# Application Performance Assessment on Wireless Ad Hoc Networks

Răzvan Beuran[1,2], Ken-ichi Chinen[2,1], Khin Thida Latt[2], Toshiyuki Miyachi[2,1], Junya Nakata[2,1], Lan Tien Nguyen[2], Yoichi Shinoda[2,1], and Yasuo Tan[2,1]

[1] National Institute of Information and Communications Technology, Hokuriku Research Centre, 2-12 Asahidai, Nomi, Ishikawa, 923-1211 Japan
[2] Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan
razvan@nict.go.jp

**Abstract.** Using ad hoc networks as alternative means of communication in disaster situations is a salutary solution. However, analysing application performance is mandatory for evaluating such a possibility. In this paper we present the two aspects of our approach to application performance assessment on wireless ad hoc networks. The first aspect refers to real-world tests in which we quantify objectively the relationship between network conditions and application performance. The second aspect is represented by the wireless network (WLAN) emulator that we design to run on StarBED, the large-scale network experiment environment of the Hokuriku Research Centre in Ishikawa, Japan. By combining these two aspects we perform experiments with real applications, while having full control of the network conditions in which the application is tested (when using emulation).

**Keywords:** Emergency networks, ad hoc networks, application performance assessment, WLAN emulation.

## 1 Introduction

Emergency communication systems are required in the preceding phase of disaster situations to issue warnings and evacuation instructions. However they are equally decisive during catastrophes and after their occurrence to coordinate the activity of rescue teams. Dependable communication is also essential in mission-critical and safety-critical systems, or even in normal business environments that require “anytime, anywhere” access to network resources.

According to a study in [1] the requirements for emergency services’ mobile communication are: resilience, coverage, access and capacity, security, regulation, group communication, fast call set-up, priority, direct mode/repeaters and gateways, integration with control room, and voice quality. The conclusion of this study is that public mobile networks, such as the GSM (Global System for Mobile Communications) standard, even with the latest proposed modifications, cannot fulfil all of the above requirements.

Public safety agencies have used specialized radio communication systems for many decades. TETRA (TErrestrial Trunked Radio) [2] is a specialist professional

K. Cho and P. Jacquet (Eds.): AINTEC 2006, LNCS 4311, pp. 128 – 138, 2006.

mobile radio used by emergency workers, such as police, fire departments, ambulance and military. Although this system was designed to meet the cited requirements, it still has several disadvantages, such as low rates of 4.8 kb/s, expensive handsets, etc. In addition its deployment is predominant only in Europe for the moment.

Unfortunately recent disasters such as the Indian Ocean tsunami in 2004, or the 2005 hurricanes Katrina and Rita in U.S.A. have shown that current communication systems, both professional and public, fail too easily under emergency conditions. Therefore new alternative means of communication are needed for emergency or calamity situations. Given the current spread of Internet one such alternative is to make a more extensive use of the technologies associated with Internet, and in particular of ad hoc wireless LANs (WLANs).

The U.S.A. federal government report on the response to hurricane Katrina [3] proposes the creation of a National Emergency Communication Strategy. One of the recommendations included in this report mentions that "there is a strong need for rapidly deployable, interoperable, commercial, off-the-shelf equipment". The same event showed how Internet-based technologies could be used to establish links with the outside world. As an anecdotal example, the only communication means between the mayor of New Orleans and the outside world, including phone calls with U.S.A. President George Bush, was a wireless Internet connection and an Internet phone account [4]. Although we will not make here a full analysis, it is easy to notice that some of the key requirements for emergency communication systems find indeed built-in support in the WLAN technology. Capacity, security, regulation, priority, direct-mode communication are all available at present in wireless LANs.

Using converged WLANs makes it possible to transmit both audio and video information, so that rescuers can communicate with each other and with remote experts. Moreover, by using WLANs emergency workers can receive data, such as street maps or building floor plans. All these means of communication that make use of a unique network infrastructure are crucial for saving lives and preventing losses.

Wireless LANs are more stable than other communication infrastructures given the fact that they are decentralized. As their potential failure is independent, ad hoc WLANs can continue functioning in emergency conditions. The nodes of WLANs have low costs, and they require little power to operate. In addition, the potential of using advanced features on WLANs that are not available in traditional public communication systems makes it possible to provide probabilistic guarantees of service for emergency responders. Using a priority-enforcement system such users could be given an assured service level, independently of the activity of regular users.

In this paper we present our approach to assessing application performance in WLAN environments, in ad hoc mode, as well as in "infrastructure" mode (i.e., access point based deployments). An objective performance assessment is a mandatory step in analyzing the dependability of applications on WLAN in view of their use as alternative means of communication in emergency situations. According to the survey we did in [5] there are several factors that currently impede application performance on wireless LANs:

1. WLAN QoS parameters (bandwidth, packet loss, delay & jitter) have a high variability in real-world environments;
2. Existing WLAN QoS mechanisms are only of limited use for managing contention when applications with different QoS requirements, such as VoIP (Voice over IP) calls and TCP-based data traffic, share the same communication channel;

3. Real-time applications such as VoIP or video communication require timely servicing of the traffic; this is a challenging task in WLANs, even when using QoS enforcement, since most currently-implemented QoS mechanisms focus only on bandwidth provisioning;
4. Roaming between access points introduces communication gaps that may even be of the order of seconds, an unacceptable situation for real-time applications.

The paper is structured as follows. First we present the analysis methodology that we propose for the study of application performance over wireless LANs. This includes tests in real-world environments as well as the use of WLAN emulation, which is a key complementary element of our approach. Following that we give some illustrative results for our study of VoIP performance on WLAN, since reliable voice communication is one main requirement for emergency communication systems. The paper ends with a section of conclusions and future work.

## 2 Performance Assessment Methodology

The first step to take in studying IP application performance and dependability on WLANs is to define a test methodology that allows assessing objectively application performance, and understanding its dependence on network conditions. The methodology that we propose can be employed for any network application, but for each application under test specific metrics have to be used to assess performance. Objective performance assessment is important since users of WLANs, in either regular or special environments, such as emergency or disaster conditions, require that applications run at a satisfactory performance level. The approach we propose makes it possible to analyse application performance in a wide range of controllable network conditions. By correlating an objective assessment of the User-Perceived Quality (UPQ) for the applications under study with the corresponding network conditions one can determine the reasons of application performance degradation and investigate mechanisms to ensure satisfactory performance. Our approach has two aspects that will be detailed next: real-world tests and WLAN emulation.

### 2.1 Real-World Tests

Tests in real-world environments are one aspect of our approach to application performance assessment. They permits us to capture and analyse the behaviour of real applications in real network conditions. Another use is in the calibration of the emulation system that we develop, which will be described in Section 2.2. Below is a typical experimental setup for real-world WLAN tests. We show here the case when access points are used, but for ad hoc networks the measurement side is identical.

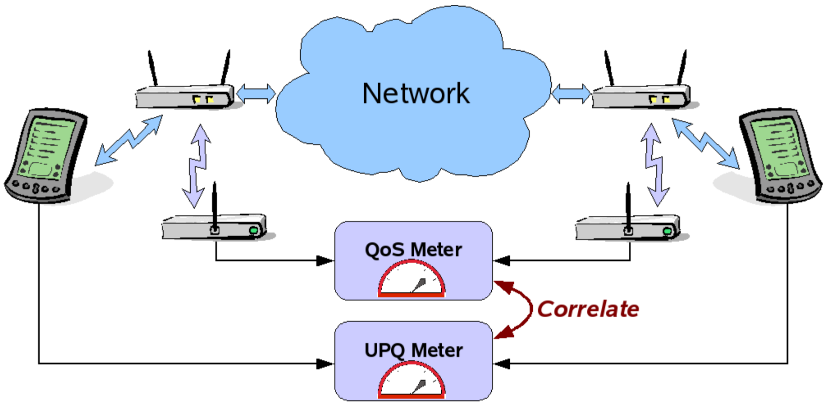


**Fig. 1.** Experimental setup for application performance assessment on WLAN

The setup in Figure 1 is adapted from the system we previously used to study VoIP performance on wired networks [6]. Using this setup we carry out a two-level analysis. At the level of the network we investigate the performance issues of WLANs, such as the dependency on signal strength, on the number of access points or peer ad hoc nodes, on quality degradation management techniques, etc. At this level there are two classes of metrics we use: (i) physical environment metrics, such as the power of the received signal, *Pr*, and the Signal to Noise Ratio, *SNR*; (ii) generic network metrics, which are the three inter-dependent QoS parameters: bandwidth, packet loss, and delay & jitter. These parameters are computed by the "QoS Meter" block that uses as input monitored WLAN traffic traces. For capturing the traffic we use the AiroPeek software of WildPackets, Inc. [7], in connection with wireless probes, such as RFGrabber of the same company, or high-end WLAN adapters, such as ORiNOCO 11a/b/g Gold or Cisco Aironet 802.11a/b/g. Special drivers are required to capture both data, and control & management packets, as well as signal and noise levels in the WLAN.

Simultaneously, at application level, we measure the User-Perceived Quality (UPQ) for the application under study. In the case of VoIP we use methods such as the ITU-T recommendations G.107 [8] and P.862 [9]. For file transfers, metrics such as the goodput and the transfer time performance are well suited [10]. In the case of video communication one can use a tool such as the Psytechnics Video Agent for Communications, from Psytechnics, Ltd. [11]. In our setup the function of measuring user-perceived quality is conceptually performed by the "UPQ Meter" block.

Correlating the WLAN-level and application-level performance permits us to establish objectively what are the requirements of a network application in order to ensure user satisfaction, as well as determine what type of mechanisms are needed to meet these requirements on WLANs.

One important aspect of this setup is its ability to capture the dynamic behaviour of the tested systems, which is made possible through the use of the above-mentioned WLAN traffic monitors (sniffers). Running averages and global assessments are not meaningful for short-term performance issues, which are nevertheless critical in disaster situations. Moreover WLANs are dynamic environments by definition: signal reception conditions fluctuate, the number of nodes and their position vary, the access points with which the nodes communicate, or their peers in ad hoc networks, change. By capturing the dynamic behaviour of the network we can follow and understand application performance fluctuations over time.

### 2.2 WLAN Emulation

Real-world experiments are only one aspect of our research. Such experiments are very useful for understanding the behaviour of real WLAN systems. However the range of conditions that can be tested in real-world experiments is limited and difficult to control. Therefore we proceeded to design a WLAN emulator that gives us direct control over the network conditions which we use to assess application performance.

The WLAN emulator that we develop makes use of StarBED [12]. StarBED is the large scale network experiment environment of National Institute of Information and Communications Technology (NICT), Hokuriku Research Centre in Ishikawa, Japan. This experiment environment is a cluster-based testbed currently employing about 700 PCs. The use of a custom-designed experiment-support software, nicknamed SpringOS, makes it possible to define complex experiments on StarBED in a straightforward manner.

A survey of existing WLAN-related real-world and simulation testbeds is available in [13]. General characteristics of WLAN emulators are summarized in [14]. Our design addresses these requirements through a scenario-driven two-stage approach inspired by [15] and [16]. The novelty of our work consists in the quality degradation view we take, the emphasis we lay on emulation realism, as well as the use of a large-scale testbed for experiments.

The setup we use when testing application performance through WLAN emulation is presented in Figure 2. Notice the similarity that exists with the test setup discussed in Section 2.1.

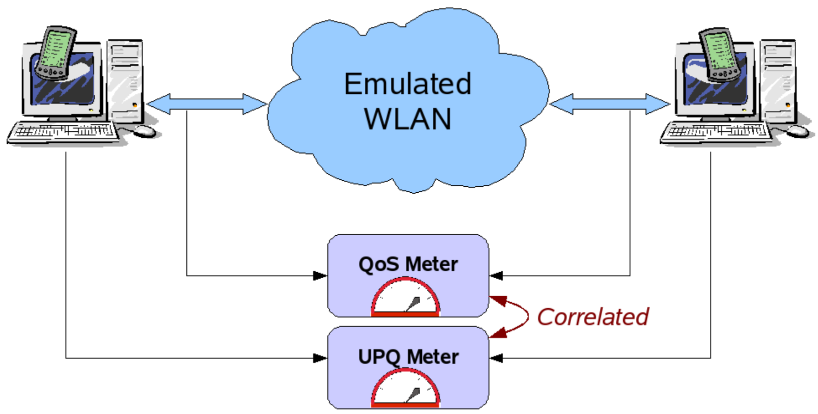


**Fig. 2.** WLAN emulation for application performance assessment

In the setup depicted in Figure 2 we use PCs to play the role of the mobile nodes in the emulated WLAN. The applications that run on the end PCs are of the same type with those running on a PDA, or a wireless phone. Therefore by using this setup we can test the performance of the same applications as those in the real-world experiments. Note that in the context of emulation too we measure the QoS parameters and the UPQ. Correlating these two measures permits in this case also to determine objectively the relationship that exists between application performance and network quality degradation under varying network conditions. The difference with respect to the real-world tests is that in this setup the full control over the scenarios and conditions that we use during the tests is readily possible.

There are several key requirements for such a methodology to be effective, and our approach addresses them. Realism of the emulation is important because it allows drawing conclusions that are thereafter useful in real-life deployment. Most application performance studies focus on data-transfer applications, typically based on TCP/IP. Models used are oversimplified in a bandwidth-oriented perspective, hence do not adequately reflect real network conditions as experienced by users, and only reproduce a simplified theoretical behaviour of WLANs. However the edge effects that do occur in reality are generally overlooked, for example between rate changes, as the WLAN cards automatically adapt operation rate to signal reception conditions using a mechanism such as Auto-Rate Fallback (ARF) [17]. Nevertheless these effects have significant consequences at application level, as our example in Section 3 shows. Such aspects are taken into account by the WLAN operation models that we employ. Moreover, in order to achieve an even higher degree of realism we propose to calibrate our system by combining observations and traffic traces of real WLANs with our analytical model of the WLAN environment, so that the calibrated model accurately describes the observed network behaviour.

The scenario-driven architecture we propose here has two stages. In the first stage, from a real-world scenario representation we create a network quality degradation ($\Delta Q$) description which corresponds to the real-world events (see Figure 3). The $\Delta Q$ description represents the varying effects of the network on application traffic. Computing the $\Delta Q$ description is the key element of our system. This step is achieved by using a model of the WLAN communication, in particular by taking into account the properties of the physical layer and the data link layer (more specifically the MAC sub-layer) of the WLAN technology. For the physical layer the main model used is the log-distance path loss model [18]. For the MAC sub-layer we deploy several models for determining first frame error rates, and then packet loss rates, delay and jitter, as well as bandwidth variation according to the ARF mechanism.

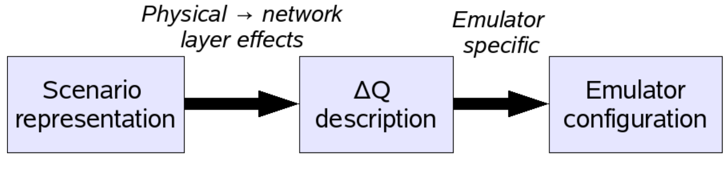


**Fig. 3.** Two-stage scenario-driven emulation

Subsequently the $\Delta Q$ description is converted into an emulator configuration which is used during the effective emulation process to replicate the user scenario. This allows the study of scenario effects on the real application under test, all this under realistic conditions.

As an example, assume that the scenario representation describes how, from an initial position, two mobile nodes move with respect to each other. As a consequence of the motion, the received radio signal strength will change. This is the WLAN physical layer effect, and it is quantified by means of the received signal power, *Pr*. We model the dependency of *Pr* on the distance between transmitter and receiver using the log-distance path loss model. At the moment attenuation in only considered from the point of view of path loss due to distance and shadowing. We are now working on integrating models of fading (e.g., Rayleigh fading) into our system.

WLAN performance is not only affected by the received signal power, *Pr*, but also by the environmental noise, which is quantified by means of the Signal to Noise Ratio, *SNR*. The effects of the received signal power and the noise are accounted for by using error models for the corresponding signal encodings, models which allow us to compute the corresponding bit error rate and frame error rate. These error models are based on typical receive sensitivities of WLAN adapters, as well as experimentally-determined characteristics of some widely-used WLAN-adapter components, such as the Intersil HFA3861B Direct Sequence Spread Spectrum Baseband Processor [19].

The variation of *Pr* and *SNR* causes quality degradation at the data-link layer as follows. The weaker signal first induces higher error rates and consequently packet loss. Simultaneously there is an increase in delay as the number of retransmissions becomes larger. Finally the WLAN adapters change the channel encoding and rate, according to a mechanism such as ARF, and the available bandwidth diminishes. These effects are modelled by taking into account the specific 802.11 MAC sub-layer, as well as the behaviour of ARF; the result is the $\Delta Q$ description associated to the emulated scenario.

The second stage of our approach is emulator specific. In this stage the generic $\Delta Q$ description calculated previously is converted into an emulator configuration. Initially the conversion target is the *dummynet* network emulator [20]. However by decoupling the two conceptually-independent emulation stages we make it possible that later on other wired-network emulators can be used as well, that are more accurate than *dummynet* and that have a richer set of features. Note that the possibility of using various network emulators running on Linux or FreeBSD is a feature that has been recently added to StarBED.

## 3 Experimental Results

In this section we show some illustrative results for VoIP performance over WLAN. We chose to study VoIP as a representative real-time application, since voice communication is one of the most important forms of communication in emergency situations.

As mentioned before, two methods are more appropriate to measure VoIP UPQ: the R-value score, which is the result of the E-model described in the ITU-T

Recommendation G.107, and the PESQ score, proposed by the ITU-T Recommendation P.862. A detailed comparison between the two metrics is available in [5]. For the purpose of this paper it is important to note that the R-value is computed uniquely based on traffic measurements, therefore it is only a *prediction* of what the user satisfaction would be if using the corresponding connection. On the other hand the PESQ score is calculated using voice recordings for the measured connection; hence it is a more accurate *estimate* of the VoIP UPQ for that connection. Note that the R-value can be converted to the Mean Opinion Score (MOS) scale [21], which is also used by the PESQ score. This makes it easy to compare results obtained using the two metrics. A value of 4.5 on the MOS scale indicates optimum quality, with good quality being associated to scores higher than 3. Quality is considered acceptable for scores between 2 and 3, whereas scores lower than 2 indicate unacceptable quality.

The results presented below were obtained using the setup in Figure 2. For simplicity reasons we consider a scenario with only two nodes. However it is possible to use the current system to study multi-hop environments as well.

In our scenario example the two nodes are at an initial distance of 10 m with respect to each other. One of them starts moving on a perpendicular direction with a speed of 0.5 m/s for a time interval of 30 s. This is a scenario fragment representative for the communication between an ad hoc network user moving in a building while making a VoIP over WLAN phone call to another user.

With the parameters $\alpha = 5.5$ (difficult reception conditions) and $Pr0 = -20$ dB in the log-distance path loss model we obtain the power of the received signal, *Pr*, shown in Figure 4 together with the distance between the mobile nodes. Horizontal dashed lines indicate the thresholds between minimum signal power levels for different operating rates of 802.11b (from top to bottom, the limit between 11 and 5.5 Mb/s, and the limit between 5.5 and 2 Mb/s, respectively). Note that *Pr* falls under the 11 Mb/s threshold after $t \approx 17.5$ s, and below the 5.5 Mb/s threshold at $t \approx 26$ s. The consequences of the *Pr* variation on the quality degradation were computed as $\Delta Q$ descriptors. Using them as the input of the ITU-T E-model we calculated the R-value, and predicted what the user satisfaction would be if communicating under such circumstances.

The next step was to use the same $\Delta Q$ descriptors to drive the network emulator *dummynet* while real voice data was sent through the network. We used a customized version of the SpeakFreely 7.6a application [22] that we modified to save the output voice signal. For this test SpeakFreely was configured to make use of the codec G.711 [23]. The voice input consists of standard voice test files supplied with the ITU-T P.862 recommendation. Based on the input and output voice signals we computed the PESQ score. Both the PESQ score and the MOS R-value (the MOS score obtained analytically from the previously computed R-value) are represented in Figure 5.

We can observe in Figure 5 that the MOS R-value has generally higher (i.e. more optimistic) values than the PESQ score, which is deemed to be more realistic. In the case when $t < 17$ s signal conditions are good, therefore the average packet loss was considered low (1%). Hence the MOS R-value is stable around a value of 4.3. However the PESQ score has a much more evident variation over the same interval. The explanation consists in the fact that while average packet loss values over a long period might be 1%, over shorter timescales and for a small number of packets the loss values change significantly, and this change leads to a important perceived quality variation, as it would be observed in a real situation.

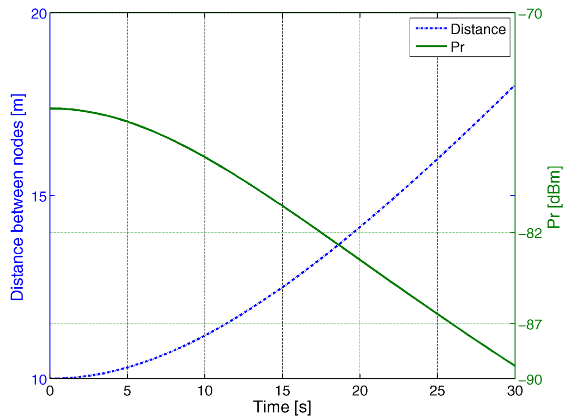


**Fig. 4.** Distance between mobile nodes and the power of the received signal versus time

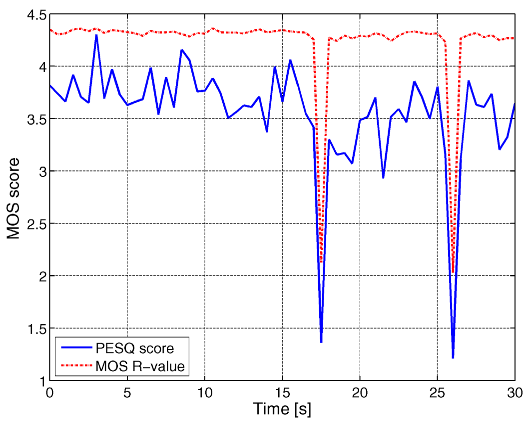


**Fig. 5.** PESQ score and the MOS R-value versus time

As remarked previously, at times $t \approx 17.5$ s and $t \approx 26$ s the decreasing level of the received signal power triggers rate transitions, with obvious effects on VoIP UPQ. At these moments the quality drops under the threshold of acceptable quality (MOS value = 2). However these effects are not caused by the rate change in itself (which only goes down to 2 Mb/s in the worst case, largely sufficient for the roughly 80 kb/s stream generated by the G.711 codec). On the contrary, these effects are produced by the packet loss and delay variation that occur just before and during the rate transition, and that are ignored in most WLAN emulation implementations. Moreover, once *Pr* falls under the optimum reception threshold for 11 Mb/s, the auto-rate fallback mechanism in 802.11 is triggered, and this causes a slight additional quality degradation, which is noticeable after the first rate transition.

## 4 Conclusions

Disaster situations impose strong requirements on communication systems, and in such cases dependability and performance guarantees are essential. This need is addressed by our research in the area of application performance on ad hoc and infrastructure mode WLANs. The two directions of our research are application performance analysis and assurance.

As a first step we started using real-world tests and a WLAN emulated environment to determine the relationship between the events in the physical world associated with WLANs, the corresponding network quality degradation and in the end the user perceived quality variation for network applications. Our methodology lays emphasis on dynamic behaviour capture and emulation realism.

The approach we propose for WLAN emulation allows the transformation of a user-meaningful real-world representation of a WLAN environment (termed "scenario representation") into a network quality degradation description (termed "$\Delta Q$ description"). The $\Delta Q$ description is sufficient to subsequently configure a wired-network emulator and effectively reproduce an environment that corresponds accurately at network level to the emulated WLAN scenario.

We illustrated the practical use of our approach through a simple real-world scenario, for which we determined the induced network quality degradation. The selected application for this experiment was VoIP, since speech is one of the most important forms of communication in emergency situations. We quantified the influence of the quality degradation on VoIP UPQ in an objective manner using ITU-T recommendations concerning expected user satisfaction for VoIP communication.

As a second step of our future research we will investigate application performance assurance. We intend to use the same setup to study the issue of application performance assurance in WLAN environments, which is vital in emergency conditions, such as disaster rescue operations and other critical situations. The framework and the specific techniques that we plan to develop will allow the creation of ad hoc WLANs under critical conditions, and the assurance of quality guarantees even under such circumstances for high-priority users (e.g., public safety teams, hospitals, etc.). The IEEE 802.11e standard for QoS on WLAN will be analysed first to determine whether it is suited for use in the context of emergency communication systems.

## References

1. Swan, D.: Analysis in the ability of Public Mobile Communications to support mission critical events for the Emergency Services. White paper, Mason Communications Ltd, Manchester England, (March 2003)
2. TETRA: TETRA Memorandum of Understanding. http://www.tetramou.com
3. Towsend, F.F. (ed.): The Federal Response to Hurricane Katrina – Lessons Learned. U.S.A. Federal Government Report (February 2006)
4. Rhoads, C.: After Katrina, city officials struggled to keep order, In: The Wall Street Journal, (9 September 2005)
5. Beuran, R.: VoIP over Wireless LAN Survey. Research report, IS-RR-2006-005, Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan (April 2006)

6. Beuran, R., Ivanovici, M.: User-Perceived Quality Assessment for VoIP Applications. Technical report (delivered to U4EA Technologies), CERN-OPEN-2004-007, CERN, Geneva, Switzerland (January 2004)
7. WildPackets, Inc.: http://www.wildpackets.com
8. ITU-T: ITU-T Recommendation G.107: The E-model, a computational model for use in transmission planning (March 2005)
9. ITU-T: ITU-T Recommendation P.862: Perceptual evaluation of speech quality (PESQ), an objective method for end to end speech quality assessment of narrow-band telephone networks and codecs (February 2001)
10. Beuran, R., Ivanovici, M., Dobinson, B., Davies, N., Thompson, P.: Network Quality of Service Measurement System for Application Requirements Evaluation. In: Proc. of International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS'03, Montreal, Canada (July 2003) 380-387
11. Psytechnics, Ltd.: http://www.psytechnics.com
12. Miyachi, T., Chinen, K., Shinoda, Y.: Automatic Configuration and Execution of Internet Experiments on an Actual Node-based Testbed. In: Proc. of Tridentcom2005, Trento, Italy (February 2005) 274-282
13. Kropff, M., Krop, T., Hollick, M., Mogre, P.S., Steinmetz, R.: A Survey of Real-World and Emulation Testbeds for Mobile Ad hoc Network. In: Proc. of TridentCom2006, Barcelona, Spain (March 2006)
14. Kojo, M., Gurtov, A., Manner, J., Sarolahti, P., Alanko, T., Raatikainen, K.: Seawind: a Wireless Network Emulator. In: Proc. of 11th Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), Aachen, Germany (September 2001)
15. Bateman, M., Allison, C., Ruddle, A.: A Scenario Driven Emulator for Wireless, Fixed and Ad Hoc networks. In: Proc. of PGNet2003, Liverpool, U.K. (June 2003) 273-278
16. Perennou, T., Conchon, E., Dairaine, L., Diazet, M.: Two-Stage Wireless Network Emulation. In: Proc. of WCC2004, Toulouse, France (August 2004)
17. Kamerman, A., Monteban, L.: WaveLAN-II: A High-performance wireless LAN for the unlicensed band. In: Bell Lab Technical Journal (summer 1997) 118-133
18. Rappaport, T.S.: Wireless Communications – Principles and Practice. Prentice Hall PTR, 2$^{nd}$ edition (2002)
19. Intersil: HFA3861B Direct Sequence Spread Spectrum Baseband Processor (2000)
20. Dummynet: FreeBSD network emulator. http://info.iet.unipi.it/~luigi/ip_dummynet
21. ITU-T: ITU-T Recommendation P.800: Methods for subjective determination of transmission quality (August 1996)
22. Wiles, B.C., Walker, J.: Speak Freely 7.6a. http://www.speakfreely.org
23. ITU-T: ITU-T Recommendation G.711: Pulse Code Modulation (PCM) of voice frequencies (1993)