MODELING AND CHARACTERIZATION
OF TRAFFIC
IN A PUBLIC SAFETY WIRELESS NETWORK

Božidar Vujičić
bvujicic@cs.sfu.ca

Communication Networks Laboratory
http://www.ensc.sfu.ca/cnl
Simon Fraser University
Vancouver, BC
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
E-Comm network: coverage and user agencies

RCMP and Police

Fire

Ambulance

Other

TG: Talk group
R: Radio device (user)
E-Comm network architecture

Users

Burnaby

Vancouver

Transmitters/Repeaters

PSTN

PBX

Dispatch console

Network switch

Database server

Data gateway

Management console

Other EDACS systems
Structure of trunked radio systems

- Cell
  - Repeater
  - Channels
  - Cell controller
- Network management system
- Central switch
- Dispatch console
- User radios
- Cell
- Cell
Network characteristics

- **EDACS**: Enhanced Digital Access Communications Systems
- **Simulcast**: repeaters covering one cell use identical frequencies
- **Trunking**: available frequencies in a cell are shared dynamically among mobile users
  - transmission trunking
  - message trunking
- **Cell capacity** (number of available frequencies in a cell):
  - one radio channel occupies one frequency
  - one call occupies one radio channel
Call establishment

- Users are organized in talk groups:
  - one-to-many type of conversations
- Push-to-talk (PTT) mechanism for network access:
  - user presses the PTT button
  - system locates other members of the talk group
  - system checks for availability of channels:
    - channel available: call established
    - all channels busy: call queued/dropped
  - user releases PTT:
    - call terminates
Erlang traffic models

Erlang B

\[ P_B = \frac{A^N}{\sum_{x=0}^{N} \frac{N!}{x!} A^x} \]

Erlang C

\[ P_C = \frac{A^N}{\sum_{x=0}^{N-1} \frac{N!}{x!} A^x N/N!} + \frac{N}{N - A} \]

- \( P_B \): probability of rejecting a call
- \( P_C \): probability of delaying a call
- \( N \): number of channels/lines
- \( A \): total traffic volume
Erlang traffic models (2)

- Erlang B model assumes:
  - call holding time follows exponential distribution
  - blocked call will be rejected immediately

- Erlang C model assumes:
  - call holding time follows exponential distribution
  - blocked call will be put into a FIFO queue with infinite size
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Previous work

- Simulation:
  - OPNET
  - WarnSim

- Traffic prediction based on user clusters
  - Seasonal ARIMA model

- Statistical analysis of traffic


Traffic data

- 2001 data set:
  - 2 days of traffic data
    - 2001-11-1 to 2001-11-02 (110,348 calls)
- 2002 data set:
  - 28 days of continuous traffic data.
    - 2002-02-10 to 2002-03-09 (1,916,943 calls)
- 2003 data set:
  - 92 days of continuous traffic data
    - 2003-03-01 to 2003-05-31 (8,756,930 calls)
Sample of processed data: 2003-03-01

<table>
<thead>
<tr>
<th>No</th>
<th>Time (hh:mm:ss)(ms)</th>
<th>Call Duration (ms)</th>
<th>System Id</th>
<th>Channel Id</th>
<th>Caller</th>
<th>Callee</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>00:00:00 30</td>
<td>1340</td>
<td>1</td>
<td>12</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>00:00:00 489</td>
<td>1350</td>
<td>7</td>
<td>4</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>29</td>
<td>00:00:03 620</td>
<td>7550</td>
<td>2</td>
<td>7</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>31</td>
<td>00:00:03 760</td>
<td>7560</td>
<td>1</td>
<td>3</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>37</td>
<td>00:00:04 260</td>
<td>7560</td>
<td>7</td>
<td>6</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>38</td>
<td>00:00:04 340</td>
<td>7560</td>
<td>6</td>
<td>6</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>
Traffic data used for OPNET simulations

- Timestamps and durations corresponding to single call differ due to discrepancies in records:
  - the smallest timestamp was chosen arbitrarily
  - the largest call duration (worst-case scenario) was used
- Original timestamp represents date and time of call start
  - in simulations: timestamp is difference between the original timestamp and arbitrary reference time
  - reference times: 0:00 on February 25, 2002 and 0:00 on March 10, 2003

<table>
<thead>
<tr>
<th>Trace (dataset)</th>
<th>Time span</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>0:00, February 25, 2002 – 24:00, March 3, 2002</td>
</tr>
<tr>
<td>2003</td>
<td>0:00, March 10, 2003 – 24:00, March 16, 2003</td>
</tr>
</tbody>
</table>
Data processing for OPNET model

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Duration (ms)</th>
<th>Caller</th>
<th>Callee</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-03-20 0:00:10.639</td>
<td>4,870</td>
<td>A</td>
<td>B</td>
<td>4</td>
</tr>
<tr>
<td>2003-03-20 0:00:10.599</td>
<td>4,830</td>
<td>A</td>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>2003-03-20 0:00:10.529</td>
<td>4,860</td>
<td>A</td>
<td>B</td>
<td>9</td>
</tr>
<tr>
<td>2003-03-20 0:00:10.510</td>
<td>4,870</td>
<td>A</td>
<td>B</td>
<td>10</td>
</tr>
</tbody>
</table>

{10.510; 4,870; 4; 8; 9; 10}
Data discrepancies

- Coarse resolution of the timestamp
  - activity data: 10 ms
  - data model: 1 s
- Example:

![Graph showing occupied channels over time with dashed line for deployed network and solid line for model.](image)
Data discrepancies: 2003

- Overlapping usage of channels

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Duration (ms)</th>
<th>Cell</th>
<th>Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-03-20 0:00:33.370</td>
<td>9,420</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>2003-03-20 0:00:42.769</td>
<td>4,290</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

- 0:00:42.769 < 0:00:33.370 + 9.420
  - channel 4 in cell 10 is occupied by **two calls** at the same time!
Traffic data used for statistical modeling

- Records of network events:
  - established, queued, and dropped calls in the Vancouver cell
- Traffic data span periods during:

<table>
<thead>
<tr>
<th>Trace (dataset)</th>
<th>Time span</th>
<th>No. of established calls</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>November 1–2, 2001</td>
<td>110,348</td>
</tr>
<tr>
<td>2002</td>
<td>March 1–7, 2002</td>
<td>370,510</td>
</tr>
</tbody>
</table>
Hourly traces

- Call holding and call inter-arrival times from the **five busiest hours** in each dataset (2001, 2002, and 2003)

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th></th>
<th>2002</th>
<th></th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Day/hour</strong></td>
<td><strong>No.</strong></td>
<td><strong>Day/hour</strong></td>
<td><strong>No.</strong></td>
<td><strong>Day/hour</strong></td>
<td><strong>No.</strong></td>
</tr>
<tr>
<td>02.11.2001 15:00–16:00</td>
<td>3,718</td>
<td>01.03.2002 04:00–05:00</td>
<td>4,436</td>
<td>26.03.2003 22:00–23:00</td>
<td>4,919</td>
</tr>
<tr>
<td>01.11.2001 00:00–01:00</td>
<td>3,707</td>
<td>01.03.2002 22:00–23:00</td>
<td>4,314</td>
<td>25.03.2003 23:00–24:00</td>
<td>4,249</td>
</tr>
<tr>
<td>02.11.2001 16:00–17:00</td>
<td>3,492</td>
<td>01.03.2002 23:00–24:00</td>
<td>4,179</td>
<td>26.03.2003 23:00–24:00</td>
<td>4,222</td>
</tr>
<tr>
<td>01.11.2001 19:00–20:00</td>
<td>3,312</td>
<td>01.03.2002 00:00–01:00</td>
<td>3,971</td>
<td>29.03.2003 02:00–03:00</td>
<td>4,150</td>
</tr>
<tr>
<td>02.11.2001 20:00–21:00</td>
<td>3,227</td>
<td>02.03.2002 00:00–01:00</td>
<td>3,939</td>
<td>29.03.2003 01:00–02:00</td>
<td>4,097</td>
</tr>
</tbody>
</table>
Example: March 26, 2003

![Bar graph showing call holding times vs. time with red arrow indicating call inter-arrival time]
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Network model

- central switch
- 11 cells
Central switch (site) model

- Reads the trace file
- Generates packets according to calls from trace file
  - one call = one packet
  - \(\text{packet\_size (bits)} = k \times \text{call\_duration (s)}\)
  - \(k\): bit rate of channels \((k=1,000 \text{ bps in simulations})\)
- Checks for availability of channels in the cells and sending packets to appropriate cells
- Collects statistics
Central switch: OPNET node model

source

dispatcher

channel_selector_1

15 statistical wires

channel_selector_11

15 statistical wires

rx_1

rx_11
Dispatcher module in the central switch: OPNET process model
Cell: OPNET node model
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Statistical concepts

- **Probability distribution:**
  - probability that outcomes of a process are within a given range of values
  - expressed through probability density (pdf) and cumulative distribution (cdf) functions

- **Autocorrelation:**
  - measures the dependence between two outcomes of a process
  - wide-sense stationary processes: autocorrelation depends only on the difference (lag) between the time instances of the outcomes
Long-range dependence: definition

- Slow decay of the autocorrelation function $r(k)$ of a (wide-sense) stationary process $X(n)$:

  \[
  \sum_{k=-\infty}^{\infty} r(k) = \infty
  \]

  \[
  r(k) = c_r k^{-(2-2H)}, \quad k \to \infty
  \]

  \[
  f(\nu) = c_f |\nu|^{-\alpha}, \quad \nu \to 0
  \]

where $f(\nu)$ is the power spectral density of $X(n)$, $c_r$ and $c_f$ are non-zero constants, and $0 < \alpha < 1$.

$0.5 < H < 1$ implies LRD

LRD: long-range dependence
Wavelet coefficients

- Discrete wavelet transform of a signal $X(t)$:
  
  \[ d(j, k) = \int_{-\infty}^{\infty} X(t) \psi_{j,k}(t) \, dt \]

  where
  
  \[ \psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k) \]

- $\psi(t)$: mother wavelet
  
  - $j$: octave
  
  - $k$: translation

- Reconstruction formula:
  
  \[ X(t) = \sum_{j=0}^{\infty} \sum_{k} d(j, k) \psi_{j,k}(t) \]
Let $X(t)$ be LRD process (wide-sense stationary)
- its power spectral density:
  $$f(v) \sim c_f |v|^{-\alpha}, \quad v \to 0$$
- Mean square value of its wavelet coefficients on octave $j$ satisfies:
  $$\mathbb{E}\{d(j,k)^2\} = 2^{j\alpha} c_f C(\alpha,\psi)$$
  where
  $$C(\alpha,\psi) = \int |v|^{-\alpha} |\Psi(v)|^2 dv$$
does not depend on $j$

LRD and wavelets

- Logarithm:
  \[
  \log_2 E\{d(j,k)^2\} = \alpha \times j + c
  \]

- Important property: for given \( j \), \( d(j,k) \) does not exhibit long-range dependence (with respect to \( k \))
  - with appropriately chosen mother wavelet

- Hence:
  - simple estimator for \( E\{d(j,k)^2\} \) is a sample mean:
    \[
    E\{d(j,k)^2\} = \frac{1}{n_j} \sum_{k=1}^{n_j} d(j,k)^2
    \]
  - \( n_j \): number of wavelet coefficients at octave \( j \)
Estimation of $\alpha$ and $H$

- Logscale diagram: plot of $\log_2 E\{d(j,k)^2\}$ vs. $j$ (octave)
- Linear relationship between $\log_2 E\{d(j,k)^2\}$ and $j$ on the coarsest octaves indicates LRD
- Estimation of $\alpha$:
  - linear regression of $\log_2 E\{d(j,k)^2\}$ on $j$ in the linear region of the logscale diagram
- $H = 0.5 \ (\alpha + 1)$
Logscale diagram: example

- call inter-arrival times: 22:00–23:00, 26.03.2003
- $\alpha=0.576$, $H=0.788$ (octaves 4–9)
Test for time constancy of $\alpha$

- $X(n)$: wide-sense stationary process
  - $\alpha$ does not depend on $n$
- Is $\alpha$ constant throughout the time series $X(n)$?
- Approach:
  - divide $X(n)$ into $m$ blocks of equal length
  - estimate $\alpha$ for each block
  - compare the estimates
- If $\alpha$ varies significantly, estimating $\alpha$ for the entire time series is not meaningful
- In our analysis: $m \in \{3, 4, 5, 6, 7, 8, 10\}$
Kolmogorov-Smirnov test

- Goodness-of-fit test: quantitative decision whether the empirical cumulative distribution function (ECDF) of a set of observations is consistent with a random sample from an assumed theoretical distribution.

- ECDF is a step function (step size 1/N) of N ordered data points \(Y_1, Y_2, \ldots, Y_N\):

\[
E_N = \frac{n(i)}{N}
\]

\(n(i)\): the number of data samples with values smaller than \(Y_i\).
Parameters

- Hypothesis:
  - null: the candidate distribution fits the empirical data
  - alternative: the candidate distribution does not fit the empirical data
- Input parameters: significance level $\sigma$ and tail
- Output parameters:
  - p-value
  - $k$: test statistic
  - cv: critical (cut-off) value
Input parameters

- **Significance level** $\sigma$: determines if the null hypothesis is wrongly rejected $\sigma$ percent of times, if it is in fact true
  - default value $\sigma = 0.05$
- $\sigma$ defines sensitivity of the test:
  - smaller $\sigma$ implies larger **critical value** (larger tolerance)
- **tail**: specifies whether the K-S performs two sided test (default) or tests from one or other side of the candidate distribution
Output parameters

- **Test statistic** $k$ is the maximum difference over all data points:
  \[
  k = \max_{1 \leq i \leq N} \left| F(Y_i) - \frac{i}{N} \right|
  \]
  where $F$ is the CDF of the assumed distribution

- The null hypothesis is accepted if the value of the test statistic is smaller than the critical value

- **p-value** is probability level when the difference between distributions (test statistics) becomes significant:
  - if $p$-value $\leq \sigma$: test rejects the null hypothesis

- If test returns **critical value** $= \text{NaN}$, the decision to accept or reject null hypothesis is based only on $p$-value
Best-fitting distributions: CDF

![Graph showing cumulative distribution function (CDF) of call holding time with traffic data and various models: Lognormal, Exponential, Gamma, and Weibull.]
Inter-arrival time: complementary CDF

![Graph showing complementary CDF of call inter-arrival times with traffic data and various models (Exponential, Lognormal, Weibull, Gamma).]
### K-S test: call inter-arrival times 2001

**Significance level \( \sigma = 0.1 \)**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>02.11.2001, 20:00–21:00</th>
<th>02.11.2001, 16:00–17:00</th>
<th>02.11.2001, 15:00–16:00</th>
<th>01.11.2001, 19:00–20:00</th>
<th>01.11.2001, 00:00–01:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>exponential</td>
<td>h</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.0384</td>
<td>0.0001</td>
<td>0.5416</td>
<td>0.0122</td>
<td>0.0135</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0247</td>
<td>0.0369</td>
<td>0.0131</td>
<td>0.0277</td>
<td>0.0259</td>
</tr>
<tr>
<td>Weibull</td>
<td>h</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.3036</td>
<td>0.0409</td>
<td>0.4994</td>
<td>0.1574</td>
<td>0.0837</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0171</td>
<td>0.0236</td>
<td>0.0136</td>
<td>0.0195</td>
<td>0.0206</td>
</tr>
<tr>
<td>gamma</td>
<td>h</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.3833</td>
<td>0.0062</td>
<td>0.3916</td>
<td>0.0644</td>
<td>0.0953</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0159</td>
<td>0.0287</td>
<td>0.0148</td>
<td>0.0227</td>
<td>0.0202</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Significance level ( \sigma )</th>
<th>0.01</th>
<th>0.04</th>
<th>0.05</th>
<th>0.08</th>
<th>0.09</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>02.11.2001, 16:00–17:00: cv</td>
<td>0.0275</td>
<td>0.0237</td>
<td>0.0230</td>
<td>0.0215</td>
<td>0.0211</td>
<td>0.0207</td>
</tr>
<tr>
<td>01.11.2001, 00:00–01:00: cv</td>
<td>0.0267</td>
<td>0.0229</td>
<td>0.0223</td>
<td>0.0208</td>
<td>0.0204</td>
<td>0.0201</td>
</tr>
</tbody>
</table>
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Simulation results: 2002

[Graph showing simulation results for 2002 with days of the week and channel occupancy]
Simulation results: 2003

[Graph showing traffic on different days of the week, with axes indicating time and number of occupied channels, and cumulative discarded calls.]
Observations

- Presence of daily cycles:
  - minimum utilization: ~ 2 PM
  - maximum utilization: 9 PM – 3 AM
- 2002 sample data:
  - cell 5 is the busiest
  - other cells seldom reach their capacities
- 2003 sample data:
  - several cells (2, 4, 7, and 9) have all channels occupied during busy hours
Discarded calls

- appear only in the OPNET simulation results (do not exist in the deployed network)
- occur during busy hours
- may be used to identify possibly congested cells

<table>
<thead>
<tr>
<th>Sample data</th>
<th>Cell no.</th>
<th>Capacity</th>
<th>No. of discarded calls</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td></td>
<td>original</td>
<td>91</td>
</tr>
<tr>
<td>2002</td>
<td>5</td>
<td>3 + 1</td>
<td>62</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td>original</td>
<td>1,812</td>
</tr>
<tr>
<td>2003</td>
<td>9</td>
<td>6 + 1</td>
<td>679</td>
</tr>
<tr>
<td>2003</td>
<td>4</td>
<td>5 + 1</td>
<td>521</td>
</tr>
<tr>
<td>2003</td>
<td>9</td>
<td>6 + 1</td>
<td></td>
</tr>
</tbody>
</table>
### Maximum and average utilizations

<table>
<thead>
<tr>
<th>Cell</th>
<th>Capacity</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Maximum</td>
<td>Average</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>11</td>
<td>2.5</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>7</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>5</td>
<td>0.3</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>7</td>
<td>0.7</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>6</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>6</td>
<td>0.4</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>3</td>
<td>0.2</td>
</tr>
</tbody>
</table>
General OPNET statistics for data samples

- 2002 sample data:
  - span: 8:00, February 1 – 8:00, February 8
  - number of calls: 403,590
  - discarded calls: 91

- 2003 sample data
  - span: 0:00, March 20–24:00, March 26
  - number of calls: 645,167
  - discarded calls: 1,812

- Discarded calls are due to discrepancies in the data
  - they appear only in simulation results
Roadmap

- Introduction
- Traffic data models
- OPNET simulation model
- Statistical concepts and analysis tools
- OPNET simulation results
- Statistical analysis of traffic data
- Conclusions and references
Statistical distributions

- Fourteen candidate distributions:
  - exponential, Weibull, gamma, normal, lognormal, logistic, log-logistic, Nakagami, Rayleigh, Rician, t-location scale, Birnbaum-Saunders, extreme value, inverse Gaussian

- Parameters of the distributions: calculated by performing maximum likelihood estimation

- Best fitting distributions are determined by:
  - visual inspection of the distribution of the trace and the candidate distributions
  - K-S test on potential candidates
Maximum Likelihood Estimation (MLE)

- Introduced by R. A. Fisher in 1920s
- The most popular method for parameter estimation
- Goal: to find the distribution parameters that make the given distribution that follow the most closely underlying data set
- Conduct an experiment and obtain \( N \) independent observations
- \( \theta_1, \theta_2, \ldots, \theta_k \) are \( k \) unknown constant parameters which

\[
L(x_1, x_2, \ldots, x_N \mid \theta_1, \theta_2, \ldots, \theta_k) = L = \prod_{i=1}^{N} f(x_i; \theta_1, \theta_2, \ldots, \theta_k)
\]

\[i = 1, 2, \ldots, N\]
Maximum likelihood estimation

Likelihood Function Surface
Call inter-arrival times: pdf candidates

![Call inter-arrival times: pdf candidates](image_url)
### K-S test results: 2003

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>26.03.2003, 22:00–23:00</th>
<th>25.03.2003, 23:00–24:00</th>
<th>26.03.2003, 23:00–24:00</th>
<th>29.03.2003, 02:00–03:00</th>
<th>29.03.2003, 01:00–02:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>h</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.0027</td>
<td>0.0469</td>
<td>0.4049</td>
<td>0.0316</td>
<td>0.1101</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0283</td>
<td>0.0214</td>
<td>0.0137</td>
<td>0.0205</td>
<td>0.0185</td>
</tr>
<tr>
<td>Weibull</td>
<td>h</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.4885</td>
<td>0.4662</td>
<td>0.2065</td>
<td>0.286</td>
<td>0.2337</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.013</td>
<td>0.0133</td>
<td>0.0164</td>
<td>0.014</td>
<td>0.0159</td>
</tr>
<tr>
<td>Gamma</td>
<td>h</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.3956</td>
<td>0.3458</td>
<td>0.127</td>
<td>0.145</td>
<td>0.1672</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0139</td>
<td>0.0146</td>
<td>0.0181</td>
<td>0.0163</td>
<td>0.0171</td>
</tr>
<tr>
<td>Lognormal</td>
<td>h</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>1.015E-20</td>
<td>4.717E-15</td>
<td>2.97E-16</td>
<td>3.267E-23</td>
<td>4.851E-21</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>0.0689</td>
<td>0.0629</td>
<td>0.0657</td>
<td>0.0795</td>
<td>0.0761</td>
</tr>
</tbody>
</table>
Call inter-arrival times: best-fitting distributions (cdf)

- Traffic data
- Exponential model
- Weibull model
- Gamma model
Call inter-arrival time: autocorrelation

![Graph showing autocorrelation function with 99% and 95% confidence intervals.](image)
Call inter-arrival times: 26.03.2003, 22:00–23:00

- **LRD**: $\alpha > 0$ \( (H > 0.5) \)
- other traces have similar logscale diagrams
Call inter-arrival times: estimates of $H$

- Traces pass the test for time constancy of $\alpha$: estimates of $H$ are reliable

<table>
<thead>
<tr>
<th>Day/hour</th>
<th>2001 $H$</th>
<th>Day/hour</th>
<th>2002 $H$</th>
<th>Day/hour</th>
<th>2003 $H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>02.11.2001</td>
<td>0.907</td>
<td>01.03.2002</td>
<td>0.679</td>
<td>26.03.2003</td>
<td>0.788</td>
</tr>
<tr>
<td>15:00–16:00</td>
<td></td>
<td>04:00–05:00</td>
<td></td>
<td>22:00–23:00</td>
<td></td>
</tr>
<tr>
<td>01.11.2001</td>
<td>0.802</td>
<td>01.03.2002</td>
<td>0.757</td>
<td>25.03.2003</td>
<td>0.832</td>
</tr>
<tr>
<td>00:00–01:00</td>
<td></td>
<td>22:00–23:00</td>
<td></td>
<td>23:00–24:00</td>
<td></td>
</tr>
<tr>
<td>02.11.2001</td>
<td>0.770</td>
<td>01.03.2002</td>
<td>0.780</td>
<td>26.03.2003</td>
<td>0.699</td>
</tr>
<tr>
<td>16:00–17:00</td>
<td></td>
<td>23:00–24:00</td>
<td></td>
<td>23:00–24:00</td>
<td></td>
</tr>
<tr>
<td>01.11.2001</td>
<td>0.774</td>
<td>01.03.2002</td>
<td>0.741</td>
<td>29.03.2003</td>
<td>0.696</td>
</tr>
<tr>
<td>19:00–20:00</td>
<td></td>
<td>00:00–01:00</td>
<td></td>
<td>02:00–03:00</td>
<td></td>
</tr>
<tr>
<td>02.11.2001</td>
<td>0.663</td>
<td>02.03.2002</td>
<td>0.747</td>
<td>29.03.2003</td>
<td>0.705</td>
</tr>
<tr>
<td>20:00–21:00</td>
<td></td>
<td>00:00–01:00</td>
<td></td>
<td>01:00–02:00</td>
<td></td>
</tr>
</tbody>
</table>
Call holding time: pdf candidates

Traffic data
Lognormal model
Gamma model
Weibull model
Exponential model
Normal model
Rayleigh model

Call holding time (s)
Probability density
Best-fitting distributions: cdf
K-S test results: 2003

- No distribution passes the test when the entire trace is tested (significance levels = 0.1 and 0.01)
- Lognormal distribution passes test (significance level = 0.01) for:
  - 5-6 sub-traces from 15 randomly chosen 1,000-sample sub-traces
  - passes the test for almost all 500-sample sub-traces
- Test rejects null hypothesis when the sub-traces are compared with candidate distributions:
  - exponential
  - Weibull
  - gamma
Call holding time: autocorrelation

![Graph showing autocorrelation function with 99% and 95% confidence intervals.](image)
Logscale diagram, call holding times: 26.03.2003, 22:00–23:00

- independence: $\alpha \approx 0$ ($H \approx 0.5$)
- other traces have similar logscale diagrams
Call holding times: estimates of \( H \)

- all traces (except one) pass the test for constancy of \( \alpha \)
- only one unreliable estimate (*): consistent value

<table>
<thead>
<tr>
<th>Day/hour</th>
<th>2001</th>
<th></th>
<th>2002</th>
<th></th>
<th>2003</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02.11.2001</td>
<td>0.493</td>
<td></td>
<td>0.490</td>
<td></td>
<td>0.483</td>
<td></td>
</tr>
<tr>
<td>15:00–16:00</td>
<td></td>
<td>01.03.2002</td>
<td>26.03.2003</td>
<td>22:00–23:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01.11.2001</td>
<td>0.471</td>
<td></td>
<td>0.460</td>
<td></td>
<td>0.483</td>
<td></td>
</tr>
<tr>
<td>00:00–01:00</td>
<td></td>
<td>01.03.2002</td>
<td>25.03.2003</td>
<td>23:00–24:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02.11.2001</td>
<td>0.462</td>
<td></td>
<td>0.489</td>
<td></td>
<td>0.463</td>
<td></td>
</tr>
<tr>
<td>16:00–17:00</td>
<td></td>
<td>01.03.2002</td>
<td>26.03.2003</td>
<td>23:00–24:00</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>01.11.2001</td>
<td>0.467</td>
<td></td>
<td>0.508</td>
<td></td>
<td>0.526</td>
<td></td>
</tr>
<tr>
<td>19:00–20:00</td>
<td></td>
<td>01.03.2002</td>
<td>29.03.2003</td>
<td>02:00–03:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02.11.2001</td>
<td>0.479</td>
<td></td>
<td>0.503</td>
<td></td>
<td>0.466</td>
<td></td>
</tr>
<tr>
<td>20:00–21:00</td>
<td></td>
<td>02.03.2002</td>
<td>29.03.2003</td>
<td>01:00–02:00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Call inter-arrival and call holding times

<table>
<thead>
<tr>
<th>Day/hour</th>
<th>Avg. (s)</th>
<th>Day/hour</th>
<th>Avg. (s)</th>
<th>Day/hour</th>
<th>Avg. (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>inter-arrival</td>
<td>02.11.2001 15:00–16:00</td>
<td>0.97</td>
<td>01.03.2002 04:00–05:00</td>
<td>0.81</td>
<td>26.03.2003 22:00–23:00</td>
</tr>
<tr>
<td>holding</td>
<td>3.78</td>
<td>04:00–05:00</td>
<td>4.07</td>
<td>22:00–23:00</td>
<td>4.08</td>
</tr>
<tr>
<td>inter-arrival</td>
<td>01.11.2001 00:00–01:00</td>
<td>0.97</td>
<td>01.03.2002 22:00–23:00</td>
<td>0.83</td>
<td>25.03.2003 23:00–24:00</td>
</tr>
<tr>
<td>holding</td>
<td>3.95</td>
<td>22:00–23:00</td>
<td>3.84</td>
<td>23:00–24:00</td>
<td>4.12</td>
</tr>
<tr>
<td>inter-arrival</td>
<td>02.11.2001 16:00–17:00</td>
<td>1.03</td>
<td>01.03.2002 23:00–24:00</td>
<td>0.86</td>
<td>26.03.2003 23:00–24:00</td>
</tr>
<tr>
<td>holding</td>
<td>3.99</td>
<td>23:00–24:00</td>
<td>3.88</td>
<td>23:00–24:00</td>
<td>4.04</td>
</tr>
<tr>
<td>inter-arrival</td>
<td>01.11.2001 19:00–20:00</td>
<td>1.09</td>
<td>01.03.2002 00:00–01:00</td>
<td>0.91</td>
<td>29.03.2003 02:00–03:00</td>
</tr>
<tr>
<td>holding</td>
<td>3.97</td>
<td>00:00–01:00</td>
<td>3.95</td>
<td>02:00–03:00</td>
<td>4.14</td>
</tr>
<tr>
<td>inter-arrival</td>
<td>02.11.2001 20:00–21:00</td>
<td>1.12</td>
<td>02.03.2002 00:00–01:00</td>
<td>0.91</td>
<td>29.03.2003 01:00–02:00</td>
</tr>
<tr>
<td>holding</td>
<td>3.84</td>
<td>00:00–01:00</td>
<td>4.06</td>
<td>01:00–02:00</td>
<td>4.25</td>
</tr>
</tbody>
</table>

Avg. call inter-arrival times: 1.08 s (2001), 0.86 s (2002), 0.84 s (2003)
# Distributions

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Expression</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>exponential</td>
<td>( f(x) = \frac{e^{-x/\mu}}{\mu} )</td>
<td></td>
</tr>
<tr>
<td>Weibull</td>
<td>( f(x) = ba^{-b}x^{b-1}e^{-(x/a)^b}I_{(0,\infty)}(x) )</td>
<td>( I_{(0,\infty)}(x) ): incomplete beta function</td>
</tr>
<tr>
<td>gamma</td>
<td>( f(x) = \frac{x^{a-1}e^{-(x/b)}}{b^a\Gamma(a)} )</td>
<td>( \Gamma(a) ): gamma function</td>
</tr>
<tr>
<td>lognormal</td>
<td>( f(x) = \frac{e^{-(\ln x-\mu)^2/(2\sigma^2)}}{x\sigma\sqrt{2\pi}} )</td>
<td></td>
</tr>
</tbody>
</table>
### Best fitting distributions

<table>
<thead>
<tr>
<th>Busy hour</th>
<th>Call inter-arrival times</th>
<th>Call holding times</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weibull</td>
<td>Gamma</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>02.11.2001 15:00–16:00</td>
<td>0.9785</td>
<td>1.1075</td>
</tr>
<tr>
<td>01.11.2001 00:00–01:00</td>
<td>0.9907</td>
<td>1.0517</td>
</tr>
<tr>
<td>02.11.2001 16:00–17:00</td>
<td>1.0651</td>
<td>1.0826</td>
</tr>
<tr>
<td>01.03.2002 04:00–05:00</td>
<td>0.8313</td>
<td>1.0603</td>
</tr>
<tr>
<td>01.03.2002 22:00–23:00</td>
<td>0.8532</td>
<td>1.0542</td>
</tr>
<tr>
<td>01.03.2002 23:00–24:00</td>
<td>0.8877</td>
<td>1.0790</td>
</tr>
<tr>
<td>26.03.2003 22:00–23:00</td>
<td>0.7475</td>
<td>1.0475</td>
</tr>
<tr>
<td>25.03.2003 23:00–24:00</td>
<td>0.8622</td>
<td>1.0376</td>
</tr>
<tr>
<td>26.03.2003 23:00–24:00</td>
<td>0.8579</td>
<td>1.0092</td>
</tr>
</tbody>
</table>
Estimates of $H$

- call inter-arrival times: $H \approx 0.7–0.8$
- call holding times: $H \approx 0.5$
Conclusions

- We created an OPNET model and simulated two weeks of network activity
- Network utilization exhibits daily cycles
- Between February 2002 and March 2003:
  - number of calls increased by ~ 60 %
  - average utilization increased non-uniformly across the network
- Several cells may become congested in future
Conclusions

- We analyzed busy hours voice traffic from a public safety wireless network in Vancouver, BC
  - call inter-arrival and call holding times during five busy hours from 2001, 2002, and 2003
- Statistical distribution functions of traffic traces:
  - Kolmogorov-Smirnov goodness-of-fit test
  - autocorrelation functions
  - wavelet-based estimation of the Hurst parameter
Conclusions

- Call inter-arrival times:
  - best fit: Weibull and gamma distributions
  - long-range dependent: $H \approx 0.7–0.8$

- Call holding times:
  - best fit: lognormal distribution
  - uncorrelated
References

References