The Need For Robust Statistical Analysis of MANET Performance Data

Bradley Wilson, Isaac Porche, JR Lockwood

June 24, 2010

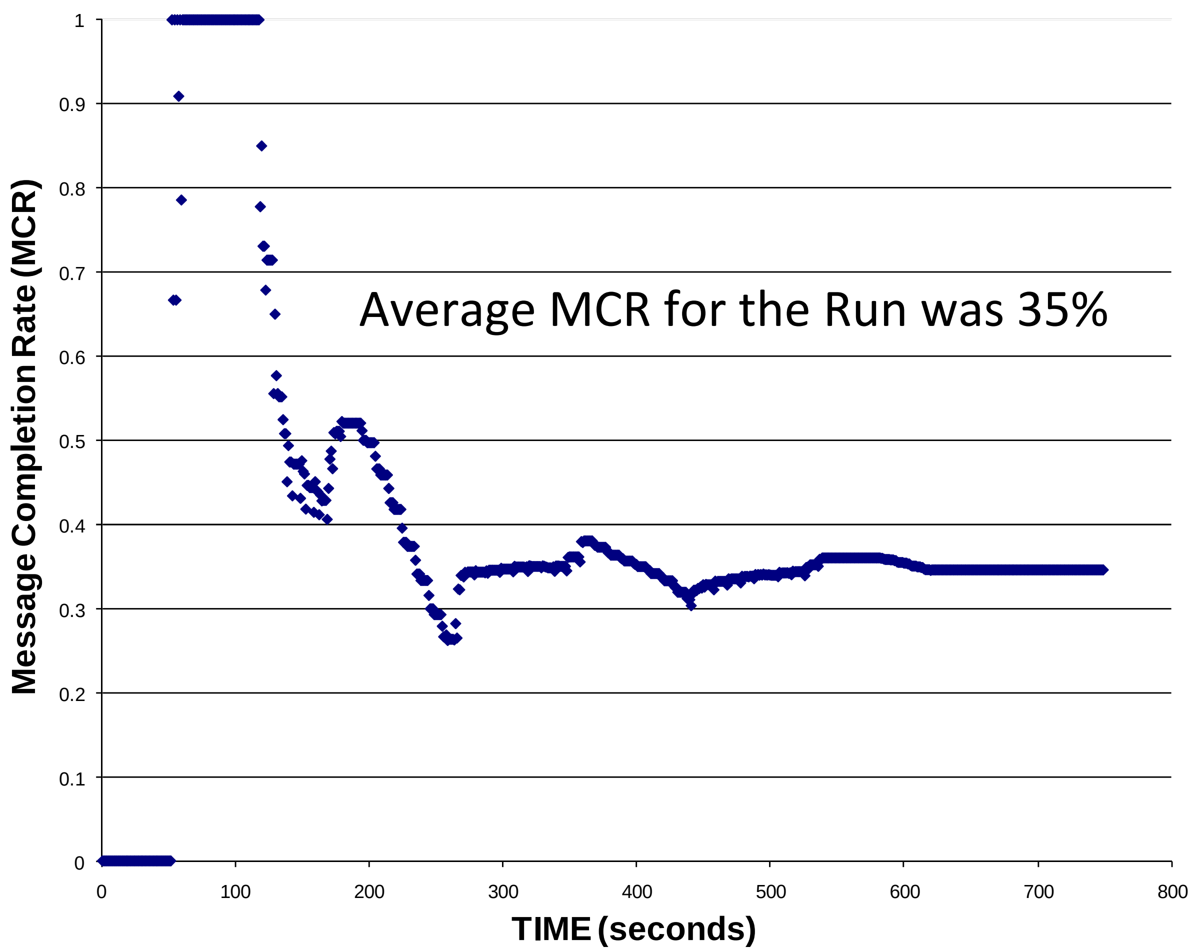
15th International Command and Control Research and Technology Symposium

Future Networks are Being Tested Today

- Our understanding of network performance depends on the quality of analysis
- The analysis can mean the difference between a functioning network and unrealistic expectations

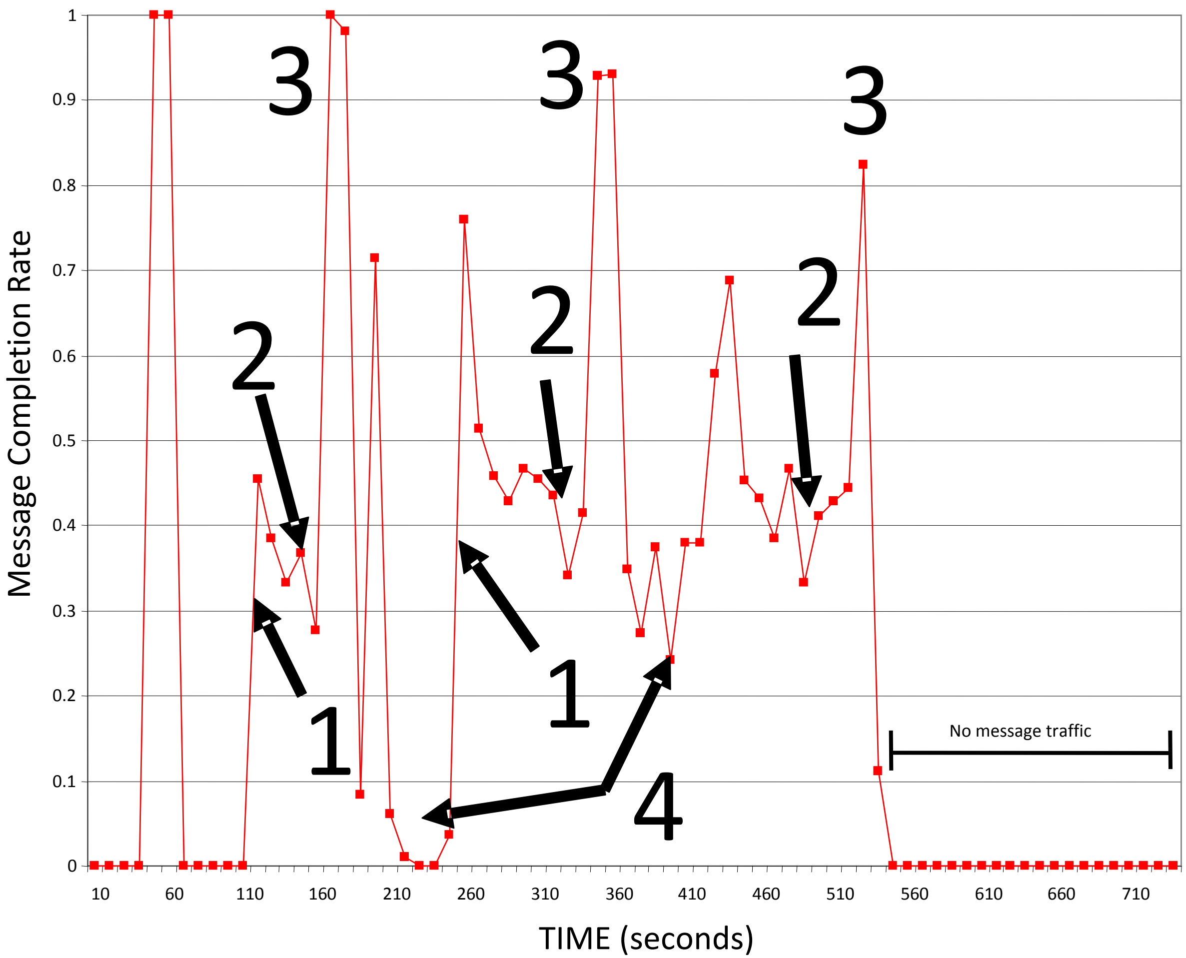
- Common metrics like Message Completion Rate (MCR) and Latency are useful, but potentially misleading because they conflate underlying variables
- We will offer several techniques to help drill down into network performance

Mean Data Can Be Misleading



Taken from "Integrating High Resolution Network Simulation with Force on Force Combat Models: Connecting MANA and QualNet", 10th ICCRTS, June 2005.

Network Throughput is Highly Dynamic



1 = UAV Moves Into Range

2 = Limited Performance Area

3 = Peak Performance Area

4 = Out of Range

Taken from "Integrating High Resolution Network Simulation with Force on Force Combat Models: Connecting MANA and QualNet", 10th ICCRTS, June 2005.

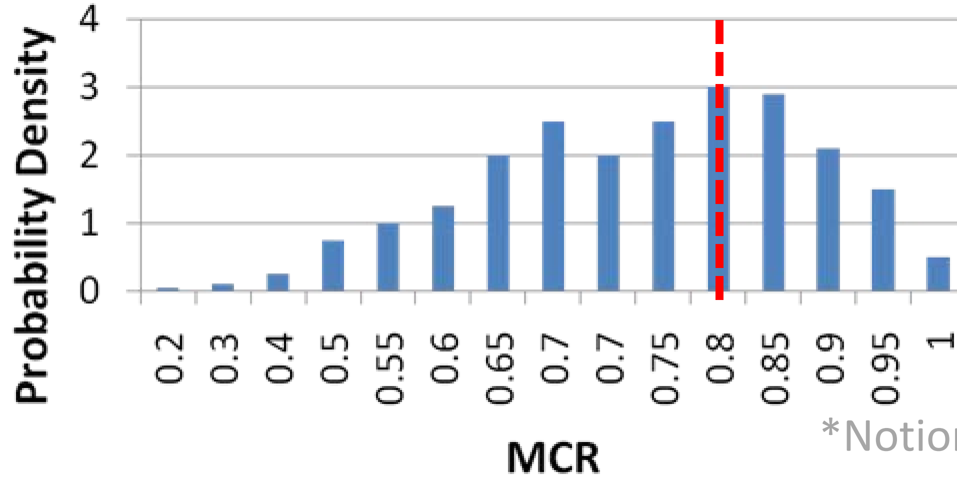
Wireless Network Performance Depends on Many Factors

- Network performance is a complex system, representing the culmination of the interaction of many variables (e.g. below):
  - Physical:
    - Antenna height
    - Line of Sight
    - Frequency and channel size
    - Mobility
    - Environment (e.g. temperature, humidity)
  - Internet / Link
    - Protocol
    - Precedence
    - Topology (hops)
  - Application
    - Offered Load
    - Message Size
  - Other
    - Human operations
    - Crypto
- Many of these factors vary <u>simultaneously</u> during a run

Data Collection is Critical

- The predictive ability of a dataset only goes as far as the data captured
- Statistical models can capture a “catch-all” variable where some results can indicate a large portion of variation that is due to unobserved effects

Distribution shows variability for MCR that averaged .8



*Notional Data

- Tracing the performance of individual packets
- Tracking node pair performance relationships, both end to end and intra-pair
- Routing information

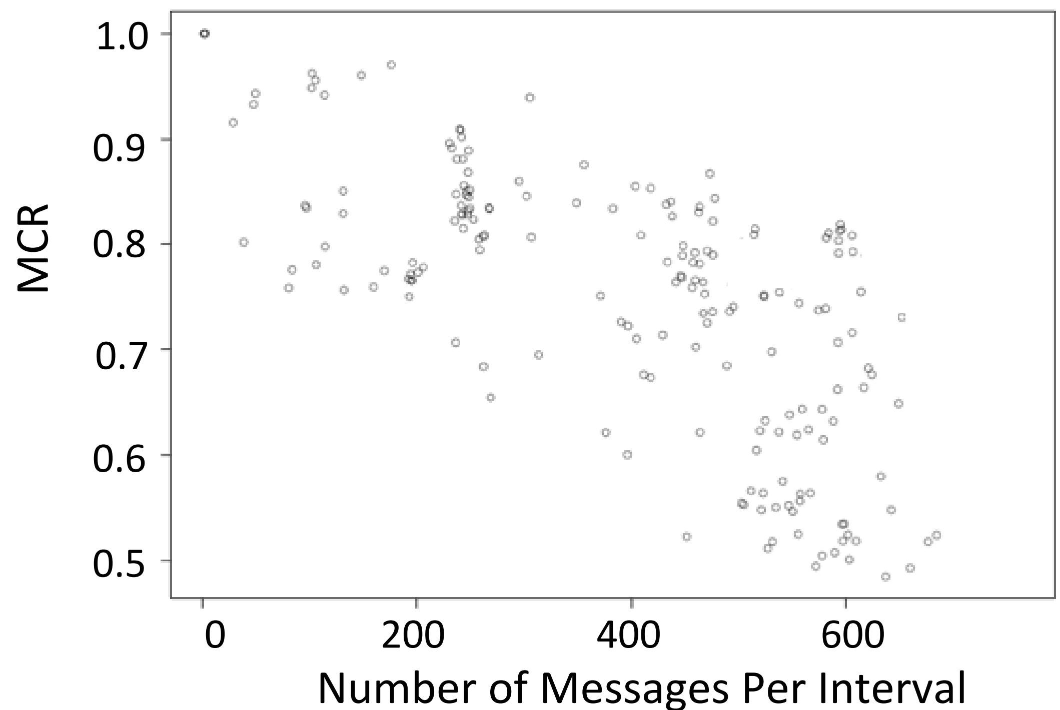
"Any Statement of Relationship Must Be Put Through Sharp Inspection" – Darrell Huff

- If B follows A, then A has caused B
- "I increased my number of network nodes and my MCR improved -> my network is scalable"

- Design of Experiments:
  - Base requirements on received load instead of offered load
  - Small message sizes
- Analysis:
  - Select only "good" runs
  - Use only raw data

A Logical First Step is to Segment the Data

- Averaged MCRs may not vary much across runs, but instantaneous traffic does
- In this case you can create a new variable that bins the traffic into time intervals (e.g. 10 second snapshots)
- Proving that experiment level data is too coarse helps MCR predictions tremendously



*Notional Data

Statistical Analysis Is Used To Determine Factor Significance and Quantify Its Impact

- Regression is designed to parse out effects of multiple variables to isolate its "true" impact
- In our analysis we tend to use two approaches:
  - Inferential Modeling: to quantify the impact of each factor
  - Predictive Modeling: to make the best possible prediction of performance (e.g. MCR, latency, etc.)

Inferential Modeling

- Standard approach for binary outcomes (e.g. MCR), Hierarchical Generalized Linear Models (HGLM)
- Allows investigations of influences of different factors on the probability that a message will be delivered successfully
  - Must account for the multi-level nature of the data with multiple sender/receiver node pairs sharing similar unobserved factors

Sample data showing the impact of factors

Factor	#exp	#pos	#neg	LCL	Estimate	UCL
Protocol – Multicast	10	1	9	-0.49	-0.35	-0.1
Precedence – Priority	10	10	0	0.15	0.21	0.39
Precedence – Immediate	10	10	0	0.45	0.5	0.6
Precedence - Override	10	8	0	0.24	0.46	0.55
Tx Antenna Height 5 m	10	0	0	-0.10	0.16	0.22
Tx Antenna Height 10 m	10	2	4	-0.41	-0.28	-0.14
Rx Antenna Height 5 m	10	1	2	-0.34	-0.19	0.05
Rx Antenna Height 10 m	10	0	7	-0.5	-0.42	-0.34
Tx Mobility 25 mph	10	1	6	-0.54	-0.34	-0.10
Rx Mobility 25 mph	10	1	8	-0.87	-0.61	-0.34
Distance	10	2	6	-0.75	-0.21	0.09
Hops	10	0	8	-0.69	-0.54	-0.29
Offered Load	10	2	8	-0.63	-0.42	-0.21

*Notional Data

Predictive Modeling

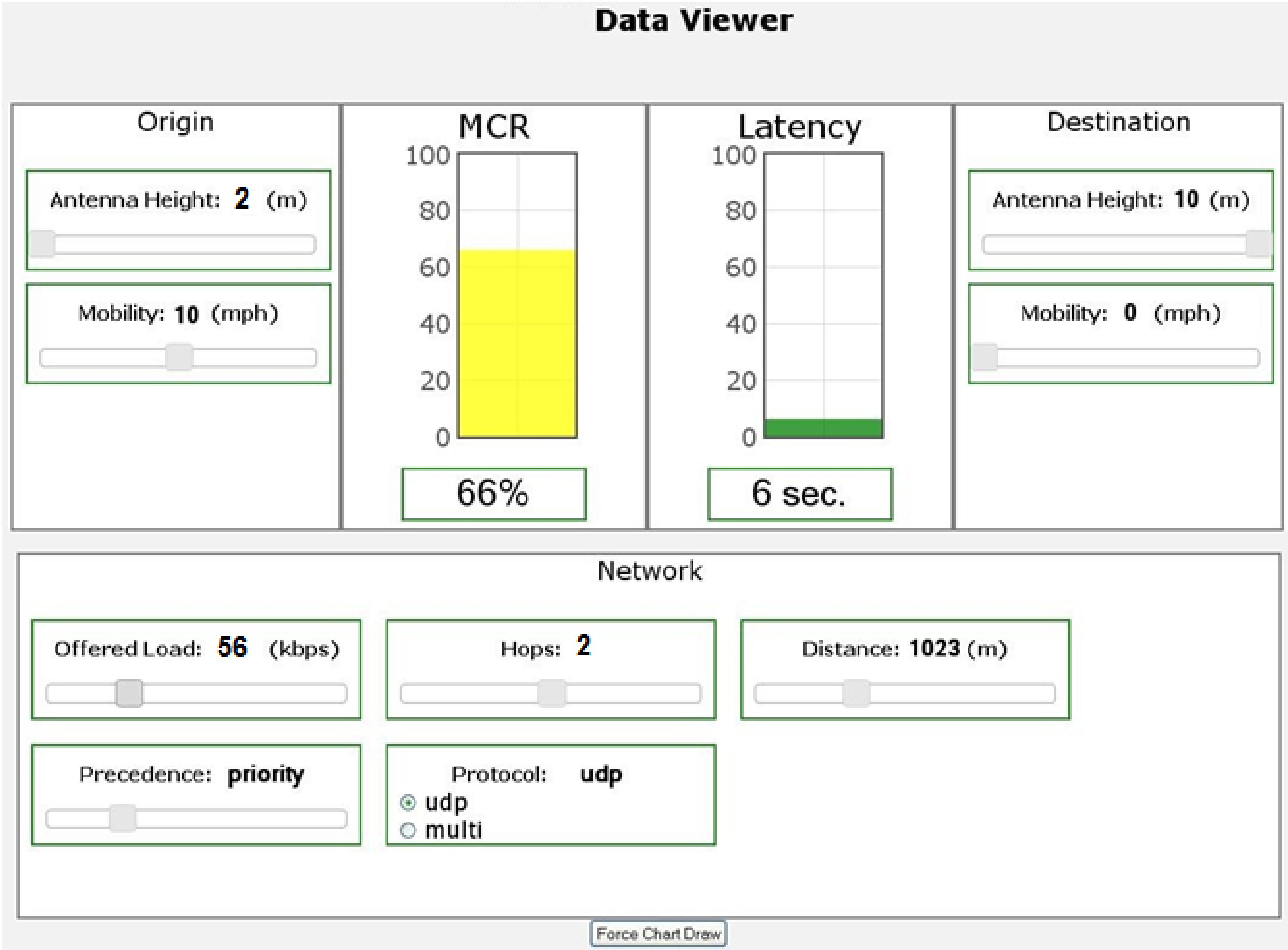
- Creates predictions of performance as a function of input variables
  - Relaxes constraints between variables and performance
  - Allows for multi-way interaction of variables
- Allows for easier comparison between different tests

The Two Techniques Are Complementary

- The inferential model allows us to better understand and interpret the individual factors on performance, holding other factors constant
- The predictive model gives us a better “black box” to predict future performance with parameters similar to those modeled

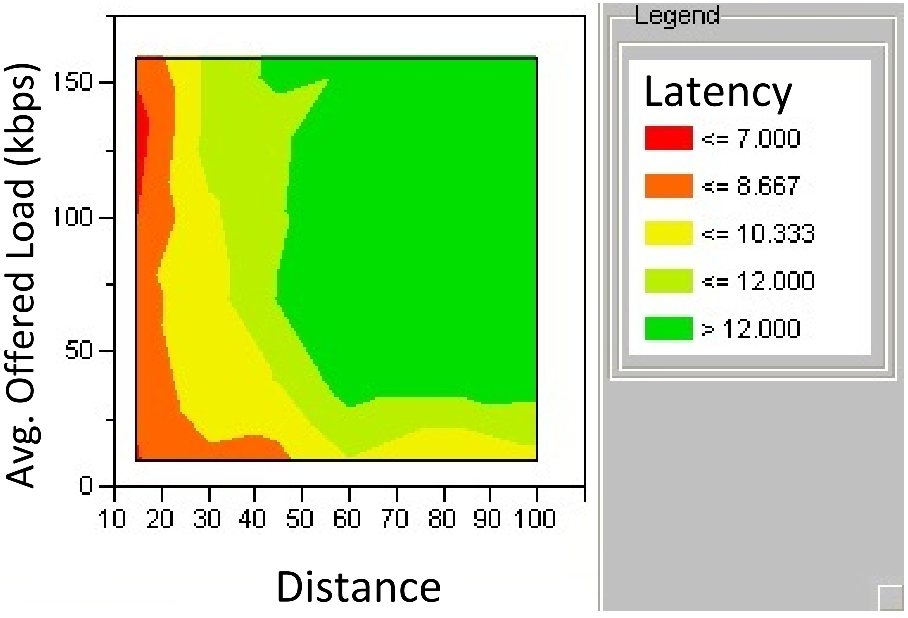
Visualization of Information is Also Important

- Predictive models are essentially "black boxes" of data that can be hard to interpret



*Notional Data

Another Useful Way to Visualize Predictive Data is Through Contour Plots



*Notional Data

Summary

- Analysis of raw and mean data can be misleading
- We're testing our radios now, specifically JTRS WNW
  - With the enormous expense of conducting live experimentation, careful statistical analysis must be done to understand the "true" nature of performance
- Successful experimentation may include:
  - Interesting DoE (i.e. the right variables and ranges of those variables)
  - Robust data collection techniques (that don't interfere with the radios)
  - Statistical analysis to fight bias

Questions

- How will network managers network performance?