Using Forecasting to Predict Long-term Resource Utilization for Web Services
Daniel W. Yoas – Advisor: Dr. Gregory Simco

Previous Research
Predictions -
1. Focused on sampling up to one second.
2. Used for OS, web, and distrusted system scheduling or resource management.
3. Deteriorated quickly under multiple steps.
5. Maximized prediction time frames under fifteen minutes.
6. Used previous readings to predict the next value of a resource.
7. Focused on predicting an exact result.
8. Considered long-term resource utilization as chaotic.

This Research
Predictions -
1. Focused on sampling every ten seconds.
2. Used to predict resource utilization for web systems.
3. Only predicted a single step ahead.
4. Used basic time-series analysis: Naïve, Simple, and Exponential Moving Averages.
5. Examined the time-frames of hours, days, and weeks.
6. Used previous readings to predict a confidence interval and the next value of a resource.
7. Focused on a prediction range for resource utilization.
8. Considered long-term resource utilization as cyclical.

These Studies
Simulation –
1. Used Windows 2008 Server with IIS 7.0.
2. 22 Linux machines running SURGE.
3. Ran for 28 days.

Live Web Services –
1. 2 Clustered Web Servers provided samples.
2. The public made the requests.
3. Ran for 8 months.

Future Research
New Areas of Studies –
1. Are network servers, distributed systems, and virtual servers predictable?
2. Which seasons provide the best predictability?
3. Can a state machine be identified for “normal”, “abnormal”, and “failure”?
4. Do the outliers create a predictable pattern?
5. How do advanced time-series analyses perform when feedback is incorporated?