Contents

In this issue

szabo@hit.bme.hu

IP mobility, co-existence of Wi-Fi networks with other systems in the license-free bands, enhancement of WiMAX protocols, location management in cellular networks are examples of hot topics in our broad area of infocommunications these days. This editor is happy that several good papers arrived from open call that address the aforementioned issues which now, after a rigorous peer review process, can be presented to our readers in this July issue of our Infocommunications Journal. I hope that our readers will find them interesting and useful for their work.

The first paper of this issue by *Balázs Héder and András Bertók* is titled "Detection of sleet attenuation in data series measured on microwave links". Radio wave propagation on terrestrial high frequency microwave point-to-point links is highly influenced by atmosphere effects, especially by the attenuation of precipitation. Usable models exist for considering rain attenuation, but the statistics of rare sleet event are less known. The measurements carried out by the authors made it possible to model the sleet attenuation. They proposed an algorithm which exploits the important differences between the second order statistics of rain and sleet attenuation time functions.

In the past years it has became a widely accepted opinion that the current network layer protocol, IPv4, suffers from several serious problems.The sixth version of the Internet Protocol (IPv6) is assumed to become the network layer of the future Internet, which unfortunately inherits some weaknesses of IPv4. For instance, there are no easy solutions of the scalability problems caused by edge site multi-homing. In the paper "Multi-Homing in IP Networks Based on Geographical Clusters" the authors, *Gábor Németh, Gábor Makrai and János Tapolcai* propose an easy to deploy multi-homing strategy that decreases the number of entries in the routing tables by aggregating together the address space of several edge networks located close to each other.

Within our series "Design Studies" we publish practical results usually achieved within the framework of large scale European projects. This time we present the achievements of an EU project SAMBA in the paper titled "A Framework for Community-Oriented Interactive Digital Television" by *O. Mayora-Ibarra et al*. This paper presents the SAMBA Framework for community-oriented idtv. The main objective of this framework is to provide local communities and citizens (including low income population) to access community-oriented content and services by means of idtv channels. In order to achieve this, SAMBA implemented several steps such as 1) development of a Content Management System for creating interactive applications that are cross-compatible between DVB-T MHP and DVB-H; 2) considering usability issues related to technology adoption by non expert users and 3) addressing low cost requirements for meeting the needs of low-income users and communities with use of power line communications for providing interactivity to remote rural areas.

Zoltán Horváth and Dávid Varga in their paper suggest a novel channel allocation technique for eliminating interference caused by RLANs on meteorological radars in 5 GHz band. Meteorological radars are used for short term weather prediction in Hungary and all over the world. These radars can be jammed by RLAN devices (e.g. home Wi-Fi routers). The authors introduce the background of this problem, and analyze the weakness of the current solution (DFS – Dynamic Frequency Selection) by modeling the radar operation and RLAN traffic, and also show its high efficiency in practice, based on well-known IEEE 802.11 RTS/CTS mechanism.

The paper by *S. Mumtaz, A. Gamerio and K. Saidul* titled "Enhanced Algorithm for WIMAX: MIESM" extends the already well studied link adaptation technique based on MIESM (Mutual Information based exponential SNR Mapping) for 802.16 based systems. The previous work on MIESM uses equal modulation order for all the subcarriers in an OFDM block. In the paper the concept of unequal modulation for the subcarriers in the single OFDM block is proposed, and a mathematical model is derived based on bivariate Gaussian distribution.

Efficient dimensioning of cellular wireless access networks depends highly on the accuracy of the underlying mathematical models of user distribution and traffic estimations. Mobility prediction also considered as an effective method contributing to the accuracy of IP multicast based multimedia transmissions, and ad hoc routing algorithms. In the paper "Markov model based location prediction in wireless cellular networks" by *Tamás Szálka, Sándor Szabó and Péter Fülöp* the authors focus on the trade-off between the accuracy and the complexity of the mathematical models used to describe user movements in the network. Markovian mobility models are proposed in order to utilize the additional information present in the mobile user's movement history thus providing more accurate results than other widely used models in the literature. The new models are applicable in real-life scenarios, because these rely on additional information effectively available in cellular networks (e.g. handover history), too.

Csaba A. Szabó, Editor-in-Chief

Detection of sleet attenuation in data series measured on microwave links

BALÁZS HÉDER, ANDRÁS BERTÓK

Budapest University of Technology and Economics,
Broadband Infocommunications and Electromagnetic Theory

balazs@mht.bme.hu

Keywords: microwave link, precipitation attenuation, sleet attenuation, first and second order statistics, fade duration, fade slope

Radio wave propagation on terrestrial high frequency microwave point-to-point links is highly influenced by atmosphere effects, especially by the attenuation of precipitation. Usable models exist for considering rain attenuation, but the statistics of rare sleet events are less known. Our measurements on microwave links makes it possible to model the sleet attenuation. Sleet events can be automatically detected with an algorithm which exploits the important differences between the second order statistics of rain and sleet attenuation time functions. This enables to easily separate and collect sleet events, which is essential for model constitution in the near future.

1. Introduction

High frequency microwave links are frequently used in the core network of mobile cellular systems or in any point-to-point or point-to-multipoint terrestrial or satellite systems, so they can be employed as physical links even in a multilink network [1]; however, because of the applied high carrier frequency (above 10 GHz) the wave propagation is highly influenced by precipitation especially by rain and, due to the high precipitation attenuation, the microwave link can even be terminated. In this frequency band mainly the rain and sleet cause significant attenuation, against which different fade mitigation techniques (FMT) must be applied as countermeasure methods. In order to design suitable FMT e.g. to plan fade margin, information about statistics of expected rain attenuation is highly important. Rain attenuation can be well predicted using the model described in detail in ITU-R recommendation [2], but no usable sleet attenuation model is known in the literature.

Until now no special attention has been paid to modelling attenuation of the rare sleet events which occur especially in winter, although sleet can cause much higher attenuation than rain in case of the same intensity [3,4]. Therefore it is essential to develop a model which can be applied during planning of microwave links.

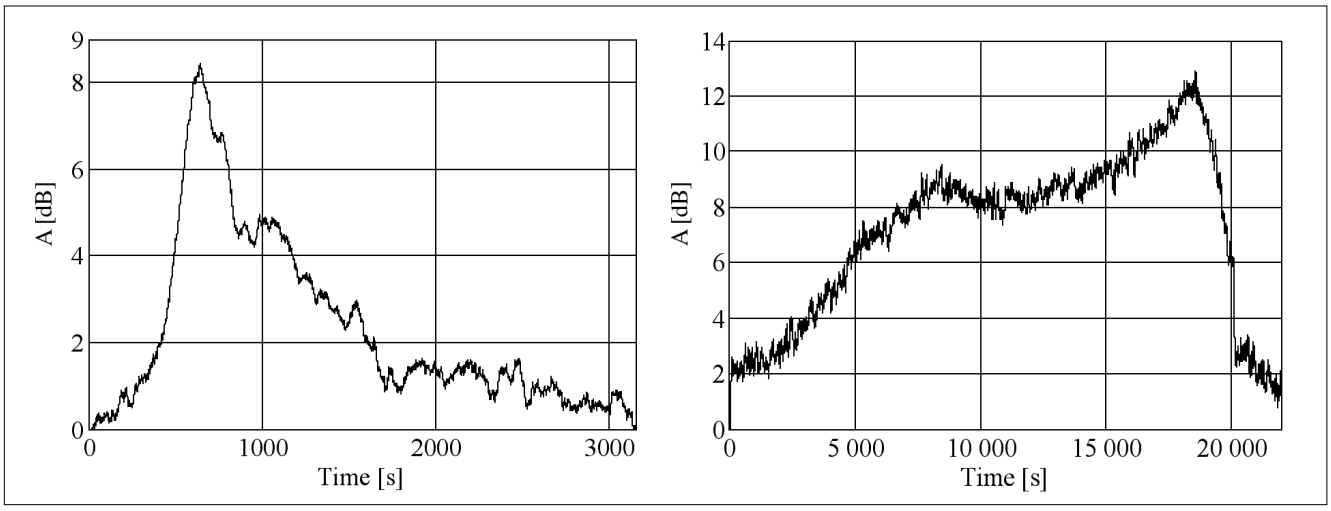
To investigate wave propagation phenomena, a country-wide rain attenuation and weather data measurement system was established in Hungary. A lot of point-to-point millimetre wave links were deployed in star topology for research purposes and meteorological stations were installed at each measuring node [5]. Parameters of links which are investigated in this paper are listed in *Table 1.*

The received IF signal level on microwave links with different parameters and some meteorological parameters, such as rain intensity, temperature, relative humidity etc. measured by the meteorological stations have been being stored since 1997, so sleet attenuation effect could be modelled using our many years measured data, if it would be possible to determine its statistical characteristics by processing appropriately high number of unique events.

It is possible to assort sleet attenuation events from measured data; because time functions of sleet and rain attenuation can be visually distinguished. An usu-

City	Name	freq. [GHz]	length [km]	pol.	k	α
Budapest	HU11	38	1.50	H	0.315	0.955
Miskolc	HU45	38	1.52	V	0.279	0.943
Győr	HU52	38	2.97	V	0.279	0.943
	HU53	23	8.49	V	0.095	1.044
	HU55	38	2.93	V	0.279	0.943
Szeged	HU61	38	3.27	V	0.279	0.943
	HU66	38	2.86	V	0.279	0.943
Budapest	Reference	23	1.00	V	0.095	1.044

Table 1. Parameters of investigated links and of the reference link



(a) Rain attenuation event measured on HU45 link *(b) Sleet attenuation event measured on HU45 link*

Figure 1. Sleet and rain attenuation as a function of time

al rain event causes lower attenuation than an usual sleet event and the duration of usual rain is smaller than that of usual sleet. Differences between the two types of events can be noticed in *Figure 1*. Both events were measured on HU45 link, so it can be observed that attenuation of sleet is indeed higher than that of rain and the sleet event is much longer than the rain event.

However, sleet events are very rare, therefore collecting appropriately high number of sleet events requires visually browsing a huge amount of measured data, which is exhausting and consumes a lot of time. Automatic detection of sleet events would make this work much faster and easier. After analytically formulating the visually noticeable differences between sleet and rain attenuation events, sleet events can be detected automatically with a suitable algorithm.

2. Data processing

To investigate characteristics of attenuation events the available measured data had to be processed. First the attenuation data series were calculated from the measured received IF signal level data series considering the median level as zero dB attenuation. Rain events with the same intensity cause different attenuation on microwave links with different parameters (Table 1).

The relative rain attenuation γ can be calculated with (1) from the rain intensity R in mm/h and the k and α frequency and polarization dependent variables [2]:

$$\gamma[dB/km] = k \cdot R^{\alpha} \tag{1}$$

Due to the link parameter dependency on relative rain attenuation, the attenuation time series measured on microwave links with different parameters cannot be compared directly. Therefore the measured attenuation values had to be transformed to a hypothetic reference link which has known parameters. Transformation was performed with (2) which is derived from the ITU-R P.530 recommendation [6] using (1). Parameters of the reference link are listed in Table 1.

$$A_h = \frac{L_h k_h}{1+\frac{L_h}{d_{0,h}}}\left(\frac{A_x\left(1+\frac{L_x}{d_{0,x}}\right)}{L_x k_x}\right)^{\frac{\alpha_h}{\alpha_x}} \tag{2}$$

The lower index x in (2) is related to the measured link whereas the lower index h is related to the hypothetic reference link; A denotes the attenuation in dB, L is the link length in km and d_0 is the path reduction factor [6]:

$$d_0 = 35 \cdot e^{\left(-0.015 \cdot R_{0.01}\right)} \tag{3}$$

In (3) the geometrical location dependent $R_{0.01}$ means the rain intensity value in mm/h which is exceeded in 0.01 percent of one year time period on the microwave link. Applying many years of measured data, ITU determined $R_{0.01}$ values for different locations on the Earth; however, the recommended $R_{0.01}$ can be refined using local measurements. Hungary is located in the H and K zones defined by ITU, for which 32 mm/h and 42 mm/h $R_{0.01}$ values are given, respectively [7,8].

By calculating the d_0 values we used interpolated $R_{0.01}$ values based on ITU recommendation and the hypothetic reference link was assumed to be located in Budapest. It must be mentioned that the measured data series contain sleet events as well and (1) cannot be applied for sleet events. However, a fictive rain with a fictive intensity can be considered which would cause the same attenuation event as caused by the sleet. With this remark sleet events can be also transformed by (2) considering $R_{0.01}$. In order to remove the scintillation effect from measured data, a moving average filtering was applied with a window length of one minute. Although the filtering mitigates the maximum measured values, it does not modify the characteristics of the events.

3. Second order statistics of attenuation

Fade duration and fade slope are relevant second order statistics of fading; they are used for the purpose of planning microwave links. Fade slope is the first derivative of fading attenuation time function [9]:

$$\zeta^{[dB/s]}(t) = \frac{dA(t)}{dt}, \tag{4}$$

where $\zeta(t)$ is the fade slope in dB/s, $A(t)$ denotes the attenuation in dB as a function of t time in seconds. Because the measured attenuation series is discrete in time, the discrete fade slope $\zeta[t_n]$ is determined by (5) where t_n is the n^{th} sampling time instance and Δt is the time interval over which fade slope is calculated.

$$\zeta^{[dB/s]}[t_n] = \frac{A\left[t_n + \frac{\Delta t}{2}\right] - A\left[t_n - \frac{\Delta t}{2}\right]}{\Delta t} \tag{5}$$

In practice the conditional probability density function $P(\zeta|A_j)$ (CPDF) is usually determined [11,12], which is defined at the attenuation level A_j and derived from $\zeta[t_n,A_j]$ discrete fade slope values calculated around dA environment of A_j:

$$[t_n, A_j] = \tag{6}$$

$$\left| \frac{A\left[t_n + \frac{\Delta t}{2}\right] - A\left[t_n - \frac{\Delta t}{2}\right]}{\Delta t} \right| A_j + dA \geq A[t_n] > A_j - dA$$

Fade duration gives the time interval during which the attenuation exceeds a given threshold [9,10].

Let $l_{e,j}$ denote the fade duration in seconds of event e at attenuation level A_j, which is the length of the interval during the e fading event in which the attenuation exceeds the A_j level. Considering the $\lambda(.)$ function, which gives the length of its argument in seconds, $l_{e,j}$ is calculated by (7). Fade duration statistics are usually demonstrated by their probability distribution:

$$l_{e,j} = \lambda\left(e : A > A_j\right) \tag{7}$$

Some differences can be noticed between second order statistics of rain and sleet events. As shown in *Figure 1*, a usual rain event has smaller length than a usual sleet event, therefore their fade duration statistics must be significantly different. At the end of the sleet event attenuation abruptly decreases to around 0 dB. Such behaviour cannot be noticed in case of rain attenuation event.

Consequently, fade slope statistics of rain and sleet attenuation must be different as well. Exploiting these differences, sleet and rain events can be recognized in a measured attenuation time series using an appropriate computer program. It must be mentioned that our goal is to detect as many sleet events as possible, therefore the false alarms, when there is a rain event in the measured data, but our algorithm decides for a sleet, are more tolerable in our case than the missed detections, when there is sleet in the measured data and it is decided for a rain event. After running the sleet detection algorithm, the miss-detected rain events can be easily filtered out manually from among the found sleet events.

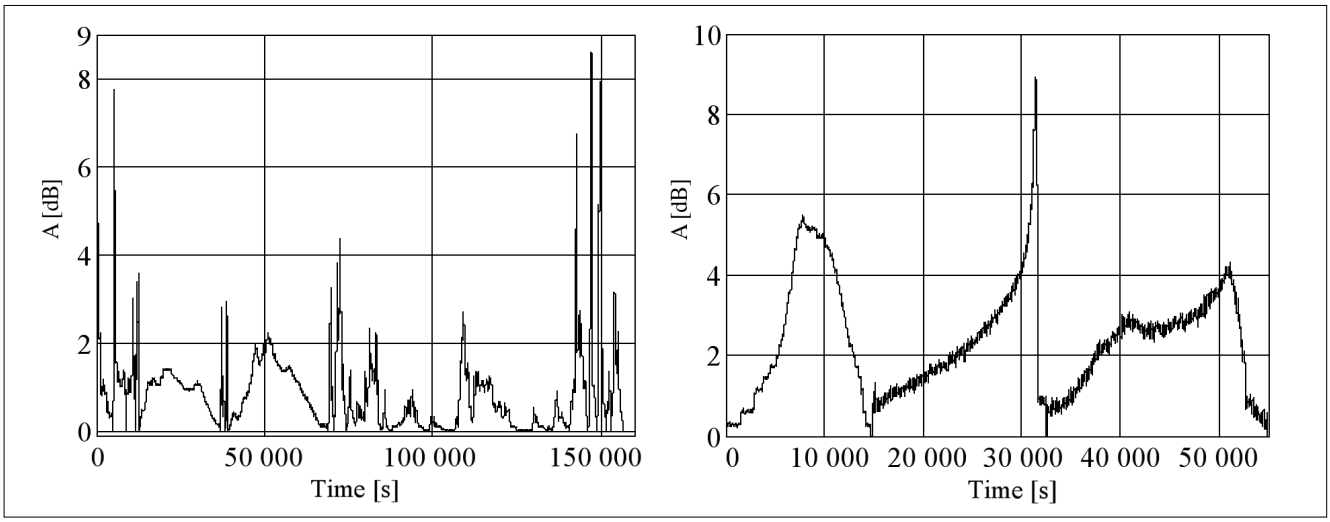
4. Determining reference statistics

Statistics which are representative of the attenuation events can be calculated from our measured attenuation time series. The event detection algorithm compares the calculated second order statistics of the found event with the pre-calculated reference statistics. In order to derive the reference characteristics, two selections of attenuation events were prepared. One of them contains only some (processed and transformed) rain attenuation events; the other contains only some sleet events *(Figure 2)*. Reference statistics were calculated from these concatenated data series.

Four attenuation levels were defined on the transformed data series where the statistical investigation of fading was performed.

Figure 2. Selected rain and sleet attenuation events

(a) Concatenated rain events *(b) Concatenated sleet events*



j	A_j [dB]	$E\{l_{r,j}\}$ [s]	$E\{l_{s,j}\}$ [s]	$l_{t,j}$ [s]
1	1	1320	3939	2500
2	1.4	1017	4719	2000
3	2	485	3796	1500
4	2.4	326	2572	1000

Table 2. Expected values of fade duration for rain and sleet events and the duration thresholds

By defining the attenuation levels two conditions have to be satisfied: (i) all the selected events in *Figure 2* have to have values at all the defined attenuation levels in order to get as accurate reference statistics as possible; (ii) the fade slope and fade duration statistics of the selected events have to be different at the defined attenuation levels.

Choosing the empirical attenuation levels A_1=1 dB, A_2=1.4 dB, A_3=2 dB, A_4=2.4 dB these necessary conditions are satisfied. First the reference fade duration statistics had to be calculated. Let $l_{r,j}$ and $l_{s,j}$ denote the length of the *r* rain event and of the *s* sleet event (i.e. the fade duration) at the A_j attenuation level, respectively, and let $E\{l\}$ denote the expected value of the duration.

The calculated expected values of fade duration for rain and sleet events are listed in *Table 2* for different attenuation levels.

Results are in accordance with our expectations: the expected length of rain events is much smaller than that of sleet events. Based on the calculated expected values a duration threshold $l_{t,j}$ was defined for every investigated attenuation level.

An event length higher than $l_{t,j}$ means that the event might be sleet, a lower length means that the event is probably a rain. By determining the thresholds special attention was paid to that the threshold must be closer to $E\{l_{r,j}\}$ at each attenuation level ensuring the minimal number of missed detections despite of more false alarms:

$$l_{t,j} < \frac{E\{l_{r,j}\} + E\{l_{s,j}\}}{2} \tag{8}$$

The exact values of $l_{t,j}$ thresholds which are listed in *Table 2* were intuitively determined so that (8) was satisfied.

Reference statistics of the fade slope were also determined from the prepared concatenated time series which are depicted in *Figure 2*. The $P(\zeta|A_j)$ conditional probability density functions of the discrete fade slope at the $A_1...A_4$ attenuation levels were calculated with (6) considering dA=0.02 dB and Δt=2 s. Maximum values and standard deviations of $P(\zeta|A_j)$ highly differ at dis-

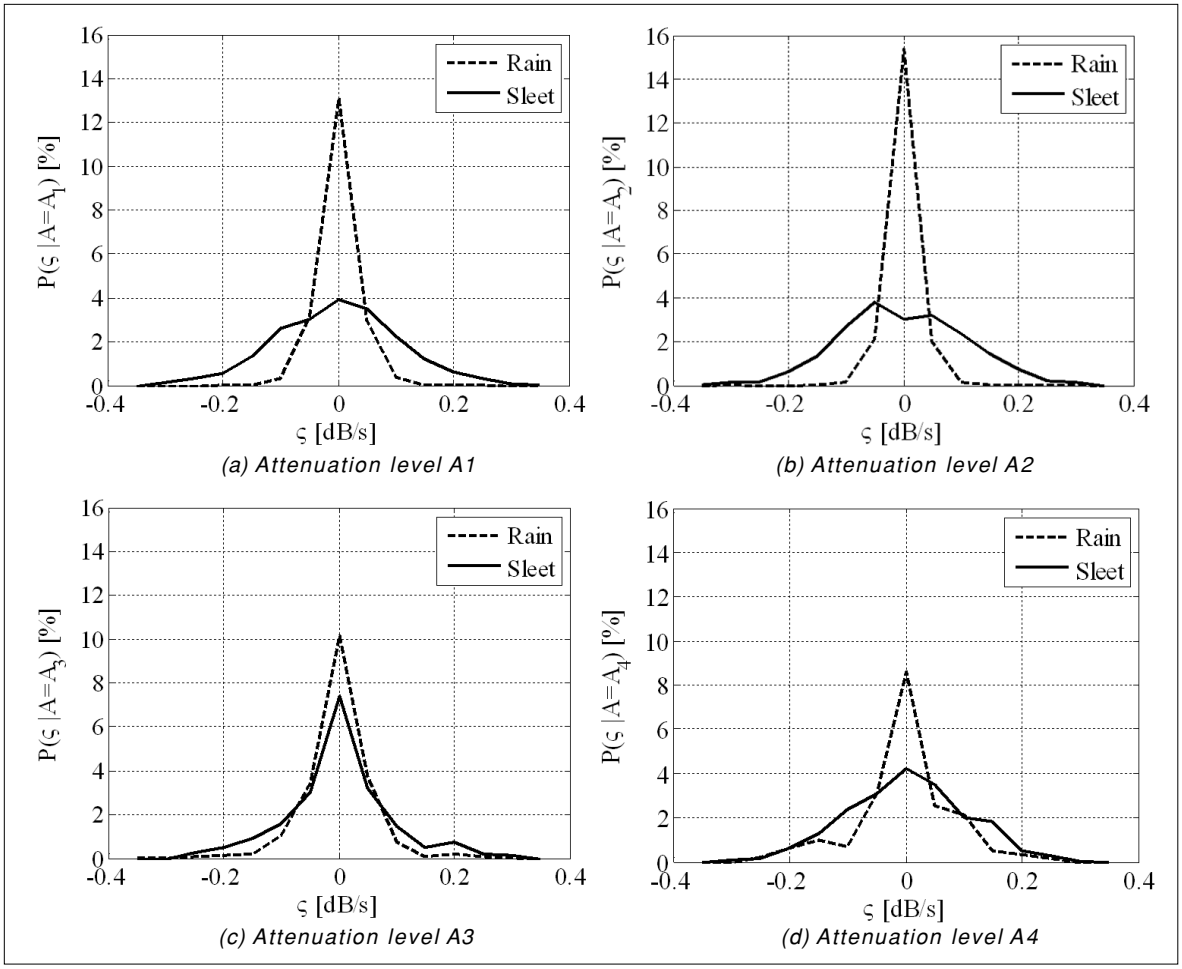


Figure 3. Conditional probability density functions of fade slope for different attenuation levels

tinct A_j levels. As shown in *Figure 3*, the calculated CPDFs correspond to the expectations. At each investigated attenuation level the maximum values of fade slope CPDFs that are calculated from the sleet event are smaller than the maximum values of fade slope CPDFs that are calculated from the rain events. Moreover, deviation of fade slope CPDF is higher in case of sleet events due to the rapid decrease of attenuation time function. It must be mentioned that the measured data series has a quantization step of 0.01 dB therefore discrete fade slope calculated with (6) has possible values of $k \cdot 0.05/2\,\text{dB/s}, k \in N$. This results in less smooth fade slope CPDF curves as can be seen in Figure 3.

Let $m_{r,j}$ and $m_{s,j}$ denote the maximum of $P(\zeta|A_j)$ in case of rain and of sleet event, respectively and let $\sigma_{r,j}$ and $\sigma_{s,j}$ denote the standard deviation of $P(\zeta|A_j)$. Let us define $m_{t,j}$ as the CPDF maximum threshold. Higher CPDF maximum than $m_{t,j}$ means that the event might be a rain, otherwise it might be a sleet. Similarly let us define $\sigma_{t,j}$ as the CPDF deviation threshold. A deviation higher than $\sigma_{t,j}$ means that the event might be a sleet, otherwise it might be a rain. To minimize missed detections (9) and (10) must be satisfied by determining the thresholds.

$$m_{t,j} > \frac{m_{r,j} + m_{s,j}}{2} \tag{9}$$

$$\sigma_{t,j} < \frac{\sigma_{r,j} + \sigma_{s,j}}{2} \tag{10}$$

The maximum and standard deviation values of fade slope CPDFs which are depicted in *Figure 3* and the applied $m_{t,j}$ and $\sigma_{t,j}$ thresholds are summarized in *Table 3*. The $m_{t,j}$ and $\sigma_{t,j}$ thresholds were intuitively determined so that (9) and (10) were satisfied.

5. Event detection algorithm

The suitable algorithm which can be applied for automatically detecting sleet events in the measured attenuation time series uses the previously determined reference statistics i.e. the $l_{t,j}$, $m_{t,j}$ and $\sigma_{t,j}$ thresholds. The flowchart of the algorithm is depicted in *Figure 4*. First of all the input data series in which we want to find sleets, have to be processed by the data processing method described in Section 2. After that the data series must be transformed to the hypothetic reference link with (2).

The computer program which uses this algorithm sweeps over the processed attenuation data series. In this paper our goal is to detect sleet events in a measured data, and as described in Section 1, sleet events have quite high duration and cause significant attenuation. Therefore in our case real attenuation events, which are denoted by *e* are considered as rain or sleet events causing remarkable attenuation and having remarkable duration. So scintillation or very short rain events with very low rain intensity are not considered as real event. Based on the previous remarks, two empirical thresholds were defined, one for attenuation and one for duration, which are denoted with A_0 and $l_{t,0}$, respectively. If the algorithm finds a data series interval

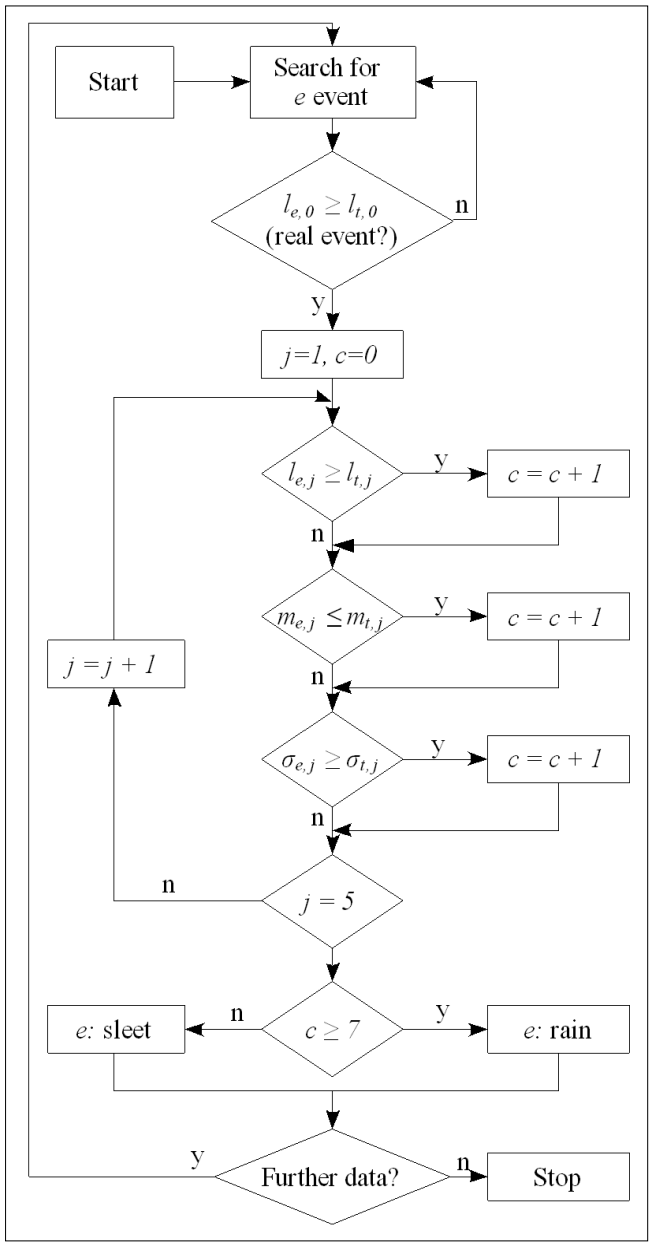
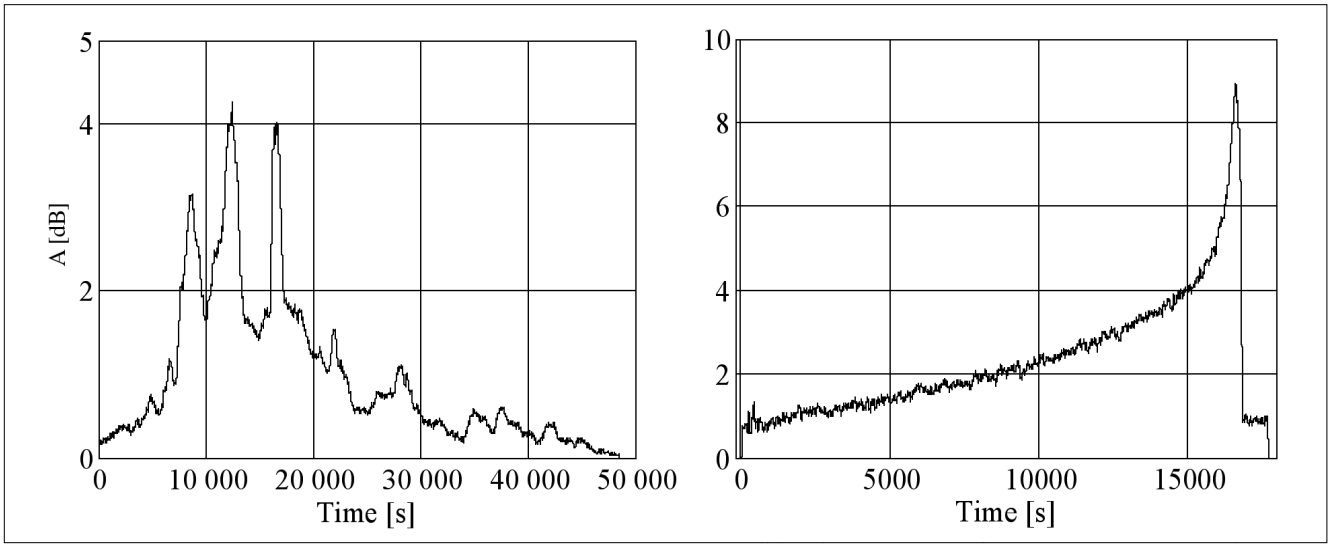


Figure 4. Flowchart of sleet detection algorithm

Table 3. Maximum and standard deviation values of fade slope's conditional probability density functions and the defined $m_{t,j}$ and $\sigma_{t,j}$ thresholds

j	A_j [dB]	$m_{r,j}$ [%]	$m_{s,j}$ [%]	$m_{t,j}$ [%]	$\sigma_{r,j}$	$\sigma_{s,j}$	$\sigma_{t,j}$
1	1	13.1	3.9	9.5	0.0036	0.0109	0.006
2	1.4	15.4	3.1	9.5	0.0031	0.0113	0.0065
3	2	10.2	7.4	9	0.0063	0.0093	0.007
4	2.4	8.6	4.2	7	0.0097	0.0108	0.01



(a) Rain attenuation event measured on HU55 in 2004 *(b) Sleet attenuation event measured on HU45 in 2004*

Figure 5. Test attenuation events

in which the attenuation exceeds the defined A_0=0.6 dB level and whose duration of $l_{e,0}$ is longer than $l_{t,0}$=300 s threshold this attenuation interval is defined to be a real *e* event. If a real *e* event is found, the algorithm starts to investigate it in details at the predefined A_j attenuation levels. The calculated $l_{e,j}$, $m_{e,j}$ and $\sigma_{e,j}$ parameters of the *e* event (i.e. the fade duration and fade slope statistics) are compared with the corresponding thresholds. One comparison at the current attenuation level is called as a test in our terminology. If the *e* event is assumed to be a sleet based on the current test, a *c* counter is incremented.

The algorithm examines the satisfaction of total 12 conditions at the 4 mentioned attenuation levels, so if $c \geq 7$ at the end of the event detection, the program decides for sleet, otherwise it decides for rain.

6. Verification results

Because of the quite small amount of the already collected sleet events, the first verification of the presented algorithm was performed on two short measured attenuation data series. One of them contains only one known and typical sleet event; the other contains only one but extraordinarily long (therefore not so typical) rain event. The time functions of the processed and transformed attenuation caused by the two events are depicted in *Figure 5*.

The rain event was registered in November 2004, whereas the sleet event was measured in January 2004.

The results of the tests are summarized in *Table 4*, where the black bullet symbol next to the test value means that the event detection passed the current test at the corresponding attenuation level, i.e. the event was recognized correctly. The black circle symbol means that the event detection failed the current test i.e. the event was not recognized. The value of the *c* counter after the event detection is also represented in Table 4. It can be noticed that although the test rain event seems to be a sleet on the basis of its length, the algorithm decided correctly for rain 8 times and decided incorrectly for sleet only 4 times (*c*=4), which resulted in a correct final detection. The test sleet event complied with all conditions (*c*=12), the program recognized it correctly. In the course of the event detection verification an extremely long rain event was chosen on purpose in order to demonstrate that theoretically in some cases the fade slope statistics are sufficient in themselves to correctly recognize the investigated event.

7. Conclusions

To be able to model sleet attenuation on microwave links it is essential to process as many sleet attenuation events as possible, which requires collecting these events from many years of measured data. Second order sta-

j	A_j [dB]	HU45, sleet						HU55, rain					
		$l_{e,j}$ [s]		$\sigma_{e,j}$		$m_{e,j}$ [%]		$l_{e,j}$ [s]		$\sigma_{e,j}$		$m_{e,j}$ [%]	
1	1	13790	•	0.0114	•	3.7	•	15564	◦	0.0029	•	13.6	•
2	1.4	11149	•	0.0113	•	3.1	•	11997	◦	0.003	•	12.7	•
3	2	7304	•	0.0113	•	2.6	•	2836	◦	0.0037	•	10.4	•
4	2.4	5935	•	0.0111	•	4.2	•	2836	◦	0.0036	•	12	•
c		12						4					

Table 4. Test results of sleet and rain detection

tistics of rain and sleet event's attenuation are highly different. With exploiting the differences sleet event can be automatically detected and collected by an appropriate algorithm.

In the work presented, some significant parameters of the event's second order statistics were determined, and then thresholds were defined in order to be able to distinguish the two types of events. The presented algorithm can detect sleet events in an arbitrary measured attenuation data series applying the pre-defined reference statistics. Our method was proven using known rain and sleet attenuation events. In the course of the verification both events were correctly recognized. It has been stated that fade slope statistics alone can be sufficient for detecting sleet events; however, in some cases fade duration statistics can make the recognition easier.

In order that our prepared program can find sleet events in arbitrary measured attenuation data with high reliability, the reference statistics need to be refined with considering as many rain and sleet attenuation events as possible.

Acknowledgement

This work was carried out in the framework of Celtic MARCH project.

Authors

BALÁZS HÉDER was born in Budapest, Hungary, in July 1980. He received in 2004 M.Sc. degree from Budapest University of Technology and Economics, in Electric Engineering. From 2004 until 2007 he was Ph.D. student at Budapest University of Technology and Economics, Department of Broadband Infocommunications and Electromagnetic Theory. From 2007 he is research assistant at the same place. He has actively participated to the Satellite Communications Network of Excellence (SatNEx) IST FP6 project. His current research interests include terrestrial microwave point to point and point to multipoint radio systems, especially broadband fixed wireless access systems (BFWA, LMDS), rain attenuation on millimetre band electromagnetic waves, diversity methods, Markov-chain modeling of rain attenuation, B3G systems.

ANDRÁS BERTÓK was born in Keszthely, Hungary, in February 1987. He will receive B.Sc. degree in 2010 from Budapest University of Technology and Economics, in Electric Engineering. He started research terrestrial microwave point to point radio systems, especially rain and sleet attenuation on high frequency electromagnetic waves from 2008 at Budapest University of Technology and Economics, Department of Broadband Infocommunications and Electromagnetic Theory.

References

[1] G. Bichot, "A multilink architecture for a global wireless Internet connectivity", Thomson, Presentation at BROADWAN Workshop: True low-cost broadband wireless access everywhere for fixed and nomadic users, Brussels, Belgium, May 2005.

[2] ITU-R P.838-2 Recommendation, "Specific Attenuation Model for Rain for Use in Prediction Methods", ITU, Geneva, Switzerland, 2003.

[3] C.J. Walden, C.L. Wilson, J.W.F. Goddard, K.S. Paulson, M.J. Willis, J.D. Eastment, "A study of the effects of melting snow on communications links in Scotland", In Proc. 12th International Conference on Antennas and Propagation (ICAP 2003), Exeter, United Kingdom, April 2003, pp.361–364.

[4] G.G. Kuznetsov, C.J. Walden, A.R. Holt, "Attenuation of microwaves in sleet", Final Report to the Radiocommunications Agency on Contract AY 3564 (51000279), Dept. of Mathematics, University of Essex – Colchester, United Kingdom, August 2000.

[5] J. Bitó, Zs. Kormányos, A. Daru, "Measurement system to investigate the influence of precipitation in a wide millimetre wave feeder network in Hungary", In Proc. of ITG 7th European Conference on Fixed Radio Systems and Networks (ECRR 2000), Dresden, Germany, September 2000, pp.335–341.

[6] ITU-R P.530-11 Recommendation, "Propagation Data and Prediction Methods Required for the Design of Terrestrial Line-of-Sight Systems", ITU, Geneva, Switzerland, 2005.

[7] ITU-R P. 837-4 Recommendation, "Characteristics of precipitation for propagation modelling", ITU, Geneva, Switzerland, 2003.

[8] CCIR Report 563-4, "Radiometeorological data", CCIR (Now ITU-R) Study Group 5, Geneva, 1990.

[9] ITU-R P. 1623-0, "Prediction method of fade dynamics on Earth-space paths", ITU, Geneva, Switzerland, 2003.

[10] B. Héder, G. Szládek, Z. Katona, J. Bitó, "Second-Order Statistics of Rain Attenuation in Hungary", In CD Proc. IEEE 5th Mediterranean Microwave Symposium (MMS 2005), Athens, Greece, September 2005.

[11] R. Singliar and B. Héder and J. Bitó, "Rain Fade Slope Analysis", In CD Proc. Broadband Europe Conf. (BBEurope 2005) Bordeaux, France, December 2005.

[12] M. van de Kamp, "Statistical Analysis of Rain fade Slope", IEEE Transactions on Antennas and Propagation, Vol. 51, No. 8, 2003, pp.1750–1759.

Multi-homing in IP networks based on geographical clusters

GÁBOR NÉMETH, GÁBOR MAKRAI, JÁNOS TAPOLCAI

Budapest University of Technology and Economics,
Department of Telecommunications and Media Informatics, High Speed Networks Laboratory
{nemethgab,tapolcai}@tmit.bme.hu, makraigabor@gmail.com

Keywords: aggregation, multi-homing, BGP, Rekhter's Law, identifier-locator split

In the past years it has became a widely accepted opinion that the current network layer protocol, IPv4, suffers from several serious problems. Besides the scalability issues of the BGP routing tables [1], several features (like mobility, security) need extensions (like MIP or IPSec, respectively). These features are available as extensions of the IPv4 protocol family. The 6th version of the Internet Protocol (IPv6) is assumed to become the network layer of the future Internet, which unfortunately inherits some weaknesses of IPv4. Such as, there are no easy solutions of the scalability problems caused by edge site multi-homing. In this paper we overview the available future internet trends, and propose an easy to deploy multi-homing strategy that decreases the number of entries in the routing tables by aggregating together the address space of several edge networks located close to each other. We sketch two possible solutions for such a scenario and investigate their performance.

1. Introduction

In the last decade, the Internet Protocol has become the leading technology for inter-machine communication. Meanwhile, a great number of different enhancements were introduced, deployed (e.g. Differentiated Services, Integrated Services), and the technological boundary of the protocol's capabilities have been reached. It turned out that the fundamental architecture assumptions that underlie the routing and addressing design of the Internet Protocol are no longer sustainable. This fact generated an increasing interest in changing or re-designing the routing and addressing architecture of the Internet [2,3], especially considering that a new version of IP, IPv6, is on its way to be globally deployed.

A fundamental design principle that has made it possible for IP to scale to the magnitude of today's Internet is that address prefixes can be aggregated at higher levels of the routing hierarchy. The classic IP aggregation scheme is based on provider-subscriber relationships, where the address space of the subscriber network is part of the larger address space of the provider network. In this case the provider does not have to announce every subscriber's address prefix separately, but it announces only its own address prefix, which trivially ensures global reachability of its subscribers.

Unfortunately, the assumption that prefixes can freely be aggregated at the provider networks is no longer valid. The two most important causes strengthening this effect are: *(i)* fast endpoint mobility and *(ii)* site multi-homing with provider independent (PI) addresses.

The growing difficulties in address aggregation results in the unprecedented increase of routing table sizes what we experience today *(Figure 1)*. The sheer volume of IP addresses to manage, and the frequent processing of update messages thereof, will inevitably lead to grave scalability problems in the long run.

A straightforward solution would be to completely remove the dependency of IP addresses from their location in the network. In such an unstructured (usually termed as "flat") routing architecture, neither endpoint mobility nor multi-homing would pose problems. Let us recall Rekhter's Law, which states *"Addressing can follow topology, or topology can follow addressing. Choose one"* [5]. Thus, according the Rekhter's Law there must be a congruency between the network topology and the addressing. Understanding the current future internet proposals we may derive a slightly different law, namely: *By removing the structure from the address space, we might need to reimplement a similar structure in the control plane.*

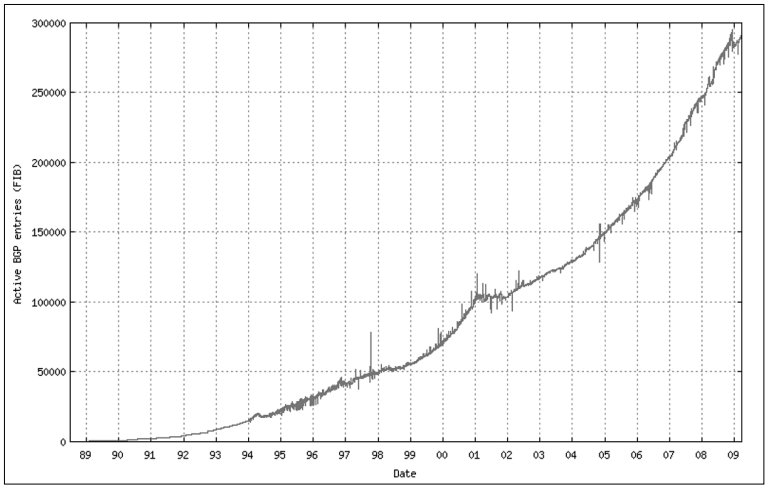


Figure 1. The BGP routing table growth [3]

In the next section we briefly describe the possible classification of the future internet proposals in order to emphasize the ideas behind the previously given law-interpretation. In Section 3 a scalable cluster based solution for multi-homed network aggregation is proposed. Later, in the same section, some simulation results are presented to characterize the proposed methods.

2. From Congruency to Structure

In this section we overview the popular trends of future internet architectures to solve the problems discussed in the previous section. An unique, usually unstructured, flat identifier (or name) is assigned for each host and/or network element, and somewhere in the network this identifier is resolved to a routeable locator (i.e., an address that can be used in a scalable routing system). The place, where this resolution is carried out can be either the host itself (like in HIP [7], shim6 [8], GSE/8+8 [9]) or somewhere deeper down in the network (like in LISP [10], eFIT [11]) (see also *Table 1*).

In the first case, users must invest in new devices, or at least update their software with a new protocol stack supporting the host based identifier-locator splitting mechanism. The majority of the home users are satisfied with the system used nowadays and may not be inclined to invest in a new one, because from their perspective the new equipment has the same capabilities. Thus, the deployment path for host-based identifier-locator splitting is somewhat obscure.

In the second case, the protocol stack of the hosts remains the same and it is the service provider that must install extra functions and (maybe) entities in its network. This approach looks more viable, however, it is still questionable whether the process of identifier-locator resolution can be made effective.

In both cases, the scalability problems are delegated to the control plane, that is, to the system coping with the identifier-to-locator and/or locator-to-identifier translation process. In the simplest case, this would amount to burdening the Domain Name System (or an appropriate successor thereof) with identifier-to-locator translation. It is yet to be seen whether the venerable DNS architecture can cope with that amount of load.

Methods differ, in addition, as to how the identifier-locator split is represented in the packets. It is possible to distinguish two different categories: map-and-encap and address rewriting (Table 1) [12].

The delivery of a packet in the map-and-encap schemes occurs as follows. A host, initially, puts the destination host's identifier into the packet. While the packet is in the transit through the core, it gets encapsulated and delivered using the locator into the destination's domain, where it is de-capsulated and delivered to its final destination using the identifier. Note that the map-and-encap schemes always append a new header to a packet instead of rewriting it, unlike the case with the rewriting methods discussed below.

	map-and-encap	*rewriting*
host-based	HIP	Shim6 Six/One
network-based	LSIP eFIT	GSE/8+8

Table 1.
The classification of the identifier-locator split proposals

The idea behind address rewriting is to divide the IP address space into two parts, and use the upper bits as the locator field and the lower bits as a unique endpoint identifier. The two parts of the address space are completely disjoint, which means that neither of the endpoints is aware of its locator. The locator part of the source address is filled in by the local egress router, while the locator part of the destination address is removed at the remote ingress router. From both sides of the communication the address of the remote host looks complete, as the locator and the identifier part is filled as it was returned by the DNS.

As a summary we may conclude, that all the introduced methods tackle the previously mentioned problems by introducing special structure in the addressing or by forcing special virtual structure of the network entities.

2.1 Push or Pull?

There are several ways to solve the problem of identifier-locator mapping: *(i)* empower the hosts to store their own bunch of locators and introduce a context establishment in the communication process (like Shim6 or Six/One) or *(ii)* delegate the mapping process to the (successor of the) DNS.

As per the enhancement of the DNS, there are two basic approaches. According to the push-based approach, the databases are pushed near the edge, i.e., the mappings are proactively fetched in order to minimize the lookup latency. In the pull-based approach, however, requests are delivered right to the authoritative server, i.e., the identifier-location mapping is performed upon the request, reactively.

Both of the push and the pull model have its own benefits:

– Pull systems have comparably low storage requirements, enabling finer and dynamically generated mappings (e.g., for mobility).
– Push systems have comparably lower latency, and hence less packet loss in the routers that buffer packets during the resolution.

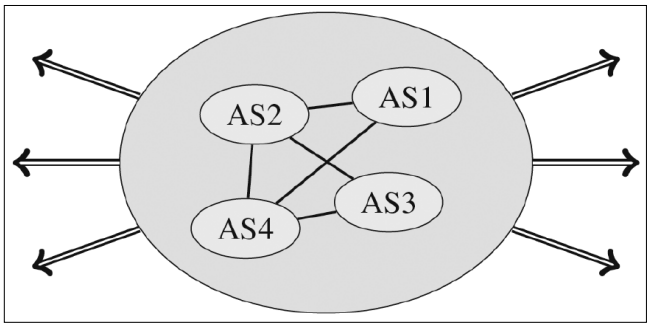
In the next section we propose a new routing architecture to deal with scalability isssues in multi-homing.

3. The proposed method: cluster based multi-homing

The scalability problems are mainly caused by small multi-homed edge networks, because the processing and storing of their small prefixes consume huge amount of resources. Thus, a new and scalable aggregation method is needed. At the same time the providers prefer not to invest much into hardware and software. Therefore, each viable method needs to be cheap and easy to deploy. Moreover, the method has to be *(i)* capable of aggregating the addresses allocated to different multi-homed edge networks and *(ii)* resistant against failure scenarios, as multi-homing was originally implemented.

These goals can be achieved by introducing special multi-homing (MH) addresses, which are completely independent from the address space of the providers' network and each multi-homed edge network gets its addresses from this special address space. Each multi-homed network has a few providers, called direct providers. These providers are able to forward packets directly to their multi-homed edge networks. However, they do not announce these addresses. These MH addresses are announced only after it is aggregated by a set of AS, called aggregation cluster (see also *Figure 2*). The edge networks aggregated together are strongly suggested to be "geographically close" to each other and to the cluster, to avoid large inter-continental hops in paths. Thus, the MH addresses have to be distributed geographically (probably by an international registry). Recall that this scheme faces only the problems caused by the small edge networks that are placed within geographically small regions.

Figure 2.
The autonomous systems called AS1, AS2, AS3 and AS4 are responsible together for a given MH address prefix. They announce the MH address together, and the incoming packets are distributed among them using IP tunneling.



Simply put, in such an architecture every packet heading to a multi-homed edge network *(i)* is first forwarded to the aggregating cluster, *(ii)* then, after detecting (one of) the real providers of the edge network, the packets are encapsulated and using IP tunnels forwarded to their real destination. After arriving to the network of the direct provider, (iii) the outer IP header is removed and the packets are forwarded using the original destination addresses.

3.1 Heuristic Models

Finding optimal aggregation clusters are almost impossible in a highly distributed and always varying system like the Internet. Instead, our focus is on simple and straightforward implementation guidelines. One can observe that the direct providers with edge networks of the same multi-homing prefix can be used as an aggregation cluster for that prefix. Also note that such a solution can minimize the AS-level path by putting the cluster as close as possible to the multi-homed edge networks. We wish to point out, that this scenario is also resistant against the same failure scenarios as the original (not scalable) architecture was. Note that, from the perspective of the edge networks, the previous cluster definition implies that they can connect only to provider networks that are part the aggregation cluster for their MH addresses. In the model presented so far the routing is performed in traditional IP forwarding until the packets reach the aggregation cluster. The routing and forwarding inside the cluster can be implemented in two different ways.

In the *Omni model* (Figure 2) each AS in a given cluster has enough information to send packet directly to its destination. That means, each AS knows every direct provider of each multi-homing edge network aggregated by the cluster. As a result, for each incoming packet only one decision is made, i.e., the selection of the direct provider's network. Thus, when a packet enters a cluster, the entry autonomous system simply matches a tunnel to it; a tunnel that ends at one of the direct providers of the destination edge network. Later in the direct provider's networks, the IP header which was used for tunneling is removed, and the packet is forwarded using the original IP address.

Special case of the Omni model can be reinterpreted as BGP Confederations [13]. In case of connected cluster, i.e., when the autonomous systems inside the same cluster are reachable through internal links, the packets can be forwarded inside the cluster without using any tunnel. This situation generally cannot be guaranteed; however from business perspective, providers may work together and deploy links among themselves. The major gain of this solution is in avoiding tunnels and so the encapsulation overhead; and the packets inside the cluster can use their own destination address, as these addresses are announced inside the cluster.

While in the Omni model each AS has sufficient information about the multi-homed edge networks in its cluster, in *MH^{2}T model* (Multi-Homing with Hash Tables) this information is reduced. The reduction is done by spreading this information among the provider networks inside the same cluster. A straightforward solution for this distribution is implementing distributed hash table (DHT).

3.2 Simulation Results

The goal for both multi-homing models was to reduce the size of BGP routing tables by improving the address aggregation procedure. First, let us theoretically estimate the reduction in the size of the routing.

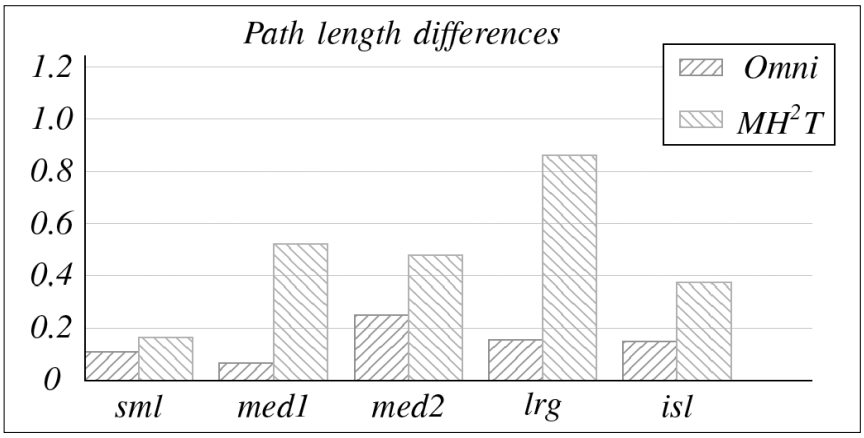


Figure 3.
Increase in the path lengths compared to the traditional BGP path lengths normalized with the size of clusters.

In the Omni model, unfortunately, the effect of the proposed aggregation model cannot be perceived in the local routing table sizes. As a result, an autonomous system participating in an aggregation cluster does not have direct benefits of the aggregation of its own edge networks, but all other autonomous systems outside the cluster do so. Thus, on the whole the size of the routing tables decreases; according to the actual configuration routing tables with 70-80-90% less entries can be reported. On the other hand, in the MH^{2}T model creating each cluster causes observable decrease in routing tables. Unfortunately, there is some data needed to maintain the DHT, which obviously increases the number of stored entries.

We have evaluated the performance of the proposed models – here we only want to show the results according to the average path length – using networks with different cluster sizes. We denoted the networks as *sml, med1, med2, lrg* and *isl*, where the cluster sizes were 3, 5, 5, 10 and 5, respectively (they were generated using reference networks in [13]). Simulation results were derived using bimodal traffic pattern.

The simulations showed that the average path length increases for both of the introduced cluster-based models. However, it never oversteps the diameter of the cluster *(Figure 3)*. Naturally, the paths observable when using the MH^{2}T model are longer than in case of the Omni model. This observation is also conforming with our expectations: the packets have to travel along the DHT to collect all the information necessary to forward them to their destination. It is important to reflect on large path length difference, mainly in case of the *lrg* network. In this case the cluster was rarely connected, i.e., packets had to enter and leave the same cluster several times. These extra hops outside the cluster have huge impact for path lengths.

We also tracked the effect of the cluster connectivity *(Figure 4)*. According to the figure, increasing the cluster connectivity the AS-level path stretch decreases rapidly. This is because inside the cluster the packets have to take shorter paths due to the dense inter-provider connectivity.

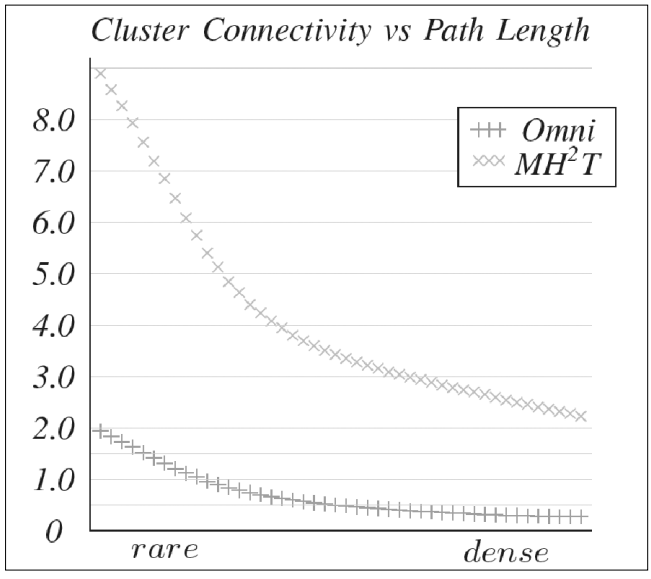
4. Conclusion

In this paper first the consequences of Rekhter's Law is discussed, as the key design principle of the future internet architectures.

In our understanding the law can be interpreted as follows: removing the structure from the address space should be performed with special care, otherwise a similar structure in the control plane must be reimplemented. It is followed by a novel proposal on scalable multi-homing solution, which is easy enough to configure to attract network administrators. The method is based on clusters that aggregate the multi-homing addresses of multiple edge network. Two different models are introduced describing the inner behavior of the cluster; in the Omni model where each AS knows every direct providers of each multi-homing edge network aggregated by the cluster, while in the MH^{2}T model this information is stored in a DHT.

While both models reduce the size of routing tables, they increase the path lengths. In the simulation the average path lengths for connected clusters vary about the original path lengths (there is an approximately 20% difference of the size of the cluster). Note that when the topology graph of cluster is not strongly connected, the packets reaching the cluster may leave the cluster and return at another AS, which cause longer paths and extra costs for the providers. In our simulations in this rarely connected case the path length increased approximately 80% relative to the size of the cluster. Unfortunately, not only the length of the paths grew, but also

Figure 4.
Increase in the path lengths compared to the traditional BGP path lengths, when increasing the cluster connectivity by adding links. The path lengths are hops between AS's, and the lengths are normalized by the size of the cluster (for lrg it is 10).



the load of the AS's. The load increase may be a drawback for the providers, thus they may avoid participating in any cluster, or even if they participate they advertise longer paths in BGP advertisements.

Our future work is to develop methods for load balancing to deny uncooperative announcements.

Authors

GÁBOR NÉMETH has received his M.Sc. degree in technical informatics from Budapest University of Technology and Economics (BME), Budapest, Hungary in 2007, where he is currently working towards his Ph.D. degree at the Department of Telecommunications and Media Informatics (TMIT). His research interests focus on multi-homing, routing and future Internet trends.

GÁBOR MAKRAI has received his B.Sc. degree in 2009 from Budapest University of Technology and Economics (BME), Budapest, Hungary, where he will continue his study and research towards his M.Sc. degree.

JÁNOS TAPOLCAI received his M.Sc. degree in Technical Informatics in 2000 and Ph.D. in Computer Science in 2005 from Budapest University of Technology and Economics (BME), Budapest, Hungary. Currently he is an associate professor at the High-Speed Networks Laboratory at the Department of Telecommunications and Media Informatics at BME. His research interests include applied mathematics, combinatorial optimization, linear programming, communication networks, routing and addressing, availability analysis and distributed computing. He has been involved in several related European and Canadian projects. He is an author of over 40 scientific publications, and is the recipient of the Best Paper Award in ICC'06.

References

[1] T. Bu, L. Gao, D. Towsley,
"On characterizing BGP routing table growth,"
In Proceedings of IEEE GLOBECOM'02,
Taipei, Taiwan, November 2002.

[2] FIND, NSF NeTS FIND Initiative Website,
http://www.nets-find.net/

[3] Internet Research Task Force Routing
Research Group,
http://www.irtf.org/

[4] BGP Routing Table Analysis Reports,
http://bgp.potaroo.net/

[5] D. Meyer, Ed., L. Zhang, Ed., K. Fall, Ed.,
"Report from the IAB Workshop on
Routing and Addressing,"
RFC 4984, September 2007.

[6] P. Savola, T. Chown,
"A Survey of IPv6 Site Multihoming Proposals,"
In Proceedings of the 8th International Conference of
Telecommunications (ConTEL) 2005,
Zagreb, Croatia, June 2005, pp.41–48.

[7] R. Moskowitz, P. Nikander,
"Host Identity Protocol (HIP) Architecture,"
RFC 4423, May 2006.

[8] P. Savola,
"IPv6 site multihoming using a host-based shim layer,"
In 5th International Conference on Networking and
the International Conf. on Systems (ICN/ICONS/MCL'06),
Mauritius, April 2006.

[9] Mike O'Dell,
"GSE – An Alternate Addressing Architecture for IPv6,"
IETF Draft, 1997.

[10] D. Farinacci, V. Fuller, D. Oran,
"Locator/ID Separation Protocol (LISP),"
IETF Draft, 2008.

[11] D. Massey, L. Wang, B. Zhang L. Zhang,
"A Proposal for Scalable Internet Routing & Addressing,"
IETF Draft, 2007.

[12] D. Meyer,
"Update on Routing and Addressing at IETF 69",
IETF Journal, Volume 3, Issue 2, October 2007.
http://www.isoc.org/tools/blogs/ietfjournal/

[13] D. Medhi, K. Ramasary,
"Network Routing Algorithms,
Protocols and Architecture",
Elsevier Inc., 2007.

A framework for community-oriented interactive digital television

O. MAYORA-IBARRA[1], D. GABOS[2], E. FURTADO[3], R. CAVALIERE[4], A.C. PASCALICCHIO[5], R. FILEV MAIA[2]

[1]CREATE-NET, Trento, Italy; [2]POLI/USP, São Paulo, Brazil; [3]UNIFOR, Fortaleza, Brazil; [4]TIS Innovation Park, Bolzano, Italy; [5]Mackenzie University, São Paulo, Brazil

oscar.mayora@create-net.org

Keywords: idtv, users content creation, power line communications, digital inclusion

This paper presents the SAMBA Framework for community-oriented idtv. The main objective of this framework is to provide local communities and citizens – including low income population – to access community-oriented content and services by means of idtv channels. In order to achieve this, SAMBA implemented several steps such as: 1) development of a Content Management System for creating interactive applications that are cross-compatible between DVB-T MHP and DVB-H, 2) considering usability issues related to technology adoption by non expert users and 3) addressing low cost requirements for meeting the needs of low-income users and communities with use of power line communications for providing interactivity to remote rural areas.

1. Introduction

In the latest years, media generation and its consumption patterns by users has been going through a major transformation. The increasing participation of users in the development of content for the web and their active involvement in online communities of interest has redefined their role of mere consumers of media to become "prosumers" [1]. This trend that started as a phenomenon in web domain is migrating to other spheres such as the mobile web [2] and more recently to other platforms such as the interactive digital television *(idtv)*. In particular in this latest platform, the involvement of users has recently shown to be present across all the production-consumption lifecycle in a similar way as happened some years ago within the web domain [3] including the emergence of virtual idtv communities dedicated to social networking [4-7].

In particular, the community-level dimension of idtv has a significant peculiarity when compared to traditional online web communities. Since idtv coverage is limited to well defined geographic areas, the communities built around idtv are intrinsically connected to particular territories and can be naturally associated to specific local services. Moreover, for territories affected with low accessibility to the Internet such as rural areas in developing countries, idtv can be presented as a potential good candidate for providing interactive services covering the local community needs at different levels such as entertainment, education and social inclusion, among others.

The access to idtv infrastructure alone cannot tackle the digital divide in marginalized areas unless it is accompanied by equipping users with the necessary skills to utilize its potential for creating and consuming inclusive services. In this paper, we present an idtv framework developed by a European Project *(SAMBA – System for Advanced interactive digital television and Mobile services in BrAzil)* for providing rural communities in Brazil with the capabilities and means to develop sustainable interactive, community-oriented services [8]. In particular, the SAMBA framework enables local communities to produce idtv content and broadcast it through community access idtv channels. In this way, citizens, including low income population, are empowered with a way to participate in the process of creating and accessing digital content and in the services derived from it [9].

The platform being developed by SAMBA allows the development of innovative interactive applications focused on social local communities including services with a strong social focus, such as T-Learning, T-Health, T-Government and T-Commerce. SAMBA framework has proven to be generic enough to be adaptable to diverse contexts by its implementation in two different environments, one in Brazilian North-East in the town of Barreirinhas and one in the Alpine region in Northern Italy in the town of Naz-Sciaves [10].

2. Framework overview and system architecture

2.1 General framework overview

The SAMBA framework consisting of three main domains is presented at a high level in *Figure 1*. The main components highlighted in the figure are: 1) *User Domain;* in which primary users consume the interactive services, 2) *Platform Domain;* where transmission of applications and reception of interactive input is performed and 3) *Content Domain;* related to the tools for producing interactive contents by local communities.

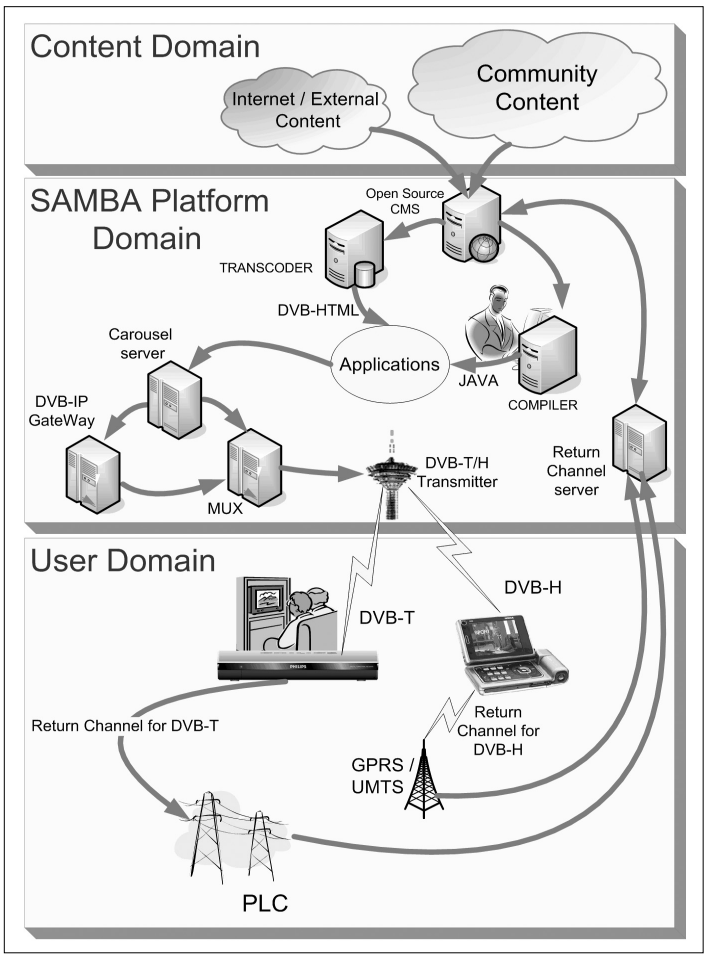


Figure 1. SAMBA framework overview

Regarding the *Users Domain,* SAMBA framework comprises two different types of users: Primary users that are those who consume content and Secondary users who are the ones that create content. Eventually, primary and secondary users could be the same individuals.

Both types of users interact with the *Platform Domain* in different ways: Primary users interact with idtv applications through a setup-box that 1) receive the idtv signal and interactive services from local broadcasters and 2) provide interactive input from users through the return channel. SAMBA framework allows for adaptation of the return channel according to the specific needs. In particular, during SAMBA project implementation the use of power line communications (PLC) was identified as the most suitable return channel to fulfill the requirements of the testbeds in remote rural areas. Specifically in the case of Barreirinhas town in Brazilian North-East, PLC was the best way to provide sustainable interactivity to idtv due to the ubiquity of power grid in Brazil even in remote rural locations.

The applications created in the *Content Domain* are developed with the use of a Content Management System (CMS). In SAMBA, the CMS can work both in a local server or in a remote one accessible through the internet that allows secondary users to create their own content and to reutilize other existing content available in the web through RSS feeds.

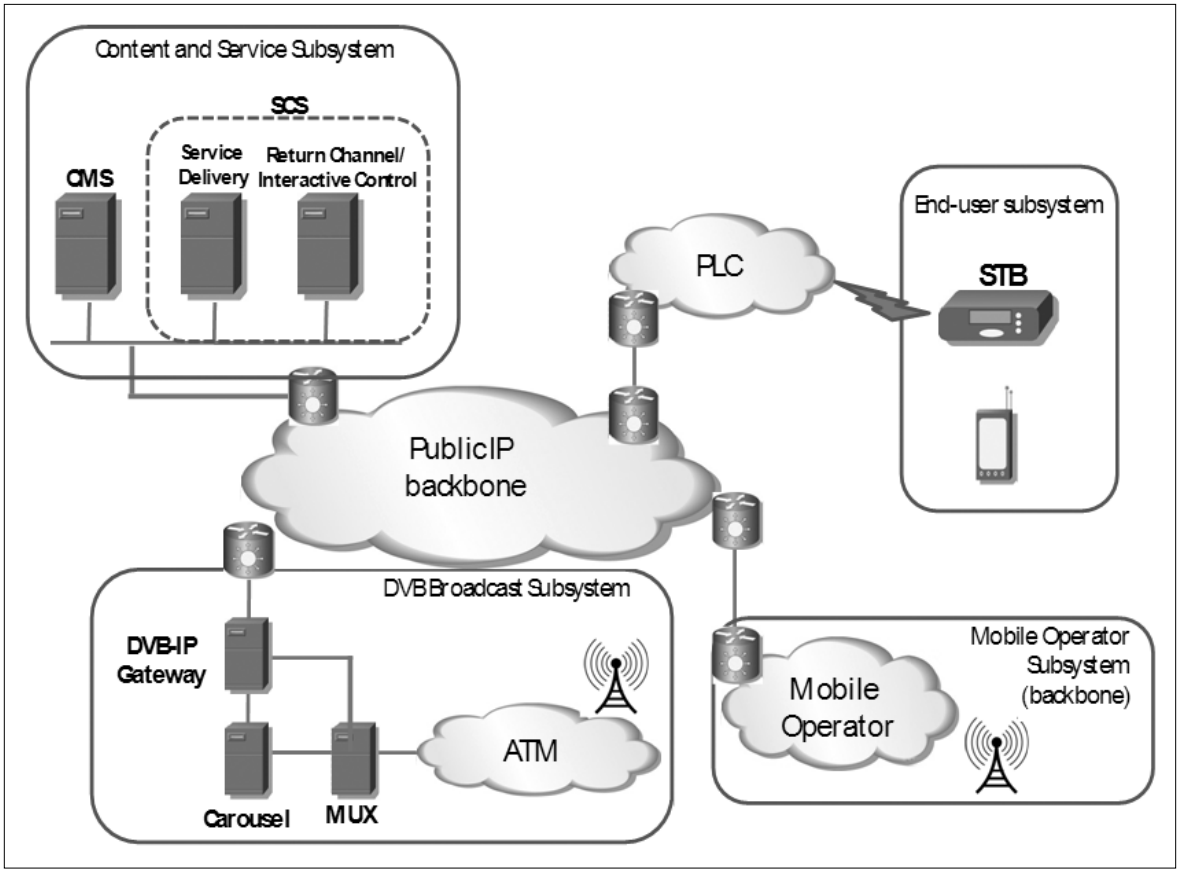


Figure 2. SAMBA architecture overview

2.2 System architecture and detailed description of components

The SAMBA architecture is split among the different system domains described in the previous section and is outlined in *Figure 2*. The next items describe the several sub-systems of the SAMBA architecture defining its functionalities and showing the set and stack of protocols that it uses to interact with the other sub-systems.

• *Content and Service Subsystem (CSS)*

The CSS in SAMBA includes the CMS and the Service Creation Subsystem (SCS). The CMS is responsible for handling the content which is provided by the end-user. The multimedia content is created and stored in the CMS to be delivered to the idtv broadcasting system. The CMS requires adaptation of the content items for each kind of user device (TV, mobile receiver, etc.). Additionally, the CMS sends all information about the content (such as duration, type of device) to the Service Creation Subsystem in order to make possible the delivery of content to the right kind of device. In SAMBA framework, the CMS enables the secondary users to create the following applications:

– *T-Photo*
  Allowing to upload and comment pictures to a given group of users.
– *T-Vote*
  Related to conducting surveys based on voting between different options.
– *T-Info*
  Consisting mainly in text information displayed to the user.
– *T-SMS*
  Enabling users to send SMS from the TV set.
– *T-RSS*
  Concerning use of web content through rss feeds.
– *T-Video*
  Application enabled only for mobile devices for uploading video (only for mobile)

Regarding the SCS its main scope is to handle the multimedia content in order to build the Electronic Program Guide (EPG) and other useful information about the content. This equipment creates and sends the EPG and information about other multimedia content in order to provide additional information to the final user.

• *DVB Broadcast Subsystem*

The DVB Broadcast subsystem is composed of two parts in order to emphasize the aspect of IP/DVB conversion. The DVB-IP Gateway converts the IP traffic, including transport stream, EPG (Electronic Program Guide) and other data, into a DVB streaming in order to be sent to the MUX. This equipment may receive data from several Carousel Servers and must organize all data in an appropriate manner in order to be transmitted by MUX.

– *Carousel Server:* In the Digital TV DVB Systems, the Data Carousel is based on the DSM-CC (Digital Storage Media – Command and Control) standard ISO 13818-6. The Carousel Server is used to send data files and content pages to the users.
– *MUX and DVB-IP Gateway:* The DVB-IP Gateway is an end system, which has the role of adapting the IP services to the DVB world. Its main functionality is to configure the MUX in order to allow MUX encapsulating data. Therefore, the DVB-IP Gateway needs to exchange information with the SCS.

The Multiplexer (MUX) is equipment composed by audio and video coders/decoders whose task is to code all content provided by SAMBA into MPEG streaming (MPEG2 or MPEG4) and send it to the DVB-T/H transmitters. MUX builds the Packetized Elementary Streams (PES) and based on this MUX generates the transport streaming because of the interaction of several DVB-IP gateways. Neither of these equipments provide any interaction with the end user or the other SAMBA systems. The only data sent by the MUX to the other systems are managerial data, such QoS information and about its operation status.

• *End User Subsystem*

The end user subsystem in SAMBA includes the utilization of both a TV set connected to a setup-box (STB) as well as a mobile TV enabled phone. Regarding the STB, it contains the hardware for DVB-T reception and PLC transmission and reception and the basic software for the DVB standard functionalities, including Transport Stream processing with the respective formats, reception and running of applications in the required standards (MHP for SAMBA case). In the PLC interface, the STB needs to support the required standard for the metropolitan access in the local network considering the physical layer and the MAC sub-layer. Concerning the management of the mobile device, it has the same requirements and constraints of the STB, considering the support of SNMP agents aimed at the mobile environment of DVB-H with a GSM/GPRS return channel.

• *Mobile Operator Subsystem*

The mobile network can be integrated into the management architecture by the same approaches of the Digital TV Broadcasters. It is possible to consider a solution with the sharing of information through the communication of the NMSs using SNMP or data base synchronization. Another possible approach is to use the integration of QoS monitoring information and control over the mobile DVB-H services.

• *PLC Network*

The PLC Network is composed of two main parts: the low voltage and medium voltage networks. The providers of equipments and solutions for PLC networks have implementations of network management. They consist of NMSs that monitor and control the network using proprietary communication and support the use of SNMP. This means that the equipments of a PLC network can be managed in three different ways:

– through the communication between the NMS of the PLC network and the SIMP using SNMP;
– through the synchronization of management data base of the NMS of the PLC network and the SIMP;
– through the direct management of the PLC network equipments by the SIMP, using SNMP.

3. Testbeds implementation and usability evaluation

3.1 Testbed implementation

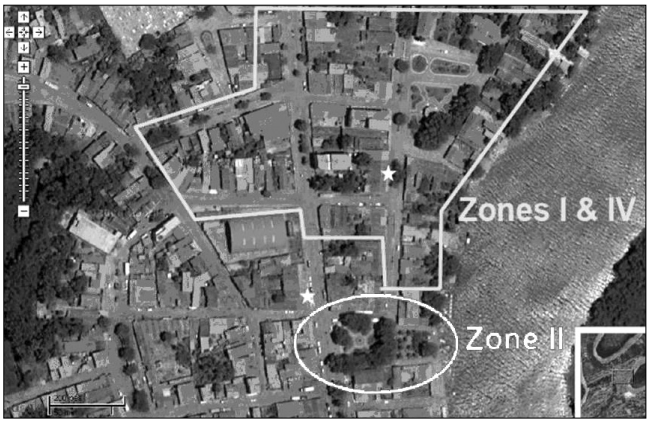
• *SAMBA Generic Infrastructure*

In order to implement the different testbeds of SAMBA Platform, a general infrastructure for the delivery of the interactive services was defined. The different components of this general infrastructure are shown in *Figure 3*. Each component in this infrastructure corresponds to the modules defined in SAMBA system architecture and the three different domains of the framework. Once defined the generic infrastructure, it was instantiated and customized for operating in both testbeds: Barreirinhas and Naz-Sciaves. A common characteristic of both testbeds was the use of power line communications for providing interactivity as the return channel.

• *Brazilian Testbed*

The testbed in the Brazilian environment was set up in the context of an ongoing project on PLC in the town of Barreirinhas. This project provides already the access to the PLC network infrastructure within the context of this small village in the Maranhao State. The testbed allowed for validation in a living environment the complete system, composed of CMS, Community Access Channel and PLC return channel for interactivity. The PLC backbone in Barreirinhas is connected to the external world through a bi-directional satellite link that provides connectivity to the ground, and a gateway takes this connection to the PLC backbone. The entire village is already connected via the power distribution system to this backbone. The system has been tested in some public buildings as well as in users' homes. *Figure 4* presents an aerial view of the area of coverage of SAMBA testbed. In this figure 'Zone II' corresponds to the place where the satellite link is located (Municipality office) and the area denoted by 'Zones I & IV' corresponds to the places covered by PLC infrastructure.

Figure 4. Barreirinhas testbed site



• *Italian Testbed*

Given its particular geomorphologic structure, South Tyrol region in Northern Italy is affected by a particular digital divide situation between the cities and the alpine valleys. Indeed outside the cities there is a lack of broadband connections and in general Internet connection is slow if available whatsoever. For this reason, idtv was

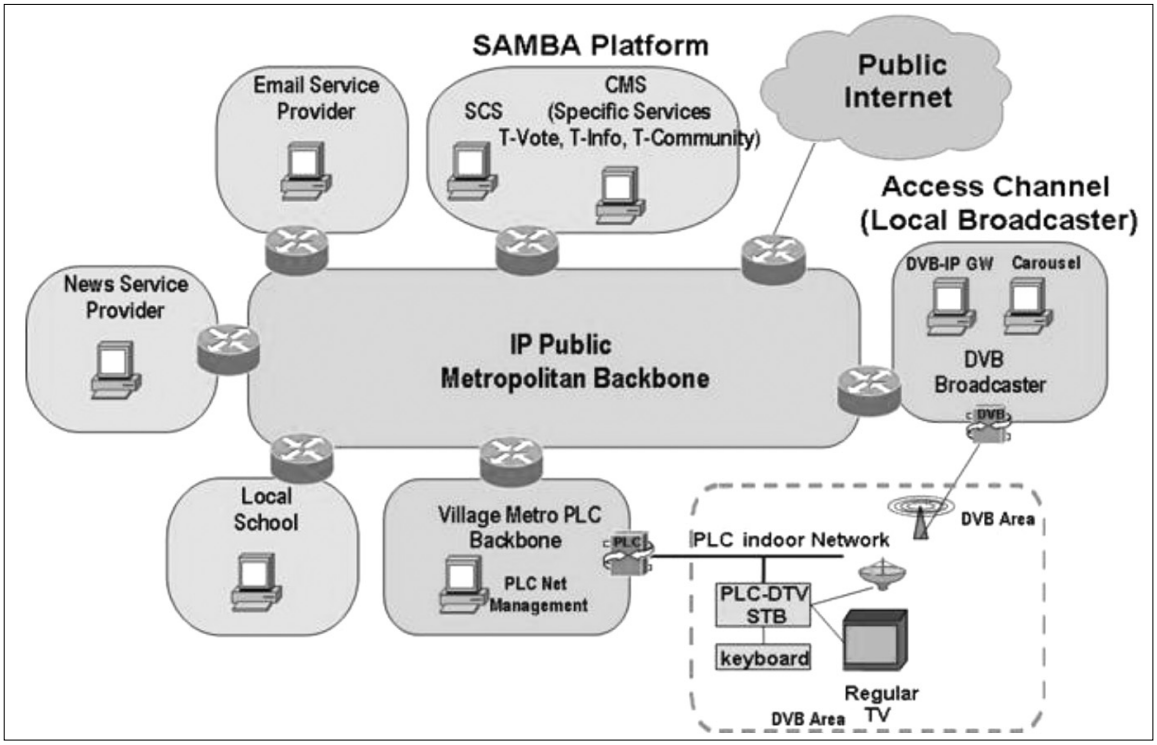


Figure 3. SAMBA basic infrastructure services

Figure 5. Naz-Sciaves testbed site

identified as a good alternative in this region for providing citizens with the means for accessing relevant information regarding more specifically the different communities of the valleys. Moreover, in order to solve the problem of low connectivity, the PLC technology was considered as the best solution for the area to give to it a broadband connection.

Due to its particular position, Naz-Sciaves was chosen as the locality of a first pilot PLC test of the local electric company located in the city of Bressanone. Naz-Sciaves has been the first town not only in South Tyrol but also in Italy in which the PLC technology is used to provide connectivity through the electric distribution network. *Figure 5* presents the areas connected with PLC distribution lines in Naz-Sciaves.

3.2 Usability evaluation

• *Verification and Validation*

Usability inspection and evaluation was conducted following first a Verification and Validation approach (V&V) [11].

The verification of SAMBA applications was done in two parts. In the first part, the verification of the applications was made by a heuristic evaluation of usability conducted by three experts. In particular, the Primary Users applications were tested through the use of a specific idtv emulator for PC. During these testing phases, the evaluators started an application and tried to simulate the typical behaviour of a user interacting with the system. Some preliminary set of usability problems were identified such as duplication of pages, inoperative buttons, low time response, text and images lost. In the second part, an ad hoc verification of the system based on a list of heuristics and checklists performed by the

Table 1. Evaluation methodology

Types of Evaluation		Focus and method (group/individual)	Artifacts and techniques	Examples
Evaluation of system usability	Indicators of user performance	Focusing on user performance as it is perceived by the HCI specialists (in group)	Observations and Checklist about **Efficacy**	Difficulties to complete a task
	Indicators of system usability	Focusing on usability as it is perceived by the HCI specialists (in group)	Observations and Checklist about **Learnability**	Compatibility with user tasks; verbalizations expressing users' judgments on the compatibility of the interface with a given task; help frequency
Evaluation of users' acceptance and system utility		Focusing on usability as it is perceived by the users - the users' perceived usefulness and their perceived ease of use (in group and by individual)	Post-experimental one to many and many-to-many interviews and focus group;	Questions will concern: general satisfaction, particular aspects of the user interface that are judged problematic and utility for social inclusion
			Observations	Observations related to users perceived usefulness (the acceptability and the way users faced problems) and observations related to users' perceived ease of use(problems concerning the use of the devices for interaction).
			Diaries	Analysis of the users' involvement with the application after having used it

team of evaluators together with one of the developers the applications were corrected.

The purpose of the validation process was to evaluate the technology acceptance by the users to attain three main objectives, namely:

– the evaluation of system usability;
– the evaluation of system utility;
– the evaluation of user acceptance.

• *Usability, Utility and Acceptance Evaluation Approach*

The evaluation of system usability in SAMBA aimed at identifying potential problems related to serviceability, difficulty to use the system, etc. These problems were identified on the basis of the usability characteristics needed to guarantee the product quality, according to the ISO9126 (ISO/IEC 9126, 1998) criteria.

The following indicators of user interactions served as a basis for the analysis of the observations of the use of the SAMBA applications in the primary users' homes.

– Indicator of user performance: an indicator related to the efficacy characteristic of ISO 9126
– Indicator of system usability: an indicator related to the learnability characteristic of ISO 9126

Based on these indicators, the evaluators designed and used a checklist as the main artifact while observing the users interactions. The evaluators also focused on usability as it was perceived by the primary and secondary users. The perceived usefulness scope was that of verifying the social acceptability by analyzing the users' behavior related to the acceptance of interaction with the system, its potential for social inclusion and for social practice intended in terms of involvement with the content.

In order to conduct the evaluation, the experts performed direct observations, questionnaires, checklists, interviews and focus groups for both, Primary and Secondary users.

A summary of the evaluation methodology is described in *Table 1*.

• *Evaluation results*

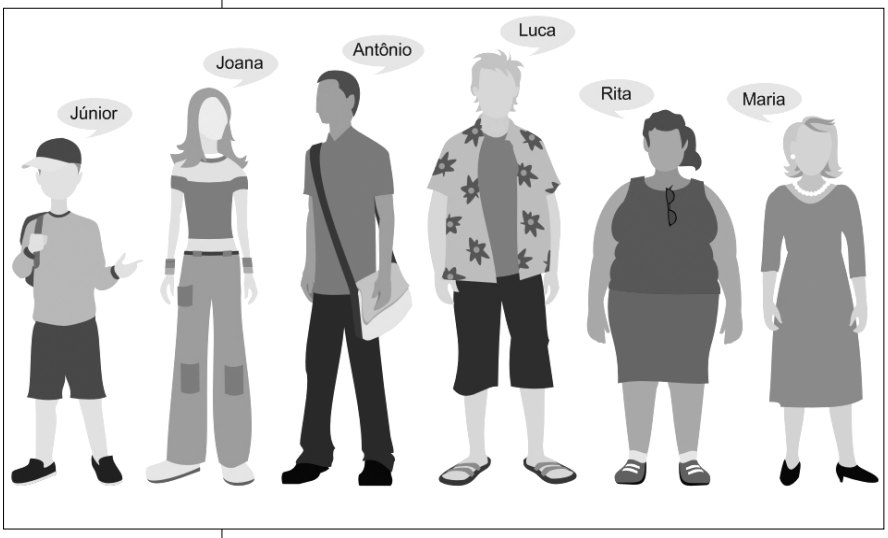
In both contexts (Brazilian and Italian), primary and secondary users were involved in testing the SAMBA system. In general, the system was perceived by users, especially in the Brazilian context, as a very relevant tool for supporting the local community. Such affirmation is related to the limitations of the town on the basis of factors such as: restrictions on access to the Internet or other forms of interactive information sharing, few secondary schools and few options of entertainment offered to the community.

The summary of the results on the verification and validation process and the usability inspections is presented in *Table 2 (see on the next page)*.

4. Sustainability of SAMBA approach

One of the main contributions of SAMBA framework stands in the fact that idtv presents an ideal way for bridging the digital divide, especially in remote areas where people have difficulty to access Internet services. SAMBA approach pursues the goal of enabling local communities in such areas to produce and broadcast idtv content and services while using accessible means such as Power Line Communications (PLC) as return channel. In this way, SAMBA aims to guarantee a sustainable model of idtv service provision even in underdeveloped areas.

Figure 6.
SAMBA users profiles represented as PERSONAS



Accordingly, preliminary studies on user-requirements and expectations were done in both locations and a set of prototypic primary users profiles were identified and modelled as Personas (*Figure 6*) [12,13]. These Personas were used for producing scenario-based representations and definition of SAMBA services.

Additionally, in order to assure sustainability of SAMBA approach, the services identified were analyzed in terms of potential business models considering the economic and cultural constraints imposed by the target regions as well as the expected actors to be involved. In general, five potential actors involved in services and business models were identified:

– Primary User
– Secondary User
– Infrastructure Provider
– Content Manager
– Service Provider.

The business model proposal is a hybrid model situated in the middle between open TV (free-to-air) and pay-per-view TV (paid broadcasting). In this model, pri-

V & V activities	Brazilian context	Italian context
General descriptions		
Products Verified in simulated context: CMS, idtv and mobile idtv applications	yes	yes
System Verified in real context	PLC network	PLC network and ADSL as complementary
Types of Verification tests	ad hoc verification in group and Heuristic evaluation	technical tests
System validated in real context	CMS applications, T-Info and T-Photo Gallery	CMS applications, T-Info, T-Photo Gallery, T-Vote and T-SMS
Methodology used for the validation		
Sample	5 users of 3 local organizations	4 users of TIS organization
Needed support	Training	no need of Training
Artifacts used	CMS tutorial, questionnaires, diary and consent term	idtv applications tutorial, questionnaires and consent term
Environment used in usability tests with primary users	tests with 10 users in primary users' homes and of their friends (n=5)	usability tests with 9 primary users in 4 homes in Natz-Shiaves
Contextualization with users	invitation by correspondence and by personal contact as well as via telephone, five meetings with gifts, snacks and certificates	invitation by correspondence and telephone and two meetings during the project
Usability tests	during three days, team at least of 3 (designer, evaluator and psychologist) and 1 technician	during four weeks, team at least of 2 (evaluator and technician)
Created content		
Type of content	education, business, tourism, news	education, business, tourism, news
Main results		
Users´ acceptance and utility of CMS applications	Users showed interest in producing content for all types of people including the desire to reach users from more far away houses. They were very positive about the access of users to the local content.	Users appreciated the potentialities of the tools to create different types of contents for different services scenarios.
Usability of CMS applications	Despite reporting an average degree of difficulty, users were able to create their contents using the applications and the CMS manual and requested eventually the assistance of the evaluators' team. Main problem: low internet access.	Users impact with the application was straightforward and did not require to much support by the evaluators. Users provided some useful suggestions to improve further the tools. Some minor bugs of the applications were found and corrected.
Users´ acceptance and utility of idtv applications	Most users had positive experiences when using the applications. "I liked more the application of text. Because as a teacher I am part of that context." They were involved by the content and by the possibility of seeing what they referred to as a "possibility" becoming "real". The utility for tackling digital divide was perceived by most users, according to the fact that TV is the most common communication mean in the city.	The primary users had a positive feedback about the tests. Most of the users appreciated the possibility to use the applications in specific scenarios of use. The highest level of appreciation was provided by older people. In general, younger users were not interested by TV oriented content, instead they preferred the PC and Internet-based technologies. However they were very interested to see Brazilian content on TV.
Usability of idtv applications	Some difficulties appeared in function of the lack of familiarity with technology, legibility (in some TV screens, the applications were cut or had a black line) and the buttons on the remote control were in English. Suggestions were mainly related to have dynamic information.	Some difficulties appeared when inputting text. The users' perceived level of interactivity provided by the applications was high. Suggestions were mainly related to have dynamic information instead of static (more animation, zoom). Others suggested to make navigation easier for the users in T-Info and to improve the quality of the service (reduce waiting times originated by slow connections).

Table 2. Evaluation results

mary user is a free of charge player. The sustainability of the model will be assured by the advertisements generated by the services providers and by the secondary users. These players are in fact mainly content providers and sponsors. *Figure 7* shows the service flow.

A set of prototypical potential services identified by the sustainable business model analysis are presented below. All the following services are mere examples that could be taken as basis for defining more sophisticated services:

S-Business – this service can provide advertisement or general information about business subjects. The information can include information about products, services or general information about business opportunities such as information about procedures to register a small business according to local regulations, quotations, taxes, etc. The candidate secondary users for this service could be any kind of service provider or product advertiser as well as governments and regulatory agencies, unions, business consultants, user and/or provider associations, etc.

S-Tourism – this service is a variation of the S-Business with focus on tourism activities. The information can include information about hotels, restaurants, crafts shops, local tours, etc. and the service would be directed to tourists staying e.g. in a hotel in the specific location.

S-Health – this service includes presentation of information related to health campaigns, available pharmacies, available doctors, information about emergencies or epidemics, etc. The service could be offered by local/national health institutions and directed to all the population.

S-News – this service includes provisioning of information related to relevant news. The service could be offered by local broadcaster and could be sustainable based on the advertisement that could be sold to different customers. The service is directed to all the population and could use interactive input from users to monitor interest in certain topics.

S-Energy – this kind of service may consist of presenting the information related to energy consumption of individual users as well as general information about energy campaigns and other tips about reduction of energy waste. The service could be relevant to all the population.

S-Government – this service could allow the local, regional or national governments to provide opportune information to citizens on different relevant issues related to government practices and services. The beneficiaries of such service include the whole population.

S-Education – through this kind of services, the population could get relevant information regarding educational materials. This service could be a very relevant one especially in remote areas where the number of teachers and the availability of educational resources is limited.

5. Concluding remarks

In this paper we presented SAMBA framework for enabling production and consumption of community-oriented idtv in a sustainable way. The framework proposed in this paper was implemented in two testbeds with different characteristics. These testbeds had their own specific requirements for using a PLC infrastructure as return channel to provide interactive services to a sample of the population. Additionally to the final outcome of providing a mean for delivering interactive services to the population, SAMBA implementation provided the construction of a modular solution. This means that each individual component utilized in SAMBA such as CMS, broadcasting technology, etc. could be easily replaced and adapted according to new available resources such as more sophisticated interactive channels or transmission means (e.g. based on GPRS, ADSL, WiMAX etc.) [14]. Within the SAMBA framework business models were defined that allow for the creation of sustainable services that could be utilized as a powerful tool towards reduction of digital divide.

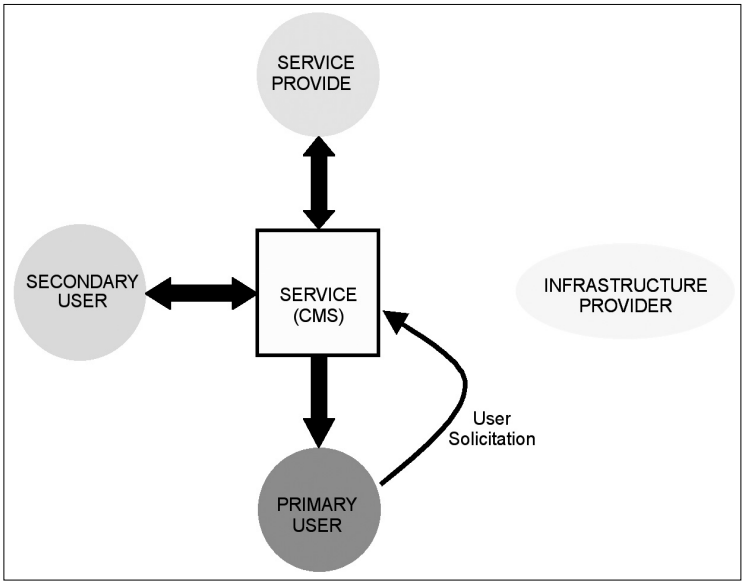
SAMBA's encouraging results allow for envisioning relevant impact in three main dimensions including societal, economical and technological domains:

Societal impact

Promoting the creation of Community Access Channels based on idtv, for the improvement of public services, local community administration reinforcement and decision making and democratic processes in developing countries. The content, created by local communities to their own citizens and users, have a good potential to foster creation of strong social-focused content, e.g. in the fields of eLearning, eHealth, eGovernment, etc.

In a broader scope, the results of SAMBA are envisioned to impact the characterization of key issues re-

Figure 7. Service flow diagram



levant for the adoption of idtv as a feasible technology for bridging the digital divide, especially in development areas. This creation practice would allow moving towards higher societal impact in a self-motivated synchronization process between people and technology.

Economic impact

SAMBA results are envisioned to favor the development of digital terrestrial TV infrastructures and services in developing areas, e.g. the Latin American region, as a successful example of the use of Community Channels in rural areas through PLC infrastructures. SAMBA introduced an easy to use front-end for fostering services and social inclusion through the introduction of idtv as a content production system to be used by local communities under specific business-cases conditions.

Moreover, by bringing together different low cost tools for content production (mobile, web etc.) integrated into Globally Executable MHP, SAMBA is open to a huge potential for new applications and services both in developing and industrialized countries. Additionally, a successful integration could also have potential positive impacts on DVB/MHP adoption in Europe and elsewhere.

Technological impact

The users studies and in specific usability tests performed in SAMBA with real users have brought clear insight to the definition of idtv services in developing countries. The technological solution of SAMBA system is compatible with GEM standard for allowing MHP applications but nothing excludes an easy transformation into other approaches (e.g. Ginga).

Additionally, the identification of relevant features to be included in the PLC equipments to enable convergent multimedia interactive traffic in future applications has been identified as a good approach for improving digital inclusion through SAMBA services; making possible the offering of interactive contents with appropriate user end-to-end QoS perception.

Authors

OSCAR MAYORA-IBARRA obtained his Ph.D. in Electronic Engineering and Informatics at DIBE, University of Genoa, Italy in 2000. In the same year, he joined the Advance Interactive Systems Laboratory at VTT Electronics in Oulu, Finland, as an ERCIM Visiting Research Fellow. In August 2001 he was appointed Associate Professor in Computer Science at Tecnológico de Monterrey Campus Cuernavaca where he became head of the Graduate Program in Computer Science in January 2002. Dr. Mayora has published more than 50 research papers and has been invited as Guest Editor of several special issues in International Journals. Dr. Mayora has chaired several scientific events in areas such as human-computer interaction, pervasive technologies for healthcare, interactive entertainment and user-centric media. Dr. Mayora has coordinated research projects at National and International level including EU FP6 project SAMBA. Since September 2004, Dr. Mayora is the head of Multimedia, Interaction and Smart Environments Group in CREATE-NET research center in Trento, Italy.

DENIS GABOS is an Electronic Engineering (ITA, 1983) and PhD in Digital Systems – Computer Networks (EPUSP, 1999). He is an Associated Researcher and teacher at LARC – Laboratory of Computer Architecture and Networks of Escola Politécnica da Universidade de São Paulo. He is also a Coordinator at Centro Universitário SENAC and Professor at FEI – Faculdade de Engenharia Industrial. He has been participating and coordinating activities in some European Comission research projects (IST) like INSTINC and SAMBA. His main research areas are: Network Management, Convergent Networks and Mobility. He received Décio Leal de Zagotis AWARD from Escola Politécnica. He has 25 years of experience in the area of data communications and computer networks. He worked in development of data communication hardware and software, design and deployment of corporate computer networks. He worked also in the area of project management, consultancy, training and development of training programs on computer networks and information technology for private companies.

ELIZABETH SUCUPIRA FURTADO is researcher in Human Computer Interaction (HCI). She did Ph.D. in Computer Science in 1997 in France and was scholar visiting at HCI laboratory in Stanford University, during 2006. She is consultant of Software Quality Process. She is advisor of the graduate and undergraduate students in the master program in Computer Science, University of Fortaleza (UNIFOR), Ce, Brazil since 1989, in the following themes: System Usability; Definition of tools to support the generation of user interfaces and contents for Interactive TeleVision (iTV) and; Evaluation of systems. She organized the 5th national conference in Human Factors in Computing Systems in 2002; the first workshop on iTV in the Brazilian Conference, in 2005; the first workshop Investigating new user experience challenges in iTV: mobility and sociability, joint with the international conference CHI'2006, in April 2006, Montreal, Canada and; the CLIHC'2009 in Mexico. She has been involved in two relevant projects for Digital TV Systems: the first one refers to the definition of the Brazilian Digital Television System (named SBTVD), financed by the Brazilian Government in 2005; the second one, an European Union-funded project, refers to the development of SAMBA system, in which she performed users' field studies and the interaction design of iTV applications as well as usability evaluations.

ROBERTO CAVALIERE received his Bachelor (B.S.) and Master degrees (M.S.) in Telecommunications Engineering (summa cum laude) from the University of Trento, Italy, in 2006 and 2008, respectively. His main competences are in the area of pattern recognition and remote sensing. Since March 2008 he has been working in the Research and Development team of the Digital Technologies Department in TIS Innovation Park, Bolzano, Italy, where he is involved in numerous projects of local, national and European scope. His current main interests are studying innovative and applied uses of the communication technologies, especially in mobile conditions.

AGOSTINHO CELSO PASCALICCHIO obtained his Bachelor degree in Economic Theory at Univesidade de São Paulo, Brazil in 1976. Later, he received a Master degree in Economic Theory at University of Illinois, USA and currently he is a PhD student at Universidade de São Paulo at Electric and Energy Institute. Currently he is professor at Universidade Mackenzie, Brazil teaching Economic Theory and Project Valuation. He is an Executive with 28 years experience in the financial industry. He has extensive experience in energy projects in particular in hydroelectric and in renewable energy. as well as experience in infrastructure services such as sea port, telecommunication and pulp and paper industry. He has been strategic leader and financial advisor in several infrastructure projects. He is experienced in financial strategies with BNDES (Brazilian National Development Bank). At AES Corp and AES Eletropaulo he estab-

lished benchmarks and international procedures which are in practice up to now. He is a valuation specialist in strategic fields of the electric power business such as the "intelligrid" or "smart grid". He has been financial leader of the first Brazilian sea port privatization: "Paranaguá, Antonina – Ponta do Felix" and financial leader in several "due-diligences" processes.

RODRIGO FILEV MAIA graduated as Electronic Engineer and Master of Electronic Engineering by Escola Politécnica – University of São Paulo (EPUSP). He is a PhD candidate and Researcher of Open System Laboratory from EPUSP and is working with distributed systems (multi-agent), quality of service in heterogeneous networks and convergent services. He worked in IST projects IST-INSTINCT, IST-BELIEVE and IST-SAMBA. He is Assistant Professor of Centro Unviersitário FEI teaching system architecture and programming, and he is the coach of FEI team in IPCP. Its areas of interest include complex system architectures, distributed systems, communication networks and multi-agent architecture.

References

[1] Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking. Source OCDE Science et technologies de l'information, OECD Organization for Economic Co-operation and Development, September 2007.

[2] M.A.M. Kramer, E. Reponen, M. Obrist, MobiMundi: exploring the impact of user-generated mobile content – the participatory panopticon. Proceedings of MobileHCI'08, September 2008.

[3] P. Cesar, K. Chorianopoulos, Interactivity and user participation in the television lifecycle: creating, sharing, and controlling content. Proceeding of the 1st International Conference on Designing Interactive User Experiences for TV and Video (UXTV'08), October 2008.

[4] E. Mantzari, G. Lekakos, A. Vrechopoulos, Social tv: introducing virtual socialization in the tv experience. Proceeding of the 1st International Conference on Designing Interactive User Experiences for TV and Video (UXTV'08), October 2008.

[5] B. Collini-Nocker, Community TV: a new dimension for immersive social networking. Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia (MoMM'08), November 2008.

[6] K. Squire, C. Johnson, Supporting Distributed Communities of Practice with Interactive Television. Journal of Educational Technology Research and Development. Vol. 48. No. 1, 2000, pp.23–43.

[7] O. Mayora-Ibarra, C. Costa, "Sociability Issues for Designing Usable Mobile idtv Applications." CHI 2006 Workshop on Investigating new user experience challenges in idtv: mobility & sociability. Montreal, Canada, April 2006.

[8] Oscar Mayora, Cristina Costa, The SAMBA Approach to Community-Based Interactive Digital Television. Proceedings of ChinaCom Conference, Shanghai, China, August 2007.

[9] Martucci M., Hirakawa A., Jatoba P., "SAMBA Project: A test bed for PLC application as a digital inclusion tool". In Proceedings of IEEE ISPLC'08 – International Symposium on Power Line Communications and Its Applications, Jeju, Korea, April 2008.

[10] Elena Bensi, Giuseppe Franco, Margarita Anastassova, Oscar Mayora-Ibarra, "User-Centric Requirements and Design of a Testbed for iDTV in Alpine Rural Areas". Proceedings of IEEE ISCE'08 Conference, Algarve, Portugal, April 2008.

[11] Lee, K., Phongpaibul, M., Boehm, B., Value-based verification and validation guidelines, Technical Report USC-CSE-2005-502, February 2005.

[12] Furtado E., Kampf T. , Piccolo L., Baranauskas C., "Prospecting the Appropriation of the Digital TV in a Brazilian Project", In Proceedings of EuroiTV'08, Salzburg, Austria, July 2008.

[13] Furtado, E., Mayora, O., Anastassova, M., Kampf, T., Vasconcelos, P., An investigation of iDTV user needs in Brazilian and Italian communities: Preliminary cross-cultural findings. INTERACT Workshop on "Innovation inspired by diversity: Perspectives, Challenges and Opportunities for Human-Computer Interaction in Latin America – CLIHC 2007", Rio de Janeiro, Brazil, September 2007.

[14] L.G. Meloni, A New WiMAX Profile for DTV Return Channel and Wireless Access. Proceedings of Mobile Wimax'07. Orlando, FA, 2007.

Channel allocation technique for eliminating interference caused by RLANs on meteorological radars in 5 GHz band

ZOLTÁN HORVÁTH, DÁVID VARGA

Budapest University of Technology and Economics, Department of Telecommunications
horvathz@hit.bme.hu, varga.david@duvinet.hu

Keywords: 802.11a, interference, meteorological radar, RLAN, WLAN, Wi-Fi, 5 GHz band, RTS/CTS, DFS, 802.11h

Meteorological radars are used for short term weather prediction in Hungary and all over the world. These radars can be jammed by RLAN devices (e.g. home Wi-Fi routers). We introduce the background of this problem, and analyze the weakness of the current solution (DFS – Dynamic Frequency Selection). We analyze it theoretically by modeling the radar operation and RLAN traffic, and we also show its high efficiency in practice, based on well-known IEEE 802.11 RTS/CTS mechanism.

1. Introduction

The introduction of modern meteorological radars has revolutionized accurate short-term forecasts. But at that time nobody thought, that the quickly spread wireless networks (further: Wi-Fi – Wireless Fidelity, WLAN – Wireless Local Area Network, RLAN – Radio Local Area Network) [4] would affect negatively the performance of radar systems – in a large number of countries worldwide [1-3 and 9-15].

In the beginning of the next section (Subsection 2.1) we show how this interference appears on the screen of meteorological radars, and discuss the serious consequences it may cause. We also specify the origins of the interference from a technical point of view. Of course, as the problem expanded, engineers tried to come up with a solution. This led to the development of DFS (Dynamic Frequency Selection), which is a standardized method introduced in IEEE 802.11h.

Of course, the WLAN devices need to comply with it, so DFS compliance tests were introduced in the ETSI 301 893 documents. The ETSI standard is still under development. Almost every year or two a newer version is revealed, trying to make the tests be more similar to real life events. The details of DFS are discussed in Subsection 2.2.

Unfortunately, the DFS still can not provide enough protection for the radar systems; many WLAN devices don't perfectly comply with the standards. We summarize the problems with DFS (identified by us) in Subsection 2.3. The next subsection presents some of the solutions we proposed, that could possibly detect and even filter WLAN interference at the radar systems. Some of these solutions are easier to manage, some are only

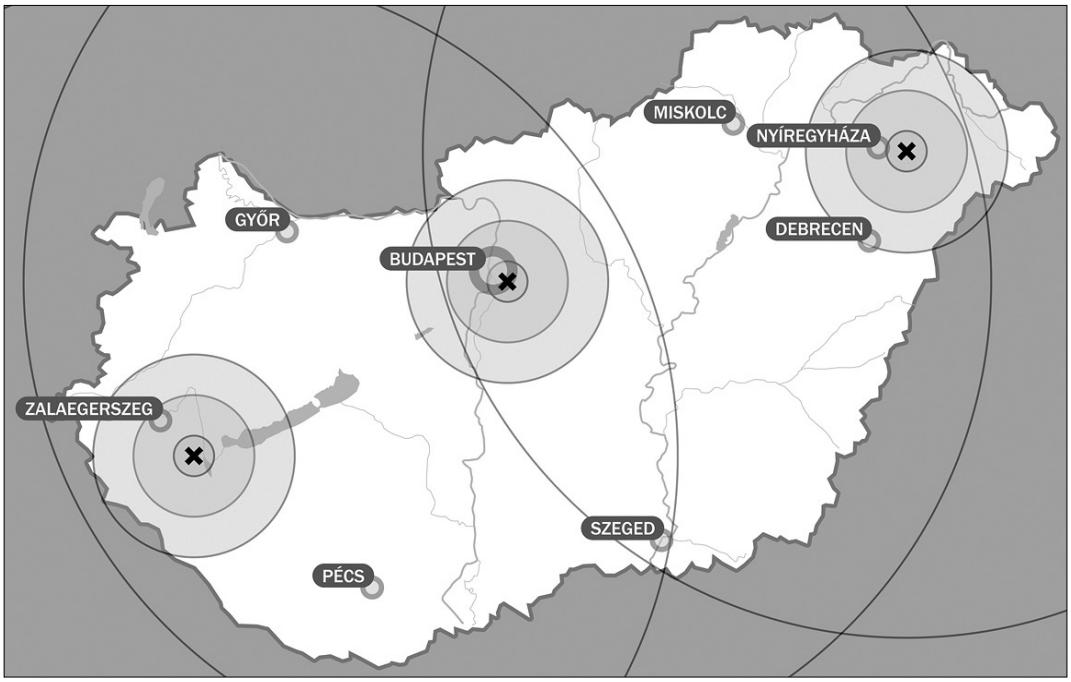


Figure 1. The three meteorological radars in Hungary

The concentric circles show affected areas with 10, 30 and 50 km radius and 240 km radius as maximum of radars observation range around the three meteorological radars in Hungary (Pogányvár, Budapest and Napkor).

theoretical, and could not be implemented because of the technical parameters of the radars.

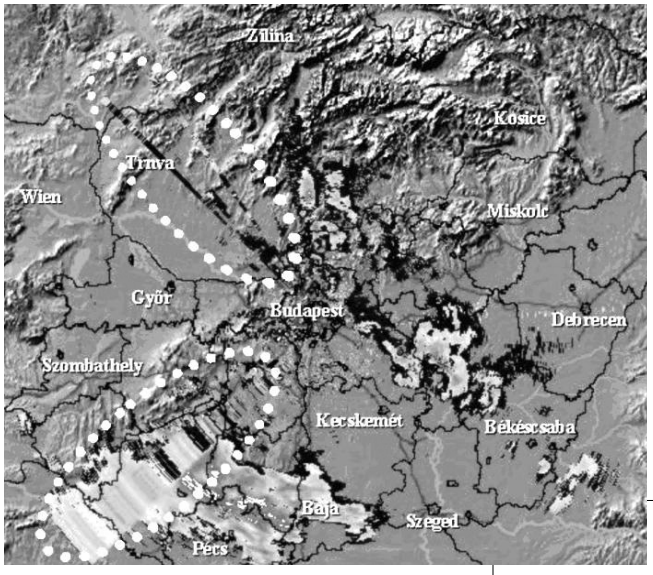
As the main topic of this paper, in Section 3 and 4 we introduce a method, which does not only detect and filter WLAN interference, but also eliminates it before it could actually happen. We present here a preventive solution, which is based on channel allocation. It can reserve the channel for the radar while the measurements are done, by silencing the WLAN transmitters in direction. In Section 3 we present the overview of the main idea and some background information for the next section (Section 4), in which we introduce the allocation technique in detail using traffic models and estimation, and present some evaluation of it. Finally, conclusions are summarized in Section 5.

2. Interference and some solutions

2.1 Introduction of the interference

As part of the European weather forecast system, there are three weather radars in operation in Hungary under the supervision of Hungarian Meteorological Service (OMSZ) and several others throughout Europe and all over the word. The locations of the Hungarian radars are shown in *Figure 1*. These radars measure the atmosphere precipitation. Radar operation is detailed in Subsection 3.3.

Figure 2.
RLAN interference in the picture of the meteorological radar
There are not only clouds in the picture but also strips and sectors are shown marked by dotted curves. They are caused by RLAN interference, and inhibit observing of the precipitation.



Based on the information and pictures provided by the OMSZ, *Figure 2* shows the influence of the strays on a rough radar image. Each shade means a different dBZ level, corresponding to the intensity of the reflected signal. If the shade represents a larger numerical value, it means higher received signal strength [16,18].

The jammed layers indicate significant quantity of rain, so their influence is rather disturbing. It is also dangerous when the signals reflected by precipitation are combined with the ones from the strays (see in the left bottom of Figure 2) and as a result we may come to a false conclusion regarding the quantity of the precipitation. This may cause significant problems in the weather forecast and pre-estimations.

The layers and sectors appearing in the images are mostly caused by IEEE 802.11a standard WLAN devices located close to ground and operating within the radar's frequency range [1-3 and 9-10]. One of the (frequently used) frequency bands where the meteorological radars may operate is between 5600-5650 MHz, which overlaps with 3 of the 802.11a channels (No. 120, 124 and 128) [17]. They are shown in *Figure 3*.

2.2 DFS to solve the problem

A method has been standardized to solve this problem. Dynamic Frequency Selection (DFS) has become the technological solution to resolve the interference issues between meteorological radar systems and WLAN devices. There are two standards related to DFS: the IEEE 802.11h standard and the ETSI EN 301 893 directives.

The IEEE 802.11h standard is [4] an amendment to the original 802.11 standard [4] which deals with the radio spectrum and power management operations in detail. It defines new processes, message types and frame types to be implemented. Although the main function of the standard is to cooperate with European radar systems, it also offers a possibility to have a uniformly used

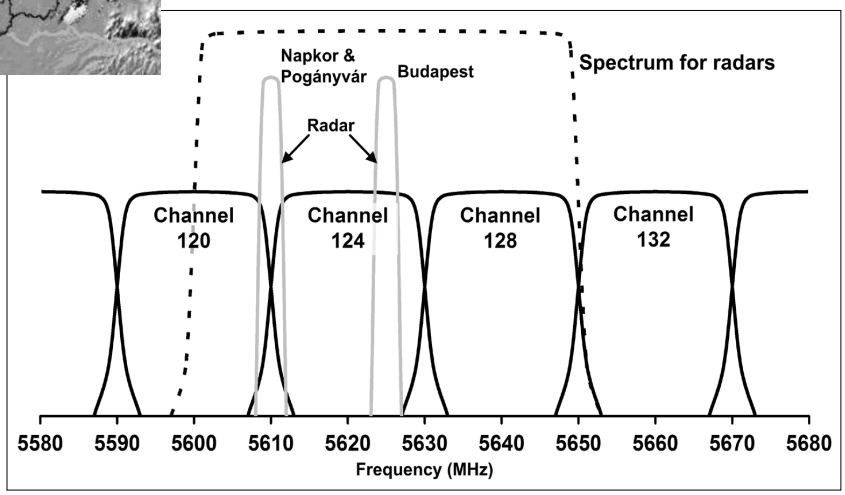


Figure 3.
Usage of the frequency band by 802.11a channels and the meteorological radars in Hungary
The 5600-5650 MHz band for meteorological radars and the three affected 802.11a channels (No. 120, 124 and 128) are shown in the picture. There are two narrow frequency bands used by the three Hungarian meteorological stations (Pogányvár, Budapest and Napkor).

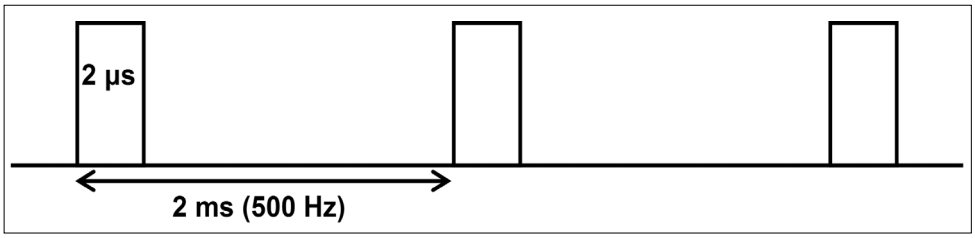


Figure 4.
A sample of radar signals for DFS compatibility tests

Similar DFS test patterns are defined in ETSI directive [5-8]. This one specifies 2 µs pulses with 2 ms repetition time (500 Hz repetition frequency).

radio spectrum, and to manage coverage or power consumption with Transmit Power Control (TPC).

The standard defines horizontal and vertical communication protocols (between stations and within a station, respectively), but it allows the manufacturers to choose their own implementation. It does not even define the conditions (e.g. radar signal detecting), that start the extended functions in the devices. The ETSI EN 301 393 documents [5-8] contain information regarding these conditions [11-15].

The ETSI EN 301 893 standards summarize the functional requirements, that every radio access network operating in the 5 GHz band has to meet. These requirements consist of the specification of transmitted signals, but also contain methods of spectrum management, such as DFS. In practice, a device is marked DFS compatible, if it passes the DFS tests of the actual ETSI EN 301 893 standard (further: ETSI). On the other hand, it is questionable whether this DFS compatibility provides enough protection for meteorological radars.

2.3 Problems regarding DFS

We examined the efficiency of DFS using ETSI v1.4.1 [7] both theoretically and in practice [3], and found that the following problems still exist. We introduce briefly these already known and those revealed by us problems here. One of the known issues is that the minimum pulse width for testing against DFS is 0.8 µs, but Hungarian and other radars also use 0.4 µs for better radar resolution, which is harder to detect [1-3, 6-9 and 11-15]. A sample of a radar signal (as ETSI DFS test pattern) is shown in *Figure 4.*

We found that Channel Availability Check time is only 60 seconds in ETSI v1.2.3 [5], v1.3.1 [6] and v1.4.1 [7], but it can be shorter than the radar rotation period [15]. (Note that this has been changed to 10 minutes in ETSI v1.5.1 [8].)

We found also that DFS Slave devices are not required to sense radar signals. In some scenarios, when a DFS Slave device faces the radar, and the radar signal is too weak at the DFS Master, the WLAN devices will not switch the channel, and the DFS Slave will continuously jam the radar [3].

We collected more than 50 certificates of 802.11a WLAN devices on the market, and most of them only complied with older, v1.2.3 [5] or v1.3.1 [6] versions of ETSI. This means, that even if the device was called DFS compatible at the time it was designed, it would not certainly pass the newer versions of ETSI. But these devices are still in operation, or even can be bought and used.

There were some devices we actually tested, and some of them let the end user enable or disable DFS or Radar signal detection, although this function should be automatically and always enabled.

2.4 Our proposed solutions

As we can see, DFS can not, and probably never will provide a perfect solution against radar interference. We came up with some ideas, which are detailed in [3]. Here we provide a quick overview of them.

If we also detect signals in the full 20 MHz wide 802.11a channel, which embraces the 1.25 MHz wide spectrum of the radar, and we sense signals there at the moment when we receive the reflected radar signals, we may say that there was also WLAN interference. In this case the result of those radar measurements can be ignored.

Interference can also be detected or filtered in time scale, if we only look for reflected radar signals in the time period when they could have returned after reflected by hydrometeors. This possible time period can be calculated from the typical minimum and maximum height of the clouds in the actual season, and the altitude angle of the radar. Interference can also be detected or filtered if precipitation maps are received from other sources, including satellites or terrestrial optical camera system, which can observe without this interference. If we use more radars to scan a selected area, then by comparing the different measurements we are able to detect or filter the interference. This can be done by specific algorithms, or majority voting in case of using at least three radars.

There is a chance to separate radar and WLAN signals, if we use some kind of modulation on the radar signals and we detect the reflected signals via an appropriate demodulator. This way WLAN interference would only cause higher noise in demodulation. Unfortunately this method would require the modification of the radar signals in a way that current magnetron based meteorological radars are unable to provide.

The possible solutions mentioned above are useful only for detecting and filtering the already existing interference. Unlike, using our proposed method discussed in Section 3 and 4, we may eliminate the interference before it even existed.

3. Background of channel allocation for interference elimination

3.1 Overview of channel allocation

The basic idea behind channel allocation is to defer the transmission of the WLAN devices for the time the

radar faces their direction thus we have to allocate that time slot to the radar. This can be done by sending out information to the WLAN stations prior to the critical interval, which forces them to be silent until the radar turns over them.

This allocation transmission should not affect the operation of the radar; therefore the idle time of the radar should be used.

This transmission can be continuous in space domain with an omnidirectional antenna or a fast rotation directional antenna not synchronized with the radar position. Alternatively, this transmission can be concentrated to the direction where the radar measures, which is better, because it does not affect the WLANs which do not jam the radar at the time. In this case the allocation transmission should be synchronized with the radar rotation spatially. The allocation beam should forerun the radar beam for silencing RLANs at time of radar measurements, or should be identical with it (see Figure 9). In this case the radar antenna can be used to transmit channel allocation indication. In this paper we discuss this possibility.

For this allocation we try to use the RTS/CTS mechanism of 802.11, which sets the NAV of the WLAN stations, thus silencing them for a time as necessary. This mechanism is mandatory implemented in all of the WLAN devices.

3.2 Overview of RTS/CTS mechanism in WLAN

The optional RTS/CTS mechanism in 802.11 [4] is basically used to prevent the hidden terminal problem. This problem refers to a scenario, where 'A' wants to send data to 'B', and 'C' is also in the range of 'B', but out of the range of 'A' *(Figure 5)*.

Without RTS/CTS it is possible, that after 'A' starts to transmit, 'C' also starts transmitting, since it senses that the media is free, and creates interference at 'B'. Using the mechanism, prior to sending the actual data,

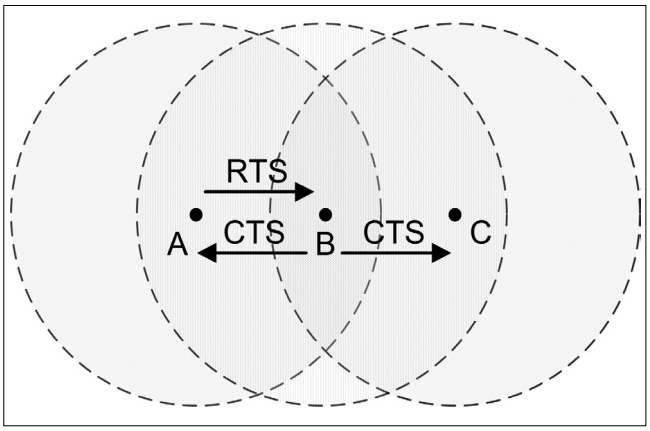


Figure 5.
Hidden terminal problem and RTS/CTS mechanism as a solution

At first node 'A' sends RTS to node 'B' requesting channel allocation. Node 'B' sends CTS indicating that it received RTS and node 'A' get the channel. Node 'C' received CTS, too.

'A' sends a RTS (Request To Send) frame, telling everyone in its range the duration of its following data frame (and belongings, e.g. ACK, SIFSs). Then 'B' sends the CTS (Clear To Send) as a reply, which stations in its range will receive. This way, every station in the range of 'A' and 'B' should not transmit, while 'A' transmits its data.

The RTS/CTS mechanism is generally used in environments where stations in the same network may exist out of each others range.

The mechanism and timing parameters can be seen in *Figure 6*.

3.3 Introduction of weather radar operation

During operation the radar rotates at a specific altitude angle (elevation) or scans a given sector and then raises the elevation.

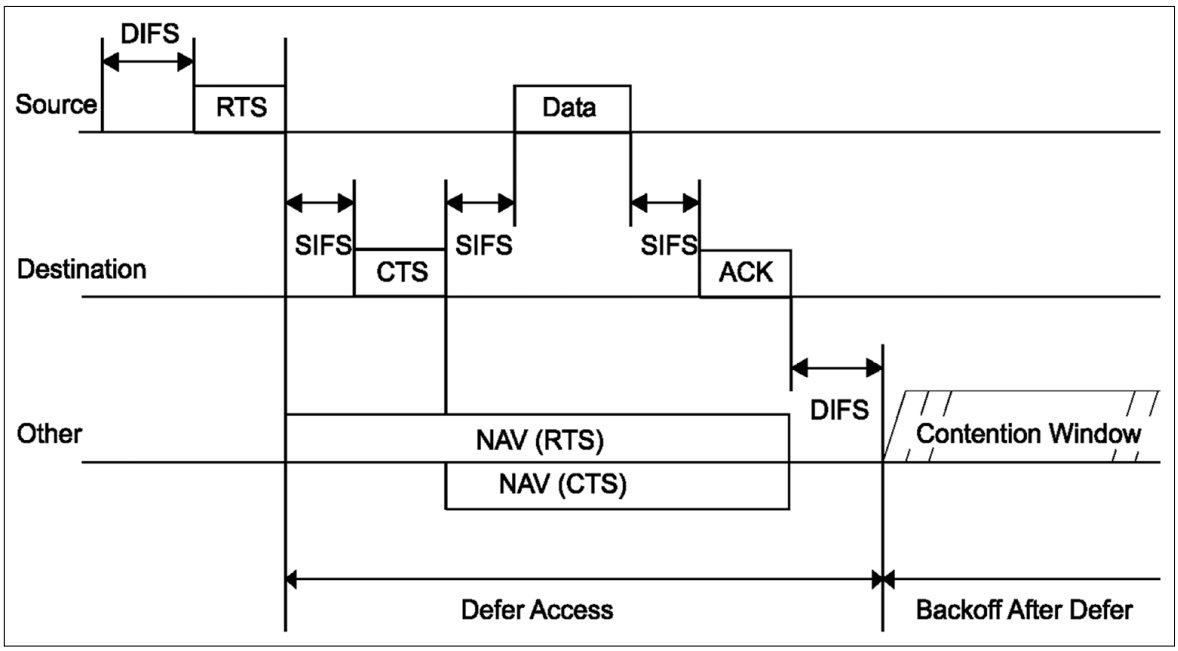


Figure 6.
Timing of RTS, CTS, data, ACK frames and NAV (IEEE 802.11, 2007, Fig. 9-7 [4])

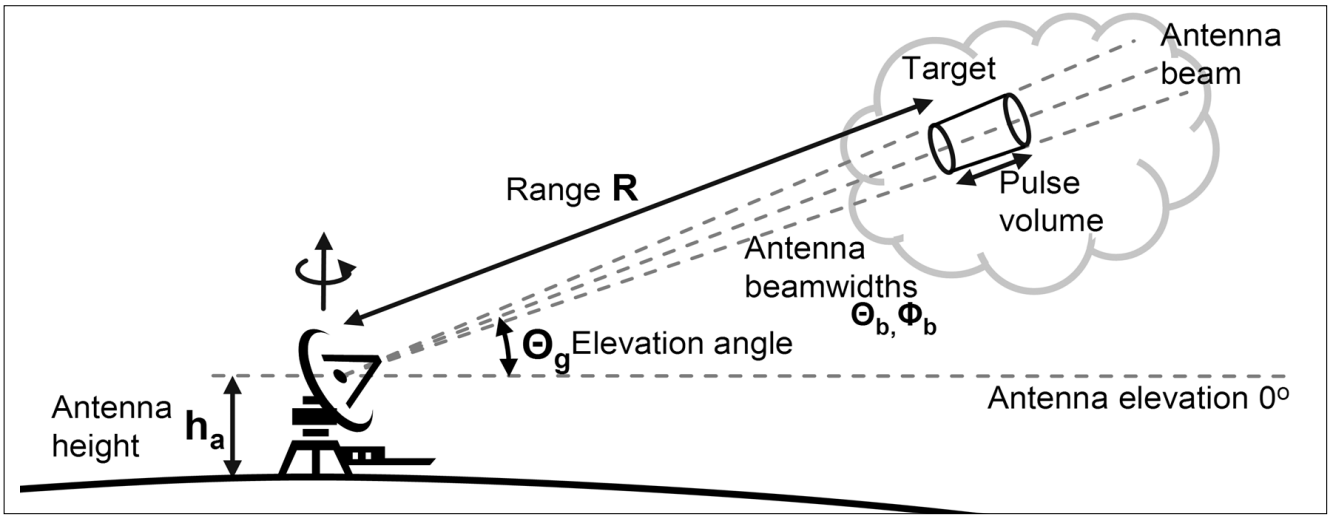


Figure 7. Parameters and operation of meteorological radars
A radar in operation continuously rotates and after each rotation increases elevation angle.

Figure 7 illustrates this operation. In the meantime it transmits radar pulses and receives echoes, reflected by hydrometeors (raindrop, ice) and attenuated by absorption and free space loss (see Fig. 8) [14,16].

Figure 8 shows radar operation in time domain. Each transmitted radar pulse is followed by the echo of it. This backscattering is limited in time by the attenuation and radar signal sensitivity. After each and before the next measurement there is an idle period that will be called 'InterMeasurement Gap' (IMG). In our channel allocation technique this gap is used, so the radar operation and functionality is not affected.

4. Modeling, analysis and evaluation

For analyzing and evaluating the proposed solution building a model is indispensable.

4.1 Modeling radar and RLAN traffic

At first radar and RLAN traffic are described by their timing and other parameters.

4.1.1 Modeling radar operation

As introduced in Subsection 3.3 the radar antenna rotates under operation. Rotation speed is given in RPM (Rotation per Minute) generally in most of the radar specifications. This value – denoted by 'β' – is needed in degree/sec (°/s) measure for further calculation.

$$\beta = \frac{RPM \cdot 360}{60} = RPM \cdot 6 \quad [°/s] \ (1)$$

Another main parameter of the radars is the horizontal beam width (α) measured in degrees (°). These parameters are shown in *Figure 9*.

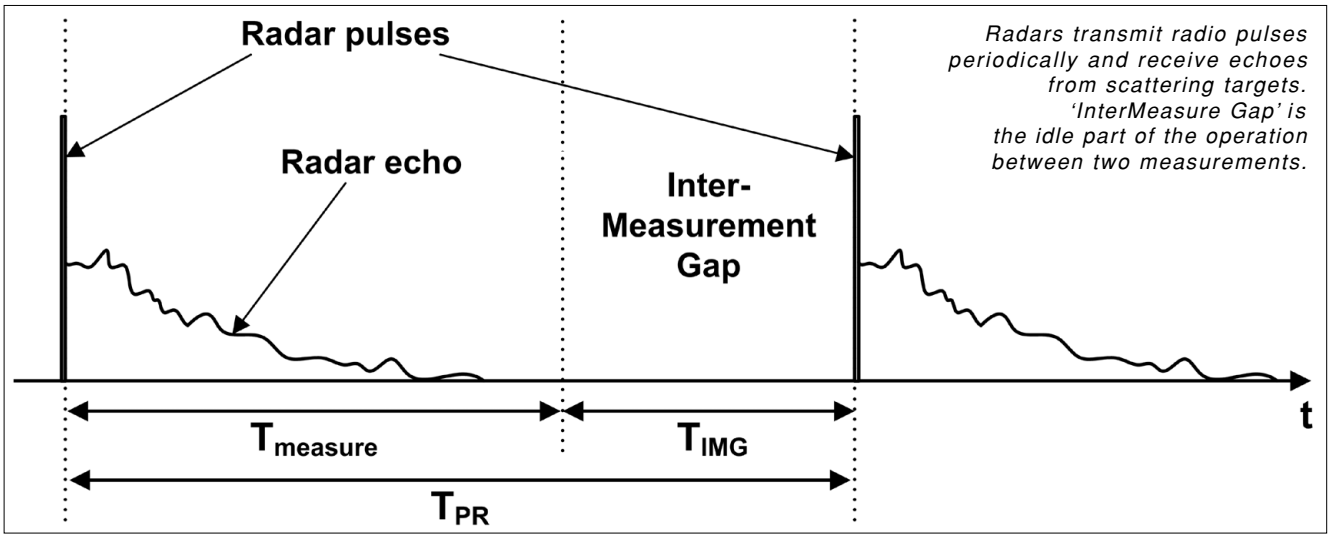
Every radar rotation has a period, when the radar scans a specific point, as described in Subsection 3.3. This 'Contacted Time' (T_{cont}) is constant for each point:

$$T_{cont} = \frac{\alpha}{\beta} \quad [s] \ (2)$$

During this period radars periodically transmit pulses. This is specified as 'Pulse Repetition Frequency' (PRF) in Hz (1/s). It has the same meaning as Pulse Repetition (PR) Time (T_{PR}):

$$T_{PR} = \frac{1}{PRF} \quad [s] \ (3)$$

Figure 8. Transmitted and received signal of radars in time domain



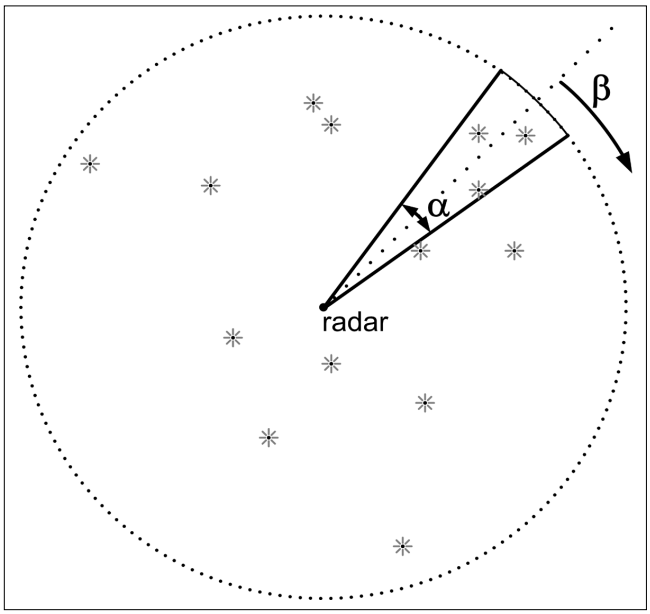


Figure 9. Radar scanning and RLAN devices
The radar rotates with α beam width and β rotation speed.

The time interval between two consecutive pulses can be divided into two periods (see Fig. 8). The first one is between the transmission of a pulse and the theoretical limit when its echo is received by the radar. This measurement time ($T_{measure}$ [s]) is calculated from maximum range of the radar (R [m]) and signal propagation speed ('speed of light') (c [m/s]):

$$T_{measure} = \frac{2 \cdot R}{c} \quad \text{[s]} \quad (4)$$

The second period is the idle time between the end of observation and the next pulse, called Time of InterMeasurement Gap (IMG) (T_{IMG} [s]):

$$T_{IMG} = T_{PR} - T_{measure} = \frac{1}{PRF} - \frac{2 \cdot R}{c} \quad \text{[s]} \quad (5)$$

The pulse length is negligible compared to other durations; therefore it is omitted in this formula. Using these parameters the utilization of the radar and the channel is:

$$U_{measure} = \frac{T_{measure}}{T_{PR}} = \frac{2 \cdot R \cdot PRF}{c} \quad (6)$$

The number of pulses (N_{P_CT}) and InterMeasurement Gaps in 'Contacted Time' (N_{IMG_CT}) is defined as:

$$N_{P\_CT} = N_{IMG\_CT} = \frac{T_{cont}}{T_{PR}} = T_{cont} \cdot PRF \quad (7)$$

4.1.2 Modeling RLAN traffic

Not only radar operation, but also RLAN traffic should be described in order to be able analyze and model the proposed solution.

We analyze two scenarios for modeling RLAN traffic, in 'Scenario I' without and in 'Scenario II' with acknowledgements (ACK).

Scenario I: RLAN traffic without ACKs

In 'Scenario I' RLAN traffic consists of data frames only without any acknowledgement (ACK), therefore RLAN transmission contains frames and idle times. In general distributions of frame size and arrival times are unknown. We use this deterministic traffic pattern for modeling, because this is the worst case: all of the frames use the maximum time duration (with maximum size) (T_{frame} [s]) and minimum interframe time (IFT) ($T_{interframe}$ [s]) consequently this case gives the most occupied channel *(Figure 10)*.

Frame time (T_{frame} [s]) consists of two parts: fixed duration for frame initialization (T_{frame_init} [s]) and the other part depending on frame size (S_{frame} [bit]) and bit rate of transmission (BR_{frame} [bit/s]):

$$T_{frame} = T_{frame\_init} + \frac{S_{frame}}{BR_{frame}} \quad \text{[s]} \quad (8)$$

Channel utilization (U_{frame}) can be calculated with these parameters:

$$U_{frame} = \frac{T_{frame}}{T_{frame} + T_{interframe}} \quad (9)$$

Scenario II: RLAN traffic with ACKs

Unlike the previous scenario, RLANs mostly use acknowledgements (ACKs) for reliable transmissions. After sending the data frame (T_{frame}) RLAN devices wait (T_{ACK_delay}) for the ACK (T_{ACK}). Similarly to the 'RLAN traffic without ACKs' model, this one simulates the worst case in deterministic way. The frame and ACK transmission periods can be grouped together (called 'Extended Frame'), supposing that the further channel allocation technique can not interrupt this frame-ACK communications *(Figure 11)*.

This grouping increases the channel utilization according to the worst case estimation. This allows a more simple way of modeling RLAN traffic with ACK, too.

Duration of frames and ACKs (T_{ACK} [s]) are calculated the same way as before:

$$T_{ACK} = T_{ACK\_init} + \frac{S_{ACK}}{BR_{ACK}} \quad \text{[s]} \quad (10)$$

And the 'Extended Frame' time ($T_{extended\_frame}$ [s]) using 'ACK Delay Time' (T_{ACK_delay} [s]) as mentioned above:

$$T_{extended\_frame} = T_{frame} + T_{ACK\_delay} + T_{ACK} \quad (11)$$

Maximum usage of channel ($U_{extended\_frame}$) can be defined in this scenario, too.

$$U_{extended\_frame} = \frac{T_{extended\_frame}}{T_{extended\_frame} + T_{interframe}} \quad (12)$$

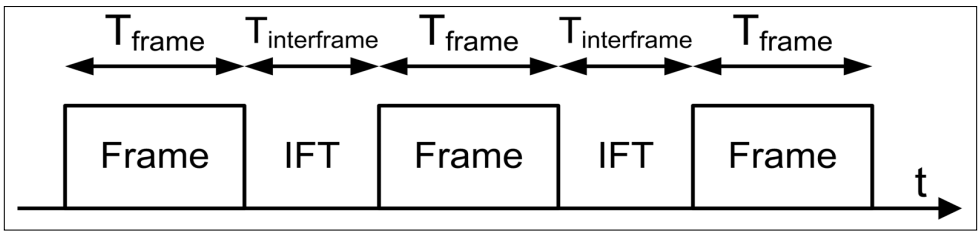


Figure 10.
Traffic scheme and timing for RLANs without ACKs

Frames with the same length (T_{frame}) and idle period (IFT) ($T_{interframe}$) alternates in our RLAN transmission model.

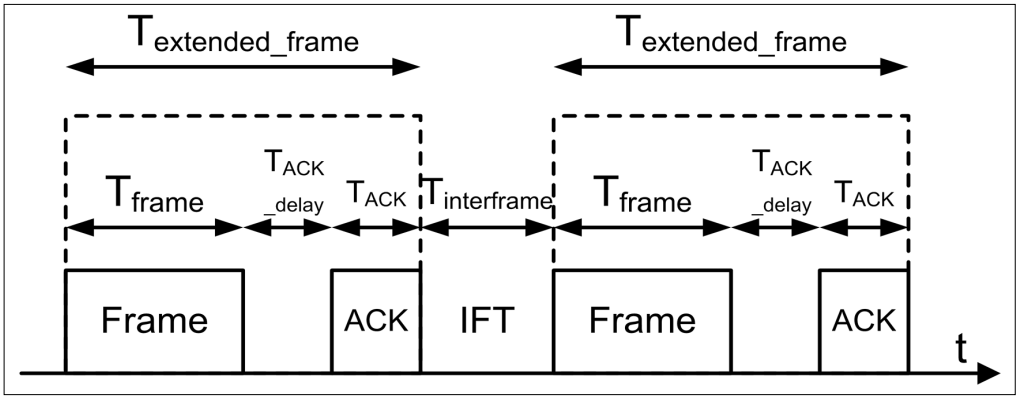


Figure 11.
Traffic scheme and timing for RLAN with ACKs

In acknowledged RLAN transmissions frames and acknowledgements (ACKs) and idle time between them (ACK delay) should be combined into 'Extended Frame'.

This is higher than U_{frame}, because 'Extended Frames' are larger than original frames but 'InterFrame Time' is equal.

4.2 Modeling channel allocation

The overview of the channel allocation technique was given in Subsection 3.1. To achieve our goal (minimizing the traffic of the RLANs during radar scan) we use Channel Allocation Frames (CAFs) (e.g. CTS) in general. RLAN traffic can be blocked for a specific duration by each CAF. But this event occurs only in the case when an RLAN device receives a CAF successfully. CAF can be successful if and only if the beginning of the CAF is in an interframe time (IFT) of the RLAN traffic. However, detecting IFT on the radar side and using detection-based adaptive transmission of CAFs can be difficult and it is unnecessary. When the radar receives signals of more than one RLAN simultaneously, it can detect fewer and shorter idle periods, due to overlapping RLAN traffics. However, CAFs can be transmitted successfully not only in these periods, because each RLAN has its own IFTs, when the allocation can occur. It can be difficult to separate traffic of RLANs, and derive when and which one has its IFT.

We decided that our proposed solution uses a simple deterministic RLAN-traffic-independent CAF transmission without using any detection and without a complex adaptive mechanism, according to the difficulty above. The rate of successful allocation can be maximized by the maximum rate of CAF frequency. This results in a deterministic structure of channel allocation transmission with using short CAFs (T_{CAF} [s]) and as short as possible idle time ('InterCAF Time') (T_{ICAF} [s]) between them *(Figure 12)*. Duration of CAF can be calculated the same way as duration of data frames with the parameters: fixed time for initialization (T_{CAF_init} [s]), size of CAF (S_{CAF} [bit]) and transmission bit rate (BR_{CAF} [bit/s]):

$$T_{CAF} = T_{CAF\_init} + \frac{S_{CAF}}{BR_{CAF}} \quad \text{[s]} \quad (13)$$

This channel allocation operation is used only in InterMeasurement Gaps of radar, as discussed in Subsection 3.3.

4.3 Analysis of proposed solution

As mentioned above, CAF can block RLAN traffic, when an RLAN device detects it. It can occur when the whole CAF is received from its beginning without overlapping with RLAN frames. RLANs using CSMA (Carrier Sense Multiple Access) do not transmit frames after beginning of any frame including CAF is detected. Therefore this successful reception of a CAF becomes a successful channel allocation, too. Applying RLAN traffic and channel allocation models described in Subsection 4.1 and 4.2, the number of successful CAFs during an interframe time (IFT), $N_{CAF\ IFT}$ can be estimated:

$$N_{CAF\_IFT} = \frac{T_{interframe}}{T_{CAF} + T_{ICAF}} \quad (14)$$

In this case we supposed that CAFs and RLAN frames are not synchronized, and one's periodicity is not exactly a multiple of other's in our deterministic model. This condition guarantees the variety of relative positions of CAFs and RLAN frames.

We supposed that CAF traffic does not affect RLAN traffic, only when successfully receiving a CAF. If this

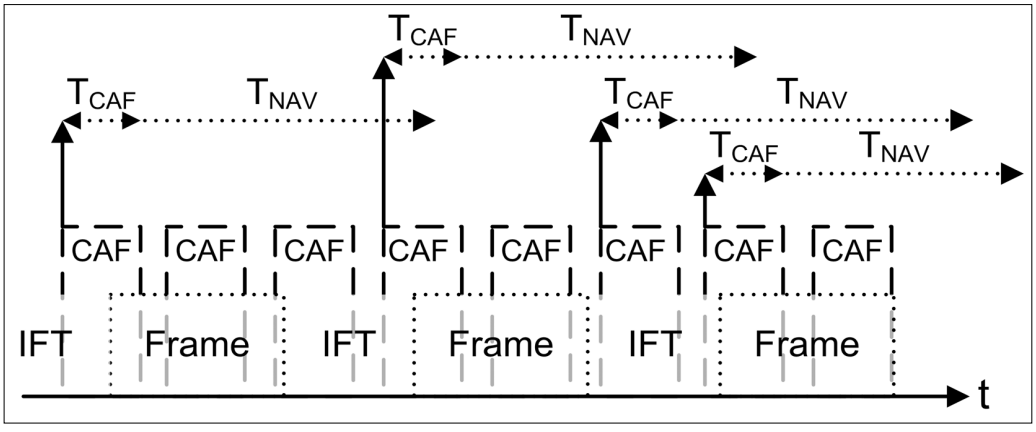


Figure 12.
Scheme and timing for channel allocation frames (CAFs) and RLAN traffic

Channel Allocation Frames (CAFs) are sent periodically and in parallel with RLAN traffic. CAFs can be detected by RLAN devices in their idle time (InterFrame Time). Receiving CAFs successfully mute RLAN devices for a time specified by CAF.

assumption is incorrect, then CAFs can cause longer idle periods and RLAN frame retransmissions (due to frame-CAF collision) in RLAN traffic, too. However, in this case utilization of the RLAN channel can not exceed the worst case limit, as described in *Scenario I* and *II*.

The interframe frequency (number of IFT in one second) (F_{IFT} [1/s]) can be calculated as

$$F_{IFT} = \frac{1}{T_{frame} + T_{interframe}} \quad [1/s] \quad (15)$$

T_{frame} can be replaced with $T_{extended\_frame}$ as necessary.

Using both values (N_{CAF_IFT} and F_{IFT}), the average frequency of successful CAFs in IFT (F_{CAF_IFT} [1/s]) can be estimated, too:

$$F_{CAF\_IFT} = N_{CAF\_IFT} \cdot F_{IFT} = \\ = \frac{T_{interframe}}{T_{CAF} + T_{ICAF}} \cdot \frac{1}{T_{frame} + T_{interframe}} = \\ = \frac{1}{T_{CAF} + T_{ICAF}} \cdot \frac{T_{interframe}}{T_{frame} + T_{interframe}} = \frac{1 - U_{frame}}{T_{CAF} + T_{ICAF}} \quad (16)$$

This result demonstrates our previous two worst case assumptions: efficiency can be increased by minimizing both RLAN traffic utilization and CAF cycle duration. This formula (16) can be used not only in the case of deterministic, but also in the case of random RLAN traffic, only the utilization of the channel should be known.

The above result is modified by usable time slot, so the more relevant value is the frequency of successful CAFs in IFTs in 'InterMeasurement Gaps' (IMG) ($F_{CAF\_IFT\_IMG}$ [1/s]):

$$F_{CAF\_IFT\_IMG} = F_{CAF\_IFT} \cdot \frac{T_{IMG}}{T_{PR}} = \\ = F_{CAF\_IFT} \cdot (1 - U_{measure}) = \\ = \frac{(1 - U_{frame}) \cdot (1 - U_{measure})}{T_{CAF} + T_{ICAF}} \quad [1/s] \quad (17)$$

Accordingly, the average number of successful channel allocations during each IMG is:

$$N_{CAF\_IFT\_IMG} = F_{CAF\_IFT} \cdot T_{IMG} = \frac{(1 - U_{frame}) \cdot T_{IMG}}{T_{CAF} + T_{ICAF}} \quad (18)$$

Another useful measure can be the number of successful channel allocations during a radar scan (T_{cont}) ($N_{CAF\_IFT\_IMG\_Tcont}$):

$$N_{CAF\_IFT\_IMG\_Tcont} = F_{CAF\_IFT\_IMG} \cdot T_{cont} = \\ = \frac{(1 - U_{frame}) \cdot (1 - U_{measure})}{T_{CAF} + T_{ICAF}} \cdot \frac{\alpha}{\beta} \quad (19)$$

Each successful channel allocation protects the radar from RLAN traffic for duration (T_{CAF_NAV}), that is the sum of the time value (NAV – Network Allocation Vector) contained in CAF (T_{NAV}) and the time of CAF itself (T_{CAF}), see Figure 12:

$$T_{CAF\_NAV} = T_{CAF} + T_{NAV} \quad [s] \quad (20)$$

With this value the minimum number of successful channel allocations in each scan period (T_{cont}) ($N_{CA\_Tcont\_min}$) can be estimated:

$$N_{CA\_Tcont\_min} = \left\lceil \frac{T_{cont}}{T_{CAF\_NAV}} \right\rceil = \left\lceil \frac{1}{T_{CAF} + T_{NAV}} \cdot \frac{\alpha}{\beta} \right\rceil \quad (21)$$

One of the most important values that can describe the efficiency of the proposed solution (ρ) is the ratio of occurred effective channel allocations ($N_{CAF\_IFT\_IMG\_Tcont}$) and the number of needed ($N_{CA\_Tcont\_min}$).

$$\rho = \frac{N_{CAF\_IFT\_IMG\_Tcont}}{N_{CA\_Tcont\_min}} = \frac{F_{CAF\_IFT\_IMG} \cdot T_{cont}}{\left\lceil \frac{T_{cont}}{T_{CAF\_NAV}} \right\rceil} = \\ = \frac{(1 - U_{frame}) \cdot (1 - U_{measure}) \cdot \frac{\alpha}{\beta}}{\left\lceil \frac{1}{T_{CAF} + T_{NAV}} \cdot \frac{\alpha}{\beta} \right\rceil} \approx \\ \approx \frac{T_{CAF} + T_{NAV}}{T_{CAF} + T_{ICAF}} \cdot (1 - U_{frame}) \cdot (1 - U_{measure}) \quad (22)$$

The approximation can be applied in case of $T_{cont} >> T_{CAF\_NAV}$. This formula clearly shows which parameters can affect the efficiency of channel allocation dominantly. This efficiency (ρ) can reach or exceed 1, if

$$1 < \frac{T_{CAF} + T_{NAV}}{T_{CAF} + T_{ICAF}} \Leftrightarrow \\ \Leftrightarrow T_{CAF} + T_{NAV} > T_{CAF} + T_{ICAF} \Leftrightarrow \\ \Leftrightarrow T_{NAV} > T_{ICAF} \quad (23)$$

Table 1. Practical parameters and constants

Parameter	Value	Comment
RPM	2	From radars specification (0-6)
β	12°/s	From radars specification (0-36°/s)
α	1°	Average radar beam width (3 dB)
PRF	400 1/s	Generally used from radar specification (250-1300)
R	240 km	Radar range: 0-240 km
c	$3 \cdot 10^8$ m/s	Speed of light
$T_{frame\ init}$	20 µs	IEEE 802.11a: preamble (16 µs) + PLCP (4 µs)
S_{frame}	1516 bytes	Supposing Ethernet traffic (64-1516 bytes – see below)
BR_{frame}	6 Mbps	IEEE 802.11a: 6-54 Mbps
$T_{interframe}$	34 µs	IEEE 802.11a: DIFS + backoff (with 1 time slot (worst case))
$T_{ACK\ init}$	20 µs	IEEE 802.11a: preamble (16 µs) + PLCP (4 µs)
S_{ACK}	14 bytes	IEEE 802.11a
BR_{ACK}	6 Mbps	IEEE 802.11a: 6-54 Mbps
$T_{ACK\ delay}$	16 µs	IEEE 802.11a: SIFS
$T_{CAF\ init}$	20 µs	IEEE 802.11a: CTS preamble (16 µs) + PLCP (4 µs)
S_{CAF}	14 bytes	IEEE 802.11a: CTS
BR_{CAF}	6 Mbps	IEEE 802.11a: CTS: 6-54 Mbps 6 Mbps for worst case and for best receiving conditions
T_{ICAF}	16 µs	We can specify it freely, but for easier implementation and compatibility we set to SIFS.
T_{NAV}	32267 µs	IEEE 802.11: CTS 15 bit value: 0-32267 in µs

4.4 Evaluation in practice

In practice most of the parameters are defined in standards and specifications. In our solution the values of the RLAN traffic parameters come from IEEE 802.11, 802.11a (including RTS/CTS introduced in Subsection 3.2) [4] and weather radar parameters as defined its specifications and using general settings [3,14,18]. For practical evaluation of this allocation technique, the worst case or default values of parameters are given in *Table 1*.

Using these values in practice (Table 1) the parameters of our model can be calculated, see *Table 2*.

Parameter	Value	Parameter	Value
T_{PR}	2500 µs	U_{frame}	86.73 %
$T_{measure}$	1600 µs	$U_{extended\_frame}$	89.06 %
T_{IMG}	900 µs	F_{IFT}	3218 Hz
T_{cont}	83.33 ms	F_{CAF_IFT}	2001 Hz
$U_{measure}$	64 %	$F_{CAF\_IFT\_IMG}$	720 Hz
T_{frame}	222.1 µs	$N_{CAF\_IFT\_IMG}$	1.8
$T_{interframe}$	34 µs	T_{CAF_NAV}	32.3 ms
$T_{ACK} = T_{CAF}$	38.67 µs	$N_{CA\_Tcont\_min}$	3
$T_{extended\_frame}$	332.8 µs	ρ	20.01

Table 2. Calculated parameters

The results of worst case calculations and estimations can be seen in Table 2. We find that since the radar is in idle state in 36% of its time, there is a 900 µs IMG for channel allocation. During an IMG, 1.8 successful channel allocations occur in average, but only 3 are needed during a 83.33 ms 'Contacted Time'. With these worst case parameters the proposed solution allocates channels at least 20 times more often than needed.

These results can be much better if the estimation is based on real parameter values, not on the worst case. For example, using a real distribution of frame sizes gives some improvement. Supposing that the frame size distribution is similar to as it was in 2000 in world wide networks, estimation can be much better. Based on an earlier publication [20], similarly to [19] and [21], the cumulative density function (CDF) of the IP packet size can be obtained. From this CDF the smoothed probability density function (PDF) (or histogram) can be derived, as shown in *Figure 13*.

We suppose these traffic characteristics describe not only Ethernet and backbone traffic, but they are valid for WLAN environment, too. This assumption can be used, because much of traffic is IP and passed through in Ethernet network, which shapes the traffic characteristic to a similar one.

We can find three modes in the histogram of packet size distribution (see Fig. 13). Only 14% of the traffic has the maximal size, 19% is around 570 bytes and almost 33% has the minimal size with 40 bytes. Due to the payload encapsulation and framing every packet gets 16-20 bytes additional overhead. Using these statistics and information, the average frame size can be around 500 bytes. In this case, the efficiency (ρ) of the solution exceeds 35, which means that CAF occur in the average 35 times more often than needed.

We can also analyze the relationship between the efficiency and RLAN bit rate, frame size and using ACKs. This comparison is shown in *Figure 15*.

The *Figure 14* shows that 20 is the lowest efficiency, but under some conditions even 115 can be reached.

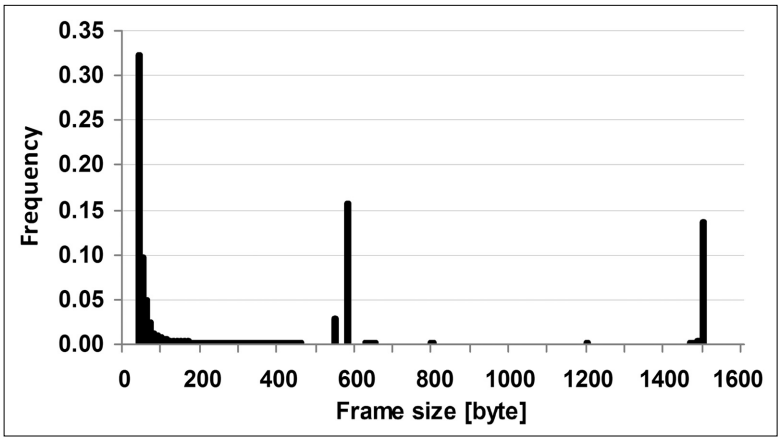


Figure 13.
Frequency of Frame Size

Smoothed probability density function (PDF) (histogram) of frame size in bytes in typical networks (based on [20])

4.5 Applicability

We can see that the solution is more efficient than needed under every circumstance. It can protect meteorological radars against

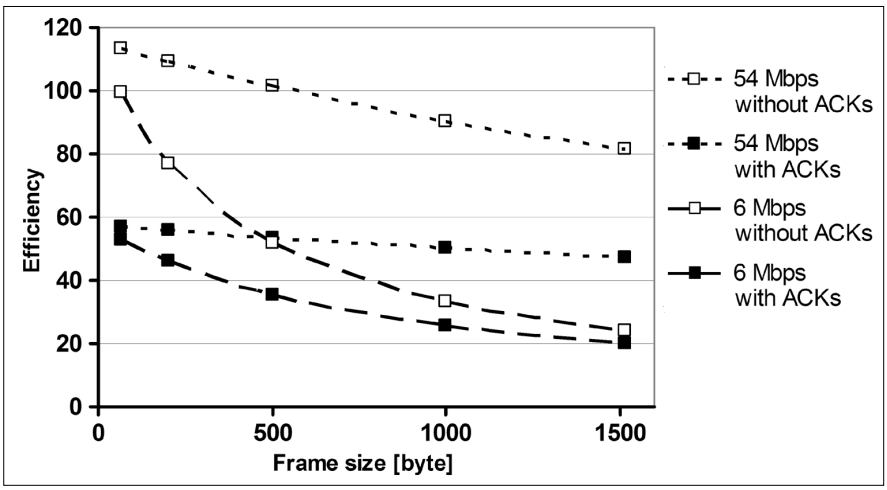


Figure 14.
The efficiency of proposed channel allocation technique

This diagram shows the efficiency connected to RLAN traffic at different frame size at 6 and 54 Mbps with or without ACKs.

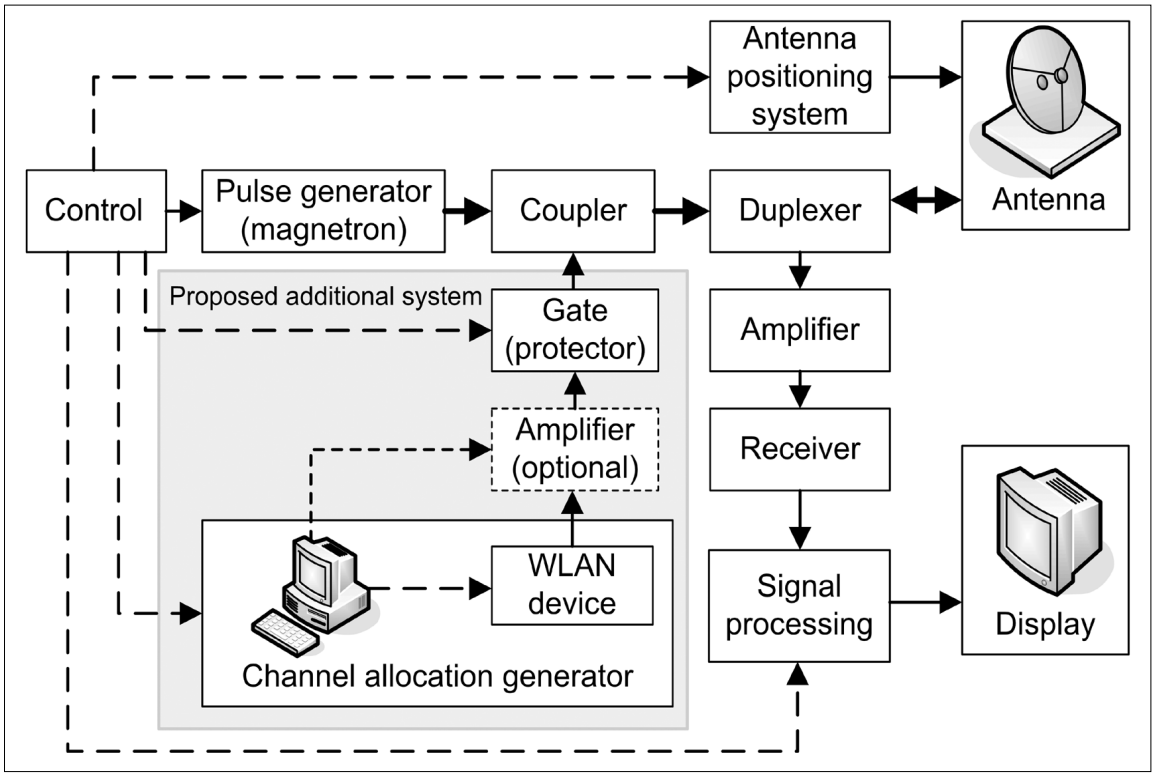


Figure 15. Radar block diagram with the proposed solution

RLAN interference using simple RTS/CTS mechanism. Thus, the implementation of this technique is not too difficult. A simple computer can send CTS frames continuously through a WLAN adapter, synchronized to radar measurement cycle. For higher efficiency we can apply an optional power amplifier. The signal can be transmitted directly into the waveguide of the radar through a coupler, which exits in most of the radars for testing and calibrating purposes. We can also use a controlled gate before the coupler to protect the amplifier against the high power radar pulses (see Fig. 15).

The proposed technique can be used as a standard-compliant solution, because it uses an ordinary WLAN device and a frame type that is specified in the standards. This can not conflict with the radar; moreover it allows undisturbed radar operation.

5. Conclusions

In this paper we have addressed the problem of interference between meteorological radars and RLAN devices. We have evaluated the current solution – DFS – and its limitation. We have proposed some new solutions, and detailed the most viable one: channel allocation based on RTS/CTS. We have given models for radar operations and RLAN traffic. We have shown that using the proposed technique the interference can be eliminated in a very efficient way, due to the mandatory and embedded functionality of RTS/CTS and parameters from standards.

We will try to test this solution in practice soon. We expect this method to be implemented all over the world and will solve the problem of 5 GHz interference.

Acknowledgement

This work has been supported by
Hungarian National Communications Authority (NHH),
Hungarian Meteorological Service (OMSZ)
and High-Speed Networks Laboratory,
Budapest University of Technology and Economics.

Authors

ZOLTÁN HORVÁTH received his M.Sc. degree in Computer Science in 2006 from the Budapest University of Technology and Economics (BME), where he currently pursues his Ph.D. studies at the Department of Telecommunications and gives Computer Networks courses. He has participated in many R&D projects related to planning, testing and building local and metropolitan area networks including WiMAX and community network technologies. He also worked as an advisor for the Hungarian National Communications Authority in applying ETSI regulations, testing devices, EMC measurements and interference examination cooperating with the Hungarian Meteorological Service. He is member of IEEE and Secretary for Project Management Division in Scientific Association for Infocommunications (Hungary).

DÁVID VARGA received his M.Sc degree in Electrical Engineering at Budapest University of Technology and Economics (BME) in 2007. Formerly he took part in building a WiMAX network for testing purposes. He developed a protocol enhancement to the 802.11 WLANs that enables direct communication between wireless stations in infrastructure mode. He worked as an advisor for the Hungarian National Communications Authority cooperating with the Hungarian Meteorological Service in applying ETSI regulations, device testing, EMC measurements and interference examination. Currently he is working on a WLAN based indoor positioning system.

References

[1] Horváth, Z., Lukovszki, Cs., Varga, D., Micskei, T., "Interference on Meteorological Radar and Wi-Fi in 5 GHz band" (in Hungarian), Híradástechnika, Vol. LXIV., No. 5., July 2009.

[2] Horváth, Z., Micskei, T., Varga, D., Lukovszki, Cs., "Interference on Meteorological Radar and WiFi in 5 GHz band", http://www.hit.bme.hu/~hotvathz/publication/radar_wifi_interference_summary_2008.pdf

[3] Horváth, Z., Micskei, T., Seller, R., Varga D., Lukovszki, Cs., "Analyzing WiFi DFS efficiency and opportunities stopping radar interference" (in Hungarian), Hungarian National Communications Authority (NHH), December 2007.

[4] "IEEE Std 802.11-2007, (Revision of IEEE Std 802.11-1999), Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", 2007.

[5] "ETSI EN 301 893 V1.2.3: Broadband Radio Access Networks (BRAN); 5 GHz high performance RLAN; Harmonized EN covering essential requirements of article 3.2 of the R&TTE Directive".

[6] "ETSI EN 301 893 V1.3.1: Broadband Radio Access Networks (BRAN); 5 GHz high performance RLAN; Harmonized EN covering essential requirements of article 3.2 of the R&TTE Directive".

[7] "ETSI EN 301 893 V1.4.1: Broadband Radio Access Networks (BRAN); 5 GHz high performance RLAN; Harmonized EN covering essential requirements of article 3.2 of the R&TTE Directive".

[8] "ETSI EN 301 893 V1.5.1: Broadband Radio Access Networks (BRAN); 5 GHz high performance RLAN; Harmonized EN covering essential requirements of article 3.2 of the R&TTE Directive".

[9] ITU-R Radio Communication Study Groups, "Studies on the effect of wireless access systems including RLANs on terrestrial meteorological radars operating in the band 5600-5650 MHz (documents 8A/103?E and 8B/65?E)", International Telecommunication Union (ITU), Geneva, Switzerland, 30 August 2004.

[10] Brandao, A. L., Sydor, J., Brett, W., Scott, J., Joe, P., Hung, D., "5 GHz RLAN interference on active meteorological radars", Vehicular Technology Conference (VTC'05), 30 May–1 June 2005, Vol.2, pp.1328–1332.

[11] European Telecommunications Standards Institute, "DFS Update: European Weather Radars – Details & Overview", BRAN#52, Sophia-Antipolis, 8-11 October 2007. http://www.ieee802.org/18/Meeting_documents/2007_Sept/BRAN52d014_European_Weather_Radar_SIgnals_-_Details__Overview.pdf

[12] Wi-Fi Alliance, Spectrum & Regulatory Committee, "Spectrum Sharing in the 5 GHz Band – DFS Best Practices", 10 October 2007. http://www.ieee802.org/18/Meeting_documents/2007_Nov/WFA-DFS-Best%20Practices.pdf

[13] ITU – Radiocommunication Study Groups, "Theoretical analysis and testing results pertaining to the determination of relevant interference protection criteria of ground-based meteorological radars", Draft new REPORT ITU-R M.2136, Working Party 5B, 3 December 2008.

[14] ITU – Radiocommunication Study Groups, "Technical and operational aspects of ground-based meteorological radars", Draft new RECOMMENDATION ITU-R M.[MET-RAD], Working Party 5B, 3 November 2008.

[15] EUMETNET, "Recommendation on C-Band Meteorological radars design to ensure global and long-term coexistence with 5 GHz RLAN", 35th EUMETNET Council, Reading, UK, 4 December 2008.

[16] Collier, C.G., "Applications of Weather Radar Systems: A guide to uses of radar data in meteorology and hydrology", John Wiley and Sons, New York, 1989.

[17] Hungarian National Communications Authority (NHH), "Broadband Data Transmission with Wireless Access Devices", 2nd Ed., Budapest, 1 October 2006. http://www.nhh.hu/dokumentum.php?cid=11887

[18] Hungarian Meteorological Service (OMSZ), Weather Radar Images, Budapest, 2007-2009, http://www.met.hu

[19] Claffy, K.C., Miller, G., Thompson, K., "The Nature of the Beast: Recent Traffic Measurements from an Internet Backbone", The 8th Annual Conference of the Internet Society (INET'98), Geneva, Switzerland, 1-24 July 1998. http://www.pdos.lcs.mit.edu/decouto/papers/claffy98.pdf

[20] McCreary, Sean and Claffy, K.C., "Trends in Wide Area IP Traffic Patterns: A View from Ames Internet Exchange", Cooperative Association for Internet Data Analysis (CAIDA), San Diego Supercomputer Center, University of California, San Diego, ITC Specialist Seminar, 2000. http://www.caida.org/publications/papers/2000/AIX0005/AIX0005.pdf

[21] Fraleigh, C., Moon, S., Lyles, B., Cotton, C., Khan, M., Moll, D., Rockell, R., Seely, T., Diot, S.C., "Packet-level traffic measurements from the Sprint IP backbone", IEEE Network, Vol.17, No.6, Nov-Dec. 2003, pp.6–16.

Markov model based location prediction in wireless cellular networks

TAMÁS SZÁLKA, SÁNDOR SZABÓ, PÉTER FÜLÖP

Budapest University of Technology and Economics, Department of Telecommunications
{szalkat, szabos, fepti}@mcl.hu

Keywords: mobility modelling, Markov model, location prediction, cellular network

The efficient dimensioning of cellular wireless access networks depends highly on the accuracy of the underlying mathematical models of user distribution and traffic estimations. Mobility prediction also considered as an effective method contributing to the accuracy of IP multicast based multimedia transmissions, and ad hoc routing algorithms. In this paper we focus on the trade-off between the accuracy and the complexity of the mathematical models used to describe user movements in the network. We propose Markovian mobility models, in order to utilize the additional information present in the mobile user's movement history thus providing more accurate results than other widely used models in the literature. The new models are applicable in real-life scenarios, because these rely on additional information effectively available in cellular networks (e.g. handover history), too. The complexity of the proposed models is analyzed and the accuracy is justified by means of simulation.

1. Introduction

Different mobility models have been proposed in the literature to cope with user mobility in different wireless and mobile networks (e.g. cellular networks, ad hoc networks, etc.). Mobility model based prediction provides useful input for dimensioning and planning of wireless mobile networks, ad hoc routing algorithms, efficient multicast transmission and call admission control [3,8].

One of the well known mobility models is *Random Walk Mobility* model, which is often used in network planning and in analyzing network algorithms, because of its simplicity [1].

In the Random Walk Mobility model [4] the node moves from its current location to a new location by randomly choosing a direction and a speed. The Random Walk model defines user movement from one position to the next one with memoryless, randomly selected speed and direction. Each movement in the Random Walk Mobility Model occurs in either a constant time interval *t* or in a constant distance travelled *d*. When a mobile node reaches a simulation boundary, it simply bounces off the simulation border with an angle determined by the incoming direction, then the node continues moving along this new path. As we mentioned, this is very easy to use, but on the other hand this very simple model presumes unrealistic conditions, like uniform user distribution in the mobile network.

To handle different levels of randomness, one can use the *Gauss-Markov Mobility Model*. In this model initially each node is assigned a current speed and direction, and at fixed intervals of time, the speed and direction of the nodes are updated. The value of speed and direction at any time instance are calculated based upon the value of speed and direction at the previous instance and a random variable [5].

However, in real-life networks, geographical characteristics such as streets and parks influence the cell residence time (dwell time) and movement directions of users in the network, and result in a non-uniform user density. While these models are appropriate for mathematical analysis, easy to use in simulations and for trace-generation, they fail to capture important characteristics of mobility patterns in specific environments, e.g. time variance, location dependence, unique speed and dwell-time distributions [2].

In order to follow user movement patterns more efficient, City Section Mobility, or Mobility Vector model can be used. In the *City Section Mobility Model* a section of a city is represented where an ad hoc network exists [6]. The streets and speed limits on the streets are based on the type of city being simulated. The streets might form a grid in the downtown area of the city with a high-speed highway near the border of the simulation area to represent a loop around the city.

If a flexible mobility framework for hybrid motion patterns is needed, one can rely on the *Mobility Vector* [7] model. A mobility vector expresses the mobility of a node as the sum of two sub vectors: the Base Vector and the Deviation vector. The base vector defines the major direction and speed of the node while the deviation vector stores the mobility deviation from the base vector. The mobility vector is expressed as an acceleration factor in different directions.

Beside the mobility models many works have discussed prediction algorithms, too. The prediction serves as an input for an optimal resource planning.

The *shadow cluster scheme* [9] estimates future resource requirements in a collection of cells in which a mobile is likely to visit in the future (as a "shadow" of the user). The shadow cluster model makes its prediction based on the mobile's previous routes. In this mo-

del, the highway traffic with various constant speeds is simulated and users travel in forward and backward directions. The shadow cluster model improves estimation of resources and decision of call admission.

User movements in a cellular network can be described as a time-series of radio cells the user visited. The handover event of active connections (e.g. cell boundary crossing) is recorded in the network management system's logs, thus the information can be extracted from the management system of cellular mobile networks, such as GSM/GPRS/UMTS networks. The user movements are described by the dwell-time and outgoing probabilities (the probability of a user leaving for each neighbouring cell, called the handover vector). These parameters can be calculated for each cell based on the time-series of visited cells of the users. However, in some cases, these parameters – dwell-time and outgoing probabilities – are not enough to capture all the information present in the time-series of user movements. In many situations the movement patterns can be estimated more precisely if the model also considers the conditional probabilities between the incoming and outgoing directions. In this paper we investigate the effect of incorporating this additional information into the mobility models on the accuracy of the models.

The results of this paper are applicable in engineering tasks of network dimensioning, can provide input for more effective Call Admission Control algorithms in order to ensure user's satisfaction and optimal resource usage in cellular wireless mobile networks [3].

2. Cell-centralized Markov mobility model

2.1 Motivation

The memoryless property of the Markov process makes the Markov models in the literature easily applicable. In many models, the Markov processes states are based simply on the physical radio cells, i.e. one state represents one radio cell. In this case, any potentially present additional information in the user movements cannot be included in the model. We propose a model, in which the states of the model are constructed according to a group of cells belonging to typical movement directions.

Our model takes into account the user history during the prediction and merges the neighbor cells into a direction group depending on the user behavior and historical distribution and patterns, yet retaining the memoryless property.

2.2 The model enhancement

In our method the direction of a user is identically distributed between 0 and 2π. The user's speed is between 0 and V_{max}. After moving in a direction with a randomly chosen speed for a given Δt time, the user changes its direction and speed. N_{cell} denotes the number of cells in the network.

A possible classification of cells can be seen in *Figure 1*. Using this case a user can be located in three different states during each time slot in simple Markov-chain based model, the stay state *(S)* and the left-area state *(L)* and the right-area state *(R)*.

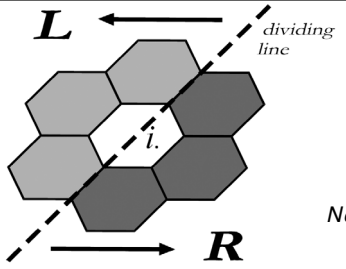


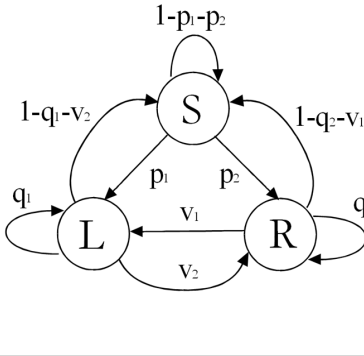
Figure 1. Neighbor cells separated into two groups

The grouping can be derived from the user behavior. If the users in right-hand side cells behave similarly from the current cell's point of view, the neighboring cells will be merged into a common cell group, which represents a state in the Markov model (*R* state). Other grouping methods can be used as well, i.e. a stand-alone cell can constitutes a group also. In our example model each of the two groups (*R* and *L*) contains three cells.

Let us define the random variable *X(t)*, which represents the movement state of a given terminal during time slot *t*. We assume that *{X(t), t* =0,1,2,...*}* is a Markov-chain with transition probabilities *p, q, v*.

If the user is in state *S* of Markov model for cell *i (current cell)*, it remains in the given cell. If the user is in state *R*, it is in range of the cells on the right-hand side, if in state *L*, it is in the left-hand side of the dividing line.

Since the transition properties are not symmetric, the left-area state and the right-area state have different probabilities. *Figure 2* depicts the Markov-chain and transition (Π) matrix.



$$\Pi = \begin{bmatrix} 1-p_1-p_2 & p_1 & p_2 \\ 1-q_1-v_2 & q_1 & v_2 \\ 1-q_2-v_1 & v_1 & q_2 \end{bmatrix}$$

Figure 2. State diagram and Π matrix of 3-state M-model (M_3)

As we mentioned before, our aim is to use the information present in the previous steps of the users for motion prediction. In this case, the transition probabilities are not handover intensities, but contain more information, specifically:

p – Probability of the event that the user stays for a timeslot in cell *i*, and in the next timeslot it moves one of the cell groups. According to the mentioned model (Fig. 2), p_1 means that the user moves in the next timeslot into one of the right (*R* state) neighbor cells.

q – Probability of the event that the user came into cell i from one of the cell groups, and in the next timeslot it moves into the same cell group where it comes from. Thus q_1 means that the user comes from the right-hand side neighbor cells *(R)* and moves into the same, *R* group (but not necessarily into the same cell of the cell group).

v – Probability of the event that the user comes into cell i from one of the cell groups, and in the next timeslot it moves into another cell group. According to the above mentioned model (see Fig. 2), v_1 means that the user comes from left neighbor cells *(L)* and moves into the right group *(R)*.

The probabilities introduced above can be calculated from real network traces.

How can we use this model?

How can we predict the user's next steps?

We assume that mobile users are in cell i at the beginning of a timeslot t. With each i cell a previously introduced Markov model is associated, which handles the users movement in this cell. The network operator's log contains the information where the users from cell i were in timeslot $t-1$ (according to the assumption that they were in one of the neighbor cells). Using this information as *initial distribution P(0)*, we calculate *P(1)* with the transition matrix of the model: $P(1) = P(0) \cdot \Pi$.

P(1) shows the predicted users distribution for the $t+1$ timeslot, in other way we predicted where the users will be from cell i *in timeslot* $t+1$.

The number of users in cell i at timeslot $t+1$, $N_i(t+1)$, is given by (1).

$$N_i(t+1) = N_i(t) \cdot P_S^{\,i}(1) + \frac{1}{3} \sum_{l \in S^i_{adj(L)}} N_l(t) \cdot P_R^{\,l}(1) + \frac{1}{3} \sum_{r \in S^i_{adj(R)}} N_r(t) \cdot P_L^{\,r}(1) \quad (1)$$

where $S^i_{adj(R)}$ means the set of cell indexes from the right-hand side cells of cell i, $S^i_{adj(L)}$ means the set of cell indexes from the left-hand side cells of cell i.

The steady state probabilities of the Markov model can represent a steady user distribution if the network parameters do not vary.

The balance equations for this Markov-chain are given in Eq. (2).

$$\begin{aligned} P_S \cdot (p_1 + p_2) &= P_L \cdot (1 - q_1 - v_2) + P_R \cdot (1 - q_2 - v_1) \\ P_L \cdot (1 - q_1) &= P_S \cdot p_1 + P_R \cdot v_1 \\ P_R \cdot (1 - q_2) &= P_S \cdot p_2 + P_L \cdot v_2 \end{aligned} \quad (2)$$

We also know that $P_S+P_L+P_R=1$, thus the steady state probabilities can be calculated.

Knowing the result we can predict the number of mobile terminals for time slot $t+1$ in a steady state:

$$N_i(t+1) = N_i(t) \cdot P_S^{\,i} + \frac{1}{3} \sum_{l \in S^i_{adj(L)}} N_l(t) \cdot P_R^{\,l} + \frac{1}{3} \sum_{r \in S^i_{adj(R)}} N_r(t) \cdot P_L^{\,r} \quad (3)$$

As we mentioned earlier, this model is based on possible cell grouping (*R* and *L* states) and performs well when the user's movement has only one typical direction, because in this case the handover intensities of the right-move (or the left-move) cells do not differ significantly.

If we try to predict the user's distribution in a city having irregular, dense road system, or in big parks, then the handover intensities could differ. From this point of view the best way is to represent all of the neighbor cells as a separated Markov state, so we create an $n+1$-state Markov model:

– stationary state *(S)*

– move to neighbor 1...n state *(M_{N1}...M_{Nn})*

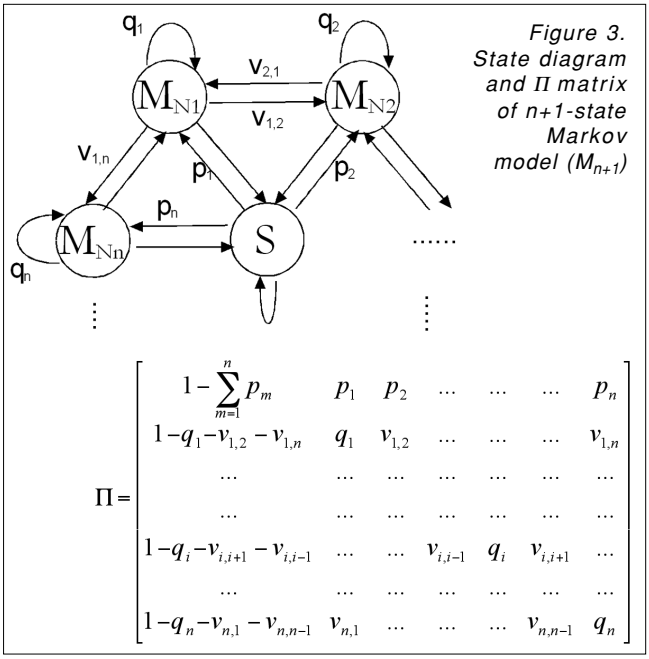
The steady state probabilities can be calculated as in the previous cases (Eq. 4).

$$\begin{aligned} & P_S \cdot \sum_{i=1}^{n} p_i = \sum_{j=1}^{n} (1 - q_j - \sum_{l \neq k}^{n} v_{k,l}) P_{Ni} \\ & \ldots \\ & P_{Nk}(1 - q_k) = P_S \cdot p_n + \sum_{i \neq k} P_{Ni} \cdot v_{i,k} \qquad 1 \le k \le n \\ & \ldots \end{aligned} \quad (4)$$

Using the steady state result the predicted number of users in the next time slot is given by Eq. (5).

$$N_i(t+1) = N_i(t) \cdot P_S^{\,i} + \sum_{j \in S^i_{adj}} N_j(t) \cdot P_{M_i}^{\,j}(j) \quad (5)$$

where S^i_{adj} means the set of cell indexes from all of the neighbors cells of cell i.



$$\Pi = \begin{bmatrix} 1 - \sum_{m=1}^{n} p_m & p_1 & p_2 & \ldots & \ldots & \ldots & p_n \\ 1 - q_1 - v_{1,2} - v_{1,n} & q_1 & v_{1,2} & \ldots & \ldots & \ldots & v_{1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 - q_i - v_{i,i+1} - v_{i,i-1} & \ldots & \ldots & v_{i,i-1} & q_i & v_{i,i+1} & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 - q_n - v_{n,1} - v_{n,n-1} & v_{n,1} & \ldots & \ldots & \ldots & v_{n,n-1} & q_n \end{bmatrix}$$

Figure 3. State diagram and Π matrix of n+1-state Markov model (M_{n+1})

2.3 Complexity and accuracy of the model

Based on the Markov model generator method introduced in the previous section, a specific model can be derived depending on the complexity limits and the precision (accuracy) demand.

The accuracy of the model increases as the number of states grows. The number of states grows when the movement history (time dimension) is increased, or when the number of direction (direction dimension) is increased. Increasing the time dimension increases the number of states exponentially, increasing the direc-

tion dimension increases it linearly. With the state-space rising, the computational complexity of the Markov steady state calculations also follows a rising curve. The question is the characteristic of these functions and the existence of a theoretical or practical optimum point.

It is assumed that each cell has *N* neighbors and the 3-state (stay, left and right-move state) model is used to determine the users movement. It is also assumed that *N*/2 cells belong to both left and right Markov-states, and the users are uniformly distributed between cells. A theoretical error can be derived from this assumption since in most cases the user motion pattern does not result in a uniform distribution in the *N*/2 cells. In the worst case the users move with probability 1 into one of the neighbor cells. The error can be measured with the difference between the uniform distribution and the worst case. This difference is given by (6).

$$1-\frac{1}{N/M}+\frac{1}{N/M}((N/M)-1)=2\left(1-\frac{1}{N/M}\right), \qquad (6)$$

where *N* means the neighbor numbers, and *M* means the direction numbers in the model.

We measured the computation complexity also as the function of state number. This enables us to compare complexity and prediction error in an easy way. Based on the 1...*M*-state model the prediction computation is calculated with the costs of Markov steady state mathematical operations and other procedures necessary for transition probabilities. The complexity can be estimated with (M_{states}^3+ M_{states}+ 1/M_{states}).

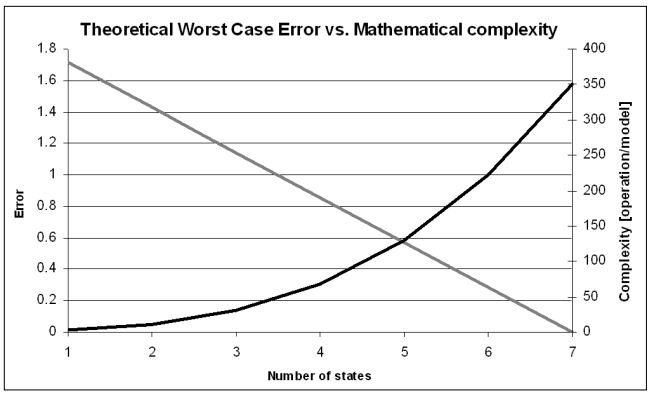


Figure 4. Complexity and accuracy of the calculation

Figure 4 shows the complexity and error characteristic. In the given model calculations the optimal point of operation is around 5 states where error is minimal at this level of error.

In the previous comparison the direction dimension sizing is used only. If we want to rely on the information of the previous steps for the estimation, then movement history has to be introduced into the model. In fact this means that every state in the Markov-chain has to be changed with *M* states. This causes exponential state number explosion that can be seen in *Figure 5*.

As we have extended our model step by step in time and in direction dimension, its precision increased along with the complexity of the model. In order to minimize

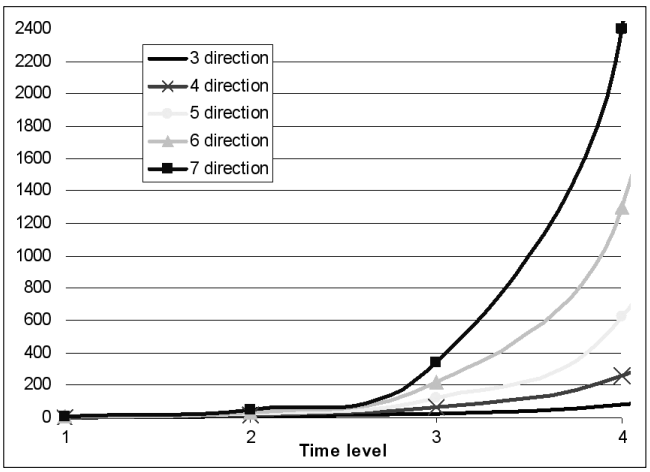


Figure 5.
The increase of number of states in the case of different direction dimensions (3,4,5,6,7 states), and different time dimensions (1,2,3,4)

the complexity, we developed a simple algorithm, which is able to minimize the number of states based on merging adjacent cells. Due to size limitation this algorithm is not discussed in this paper.

2.4 The effect of information of previous visited cells in the model

Neglecting the recent transition series of users in the cluster, i. e. when the model does not take into account in model buildup the previous steps of mobile users, the estimation is less precise.

In this section we show a simple example of the effect of previous user steps on the accuracy of the prediction. Let us consider two routes as shown in *Figure 6/b*. The accuracy of transition probability estimations is better if the model knows where the users came from, compared to the RW-like estimation which cannot differentiate the users on the two routes *(Figure 6/a)*.

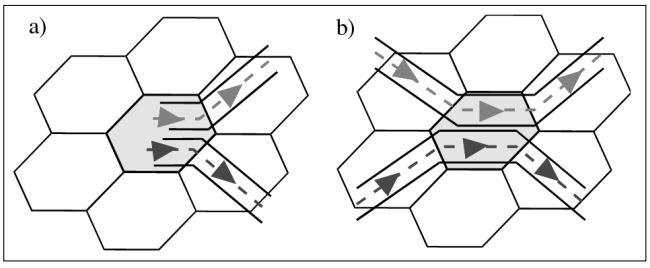


Figure 6. User prediction methods
a. Model without memory, unknown where is the users come from
b. Model with memory, the previous steps of the users taken in account

To find the error rate of a movement-history-less RW model compared to an algorithm that possesses user distribution from history, we use the example cells shown in Fig. 6. Let us define the following parameters:

– the number of incoming users on the upper route at timeslot *t* is $in1_t$,
– the number of incoming users on the lower route at timeslot *t* is $in2_t$,

– similarly the number of users leaving on the upper route at timeslot *t* is $out1_t$,
– the number of users leaving on the lower route at timeslot *t* is $out2_t$,

and the user movement directions with a simple transition matrix $\mathbf{P}=\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$.

The RW model in Fig. 6/a calculates the number of leaving users *(out1, out2)* with a history estimation of the **P** matrix, and with the sum of *in1* and *in2* (the total number of users in the observed cell), but without the knowledge of *in1, in2*.

In our proposed algorithm, the use of *in1, in2* means that the model calculates based on the information of the incoming users. The incoming users can be considered uniform or different (marked users). Since the latter is more accurate, in this comparison we use marked users.

Assume that the history estimation of the RW model's **P** matrix is based on the previous timeslot. That is the $P(out1_{t+1})$ and $P(out2_{t+1})$ probabilities can be substituted with the relative frequency of $out1_t/(out1_t+out2_t)$ and $out2_t/(out1_t+out2_t)$, respectively.

Applying the same assumption on the algorithm with history, the number of leaving users can be calculated with the **P** matrix itself, that is $out1_{t+1}= in1_t^*p_{11}+ in2_t^*p_{21}$ and $out2_{t+1}= in1_t^*p_{12}+ in2_t^*p_{22}$. At a given and constant **P** matrix let us assume that the incoming user distribution varies, that is the $in1_t/in2_t$ ratio (Incoming Distribution – ID) changes.

Figure 7 shows the error of RW compared the estimation using history.

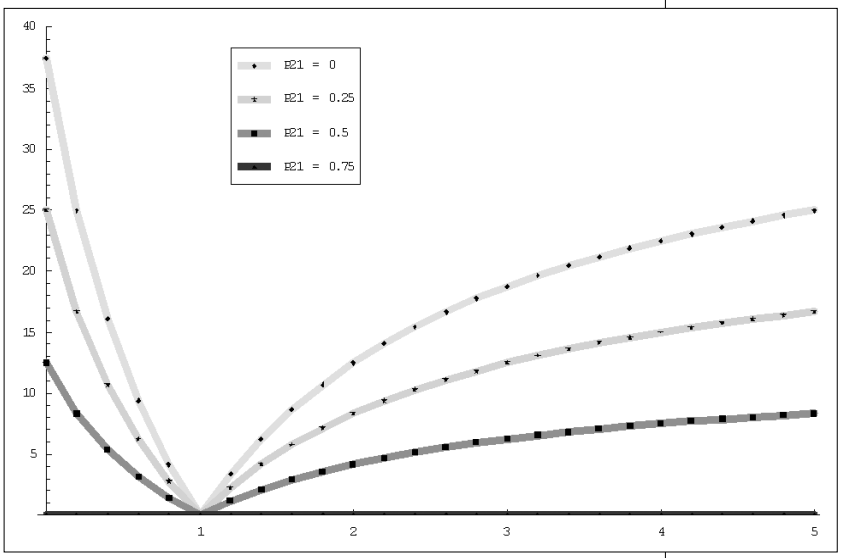


Figure 7. HOV prediction error in percents – Given $in1_{t-1}/in2_{t-1}=1$ and $P=\{\{0.75, 0.25\}.\{p_{21}, 1-p_{21}\}\}$, p_{21} plotted with four different values

The RW model works with error if ID_{t-1} is different from ID_t which is caused by the fact that the RW history **P**-estimation in this special case equals the number of leaving users of the previous timeslot. That is it does not include the actual ID_t value. Contrarily, the history-model calculates with the actual number of incoming users and the **P** matrix itself, which gives the exact probabilities of the leaving users distribution. The error rate caused by the lack of history increases as the variance of ID increases that is the $in1_t/in2_t$ ratio changes.

Using history cannot enhance further the accuracy of the estimation if $p_{21} = p_{11}$ and Fig. 7 shows a constant zero error rate ($p_{11} = p_{21} = 0.75$). In this case the outgoing direction of each user is independent of the incoming direction and the history is useless, since users arriving from each direction are leaving towards a given direction with the same probabilities.

The results show that our proposed movement history significantly increases the accuracy of the model in cases when the ID distribution in an arbitrary cell has high variance, or has periodicities without stationary distribution.

To apply the accuracy with the use of information about the previous user visited cells in the mobility model, we introduce our estimation model discussed in the previous section (M_n, where *n* denotes the number of Markov states). The states of the model show and store where the users came from. The memory can be interpreted in two meanings. Time dimension memory shows the number of timeslots in the past that the model considers. Thus a model with *m* time dimensions in time *t* can calculate the next transition based on the user position in *(t-m, ... t-2, t-1)*. Direction dimensions memory shows the number of directions that the model can differentiate. In general a cell cluster consists of hexagonal cells. The direction dimension of a model on this cluster is maximum 6. If transitions from the central cell to two adjacent cells are not differentiated then the direction dimension is decreased by 1.

If we use more information from user movement history in prediction than our model does, the complexity increases exponentially. Obviously the direction numbers in the model influence the complexity as well. We come to a simpler model if we use only left and right directions instead of usage of all neighbor cells as a state in Markov model as it was shown in the previous sections.

Our work was motivated to find an optimal size of estimation parameter space with the highest accuracy beside tolerable complexity.

3. Simulation and numerical results

The inaccuracy of the RW-based mobility models depends on the properties of the transition probabilities. The RW model is only capable of accurate prediction of user movements in case of uniform movement distributions (e.g. all elements in the transition probability vector are equal to 1/6). In the simulation we applied

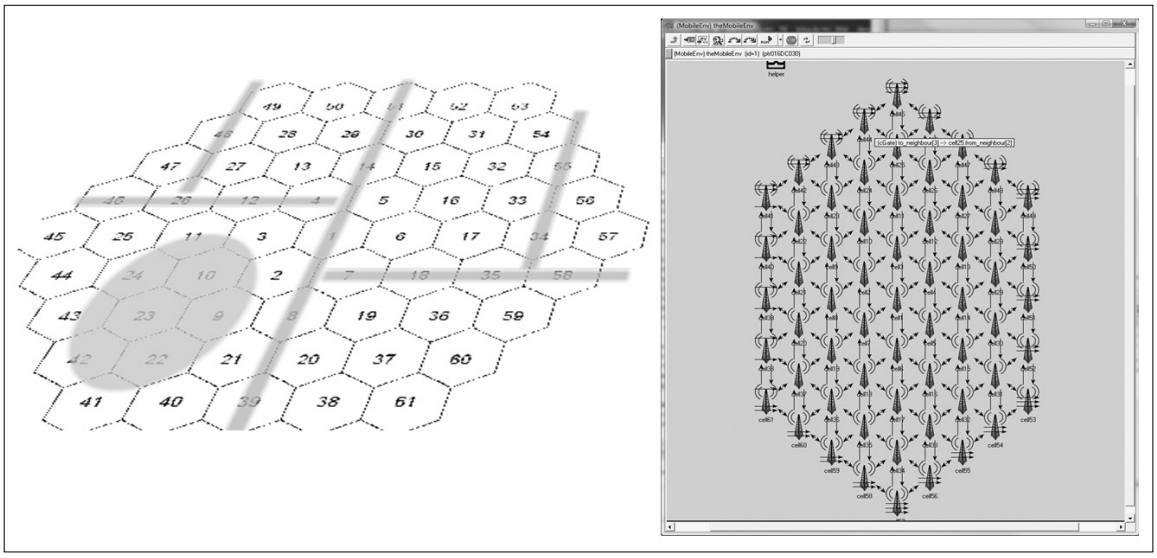
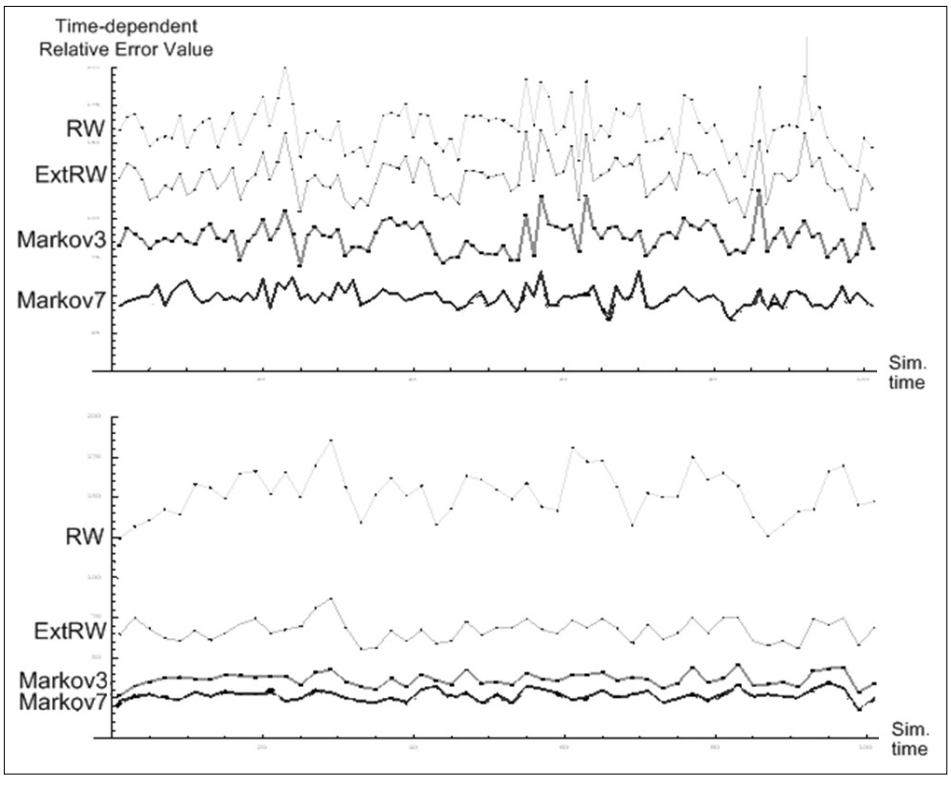


Figure 8.
a. Cell cluster with streets and park
b. OMNET++ simulation

an extended RW model (ExtRW) which was capable of tracking different cell dwell times. ExtRW can model better different user velocities with the correct dwell time parameters than RW which uses a fixed dwell time. With the variable dwell time parameter, the ExtRW simulation model can adapt to different movement velocities, i.e. a user with slow motion spends more time in each cell before he/she initiates handover to one of its neighbors. ExtRW does not cope with different directions of stepping forward. Thus stepping into each neighboring cell has an equal conditional probability. On the condition that at a specified moment a handover is done, the direction is uniformly distributed on the set of neighboring cells.

Figure 9.
TREV values in RW, ExtRW, M_3, M_7 models with λ=(1, 4)



The estimation procedure was validated by a simulation environment of a cell cluster shown in *Figure 8*. The cluster consisted of 61 named cells, the simulation environment included geographical data that are interpreted as streets on the cluster area. The drift of the movement is heading to the streets from neutral areas.

The simulation used 610 mobile terminal (10 for each cell), in the initial state uniformly distributed in the cluster. The average motion velocity of the users is parameterized with a simple PH cell dwell time simulator (reciprocal of exponentially distributed values).

The simulation consists of two parts. The reference simulation is the series of the transitions that the mobiles have initiated between cells. It produces a time-trace that contains the actual location data for each mobile terminal in the network. We have used this reference simulation as if it was a provider's real network trace.

The estimation procedure uses the past and the current reference simulation results to estimate future number of users in each cell. The estimation error is interpreted as the measure of accuracy of each mobility model in this paper.

The prediction starts 100 timeslots after the reference simulation initiation. During the warm-up process the reference simulation produces enough sample data for the correct estimation which uses the previous reference results as an input to estimate the future user distribution. Each user-transition in the 100-timeslot reference period is used to derive transition probabilities, motion speed and

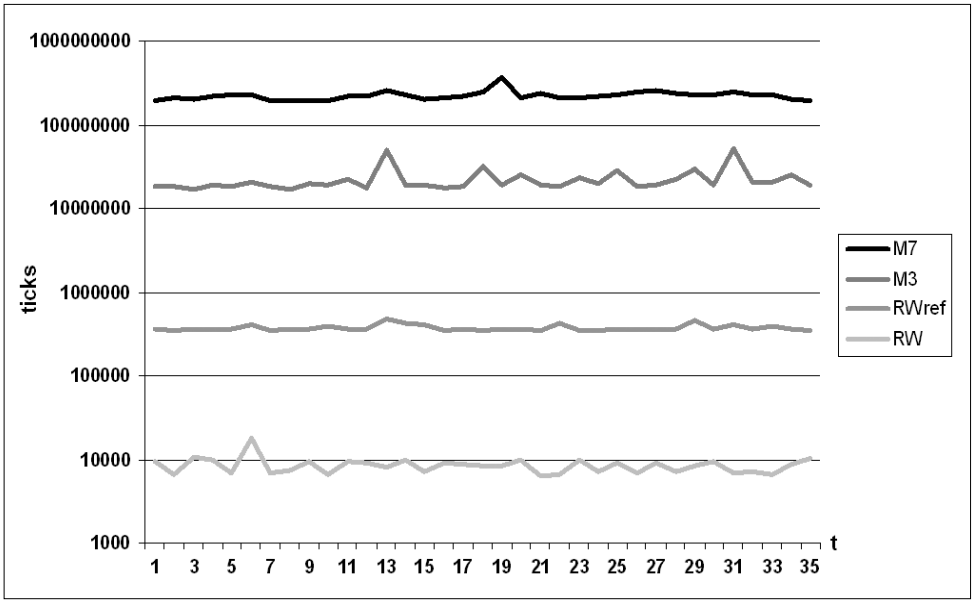


Figure 10.
CPU usage of the models on logarithmical scale

patterns in the simulation cell-space. These patterns serve as an input for the simulation threads of each mobility model. The models have the same input throughout the simulation process so that the results are comparable.

The estimation errors of the models in the simulation were measured with the average error of the cells in each timeslot. It produces a time-dependent relative error value (TREV) in each timeslot for the cell-cluster. TREV shows the average error compared to the actual user number in the cells. It can be seen in *Figure 9* that TREV depends on the dynamics of the motion, basically on the cell dwell time. The generic lambda (λ) parameter affects the motion velocity of the simulated users, the higher value means longer dwell time thus slower motion.

The relative performance of the models can be seen in *Figure 10*. The execution time of the methods of the model is plotted in each timeslot on logarithmic scale.

4. Conclusions

In this paper we proposed an alternative Markov-chain based method. The simulation results proved the analytical properties of the proposed mobility models.

The algorithm with the simple RW model is not capable of precise adaptive location prediction due to its inflexible parameter set. Since the user movement patterns in the simulation are not completely random due to the streets and geographical circumstances, the uniformly distributed Random Walk pattern cannot model it.

The Markov mobility model is the most accurate in the estimation process since it has the ability to calculate with motion direction, speed and the recent handover event (user history) also. The three-state model focuses on cell dwell time since it differentiates only two motion directions which cannot follow general drifts. The n-state model is more sophisticated in terms of both the cell dwell time and motion direction since it is capable of following for example six different drifts in the cell cluster.

The network operator may use a seven-state Markov model to make predictions on the future distribution and location of users among radio cells to justify CAC or other QoS decisions.

Authors

TAMÁS SZÁLKA was born in Budapest in 1982. He received the M.Sc. degree in Computer Science from Budapest University of Technology and Economics (BUTE) in 2006 then he started his Ph.D. studies at BUTE, on the Department of Telecommunications. Beside his Ph.D. studies, he is doing engeneering research as Ph.D. student and fellow researcher in the Mobile Innovation Centre of BUTE. As planned, he will obtain his Ph.D. degree in 2011. His work is relatad to the area of mobililty modeling and location based services of cellular networks.

PÉTER FÜLÖP was born in Mosonmagyaróvár in 1980. He received his M.Sc. degree in Technical Informatics from the Budapest University of Technology and Economics in 2005. Currently he is a PhD student at the Department of Telecommunication and fellow reserarcher in the Mobile Innovation Centre of BUTE. He is working as a system test engineer at Research and Development Department of Ericsson Telecommunications Hungary. His research interest is IP mobility, interworking of heterogeneous mobile networks and movement modeling in cellular, mobile networks. The PhD thesis is going to be written on Complex Mobility Management Systems and Applications.

SÁNDOR SZABÓ was born in 1977. He received the M.Sc. degree in Electronic Engineering from the Budapest University of Technology (BUTE) in 2000 then he started his Ph. D. studies at BUTE. Currently he is Assistant Professor at the Dept. of Telecommunications of BUTE. He leads research projects, and also a project leader at the Mobile Innovation Centre of BUTE. His research interests include mobile and wireless systems, mobility modeling, mobility management and the integration of heterogeneous wired and wireless networks.

References

[1] A. Jardosh, E.M. Belding-Royer, K.C. Almeroth, S. Suri, "Towards Realistic Mobility Models for Mobile Ad Hoc Networks", In Proceedings of ACM MobiCom, pp.217–229, September 2003.

[2] Jungkeun Yoon, Brian D. Noble, Mingyan Liu, Minkyong Kim,
"Building realistic mobility model from coarse-grained wireless trace",
In Proceedings of the 4th International Conference on Mobile Systems, Applications and Services (MobiSys), Uppsala, Sweden, June 2006.
[3] S. Michaelis, C. Wietfeld,
"Comparison of User Mobility Pattern Prediction Algorithms to increase Handover Trigger Accuracy",
IEEE Vehicular Technology Conference,
Melbourne, May 2006.
[4] M.M. Zonoozi, P. Dassanayake,
"User mobility modelling and characterization of mobility patterns",
IEEE Journal on Selected Areas in Communications, Vol. 15, No. 7, September 1997.
[5] B. Liang, Z. Haas,
Predictive distance-based mobility management for PCS networks,
In Proceedings of the Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), March 1999.
[6] V. Davies,
Evaluating mobility models within an ad hoc network,
Master's thesis, Colorado School of Mines, 2000.
[7] X. Hong, T. Kwon, M. Gerla, D. Gu, G. Pei,
"A Mobility Framework for Ad Hoc Wireless Networks",
Proceedings of ACM 2nd International Conference on Mobile Data Management (MDM 2001),
Hong Kong, January 2001.
[8] Vincent W.-S. Wong, Victor C.M. Leung,
"Location Management for Next-Generation Personal Communication Network",
IEEE Network, September/October 2000.
[9] Levine et al,
"A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks Using the Shadow Cluster Concept",
IEEE/ACM Trans. on Networking, Februar 1997.
[10] R. Chellappa, A. Jennings, N. Shenoy,
"A Comparative Study of Mobility Prediction in Cellular and Ad Hoc Wireless Networks",
Proceedings of the IEEE International Conference on Communications 2003 (ICC2003),
Alaska, USA, May 2003.

Renewal of Sister Society Agreement between HTE and IEEE ComSoc

ROLLAND VIDA

Head of International Affairs, HTE
vida@tmit.bme.hu

The Scientific Association for Infocommunications (HTE) is currently renewing the Sister Society Agreement with the IEEE Communications Society (IEEE ComSoc). The terms of the agreement were settled already, and the agreement is to be signed in the following weeks. But what does it really mean to be a Sister Society of IEEE ComSoc?

The Communications Society, established in 1952, is one of the largest societies inside IEEE. Its main task is to bring together industry professionals and academics with a common interest in advancing all communication technologies, and enable them to interact across international and technological borders, in order to produce publications, organize conferences, foster educational programs, and work on technical committees.

IEEE ComSoc has 24 Sister Societies, in countries such as Japan, Brazil, South Africa, Germany or France, to name just a few. HTE is the only Sister Society from Hungary, with the agreement being now renewed for the period 2009-2012. The main goal of these bi-lateral agreements is to foster participation at scientific conferences, or publication in technical journals, with equal terms for the members of the signing parties. To be more specific, members of HTE will be able for example to register for ComSoc sponsored meetings and conferences (e.g., Globecom, ICC, WCNC) at IEEE member rates. Also, members of HTE can subscribe to ComSoc publications (e.g., IEEE Transactions on Communications or IEEE JSAC) at reduced Sister Society rates, they can participate in the ComSoc Technical Committees (TC), being eligible to vote and to be elected as TC officers. IEEE ComSoc may also provide technical co-sponsorship for established HTE-sponsored meetings and conferences, at no additional fee.

And last but not least, the most important point as far as our journal is concerned is the agreement that IEEE ComSoc will review for publication yearly the best one or two articles previously published in the Infocommunications Journal edited by HTE. These articles will then be published in the appropriate ComSoc journal or magazine. As all the other points from the agreement, this one is also reciprocal, meaning that the Infocommunications Journal will also include yearly one or two papers previously published in an IEEE ComSoc journal or magazine.

We hope that this agreement will represent a further incentive for the Hungarian but also for other researchers to join HTE, and to publish their work in our journal, helping us in achieving our goal to continuously improve the scientific level of our publications.

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest V., Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu
Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 0018-2028 • Layout: MATT DTP Bt. • Printed by: Regiszter Kft.