NOC logs indicate no major network events during the two-month period

"Event-laden" Delay Measurement Data Set

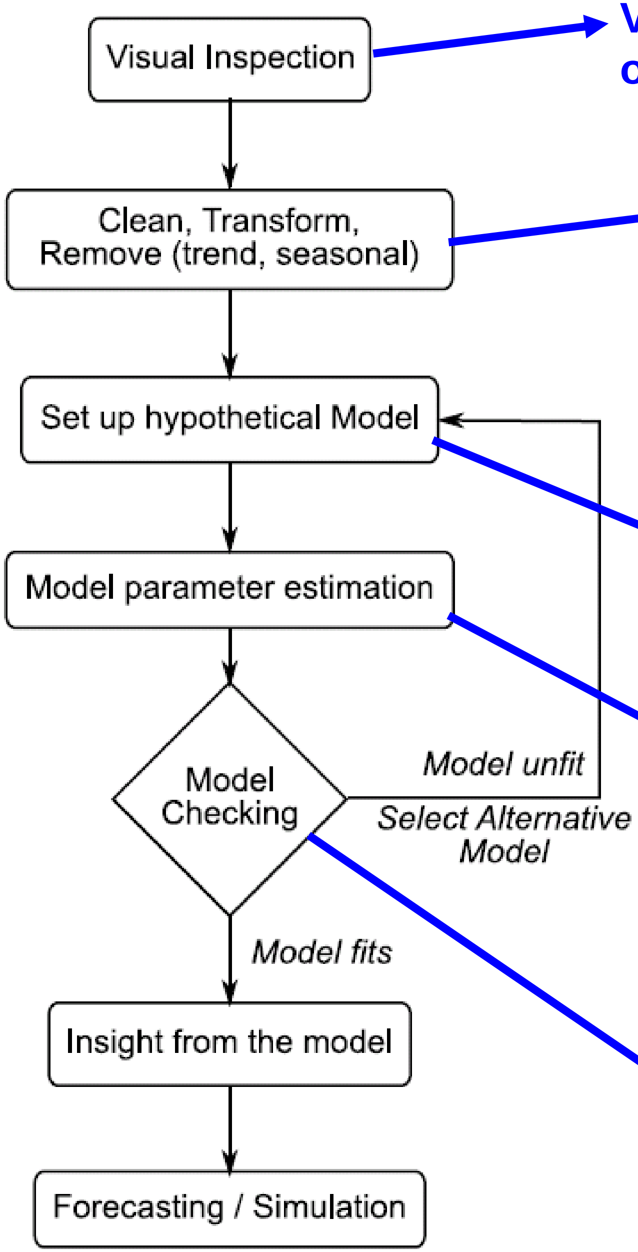


- Collected between OSU border and OSU CS Dept. routers
- Ping tool measurements over a six-month period
- Ping probing comprised of four 32 byte ICMP packets
- NOC logs indicate four route-changes due to network management activities

Topics of Discussion

- Background
- Measurement Data Sets
- Time-series Analysis Methodology
- Results Discussion
- Conclusion and Future Work

Classical Decomposition (Box-Jenkins) Procedure



Verify presence of any seasonal or time-based trends

Achieve data stationarity using techniques such as “Differencing” where you difference consecutive data points up to *N-lag*

$$\Delta X_t = X_t - X_{t-1}$$
$$\Delta X_t = (1-L)X_t$$
$$\Delta^i X_t = (1-L)^i X_t$$

Use sample Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to see if the data follows Moving Average (MA) or Auto-regressive (AR) process, respectively

$ARIMA(p,d,q)$

- ***p* – *MA* order**
- ***d* – Differencing order**
- ***q* – *AR* order**

“Goodness of Fit” tests (e.g., Akaike Information Criterion) on the selected model parameters to find model fits that are statistically significant

Two-phase Analysis Approach

- Separate each data set into two parts:
  1. Training data set
     - Perform time-series analysis for model parameters estimation
  2. Test data set
     - Verify forecasting accuracy of selected model parameters to confirm model fitness

- Routine jitter measurement data set observations
  - Total: 493; Training: 469; Test: 24

- Event-laden delay measurement data set observations
  - Total: 2164; Training: 2100; Test: 64

Topics of Discussion

- Background
- Measurement Data Sets
- Time Series Analysis Methodology
- Results Discussion
- Conclusion and Future Work

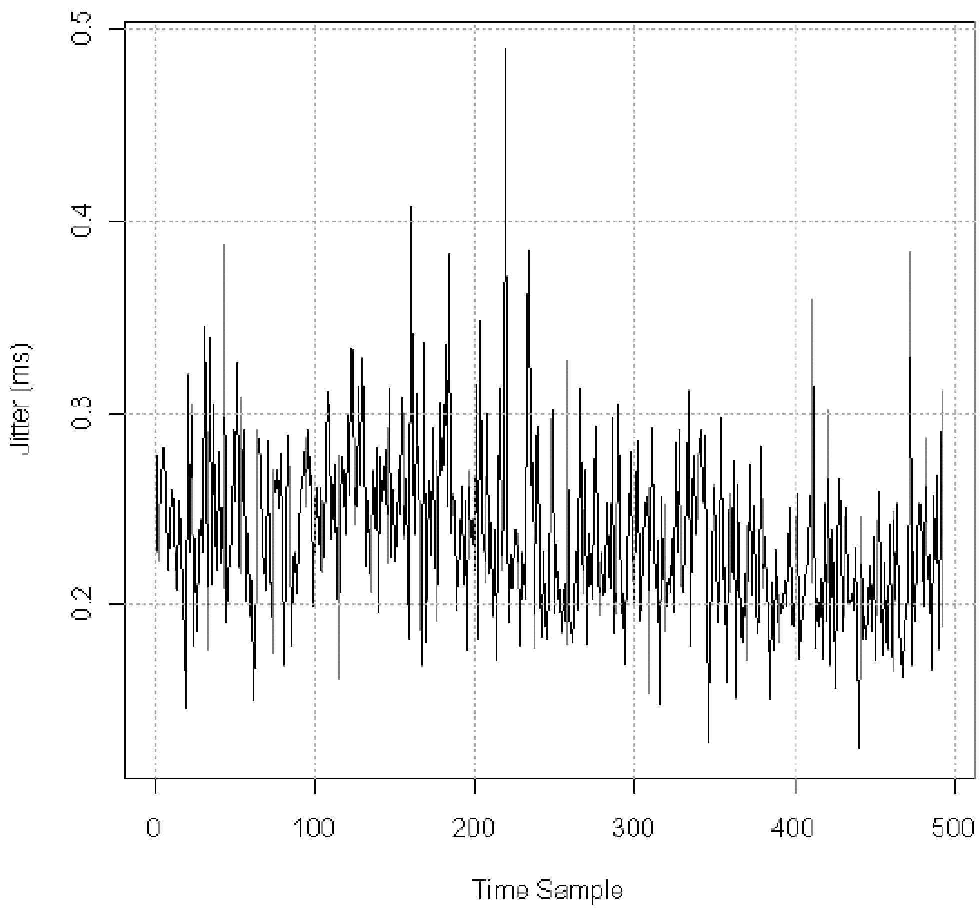
Results Discussion

- Part I: Time-series analysis of the routine jitter measurement data set
- Part II: Time-series analysis of the event-laden delay measurement data set
- Part III: "Parts versus Whole" time-series analysis of the two data sets

Results Discussion

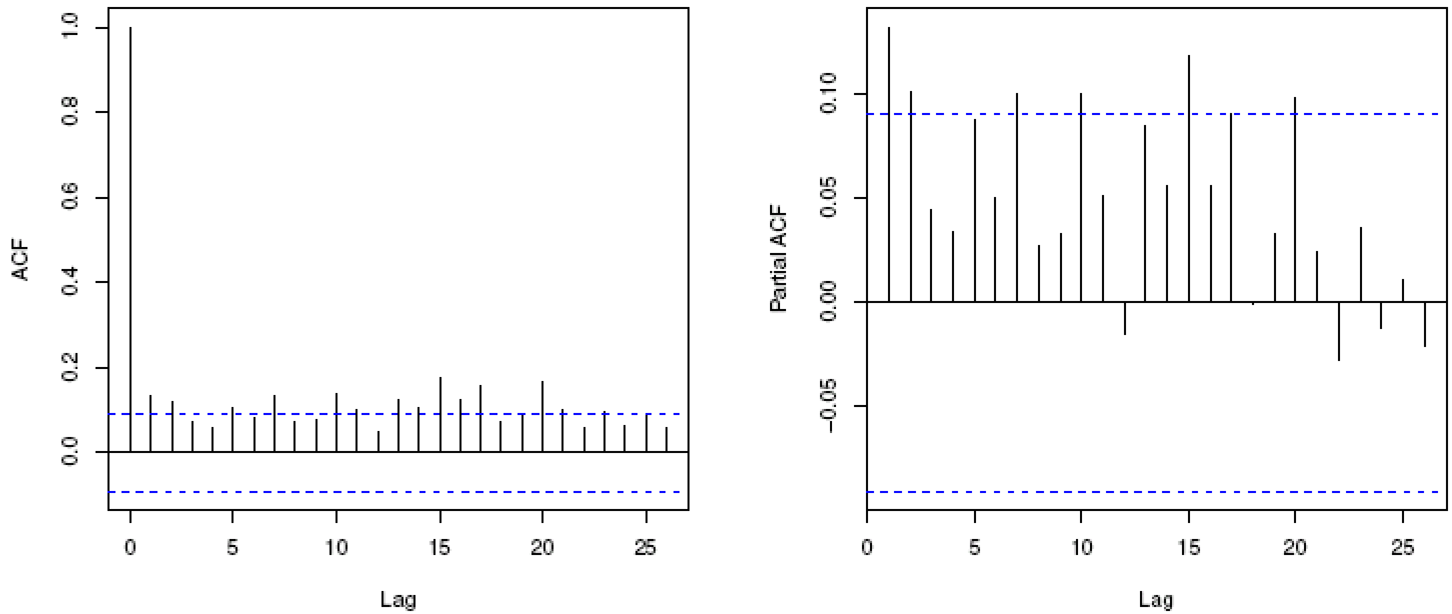
- Part I: Time-series analysis of the routine jitter measurement data set
- Part II: Time-series analysis of the event-laden delay measurement data set
- Part III: "Parts versus Whole" time-series analysis of the two data sets

Preliminary Data Examination



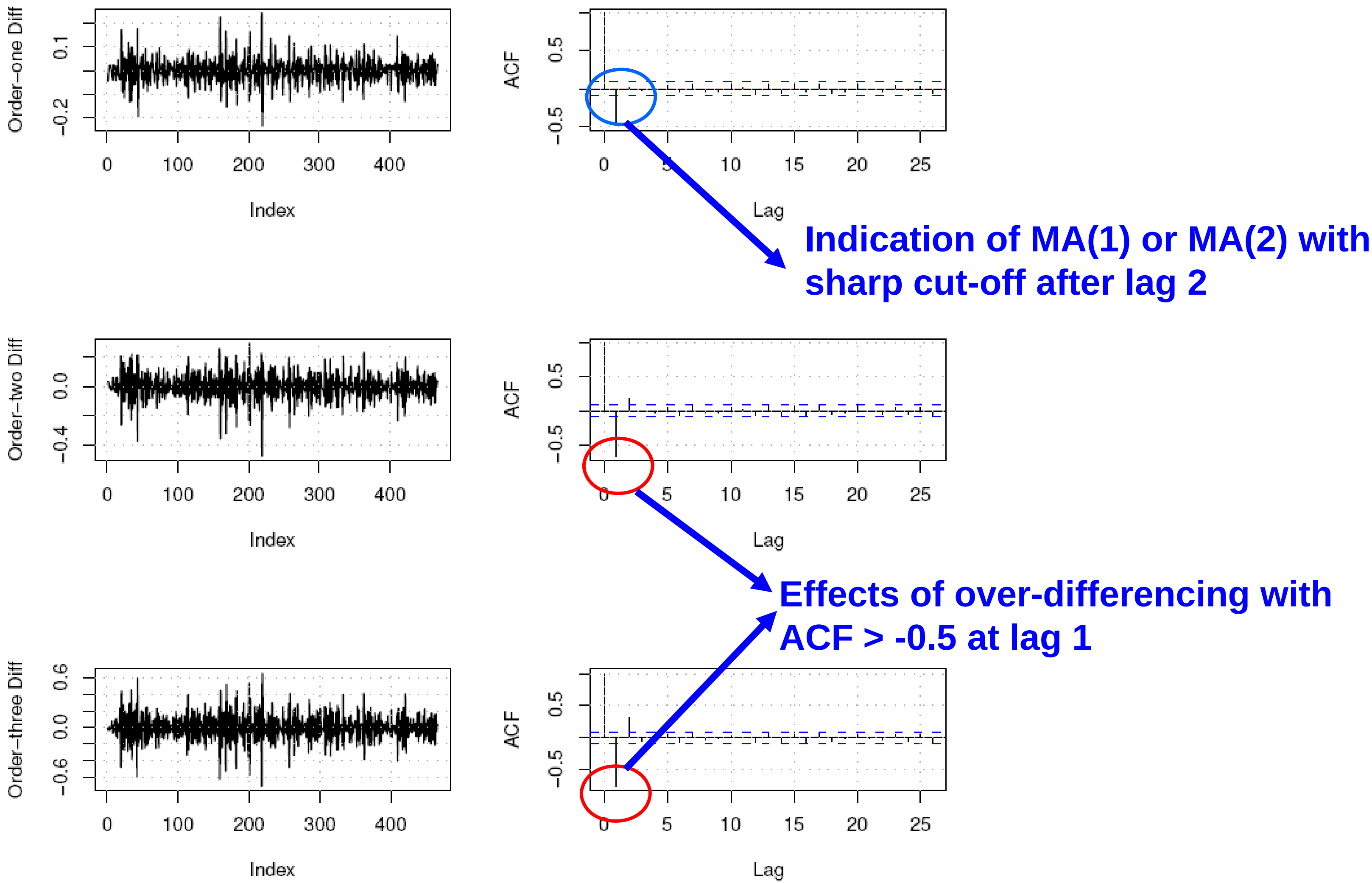
- No apparent trends or seasonality
- Frequent spikes and dips without any specific patterns

ACF and PACF

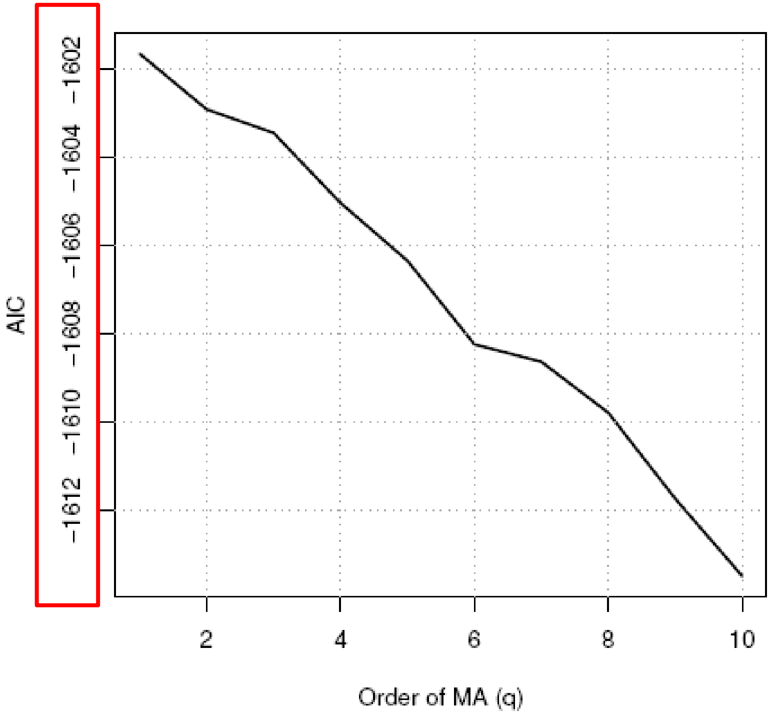


- ACF does not indicate MA
  - No clear cut-off at any lag; ACF is not decaying exponentially
- PACF does not indicate AR
  - PACF is not decaying exponentially
- Inherent trend in data present that is not visually noticeable

ACF after 1-Lag Differencing



Model Fitting

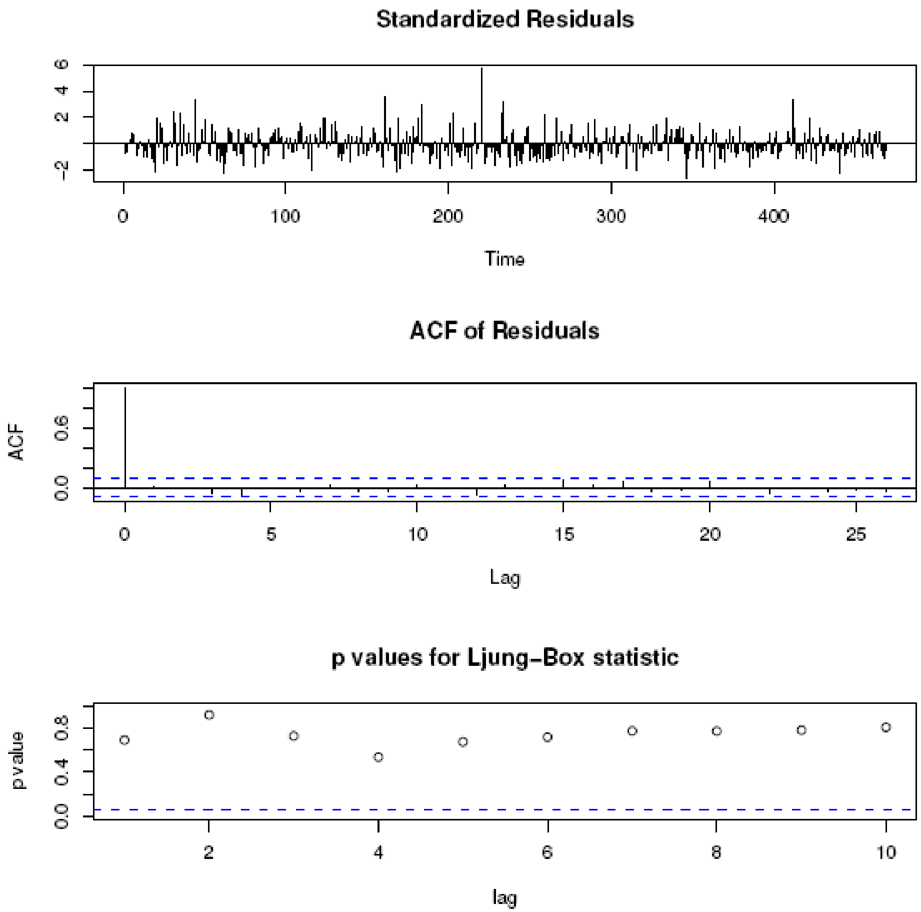


- To verify, we calculate AIC for increasing MA order and see MA(1) has minimum AIC
  - Dip in AIC is not notable for higher model orders i.e., for higher model complexity

Parameter	Value	95% CI
θ_1	-0.9440	$(-0.85286, -1.03514)$
θ_2	-0.0123	$(-0.135584, 0.110984))$
θ_3	-0.0114	$(-0.10744, 0.08464)$

- To further verify, we compare statistical significance of MA(1) parameter value i.e., θ_1 with higher order values θ_2 and θ_3
- We inspect whether 95% CI values $(\theta_x \pm 1.96 \times \sigma_{\theta_x})$ contain zero
  - 95% CI values of θ_1 are significant because they do not contain zero
  - Thus, we cannot reject the null hypothesis that MA(1) is not the suitable model

Diagnostic Checking of Fitted Model



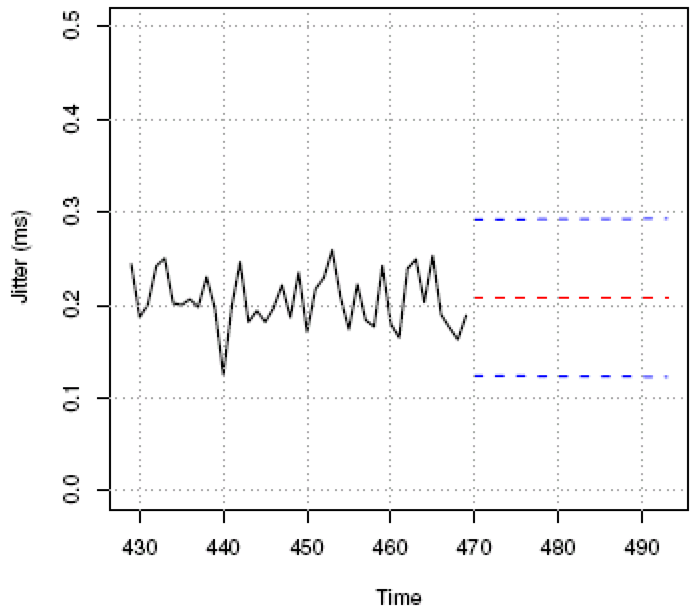
- Residuals look like noise process
- ACF of residuals resembles a white noise process
- Ljung-Box plot shows model is significant at all lags

Selected MA(1) Model

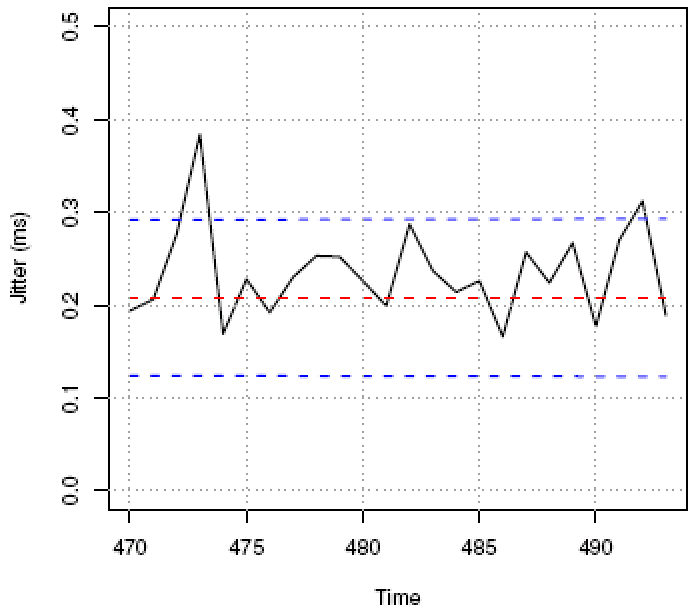
$$X_t = Z_t + (-0.9440)Z_{t-1}$$

$$\text{where } Z_i \sim \text{White Noise}(0, 0.01028) \text{ and } X_i = \text{diff}(jitter_i)$$

Prediction Based on MA(1) Model Fitting



(a) Training Data with MA(1) Prediction CI



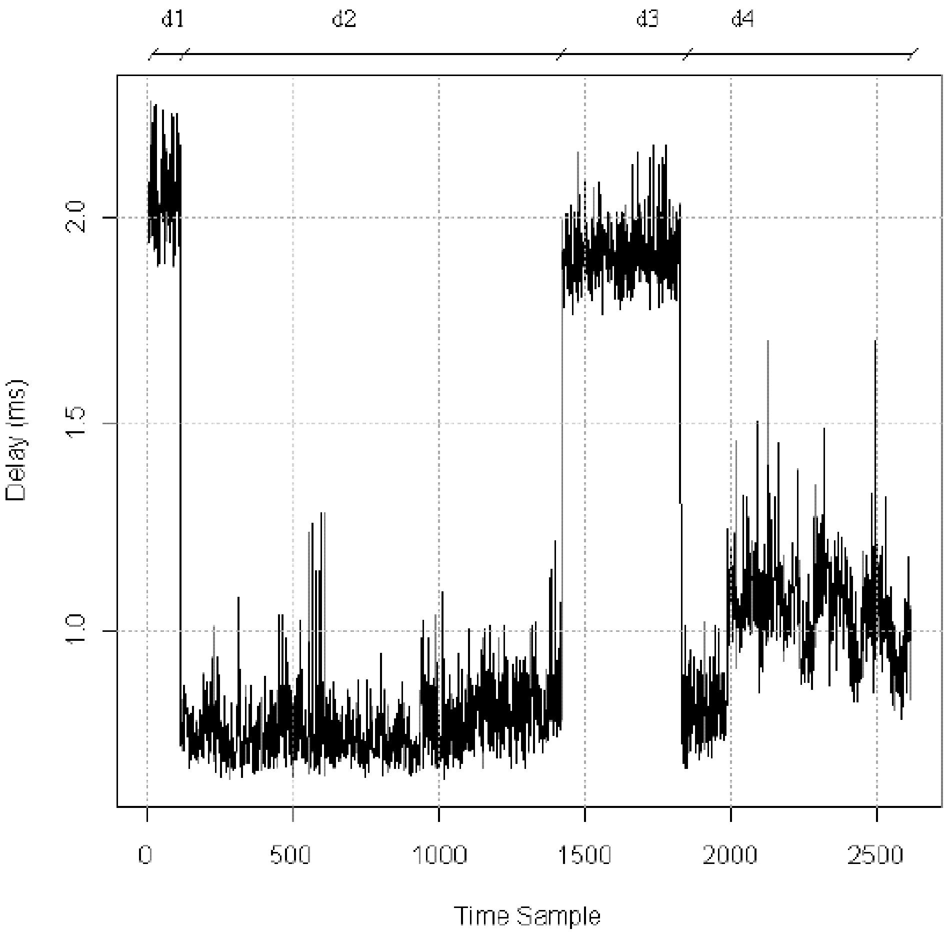
(b) Test Data with MA(1) Prediction CI

- Model prediction is close to reality
  - Most of the test data, except couple of observations, fall within the MA(1) Prediction CI

Results Discussion

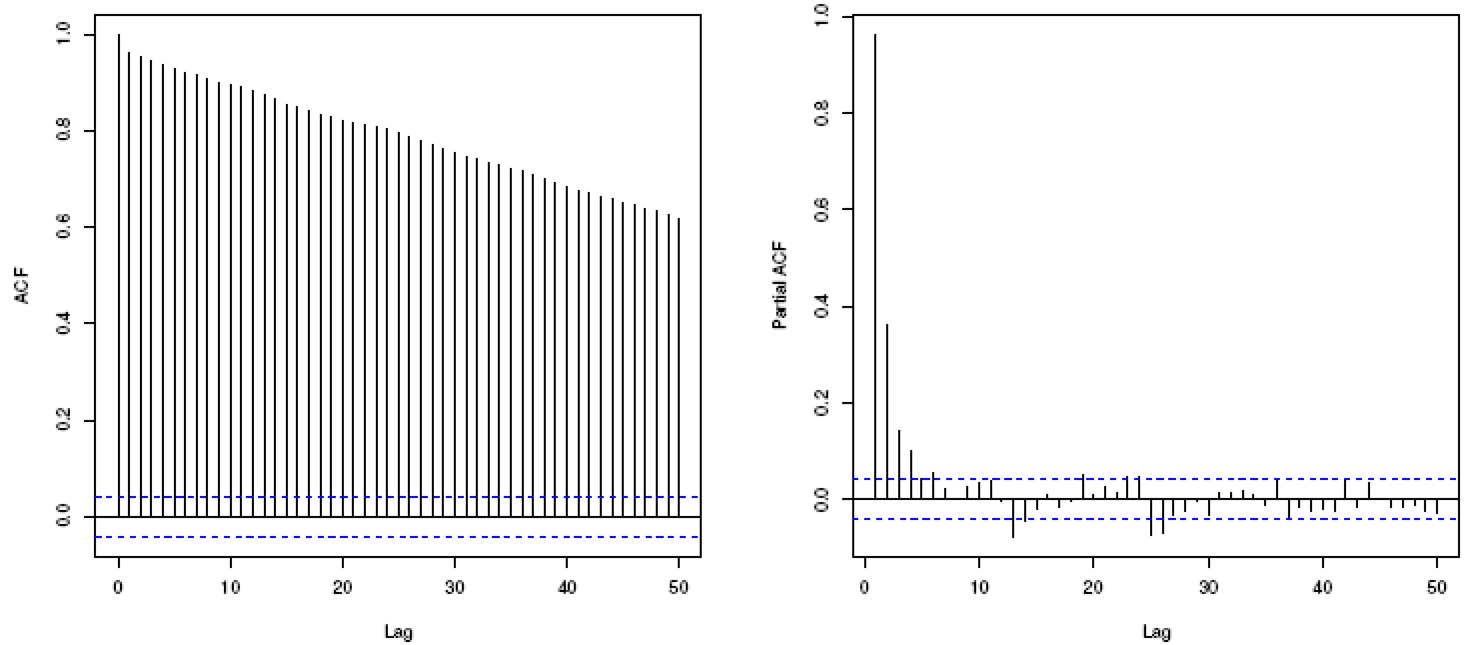
- Part I: Time-series analysis of the routine jitter measurement data set
- Part II: Time-series analysis of the event-laden delay measurement data set
- Part III: "Parts versus Whole" time-series analysis of the two data sets

Preliminary Data Examination



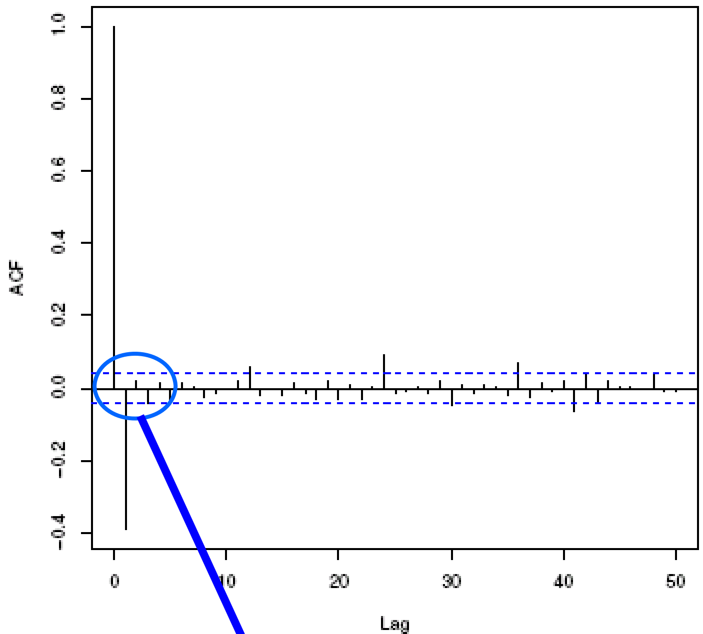
- Four distinct plateaus due to network route changes
- Frequent spikes and dips within each plateau without any specific patterns

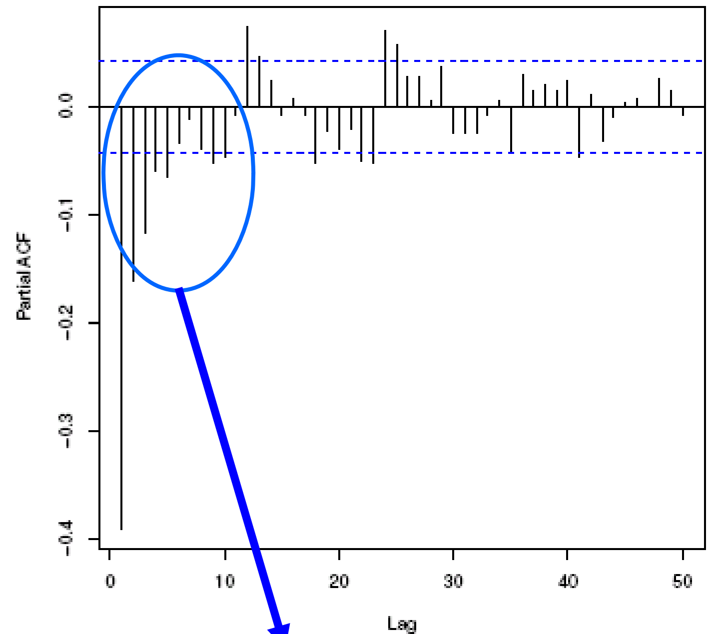
ACF and PACF



- ACF does not indicate MA
  - No clear cut-off at any lag; ACF is not decaying exponentially
- PACF indicates possibility of AR
  - PACF is decaying exponentially
- Inherent trend in data present that is not visually noticeable

ACF and PACF after 1-Lag Differencing

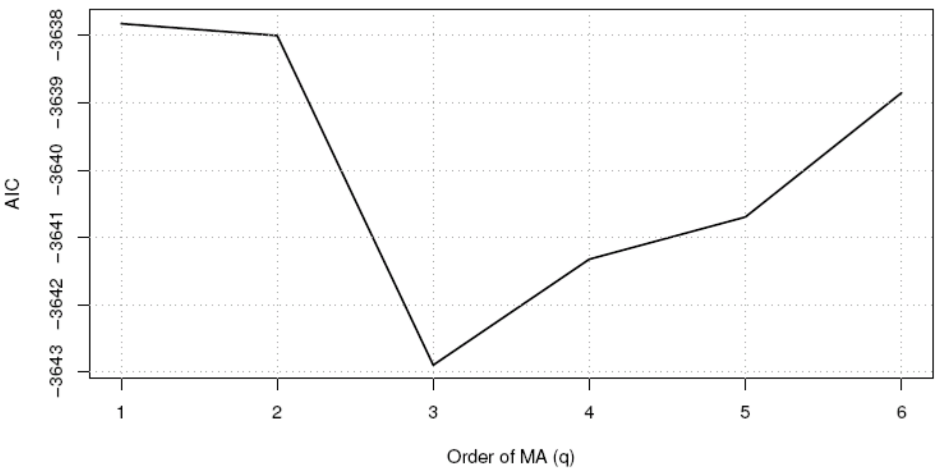




Damping pattern eliminates AR possibility

Indication of MA(1) or MA(2) with sharp cut-off after lag 2

Model Fitting

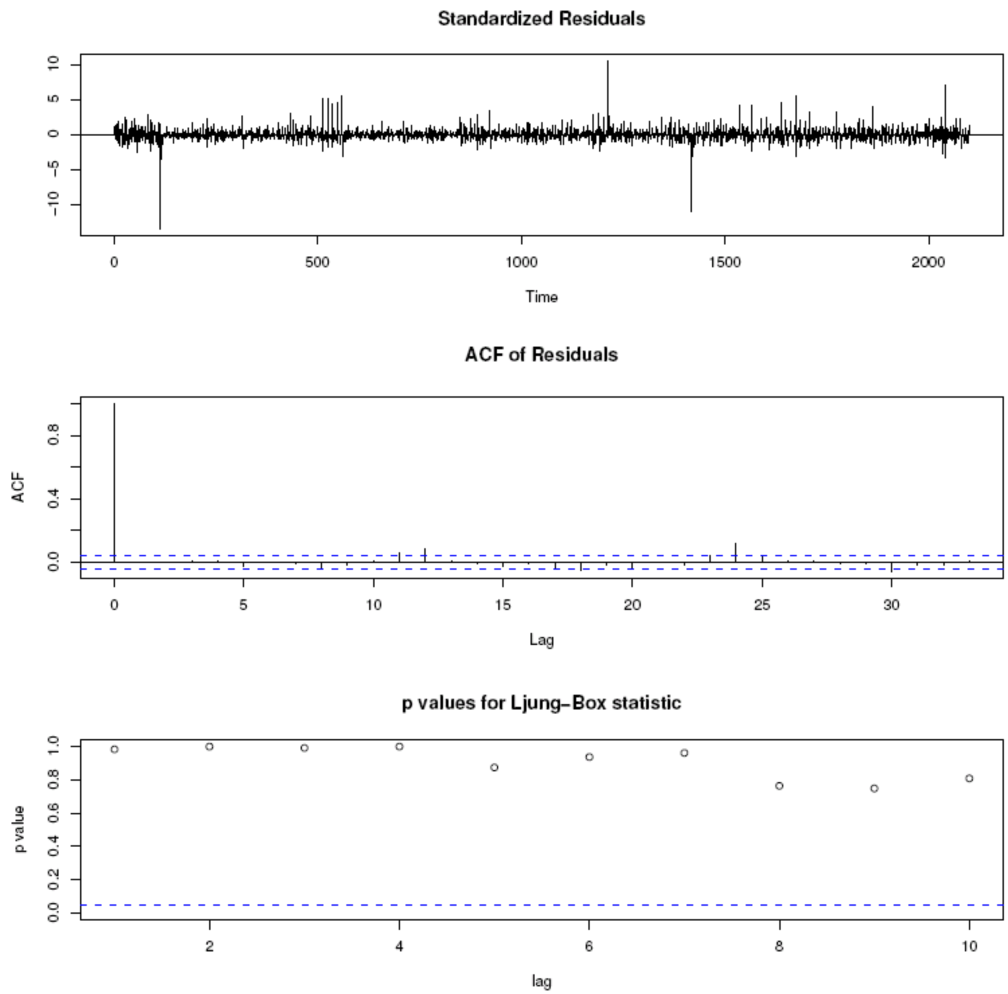


- To verify, we calculate AIC for increasing MA order and clearly see MA(3) has minimum AIC

Parameter	Value	95% Conf. Interval
θ_1	-0.4876	$(-0.5303, -0.4449)$
θ_2	-0.0064	$(-0.0552, +0.0424)$
θ_3	-0.0564	$(-0.0983, -0.0145)$

- To further verify, we compare statistical significance of MA(3) parameter values i.e., θ_1, θ_2 and θ_3
- We inspect whether 95% CI values $(\theta_x \pm 1.96 \times \sigma_{\theta_x})$ contain zero
  - 95% CI values of θ_3 are significant because they do not contain zero
  - Thus, we cannot reject the null hypothesis that MA(3) is not the suitable model

Diagnostic Checking of Fitted Model



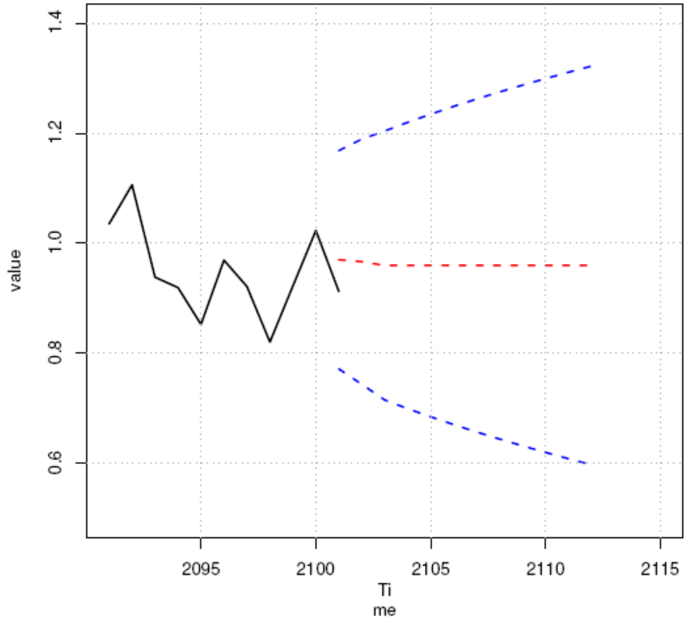
- Residuals look like noise process
- ACF of residuals resembles a white noise process
- Ljung-Box plot shows model is significant at all lags

Selected MA(3) Model

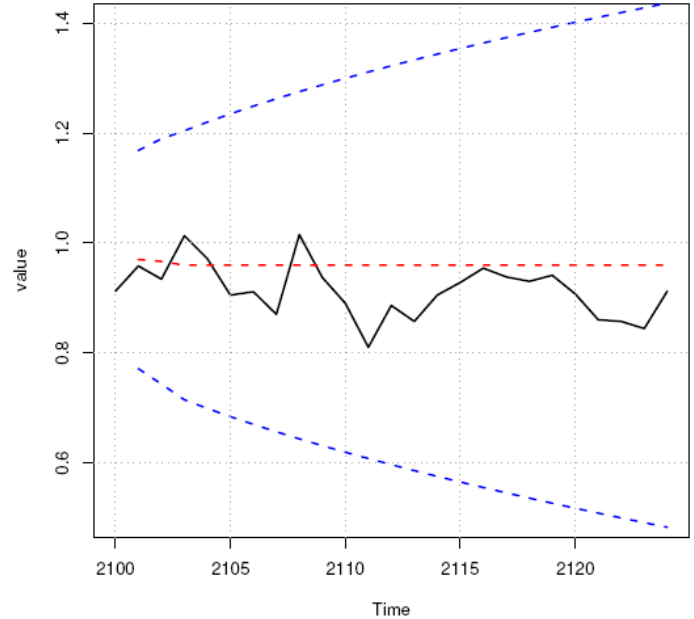
$$X_t = Z_t + (-0.4876)Z_{t-1} + (-0.0064)Z_{t-2} + (-0.0564)Z_{t-3}$$

$$\text{where } Z_i \sim \text{White Noise}(0, 0.01028) \text{ and } X_i = \text{diff}(delay_i)$$

Prediction Based on MA(3) Model Fitting



(a) Training Data with MA(1) Prediction CI



(b) Test Data with MA(1) Prediction CI

- Model prediction matches reality
  - All the test data fall within the MA(3) Prediction CI

Results Discussion

- Part I: Time-series analysis of the routine jitter measurement data set
- Part II: Time-series analysis of the event-laden delay measurement data set
- Part III: "Parts versus Whole" time-series analysis of the two data sets

“Parts Versus Whole” Time-series Analysis

- Routine jitter measurement data set
  - Split into two parts and ran Box-Jenkins analysis on each part
  - Both parts exhibited MA(1) process
- Event-laden delay measurement data set
  - Split into four parts, separated by the plateaus viz., $d1$, $d2$, $d3$, $d4$ and ran Box-Jenkins analysis on each part
  - $d1$ and $d3$ exhibited MA(1) process; $d2$ and $d4$ exhibited AR(12) process

Topics of Discussion

- Background
- Measurement Data Sets
- Time Series Analysis Methodology
- Results Discussion
- Conclusion and Future Work

Conclusion

- We presented a systematic time-series modeling of multi-resolution active network measurements
  - Analyzed Routine and Event-laden data sets
- Although limited data sets were used, we found –
  - Variability in end-to-end network path performance can be modeled using ARIMA (0, 1, q) models, with low q values
  - End-to-end network path performance has “too much memory” and auto-regressive values that are dependent on present and past values may not be pertinent
  - 1-Lag differencing can remove visually non-apparent trends (jitter data set) and plateau trends (delay data set)
  - Parts resemble the whole in absence of plateau network events
  - Plateau network events cause underlying process changes

Future Work

- Apply similar methodology to:
  - Other ActiveMon data sets
  - Other group data sets (e.g., Internet2 perfSonar, SLAC IEPM-BW)
- Lower anomaly detection false-alarms in the plateau detector implementation in ActiveMon
  - Balance trade-offs in desired *sensitivity*, *trigger duration*, *summary window* <u>dynamically</u> based on the measured time-series

Thank you!