Modelling Voice Communication in Disaster Area Scenarios

Nils Aschenbruck, Michael Gerharz, Matthias Frank, Peter Martini
University of Bonn - Institute of Computer Science IV
Roemerstr. 164, 53117 Bonn, Germany
{aschenbruck, gerharz, matthew, martini}@cs.uni-bonn.de

Abstract

This paper deals with voice communication models for disaster area scenarios. The goal is to design models that can be used to generate realistic push to talk traffic for single talk groups. The modelling is based on an analysis of empirical measurements during a catastrophe maneuver. The analysis shows that the time series comprise heavy load periods and significant correlations. Based on these characteristics, different Markov and Semi-Markov models are considered. Synthetic traffic streams for the different models are generated and evaluated by visual and statistical analysis. Finally, a case study outlines the impact of the different traffic models in network performance simulation.

1 Introduction

In catastrophe situations, public safety units need reliable communication systems. Due to the fact that any kind of infrastructure may have been destroyed by the catastrophe, there is a demand for communication systems independent from infrastructure in this "disaster area scenario". Especially for Mobile Multi-hop Ad hoc NETworks (MANETs) this scenario follows as a quasi canonical scenario. MANETs satisfy the requirement of being independent of any kind of infrastructure by their definition.

For the evaluation of network performance aspects in a disaster area scenario, modelling the data traffic realistically is one important issue. Today, the main application during catastrophes is voice communication: different users communicate via push to talk voice calls. Each call is done by one sender that starts speaking and stops after a certain amount of time. There is only a half-duplex connection (unlike a telephone call): while one user speaks, the others have to listen. Different calls with semantic connection (e.g. question and answer) may be regarded as one conversation or session, where a conversation consists of an arbitrary number of calls between two callers, and, typically, the callers alternate in calling each other. The users (*talk group*) that communicate with each other share one broadcast voice communication channel. Technically this broadcast voice communication channel may be realized e.g. as a separated physical channel or as a multicast group. The term *talk group* abstracts from the technical realization.

The voice communication may evolve in future systems e.g. the voice codecs may change, the technical realization of a talk group may change, or the voice traffic may become video traffic (video-phone). However, the message to be said (holding time) and pause between two messages (idle time) will still be interesting. Thus, the analysis of voice traffic in disaster area scenarios focussing on appropriate traffic models is a relevant research area.

In this paper, we model the push to talk traffic typical for a disaster area talk group. The model is based on an analysis of traffic measured during a civil protection maneuver where the traffic of different talk groups is analyzed. For each talk group *call holding times* (time one user speaks) and *idle times* (the pause time between two messages) are determined.

The remaining part of this paper is structured as follows: Section 2 describes related work. In section 3 we describe the concrete scenario in which the data was acquired, the measurement architecture, as well as the generation of time series. Next, the time series are analyzed with respect to dependencies and heavy load periods (section 4). Based on this, appropriate traffic models are derived (section 5). Next, we study the traffic generated by different models to analyze their impact (section 6). Finally, we conclude the paper and point out topics for future work (section 7).

2 Related Work

Paul T. Brady analyzed telephone voice conversations in the 1960s and discovered that on and off periods of voice are exponentially distributed ([2], [3]). Based on his work [3], the ITU-T standardized (ITU-T P.59 - 1993) a commonly accepted model for artificial conversational speech [10] where a voice channel is modelled by a two-state markov model (cf. figure 1) assuming one state as talk spurt

1-4244-0419-3/06/$20.00

(*ON-state*) and the other as silence (*OFF-state*). In analogy to Brady's analysis the holding times in ON- and OFF-state are exponentially distributed.

The analysis of group communication in land mobile radio systems began in the 1980s. The purpose was to derive models to be used in the design of the new trunked radio systems. The analyses of Hess ([7], [6]) characterize session length and interarrival times as exponentially distributed (similar to Brady cf. figure 1) and suggest models for peak load. He finally recommends to use the Erlang-C model. His results were confirmed by the Public Safety Advisory Committee [15]. Other studies (e.g. [5]) examine the behavior of different talk groups to the system and suggest more complex models. However, the traffic is still modelled with exponentially distributed session length and interarrival times.

New studies by Barceló et al. ([1], [12], [11]) show that modelling the channel holding and idle times as exponentially distributed is inaccurate. They suggest to use lognormal and Erlang-jk distributions, respectively. Recent studies by Trajković et al. ([14], [17]) found Weibull or Gamma distributed call interarrival times and lognormally distributed call holding times. Furthermore, correlations between calls have been examined. The call holding times show no correlation whereas interarrival times showed dependencies. Both studies imply a two state model as presented in figure 1 with different (non-exponential) state holding time distributions (two state semi-markov model).

In older studies, series of calls where regarded as one message in contrast to new ones where single calls are analyzed. A reason for this may be that new communication systems release the channel between different calls. Thus, it is necessary to model single calls (cf. [14]).

Similar to this, former studies consider the traffic of a single channel (talk group) whereas new studies consider the traffic of the complete trunked radio systems (multiple talk groups mixed). However, the lack of examining single channels results in less accuracy when modeling the traffic of one talk group. Single talk groups may only have little impact when examining infrastructure based networks and their components. However, when evaluating (infrastructure-less) MANETs the traffic of single talk groups is of major interest.

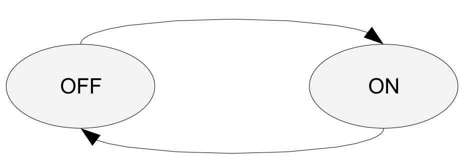

Figure 1. two state (On-/Off-Model) (semi-) markov model

Furthermore, all existing studies consider normal (non-catastrophe) load on a public safety wireless communication system. In disaster area scenarios (during catastrophes) where the infrastructure is destroyed and civil protection units join the communication channels the traffic can not a priori be expected to be equal to the general (non-catastrophe) case. Thus, traffic models for disaster area scenarios should be based on talk group separated traffic of heavy load catastrophe situations.

3 Measuring Architecture

Ideally, the traffic model should be based on load of real catastrophe situations. (Un)Fortunately catastrophes are quite rare, happen all over the whole world, and most importantly can not be planned. Thus, we decided to measure in a large catastrophe maneuver. The maneuver we base our analysis on took place in May 2005 in Cologne in preparation of the World Youth Day 2005 and FIFA Soccer Worldcup 2006. The scenario was that more than 250 people were injured by a catastrophe in an event hall. More than 955 disaster area units (firefighters, paramedics, etc.) and 279 vehicles of four administrative districts joined the maneuver over the time. As communication system, more than ten channels of the analog German national radio system, called BOS-system (68-87.5 MHz and 146-174 MHz), were used. Different talk groups were separated by different physical channels.

During the maneuver we observed the traffic on five channels. First of all we wanted to determine On- and Off-times on each channel. Based on short term fourier transformation (stft) we filtered out the relevant frequencies for human speech (100Hz - 7kHz) and summed up the intensities. When the sum was above a threshold, the channel was considered as being used (On-Time). We chose parameters that the granularity of the resulting On- and Off-times of 10ms was achieved. A smaller value makes no sense, because voice coders for tactical environments like the enhanced Mixed Excitation Linear Prediction (MELPe)[4] split the audio signal into frames of e.g. 22.5ms. Thus, assuming such a vocoder to be used, more accurate On- and Off-Times have no impact on the traffic modelled.

After this analysis there are still a lot of small calls, that only comprise spurious noise. These small calls need to be discarded. As threshold value a throwaway time had to be found to discard calls smaller than it (cf. [2]). We performed two stages to find it and considered values of up to 1s. During the first automatic stage we examined the amount of calls over different throwaway times. Values between 200 and 400ms seemed to be acceptable. As a second stage, we sorted the calls by their length and checked manually until what time calls comprise no noise but content. This test yielded a value of 300ms. Thus, all calls smaller

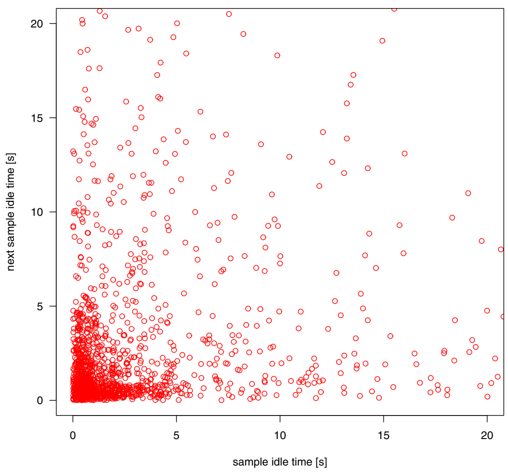

Figure 2. Dependencies idle times - scatter plot

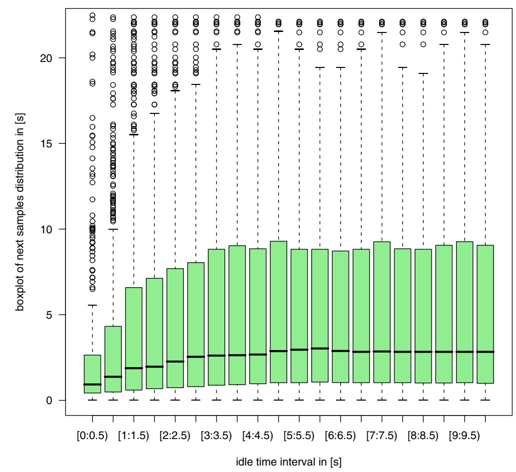

Figure 3. Dependencies idle times - box plots for intervals over time

than 300ms are discarded.

Finally, we base our analysis in the following section on 4914 calls of five talk groups. For each talk group we calculate idle and channel holding times and obtain the time series.

4 Analysis of Time Series

In this section we analyze the time series obtained. First, we focus on dependencies in one talk group. These could be expected due to the conversational manner of the communication. Next, we examine whether the load is varying over time. Different units join the maneuver at different times. Thus, a variation in load may be expected.

4.1 Dependencies

The call channel holding times do not show any indications for dependencies. However, the idle times show some characteristic indications. Figure 2 contains an example of an idle scatter plot. The scatter plot included samples $x_1, x_2, .., x_n$ as (x_i, x_{i+1}) for $i = 1, 2, .., n-1$ (cf. [13])). There are many short idle times that are followed by a short one (calls in one conversation), short idle times that are followed by a long one (last call of the conversation), and long idle times that are followed by a short one (first call of the conversation). But long idle times followed by long ones are rare (single call conversations).

Box plots (see figure 3 for an example) confirm these characteristics. Box plots are another visualization of the data of the scatter plots. The data of the x-axis x_i's are pooled in intervals (of 0.5s in figure 3). The y-axis shows box plots for the intervals. The box plot shows median, quartiles, octiles, and extremes (for further details see [13]). The figure shows that both the boxes and the medians for the first intervals are much smaller when compared to later ones. This is a confirmation of the result of many short calls that are followed by short ones.

Figure 4 shows the call holding times over the summed idle times for a short piece of trace of one talk group. There are groups of calls - one group of calls is one conversation. These conversations produce dependencies between the calls and induce the indications examined in the scatter and box plots.

The autocorrelation does not show similar results, because the dependencies observed above do not seem to be linear.

To allow for an analysis of conversations, these conversation have to be identified in the trace files. There are several possible approaches to achieve this:

- *Speaker recognition*: One possible way is to use methods of speaker recognition [16]. A new conversation starts when the speakers or at least one speaker changes. This approach appears promising, but when looking into the details there are a lot of challenges. The trace we observe contains an unknown set of speakers of unknown size. Furthermore, the signal is sometimes quite noisy and the speakers are sometimes quite stressed. So it is basically a worst case for speaker recognition, which is still quite error-prone with current algorithms even under good conditions.

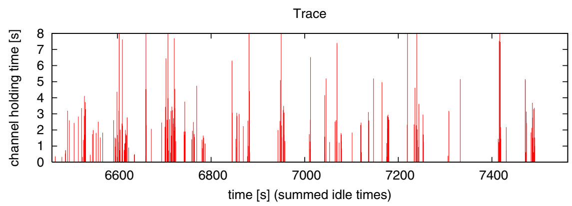

Figure 4. Call holding times over the summed idle times

- *Manual analysis*: Another approach to group the calls into sessions is the manual analysis by listening to the audio-file. Besides the disadvantage that this approach is very time-consuming and error-prone, it is only possible when the quality of the trace allows for a manual analysis.
- *Time Threshold Based Approach*: In this approach, a new conversation is assumed when the idle time is larger than a certain threshold. The problem of this approach is that it fails if the idle time between two conversations is smaller than the threshold. This may especially be the case under heavy load. However, even in heavy load cases we experienced a kind of human backoff before a new conversation was started. When wanting to access the broadcast voice channel, the caller waits for some time being sure there is no other conversation and the previous conversation has ended, before starting a new conversation.

We decided to use the time threshold based approach performing a two-stage approach to find an appropriate threshold. As a first stage we examined the count of conversations, and calls per conversation for thresholds up to 10s. Then as a refinement, we tried to determine the human backoff value manually. Finally, we decided to set the threshold to 3s, because it provided acceptable results examining count of conversations as well as calls per conversations.

Grouping calls to conversations results in four parameters that can be gained from the analysis of the time series:

call channel holding time:	the time one call lasts
call idle time:	the time between two calls of a conversation
conversation idle time:	the time between two conversations
calls per conversation:	the number of calls per conversation

4.2 Heavy Load Periods

Units are joining and leaving the maneuver and within this the talk groups as well. Incidents (e.g. another group of injured was found) cause a lot of communication. Thus, the load of one talk group is varying. For performance analysis of communication systems heavy load periods are of particular interest. To achieve reliable models the traffic of heavy load periods should be used for parameterization. Thus, our goal in this section is to find the overall heavy load period for our communication system consisting of five talk groups.

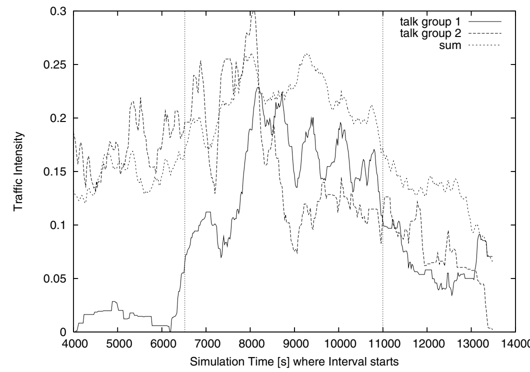

Figure 5. Traffic intensity over time

To examine the variation of the load for the different talk groups over time, we observe the traffic intensity. The traffic intensity for a single talk group is defined as:

$$TI_{[i;i+t)} = \frac{\text{time medium allocated}_{[i;i+t)}}{\text{time}_{[i;i+t]}}$$

and for a set of k talk groups it is defined as:

$$TI^{k}_{[i;i+t)} = \frac{\sum_{\forall k} \text{time medium allocated}_{[i;i+t)}}{\text{time}_{[i;i+t]} * k}$$

It can be calculated for different interval sizes t. For $t = 1h$ the metric traffic intensity is also known as Erlang[1]. Figure 5 shows the traffic intensity for two single talk groups and the all five talk groups (TI^5) over time for an interval size of $t = 900s$.

For each of the two single talk groups the load is varying. There are periods of heavier load that correlate with the tactical usage of the talk group. For example talk group 2 is the group of the first responders and casualties waiting for treatment area whereas talk group 1 is the group that is used by a casualties treatment station which has first to set up. The overall main load for the communication system can be seen at the plot of TI^5 (sum). We decided to choose the heavy load period in the interval [6390s;11900s]. The start value was motivated by the background information

[1] named after the Danish telephone engineer Agner Krarup Erlang (1878-1929)

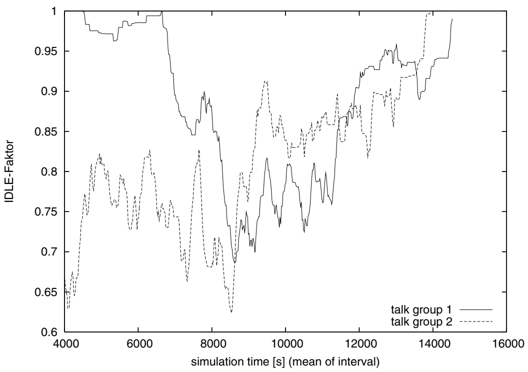

Figure 6. Idle fraction over time

that the last units join the maneuver at the interval start time (cf. load of talk group 1). From this start value to the end value $TI^5 > 0.138$ holds all the time. Note: the last interval (length 900s) starts at 11000s.

For traffic modelling it is important to know which parameter (cf. end of section 4.1) excites the heavy load period. To determine this parameter in a traffic model the heavy load periods have to be considered.

From the previous section we know that the calls are grouped in conversations. Thus, the conversation idle time may be the influencing parameter. We evaluate it by calculating the idle-fraction over time for a talk group.

$$IF_{[i;i+t)} = \frac{\sum \text{Conversation idle times}_{[i;i+t)}}{\text{time}_{[i;i+t]}}$$

Figure 6 shows the conversation idle fraction over time for the talk groups 1 and 2 for an interval size of $t = 900s$. The analysis of correlation between conversation idle fraction and traffic intensity over time shows a strong negative correlation (for talk groups 1 and 2, -0.9984364 and -0.9898027, respectively). This correlation shows that the conversation idle time is the main parameter that has a strong impact on the heavy load periods.

5 Models

Based on the information gained in the previous section, in this section we present three models for traffic generation of a single talk group differing in the level of detail:

5.1 Two-State Semi-Markov Model

The two-state markov or semi-markov model (see figure 1) is the model used for voice communications in the literature, as already mentioned in section 2. However, this approach does not model the conversational correlations (cf. section 4.1). Therefore, we use this approach mainly for reference purposes.

Distribution	Parameters	K-S dist.	p-value
CHT			
expo. mean	rate = 2.766134	0.7832	< 2.2e-16
exponential	rate = 0.3613281	0.1435	7.061e-14
lognormal	meanlog = 0.7515069 sdlog = 0.7473572	0.0387	0.2089
lognormal3	prob1 = 0.1229589 meanlog1 = -0.2750504 sdlog1 = 0.4088205 prob2 = 0.2504408 meanlog2 = 0.7426892 sdlog2 = 0.7747907 meanlog = 0.956361 sdlog3 = 0.6057901	0.0175	0.975
weibull	shape = 1.422493 scale = 3.064768	0.0606	0.008028
gamma	shape = 2.0320155 rate = 0.7346204	0.0444	0.1028
idle			
expo. mean	rate = 4.572787	0.6214	< 2.2e-16
exponential	rate = 0.2186523	0.3057	< 2.2e-16
lognormal	meanlog = 0.2556660 sdlog = 1.5288223	0.0568	0.01563
weibull	shape = 0.6235075 scale = 2.8384608	0.1068	7.369e-08
gamma	shape = 0.5021595 rate = 0.1097927	0.1683	< 2.2e-16

Table 1. Fitting of channel holding and idle times for two-state model

We fitted the commonly (cf. section 2) used distributions to our data. We restricted the data to the heavy load period to reduce the impact of variations and show only results for the talk group with the highest TI. For estimating optimal parameters we used the Maximum-Likelihood-Method. Having found the optimal parameters, the quality of the fitting to the empirical data was evaluated using the Kolmogorov-Smirnov (K-S) test. The results obtained for exponential, lognormal, weibull, and gamma distributions for the talk group with the highest traffic intensity can be found in table 1. We show results for the lognormal3 distribution as well, because it achieved significantly smaller K-S distance and larger p-values. For the exponential distribution, in addition to the values obtained by Maximum-Likelihood-Method, we also show results for the mean equally to the overall sample mean, because this is often used as standard parameter.

For the call holding time the lognormal and lognormal3 distributions shows the best fit. These results are similar to the ones Barceló et al. and Trajković et al. (cf. section 2) observed for non-catastrophe communication. For the idle times none of the distributions shows a satisfying fit. The reason for this is the conversational dependence we discovered in section 4.1. For the evaluation in section 6 we use the two-state model for reference purposes with exponential and lognormal (lognormal3 for channel holding times) distributions.

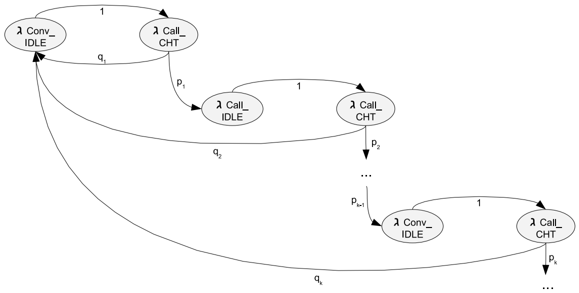

Figure 7. n-state (On-/Off-Model) (semi-) markov model

5.2 N-State Semi-Markov Model

To consider the conversational correlations observed in section 4.1 the traffic can be modelled by an n-state markov or semi-markov model (see figure 7). A new conversation starts after a conversation idle time λ_{Conv_IDLE} with at least one call with channel holding time λ_{Call_CHT}. After this call there may be either a short idle time λ_{Call_IDLE} (with probability p_1) or a longer one λ_{Conv_IDLE} (with probability $q_1 = 1 - p_1$). In case there is a long idle time we return to the first state (λ_{Conv_IDLE}). In case there is a short idle time, there is another call with channel holding time λ_{Call_CHT}. After this there may be again either a short idle time λ_{Call_IDLE} (with probability p_2) or a longer one λ_{Conv_IDLE} (with probability q_2) and so on (with probability p_k and q_k, respectively).

The transition probabilities p_k and q_k can be determined from the measured conversations calculating conditional probabilities as follows:

$$COL_k := \begin{cases} \emptyset & k \leq 0 \\ \{\text{Conversations} : \text{CallCount} \geq k\} & k \geq 1 \end{cases}$$

$$COE_k := \begin{cases} \emptyset & k \leq 0 \\ \{\text{Conversations} : \text{CallCount} = k\} & k \geq 1 \end{cases}$$

$$q_k := P\{\text{CallCount} = k \mid \text{CallCount} \geq k\} = \frac{\mid COE_k \mid}{\mid COL_k \mid}$$

$$p_k := 1 - q_k$$

The probability for a conversation containing at least k calls is the product of the probabilities $p_1, .., p_{k-1}$, whereas exactly k calls is the product of the probabilities $p_1, .., p_{k-1}, q_k$.

We calculated the probabilities based on our measurements. We mixed all talk groups to achieve a reliable sample size. Table 2 shows the results. The maximum amount of calls per conversation we observed was 28. However, the amount of conversation with a larger amount of calls is quite small. More than 97% of all calls are part of a conversation with a call count up to ten.

i	p_i	q_i
1	0.523972602739726	0.476027397260274
2	0.722875816993464	0.277124183006536
3	0.775768535262206	0.224231464737794
4	0.701631701631702	0.298368298368298
5	0.714285714285714	0.285714285714286
6	0.725581395348837	0.274418604651163
7	0.576923076923077	0.423076923076923
8	0.8	0.2
9	0.708333333333333	0.291666666666667
10	0.784313725490196	0.215686274509804
11	0.65	0.35
12	0.769230769230769	0.230769230769231
13	0.8	0.2
14	0.6875	0.3125
15	0.909090909090909	0.090909090909091
16	0.6	0.4
17	0.666666666666667	0.333333333333333
18	0.75	0.25
19-22	1	0
23	0.333333333333333	0.666666666666667
24-26	1	0
27	0	1
mean	0.74812991038259	0.25187008961741
mean(1:10)	0.703368590200826	0.296631409799174
mean(1:11)	0.698516900182569	0.301483099817431

Table 2. Probabilities for n-state model

After having found probabilities for the transitions between the different states of the model, we want to find distributions for the holding times in the different states. In theory, it is possible to use different distributions for the holding time in different steps k. However, for this approach a large enough amount of samples for each state in each step is necessary. But, especially for larger k the amount of samples is naturally smaller. The amount of samples for larger k we observed was too small to perform an appropriate fitting. Thus, we used equal distributions in each step k. Due to first tests, we only expect slight changes using different distributions. However, we plan to evaluate it for future work.

Furthermore, the distributions of call channel holding times and call idle times for different talk groups were quite similar. This could be expected, as different talk groups differ mainly in the traffic intensity and this is mainly influenced by the conversation idle time. Thus, for the advantage of larger samples we merged the samples of different talk group's call holding and call idle times.

For the conversation idle time we again use the talk group with the highest traffic intensity. It makes no sense to merge the samples for this parameter, because it depends on the traffic intensity which differs for the talk groups. As explained in section 4.2 this parameter is also influenced by the heavy load periods. Thus, we decided to restrict only the conversation idle time samples to the heavy load period.

Fitting all three samples using similar methods as in the previous section the results in table 3 are achieved. For the

Distribution	Parameters	K-S dist.	p-value
Call CHT			
exponential	rate = 0.4708984	0.1383	$< 2.2e\text{-}16$
lognormal	meanlog = 0.4438584 sdlog = 0.7766512	0.0187	0.1041
weibull	shape = 1.264021 scale = 2.308234	0.0758	$< 2.2e\text{-}16$
gamma	shape = 1.7639609 rate = 0.8304458	0.0627	6.328e-15
Call idle			
exponential	rate = 1.049609	0.1185	$< 2.2e\text{-}16$
lognormal	meanlog = -0.4093861 sdlog = 0.9831805	0.0628	5.756e-10
weibull	shape = 1.311591 scale = 1.033244	0.038	0.0006481
gamma	shape = 1.530391 rate = 1.606532	0.0314	0.008286
Conv idle			
exponential	rate = 0.06850586	0.1865	9.686e-07
lognormal	meanlog = 2.2168122 sdlog = 0.8864502	0.1163	0.007034
weibull	shape = 1.016635 scale = 14.725900	0.1807	2.368e-06
gamma	shape = 1.21680036 rate = 0.08331615	0.1505	0.0001539
Conv idle	**shifted**		
exponential	rate = 0.08618164	0.2045	5.158e-08
lognormal	meanlog = 1.390844 sdlog = 1.628825	0.0521	0.6225
weibull	shape = 0.6906014 scale = 8.8397344	0.0606	0.4267
gamma	shape = 0.58495578 rate = 0.05040292	0.0806	0.1320

Table 3. Fitting of call holding, call idle, and conversation idle times for n-state model

call holding time the best fit was achieved by the lognormal distribution. The results differ from the ones in the previous section, because this time we consider the samples of all talk groups.

For the conversation idle time it was not possible to achieve acceptable fitting due to the fact that values smaller than three seconds do not exist. Therefore, we shifted the values for three seconds and fitted the distributions. The best fit was achieved by the lognormal distribution with acceptable p-values. Note, that the random variates produced by this distributions have to be reshifted (plus 3 seconds). For the call idle times the best fit is achieved by the gamma distribution, but the p-value is quite low. One reason for this may be that the amount of samples is still too small. However, the impact of the call idle time is rather small. Thus, we used the gamma distribution which showed the best fit.

5.3 Three-State Semi-Markov Model

As the channel holding and call idle times are modelled identical, the question is whether a simpler model provides already satisfying results. Based on the model in the previous section, it is also possible to model the traffic using a model with reduced complexity. Figure 8 shows the three-state markov or semi-markov model. The basic idea is sim-

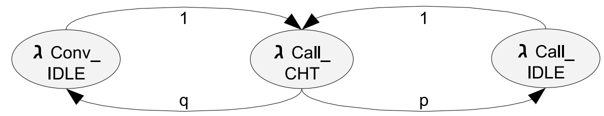

Figure 8. Three-state (semi-)markov model

ilar to the model described in the previous section. After a call with call holding time λ_{Call_CHT} (state in the middle) there may be either a short idle time or a longer one. Thus, the correlations are still considered. But, the difference is that the count of short calls in a row is not modelled due to a memoryless property of this model. The probabilities p_k and q_k do not change for different k. After k short idle times it is as probable that a long one follows as it is after the first short idle time. This model implies independence.

As transition probabilities p and q for the three-state model a mean value over $p_1, .., p_k$ respectively $q_1, .., q_k$ can be chosen. Considering also q_i and p_i for larger i has a strong impact on the probabilities p and q and may falsify the results (cf. mean values in table 2). As mentioned above more than 97% of the calls are part of a conversation with a call count smaller than 11. Thus, we chose q = 0.3 and p = 0.7 for this model.

Note: in theory there is no maximum amount of calls per conversation in the three-state model in contrast to the n-state one. However, the probability for a larger amount of calls per conversation is arbitrary small.

6 Evaluation

In this section the different models and their impact are examined. First we show visual and statistical analysis. Finally, we perform exemplary simulation to estimate the impact of the different models on simulative network performance analysis.

6.1 Visual Evaluation

The goal of the visual evaluation is to show, whether the generated traffic produced by the different traffic models differs. A further question is which model fits best to the measured data.

We generated a traffic trace of 1,000,000s length for each model with the parameters described in the previous section. To examine statistical properties we calculated traffic intensities over intervals of 1, 10, 100, 1,000 and 10,000 seconds. Figure 9 shows the traffic intensity for t = 100s for the different models. It can be seen that the two-state markov model (exponential distributions) is less varying and has a significantly higher average value. The two-state

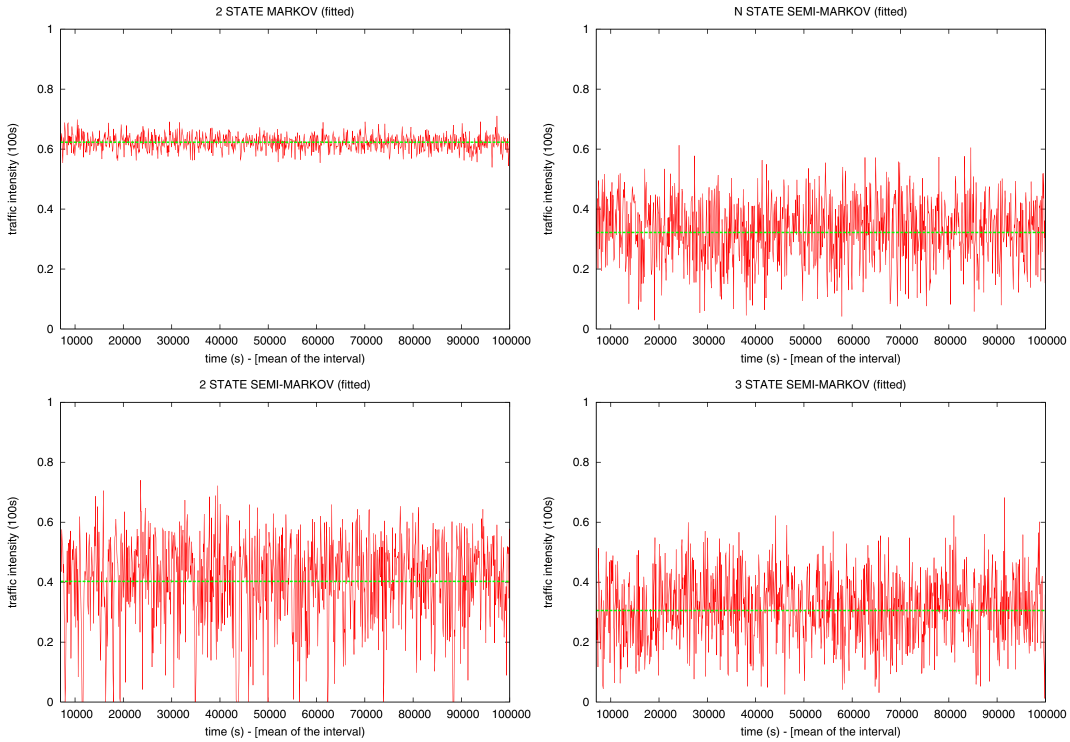

Figure 9. Traffic Intensity for $t = 100$

semi-markov model (lognormal distributions) also shows a higher average value than the other two which consider correlations.

After this visualization in large scale, we take a look at the details. Figure 10 shows call holding times over the summed idle times for a short piece of trace. The three-state and n-state models show conversational characteristics more clearly compared to the two-state models. Compared to the measured data, there is hardly any visual difference between three-state and n-state model.

As statistical measure we looked at the call count per conversation. Table 4 shows the average call counts per conversation and the respective 0.95 confidence intervals. The observations of the visual comparisons are confirmed. The model that achieves the value closest to the measured data again is the three-state semi-markov model.

Model	average	0.95 conf. int.
2 markov (fitted)	719908.000000	0.000000
2 markov (means)	1.919662	0.009419
2 semi (fitted)	3.450146	0.026590
n semi (fitted)	3.770077	0.014317
3 semi (fitted)	3.334364	0.026240
trace	2.902873	0.146910

Table 4. Count of calls per conversation

6.2 Simulative Evaluation

In this section we study whether the different models have an impact on simulative network performance analysis. For this purpose, we create a simple scenario and perform simulations using ns-2.29. The scenario contains 100 nodes in a 350mx200m area moving according to the Reference Point Group Mobility Model [8] with a speed between 0.5 and 3 m/s, a group size of five nodes without any migration from one group to the other, and 10m max. distance from the group center. The transmission range was set to 100m and we used the TwoRayGround propagation model.

As routing protocol we used simple flooding. We modelled the voice traffic according to the MELPe [4] codec with 1.2kbps and Forward Error Correction (FEC) of 1:2. This results in 21 byte IP payload every 67.5ms. For each new call a sender was selected randomly (equally distributed).

This distribution is certainly unrealistic concerning our impressions from the maneuver. However, this is not expected to have an impact on the results of the analysis. Although, the distribution of the senders is expected to have an impact in some scenarios e.g. when a reactive routing protocol is evaluated. While searching for a new route an existing one can possibly be reused. But, in our simulation we use simple flooding.

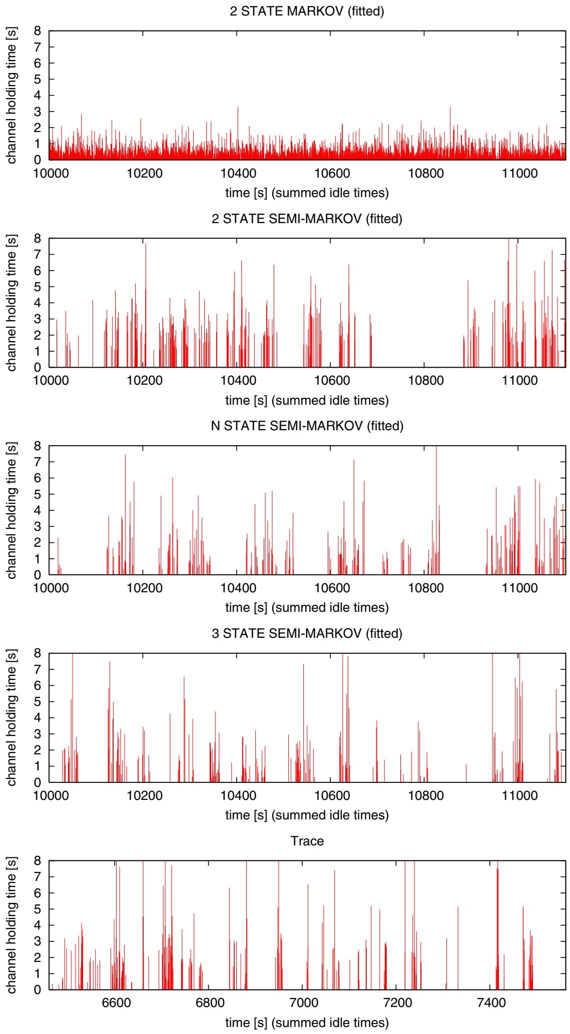

Figure 10. Call holding times over the summed idle times for different models

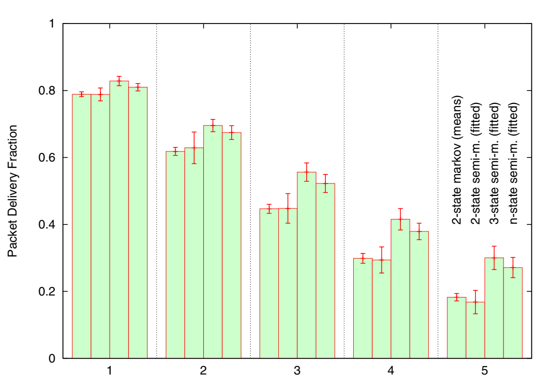

Figure 11. Packet delivery fraction for different models

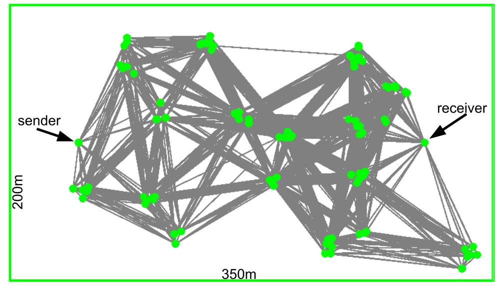

Figure 12. Snapshot of the scenario

We performed simulations with up to five parallel talk groups for each model. As simulation time we chose 3,000s and performed ten replications for each model. We chose such a large simulation time, because we wanted to simulate a large part of the generated traffic continuously. Taking shorter simulation time and choosing traffic randomly from the traces might have negative impact concerning the correlations.

It was not possible to perform similar comparative trace-based simulations. The stationary periods of the traces we observed are too short. It would only have been possible to perform shorter simulations. However, these results would not be comparative and have equal problems of shorter simulations described above.

To examine whether the different modelling of voice streams has an impact on the performance of the network, the impact on the streams themselves has to be evaluated. Thus, we examine the impact of the different models on a constant traffic stream. Therefore, we added two additional static nodes: one at (50;100) and another at (300;100). (cf. figure 12). The first sends a constant voice stream modelled after MELPe for the whole simulation time. These two nodes (and also the stream) can be regarded as a reference talk group with fixed parameters. This approach helps to avoid negative impacts of other characteristics and enables to examine the impact of the different traffic models to one talk group. By choosing any stream the results would be influenced by the specific characteristics of this stream (distribution of traffic, position of the senders, idle times, etc.).

We examined the packet delivery fraction on the reference stream in combination with the transmission delay. The packet delay has a decisive impact on the packet delivery fraction. A packet that is too late will not be of any use for the voice data communication. Thus, we assumed a packet with a transmission delay larger than a threshold as lost. According to [9] a time of 150ms is advised as a threshold.

Figure 11 shows the average packet delivery fraction of the reference stream and 0.95 confidence intervals for ten replications for each model and amount of talk groups. In general, it can be seen that the more talk groups are simulated the smaller the packet delivery fraction gets.

However, the different traffic models have different impact, too. The two-state models cause a smaller packet delivery fraction at the background stream. The reason is that the traffic of the two-state models is less bursty. Therefore, the background stream is influenced permanently and shows a larger delay. Using the 3-state and n-state model (with more bursts and gaps) results in better packet delivery fractions because the gaps of one talk group are filled up by the bursts of another. Thus, the overall load is smaller on average. The difference between the two-state models in contrast to the three- and n-state models is clearly visible. For more than two talk groups the results are significant (confidence intervals). There is no significant difference between the traffic of the three- and n-state models.

The aim of this simulative evaluation was to examine whether the different models have an impact when used in network simulation. The simulations show that the traffic has an impact on the packet delivery fraction of the background traffic especially when multiple talk groups are present.

7 Conclusion and Future Work

In this paper, we analyzed traffic that was measured during a catastrophe maneuver. Based on this traffic, we calculated time series for single talk group push to talk traffic. The time series were analyzed with respect to dependencies and heavy load periods. The results of this analysis led to three traffic models with different degree of abstraction. We fitted standard distributions to the different models using the samples of our measurements. Exponential distributions do not show a good fit. Lognormal distributions provide a better fit.

Finally, we generated traffic streams for the different models. These streams were evaluated visually and statistically. Furthermore, the impact of the different models on network simulation was examined. The two-state models show larger differences to the measured data than three-state and n-state model because the conversational correlations observed are not considered. There is no significant difference between the n-state and the simpler three-state model. Thus, there is no need to model the count of calls per conversation in detail. To model voice communications in disaster area scenarios considering conversational correlations, we propose to use a three-state semi-markov model with lognormal holding times.

In the future, we plan to perform further (especially longer) measurements to examine further statistical measures like long range dependency (self-similarity). Furthermore, we plan to examine the impact of the conversation idle fraction as a measure of load. We assume to be able to categorize different talk groups and scenarios using this measure. Additionally, we want to carry out a simulative performance analysis e.g. of routing protocols in disaster area scenarios.

References

[1] F. Barceló and S. Bueno. Idle and Inter-arrival Time Statistics in Public Access Mobile Radio (PAMR) Systems. *IEEE Globecom*, 1997.

[2] P. T. Brady. A Technique for Inestigating On-Off Patterns of Speech. *The Bell System Technical Journal, Volume XLIV, Number 1*, pages 1–22, 1965.

[3] P. T. Brady. A Model for Generating On-Off Speech Patterns in Two-Way Conversation. *The Bell System Technical Journal, Volume 48, Number 9*, pages 2445–2472, 1969.

[4] J. Collura and D. Rahikka. Interoperable Secure Voice Communications in Tactical Systems. *IEE Collquium on Speech Coding Algorithms for Radio Channels*, 2000.

[5] N. Haslett and A. Bonney. Loading Considerations for Public Safety Dispatch on Trunked Radio Systems. *IEEE Transactions on Vehicular Technology*, pages 24–31, 1987.

[6] G. Hess. *Land-Mobile Radio System Engineering*. Artech House, 1993. pages 269-273.

[7] G. Hess and J. Cohn. Communications load and delay in mobile trunked systems. *Proc. IEEE Vehicular Technology Conf.*, pages 269–273, 1981.

[8] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang. A Group Mobility Model for Ad Hoc Wireless Networks. *MSWiM*, 1999.

[9] Transmission Systems and Media - General characteristics of international telephone connections and international telephone circuits. ITU-T G.114, 1996.

[10] Artificial Conversational Speech. ITU-T P.59, 1993.

[11] J. Jordán and F. Barceló. Statistical Modelling of Channel Occupancy in Trunked PAMR Systems. *Proc. 15th Int. Teletraffic Conference*, pages 1169–1178, 1997.

[12] J. Jordán and F. Barceló. Statistical Modelling of Transmission Holding Time in PAMR Systems. *IEEE Globecom*, 1997.

[13] A. M. Law and W. D. Kelton. *Simulation Modeling and Analysis*. McGraw-Hill, 3rd ed. edition, 2000.

[14] D. Sharp, N. Cackov, N. Lasković, Q. Shao, and L. Trajković. Analysis of Public Safety Traffic on Trunked Land Mobile Radio Systems. *IEEE J-SAC, vol.22, no.7*, pages 1197–1205, 2004.

[15] G. Stone. Public safety wireless communications user traffic profiles and grade of service recommendations. *U.S. Dept. Justice, SRSC Final Report, Appendix D*, 1996.

[16] P. Vary, U. Heute, and W. Hess. *Digitale Sprachsignalverarbeitung*. Teubner, Stuttgart, 1998. [in German].

[17] B. Vujicić, N. Cackov, S. Vujicić, and L. Trajković. Modeling and Characterization of Traffic in Public Safety Wireless Networks. *Proc. SPECTS*, pages 214–223, 2005.

The authors would like to thank the Berufsfeuerwehr Köln (Fire Department Cologne), as well as the Johanniter-Unfall-Hilfe Bonn for providing valuable support during the traffic measurements.