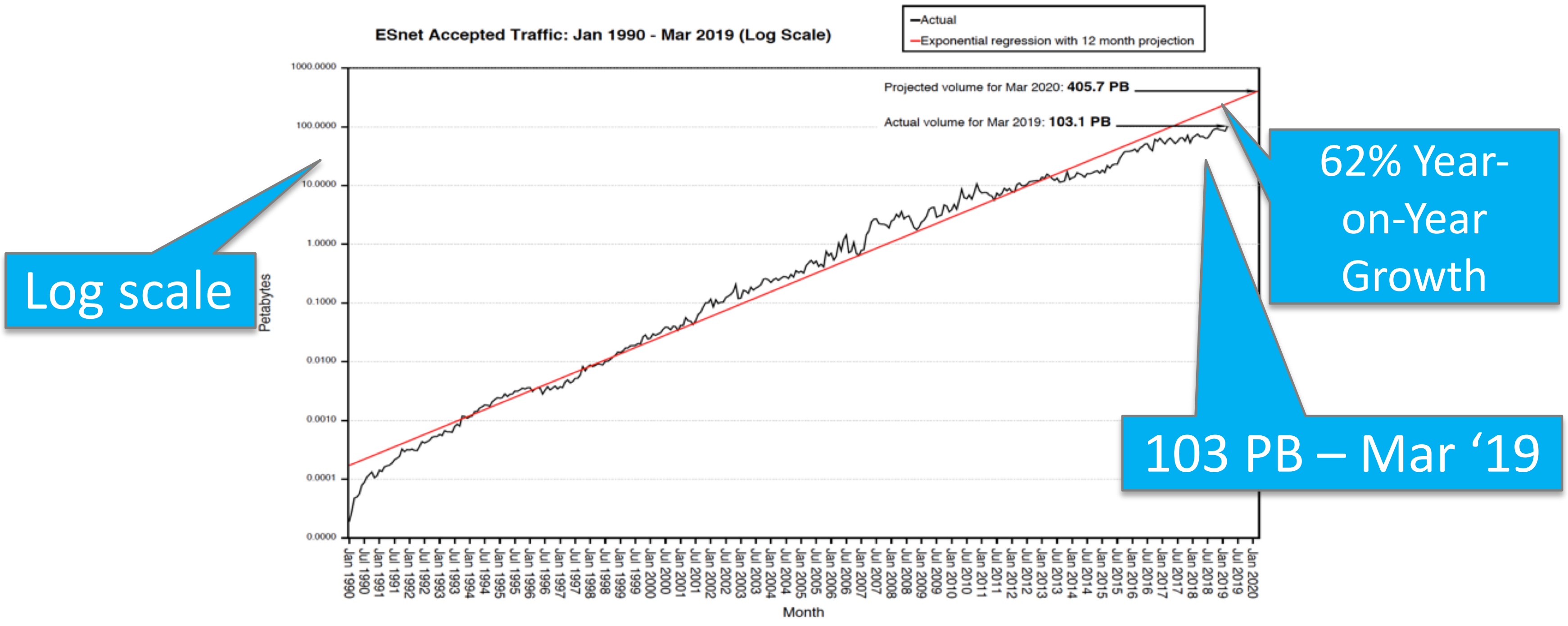
Data-driven Learning to Predict Wide Area Network Traffic

Nandini Krishnaswamy

Lawrence Berkeley National Lab

SNTA 2020

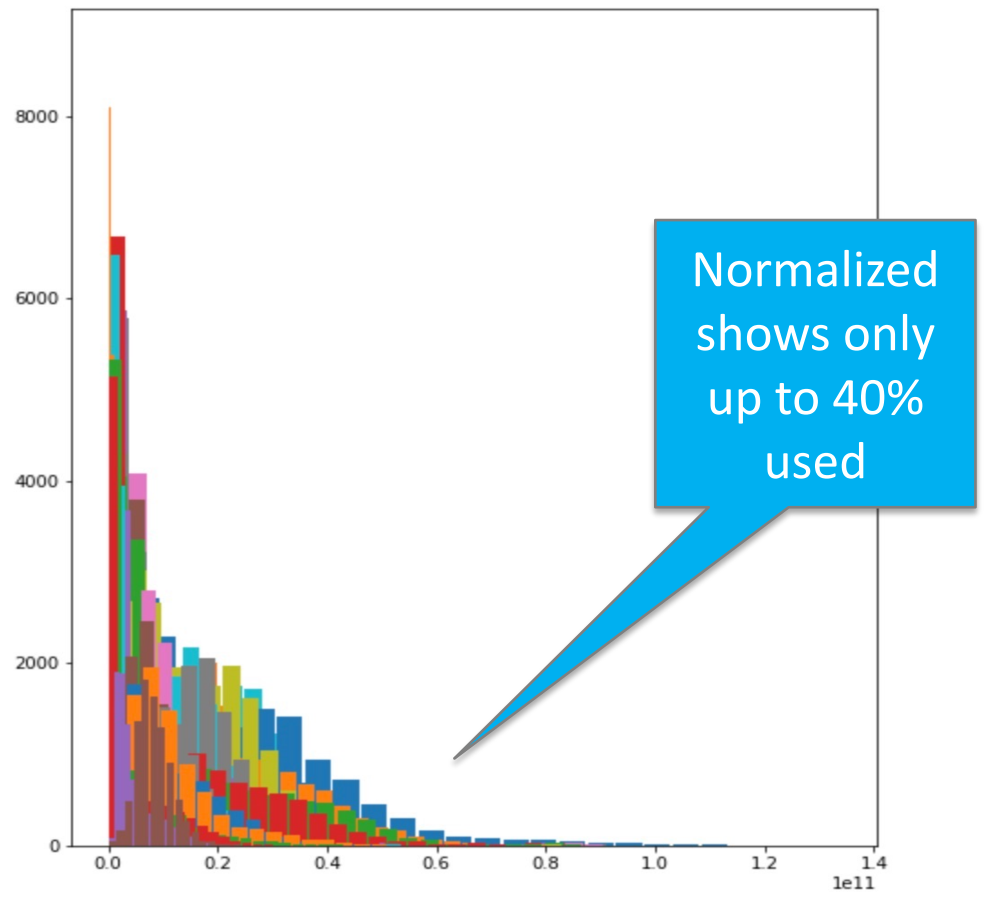
Network Traffic Growth



- This diagram illustrates the growth rate of traffic on ESnet backbone (The Department of Energy's dedicated science network).
- Projected 62% growth every year.

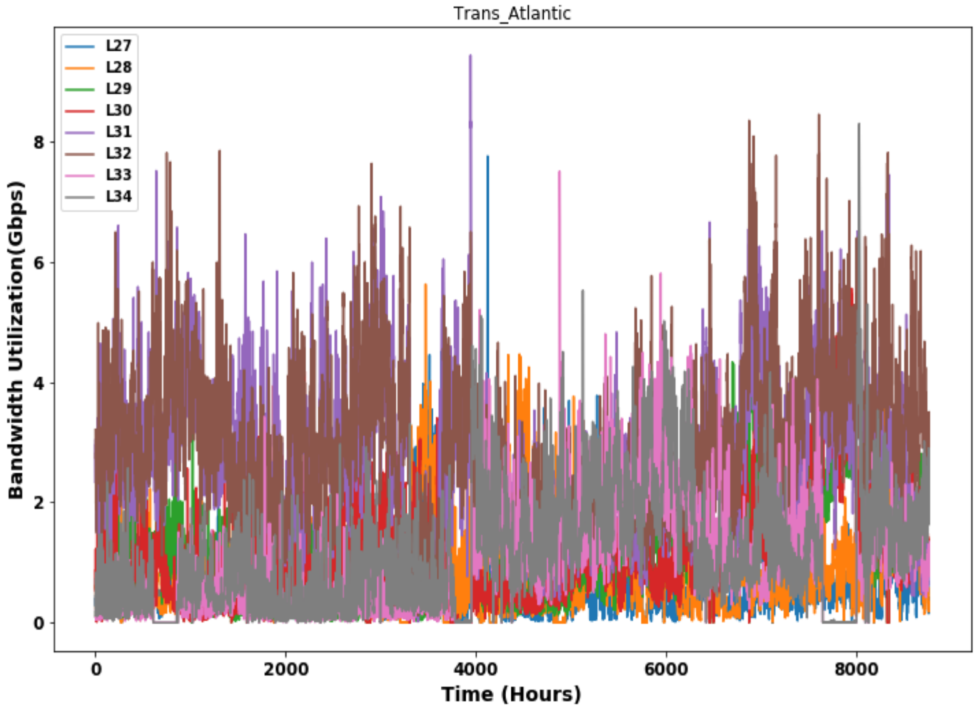
DOE Networks Link Utilization

Year 2019 Bandwidth Usage



- Links are designed to be used at 40% capacity for unanticipated traffic surges.
- How can we improve utilization?
  - Proposed solution: Predict future network traffic.

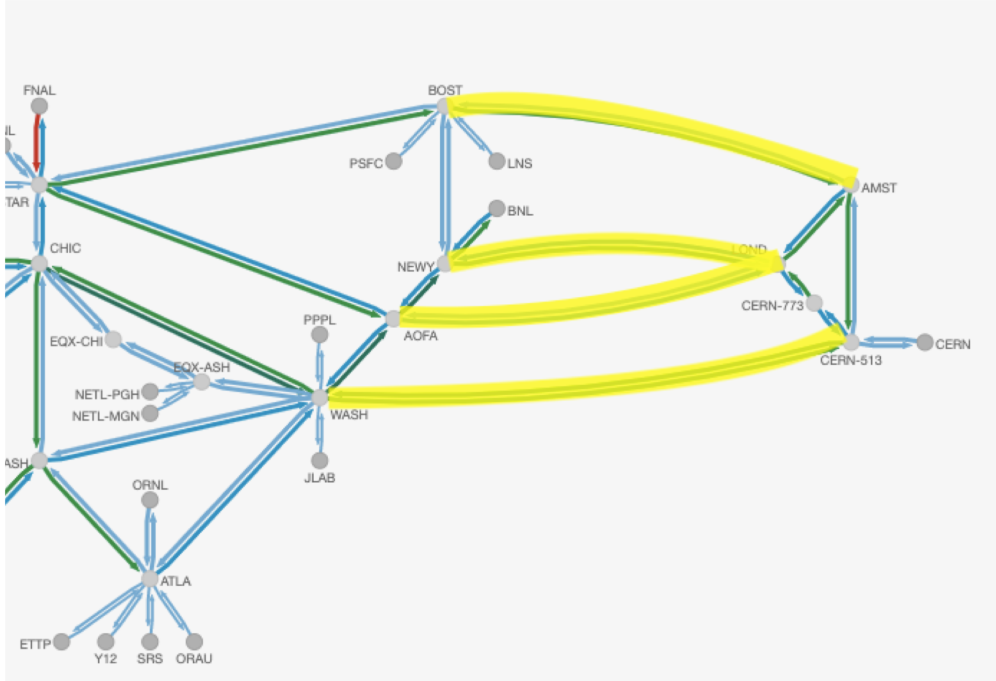
Challenges posed by Traffic Prediction



- Noisy data
- Missing data
- Multiple hour forecasts

Traffic Data Used

- SNMP data collected at router interfaces
  - Traffic volume in GBs
  - 30 second intervals (aggregated to 1 hour intervals)
  - 1 year in total
- 4 Bidirectional links (8 traces)
  - ESnet Trans-Atlantic links



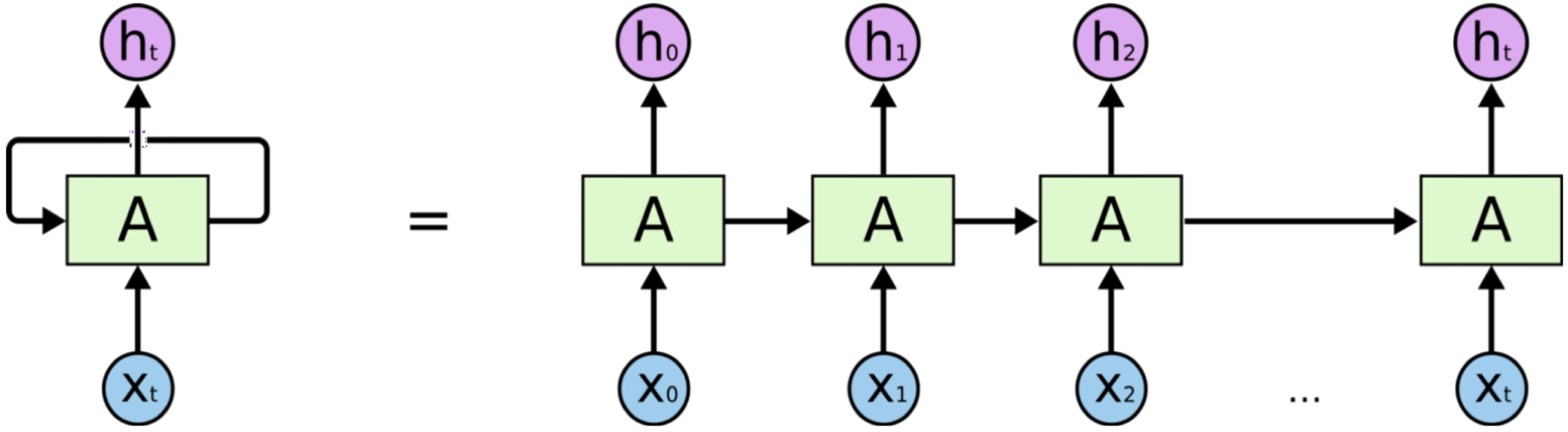
Justification of Chosen Links

- Fourier analysis
- Correlation heat map
  - file:///Users/nandinik/Desktop/2018-Jan-Dec(1).gif

Classical Time Series Algorithms

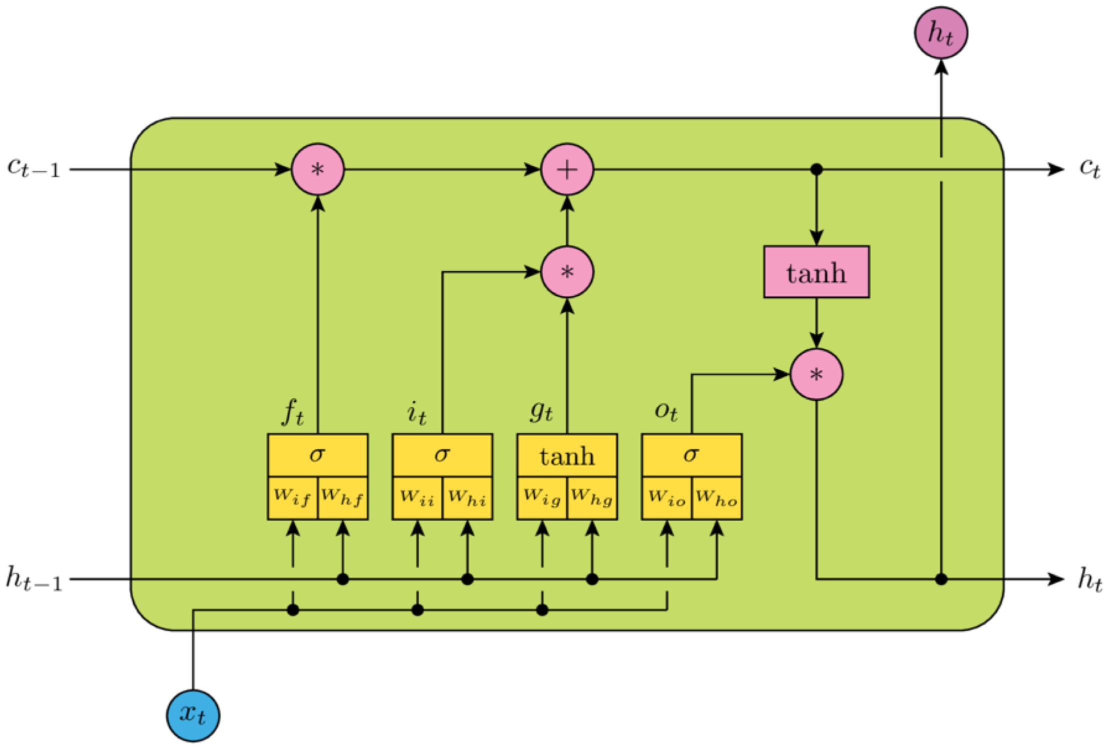
- ARIMA
  - Autoregressive Integrated Moving Average
  - Requires stationary series as input (can make series stationary through differencing)
- Holt-Winters
  - Triple exponential smoothing
  - Smoothing equations correspond to:
    - Level
    - Trend
    - Seasonality

Traditional Recurrent Neural Networks (RNNs)



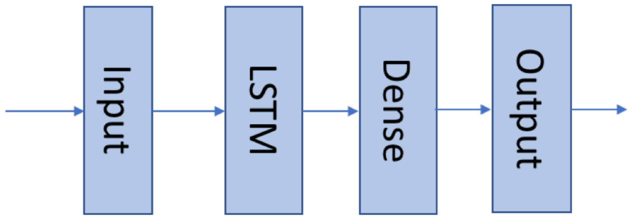
- Feedback loop -> during training, RNN will unfold into deep feedforward network
- Vanishing gradient problem -> cannot capture long-term dependencies

Long Short-Term Memory (LSTM) Network

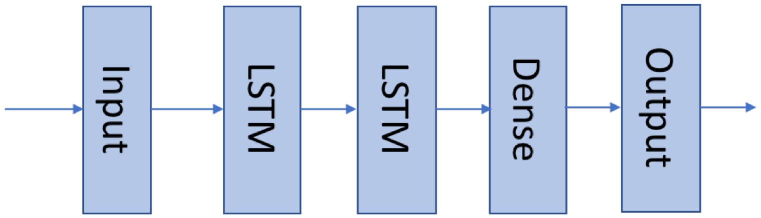


- Variant of RNN
- Memory to track long time period
- Can learn long-term dependencies

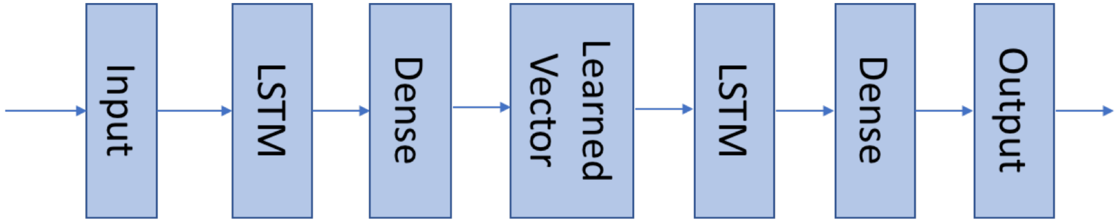
Three LSTM Variants



Simple LSTM (one LSTM layer)



Stacked LSTM (two LSTM layers)



Seq2Seq LSTM

Model Parameters

- ARIMA:
  - Inspect AC and PAC plots
- Holt-Winters
  - Trial-and-error/grid search
- LSTM
  - Tested different # of nodes in hidden layers
  - Tested different activation functions

Performance Comparison

Model	Ams-Bst-O	Ams-Bst-I	Af-Lnd-O	Af-Lnd-I	Wsh-Crn-O	Wsh-Crn-I	Lnd-NY-O	Lnd-NY-I
ARIMA	0.1630	0.1142	0.1575	0.1244	0.0418	0.0538	0.1087	0.1212
Holt-Winters	0.2172	0.3624	0.2473	**0.0626**	0.0340	**0.0439**	0.1098	0.1052
Simple LSTM	0.0502	0.0771	0.0791	0.1048	**0.0268**	0.1156	**0.0557**	0.0320
Stacked LSTM	**0.0407**	**0.0764**	**0.0660**	0.0992	0.0290	0.1129	0.0885	**0.0293**
S2S LSTM	0.0531	0.0801	0.0814	0.0949	0.0307	0.1802	0.0580	0.0293
Summary: LSTMs better than ARIMA [AR] and HW (%)								
SimpleLSTM over AR	69%	33%	50%	16%	36%	-114%	49%	74%
StackLSTM over AR	75%	33%	58%	20%	31%	-109%	19%	76%
S2S LSTM over AR	67%	30%	48%	23%	26%	-234%	47%	76%
SimpleLSTM over HW	76%	79%	68%	-67%	21%	-163%	49%	70%
StackLSTM over HW	81%	79%	73%	-58%	15%	-157%	19%	73%
S2S LSTM over HW	76%	78%	67%	-51%	10%	-310%	47%	72%

Table 4: RMSE of predicted next-24-hours Transatlantic links.

- Af-Lnd has strong seasonality
- Wash-Cern problematic data collection
- All LSTM approaches are better
- Each link has different behavior

Conclusion

- Deploy prediction tools to inform network engineering.
- Further research:
  - Extend prediction periods
  - Experiment with different NN architectures

Thank you!

Email me at nk2869@columbia.edu with any questions!