

# SAFETY AND SECURITY INCREASE FOR AIR TRAFFIC MANAGEMENT THROUGH UNNOTICEABLE WATERMARK AIRCRAFT IDENTIFICATION TAG TRANSMITTED WITH THE VHF VOICE COMMUNICATION

*Horst Hering, Eurocontrol Research Centre, Bretigny, France*

*Martin Hagmüller, Gernot Kubin, Graz University of Technology, Graz, Austria*

## 1 Abstract

Controller-pilot Very High Frequency (VHF) voice communication in Air Traffic Control (ATC) relies on amplitude modulation by a carrier frequency of the transmitted analogue voice pattern. This technology is known for its poor voice quality and to be highly sensitive for any kind of noise on the transmission/ reception path. On this party-line pilots have to identify themselves with their call-sign. Human's imperfections in speaking and understanding added to the low channel quality may cause problems to identify the originator of an Aircraft (AC) message without any ambiguity for the controller.

The proposed Aircraft Identification Tag (AIT) technique aims at overcoming this security lag. The automatic insertion of the supplementary AIT in the transmitted voice communication will secure and simplify the association of a voice message to the specific AC. Supplementary safety increase may be achieved by the mental reinforcement of the audible stimulus (pilot's voice) with a simultaneous visual stimulus in form of e.g. highlighted AC track on the Radar screen. Further security increase may be attended from the fact that fake AC VHF transmissions (e.g. jokers, terrorists, ...) becomes more difficult or even may be impossible by special watermark coding. Simple airfield structures and oceanic High Frequency (HF) communication could benefit from such a system, too.

AIT technique, therefore, uses proven audio watermark techniques (e.g. used for intellectual property rights for audio CD and digital video). These watermarks are introduced in the voice signal such that, the watermarks are not noticeable. AC and ATC VHF-transceiver need no change. The technique is adapted to the reduced VHF channel spacing of 8.33kHz. The prototype demonstrator shows very high robustness with a watermark

datalink capacity of about 100 bits per second including security data to ensure payload data integrity.

## 2 Introduction

Tactical ATC guidance over continental areas is generally performed through controller-pilot voice communication over VHF. On the VHF working frequency of a controlled sector the responsible controller communicates with all pilots flying in the sector on a timeshared manner, called party line communication. The controller pilot VHF communication is based on amplitude modulation of a carrier frequency by the transmitted analogue voice pattern. This technology is known for its poor voice quality and to be highly sensitive for any kind of disturbance on the transceiver path.

On this party line, the pilots have to start every message with their identification (aircraft call-sign e.g. Lufthansa 3501, spoken: 'Lufthansa tree five zero one'). Pilot messages without this identification (inaudible or missing) may be obsolete or may make sense only with the contextual knowledge of the controller. *Fait accompli* is, this increases controller's mental workload.

There may be many reasons why the call-sign of an aircraft message creates difficulties for the controller. The most common cases are:

- Disturbance or low quality of the transmission from the aircraft;
- Call-sign is shorted at the beginning by pilot's delay in pressing the 'push to talk' switch to activate the AC transmitter;
- Pilot uses a shortened call-sign (e.g. 'Lufthansa' only or '3501' only instead of 'Lufthansa 3501');

- Controller being occupied and omits the beginning of the message;
- Error of speaking or understanding of the call-sign;
- Very similar call-signs for different AC in the sector at the same time (e.g. Airfrance 4740 or Airfrance 4704);
- The pilot omits his AC identifier;
- Short aircraft transmission without any voice modulation (mainly interpreted by the controller as acknowledgment for his last message).

In all these cases the introduction of an electronic AIT could eliminate such identification ambiguity for the controller. Therefore the AIT uses the voice frequency spectrum and is triggered for automatic insertion by the pilot's 'push-to-talk' (PTT) switch. The tag is not noticeable. In the ATC ground system the introduced AIT is available for decoding from the speech frequency spectrum with a short delay (in the prototype demonstrator: ~900ms) after the pilot starts speaking.

### 3 Technical Background

From the early beginning the 'Direction Finder' (DF) existed in the ATC to help controllers to identify an AC. The DF takes the bearing of an aircraft during its VHF communication relative to the receiver position. This information is then displayed on a special pointing device with a compass-rose (mainly used in a tower environment: controllers working place and DF origin should be identical) or introduced as a beam on the radar screen while the AC pilot is transmitting. The displayed ray represents the DF heading and will go through the speaking AC symbol. So the DF supports controller's mental relation between audible pilot speech and AC's radar position. The DF information is often strongly jittering and is not selective for other aircrafts seen in the same direction of the DF (several AC may be crossed by the DF-ray). Over the years, the DF lost its importance for approach and tower with the upcoming aircraft identification (call-sign indication with track label) on the raster scan Radar screen through Secondary Surveillance Radar (SSR).

Mentally, a controller has to make the correlation between the voice (pilot/AC) coming

out of the speaker and the Radar track displayed on a screen. If the controller is not able to make the correlation or makes a wrong one, serious incidents/accidents may occur.

The basic idea for this project was to give ATC radio communication a similar feature like in our days most Frequency Modulation (FM) broadcast receivers have, showing the station name on a small display. Therefore the Radio Data System (RDS) for broadcast radios was developed. The RDS system is able to downstream an important quantity of data [1] to the receiver via unused sidebands.

But besides being in the same VHF band, both means of radio communication (ATC voice communication and FM broadcast) have nothing in common. So it's obvious that an 'ATC data system' may not taking any benefit from RDS findings. The RDS technique is based on technical modifications in their transmitters and receivers. A modification of the aircraft transmitter and receiver would be time consuming and extremely expensive in view of the required certification. So it is evident that a solution has to be sought in the speech frequency band.

Three different investigation were conducted:

- Using the telephone Dual Tone Multi Frequency (DTMF) technique which is used for dialling in modern telephone systems. The DTMF sends simultaneously a tone pair out of 8 defined tones (697 – 1633Hz) to represent 4 bits of data (similar systems use 3 or 4 simultaneous tones to increase the data rate). Sending these 4 bits takes about 180 ms (90 ms tones and 90 ms pause). With special filters it would be possible to make the signal nearly inaudible but it would cause some strange 'beeps' in the unequipped receivers.
- In ATC it is known that the transmitted human voice has less spectral energy around 2 kHz. This circumstance is used in some countries for inline commands for the ATC ground transmitters. Therefore from the voice frequency spectrum a bandwidth of about 300 Hz is cut out, and replaced by coded

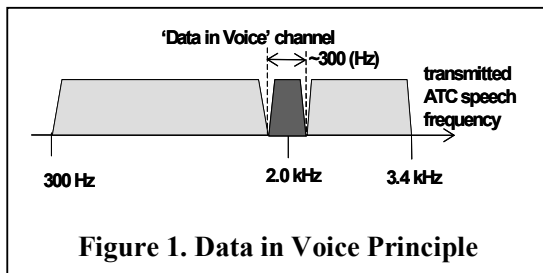


Figure 1. Data in Voice Principle

instructions for the remote ATC transmitters. Prior to the transmission the coded instructions are removed from the transmitted voice spectra. Reducing the voice spectrum around 2 kHz is operational since decades in some countries. Recently an important player in the international ATC communication promoted the idea to place for a short time digital data coded by a modem [2] in this bandwidth gap. A prototype was developed from this firm and presented at the 2002 ATCA exhibition at Washington. The proposed 'Data in Voice' (DiV) (see figure 1) implementation has a capacity of about 200 bits per second. The only disadvantage of this proposal is that every AC (not only those with the DiV) has to be equipped with a filter to remove the DiV path, otherwise it may be heard on the party line and perturb understanding.

- Another solution, finally adopted in this project, is the so called 'watermark' technology. Watermarks for papers are nearly as old as the paper. In those days the watermark certified the authenticity of the document in a normally invisible manner. Nowadays the watermark idea (imperceptible and inerasable marks on an object) is used for copyright and Intellectual Property Right (IPR) protection for different types of digital media (pictures, films, audio: analogue /digital). Figure 2 shows the principle of the digital watermark techniques used for high quality audio [3].

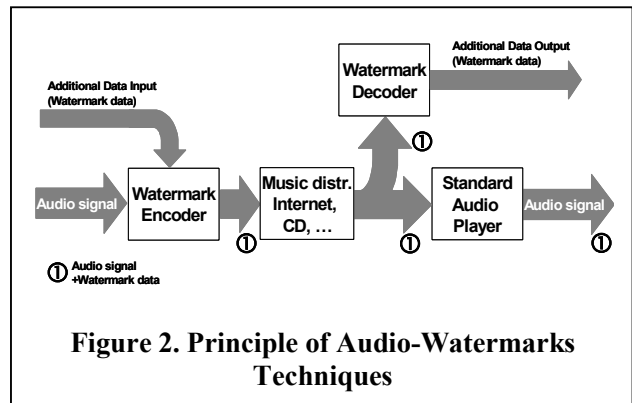


Figure 2. Principle of Audio-Watermarks Techniques

#### 4 AIT Technology

For the AIT project the digital watermark technique, was adapted to the special needs of low quality ATC communication over VHF. Watermarks for these speech signals have to be different than the usual audio watermarks due to the much narrower signal bandwidth. Compared to the 44.1kHz sampling rate for CD-audio, telephony speech is usually sampled at 8kHz. Therefore, less information can be embedded in the signal. For perceptual hiding usually the masking levels have to be calculated. The common algorithms used are optimised for CD-audio bandwidth and are computationally very expensive. Cheng et al. [4] uses 'Linear Prediction Coefficients' (LPC) for the calculation of the masking levels. Another difference is the expected channel noise. For CD-audio the channel noise is usually rather low. On one hand the audio signal loses its commercial value if the channel noise goes beyond a certain threshold. On the other hand, the distortion allowed for a watermark signal is very low as well. Both artist and consumer, would not trade in a reduced recording quality for better copyright protection.

ATC speech is generally transmitted over noisy channels. The channel noise is a disadvantage, but allows much more power for the watermark signal, since the channel noise will cover it anyway. The listener expects a certain amount of noise in the signal. A summary of the difference requirements for watermarking of CD's and ATC speech can be seen in Table 1.

Figure 3 shows the AIT functional block diagram. The grey boxes represent AIT functionality and have to be added to the existing aircraft equipment (microphone, PTT-switch,

**Table 1. CD-Audio vs. ATC-Speech Watermarks**

	CD-Audio Watermarking	ATC VHF Speech Watermarking
channel noise	very low	can be high
band-width	wideband (20kHz)	narrowband (<4kHz)
allowed distortion	not perceptible	low

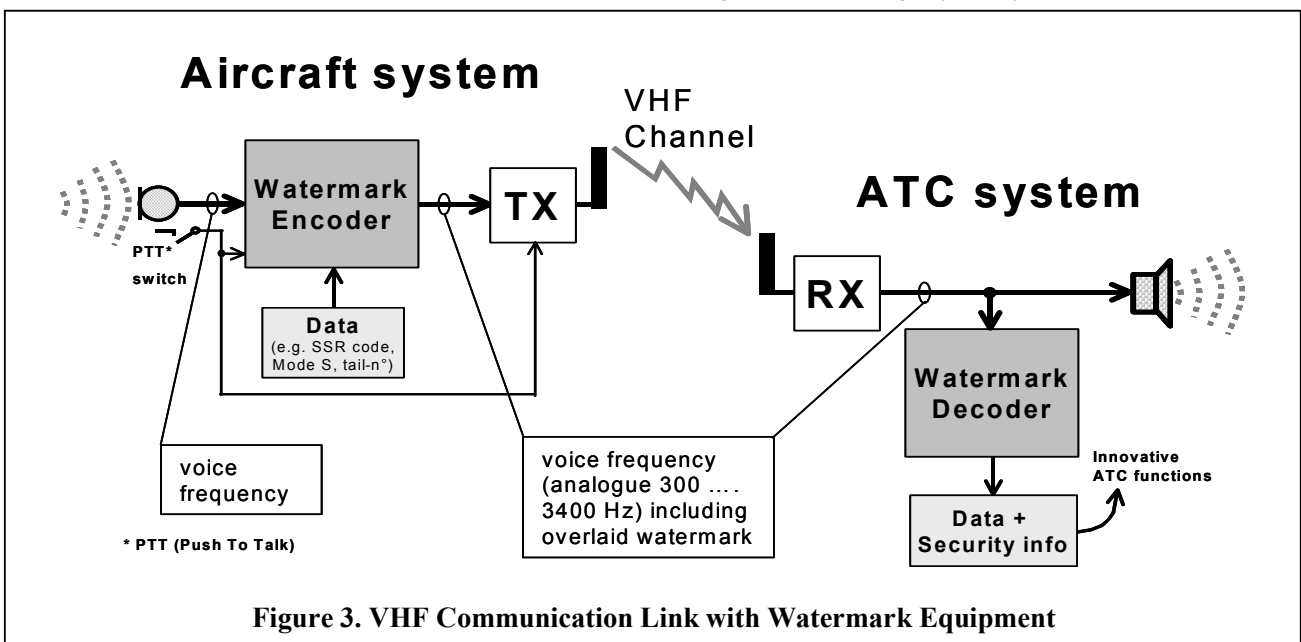
transmitter, VHF antenna) and ground system (VHF antenna, receiver, speaker).

To embed the data into the pilot's speech, the watermark encoder digitalises the voice from the microphone and adds the digital watermark. The resulting data (voice + watermark) is converted into an analogue signal again and fed as output of the watermark encoder into the AC transmitter (TX). No modification is required to the existing certified AC equipment. In the ground receiver (RX) a demodulator extracts the analogue audio from VHF for the controller. The included hidden watermark is not noticeable, as it is below the channel noise level. This audio signal may be fed in a watermark decoder for digitisation and extraction of the watermark data. New innovative ATC functions (e.g. highlighting of the speaking AC on the Radar) may now use the watermark information.

The estimated watermark data capacity for a VHF communication channel is about 100 to 200 bits per second of raw data. The raw watermark data consist of information bits (payload), synchronisation bits and data security bits. As payload data for the AIT system the SSR code (12 bits), the 'Mode S' identifier (24 bits) or the AC tail-number (36 bits) for general avionics are proposed. The use of 'Mode S' and AC tail number would be very simple, as they are mainly constant for the lifetime of an AC, they could be 'burned' in the AIT encoder. A usage of the short 12 bit SSR code would favour watermark robustness (more place for security bits). Technically, a wireless capture of the own SSR code would be possible, inside of the AC from the replies for the ground interrogation. All the proposed payload codes are well known by the ATC ground authority and therefore could easily associated with the specific flight plan of the speaking aircraft.

AIT's aim is to identify automatically the originator of an AC voice communication without any pilot's intervention, in real-time, in a safe and secure way. No modifications of the certified AC equipment.

In Figure 4, the principal timing relation of voice and watermark signals for the transmitter (TX) related to the PTT signal is given. For technical transmission reasons, the watermark signal will be slightly delayed (~100 ms) from the



**Figure 3. VHF Communication Link with Watermark Equipment**

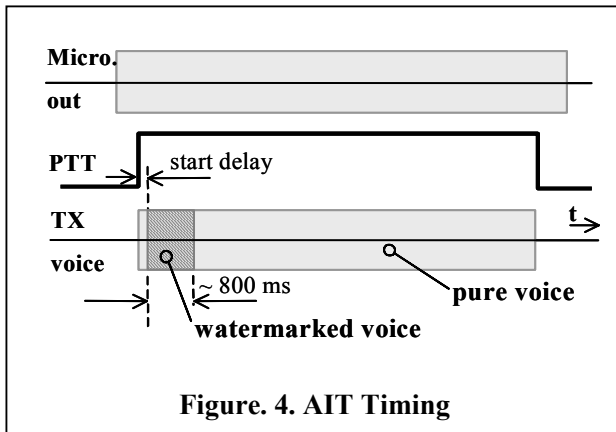


Figure 4. AIT Timing

PTT signal start. The watermark payload data will be available in the ATC ground station in less than one second after the start of receiving pilot's voice. After code correlation with other ATM data, new innovative functions (e.g. 'highlighting' of the speaking AC) could be executed to support the acting controller.

The availability of the payload data in the ground system within less than one second was chosen in a deliberate manner based on an ERC study of controller VHF voice communication for pilots [5]. Table 2 shows results for the mean duration of the controller communication, the duration of the utterance, in case the call-sign was omitted and if the call-sign was the only message. The study showed as well a calculated speaking speed of about 12 printable characters per second (e.g. 'Lufthansa tree five zero one' = 28 character).

Hypothetically, the pilot's behaviour will be similar. So, the AIT payload data will be available for the ATC system before the pilot terminated the pronunciation of its own call-sign.

Table 2, Duration of spoken Controller Utterances

	general	utt. without AC call-sign	utt. AC call-sign only
duration (s)	5.0	2.7	1.9

## 5 Developed Watermark Technology

### 5.1 Watermark Encoder

After study of the literature, an approach has been chosen which combines the spread spectrum approach with a simplified frequency masking. Spread spectrum was chosen, since it is well studied and appears rather robust against channel noise. The block diagram for the watermark encoder can be seen in Figure 5. The system can be split into three main parts. First the error control coding, which performs channel coding to get increased reliability of the results. Next is the spreading of the watermark signal, over the available frequency band. Finally the watermark is embedded into the speech signal using perceptual considerations.

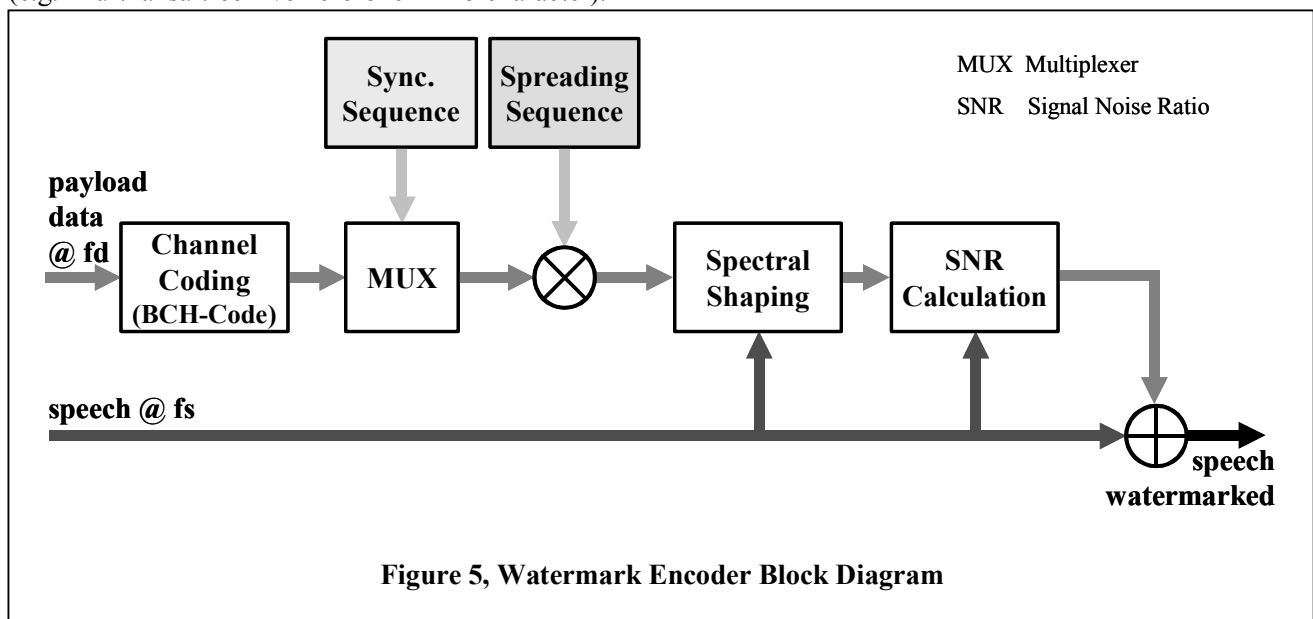


Figure 5, Watermark Encoder Block Diagram

### 5.1.1 Error Control Coding

For error control coding a BCH-code was used. BCH-codes are cyclic linear block codes and allow a large selection of block length, code rates, alphabet sizes and error correction capability [6]. The BCH decoder provides information about the number of bit errors in a received block. If this number is within the correction capabilities of the code, it will be corrected. If the detected errors are more than the correction capability the code offers, only the information about the number of errors, without a error correction is given. In this case the received data cannot be trusted and has to be discarded. In table 3 the number of bits in the code, parity bits and the correction capability for the BCH-code is presented for three different payload data sizes.

Data-bits k	Code-bits n	Parity-bits (n-k)	Correction capability t
12 bits = 2*6 bits	62 bits = 2*31 bits	50 bits = 2*25 bits	14 bits = 2*7 bits
24 bits	63 bits	39 bits	7 bits
36 bits	63 bits	27 bits	5 bits

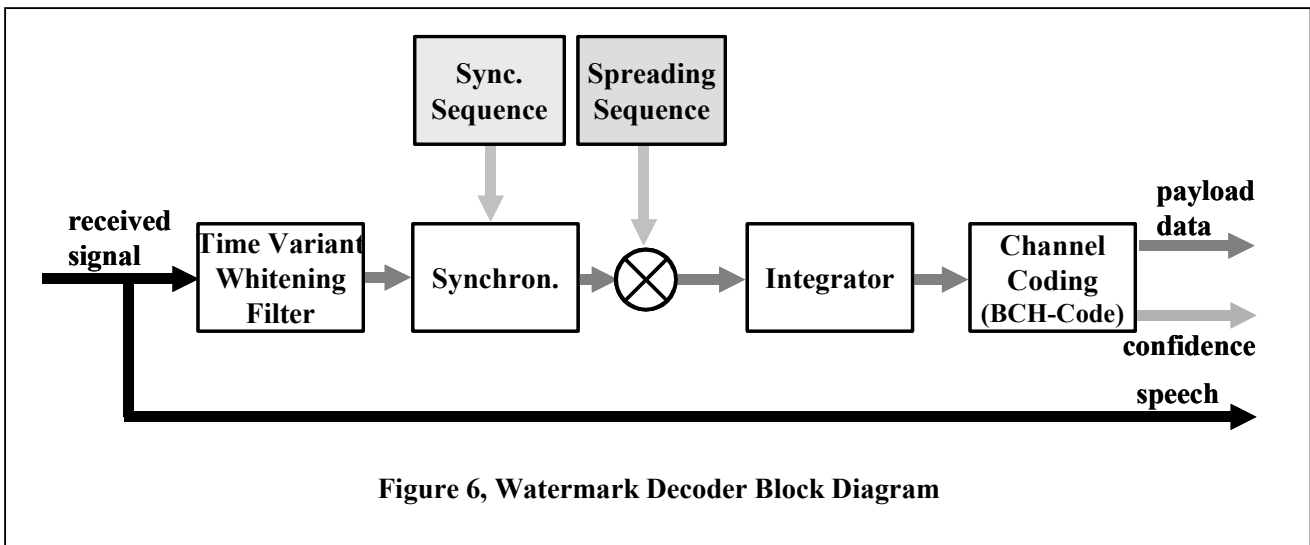
### 5.1.2 Data Spreading

The chosen approach uses spread spectrum technology to spread the data over the available bandwidth. The signal is spread by modulation with a white pseudo-noise binary sequence. The length of the sequence equals the number of samples used for one data word. In case of 8 kHz sampling rate

and a symbol rate of 100 bit/s and a total data word size of 80 bits (including 62 (63) code bits plus extra synchronization bits) this means a spreading sequence length of 6400 samples. In [7] it is claimed that the spectrum of the watermark signal should be close to the host signal spectrum to achieve robustness for synchronization. Therefore the spreading sequence is low-pass filtered to match the voice channel bandwidth.

### 5.1.3 Watermark Embedding

The spread spectrum watermark signal should next be embedded into the speech signal. Simply adding the watermark signal would result in a high interference from the host signal depending on the widely changing speech power. This distortion can be better controlled by using masking techniques both in terms of temporal energy and spectral shape. The goal of the data embedding is to put the watermark in the speech signal with maximum energy but the least possible perceptual distortion. Many watermarking algorithms use psychoacoustics knowledge about the human auditory system to minimize the distortion caused by the watermark [8, 9, 10, 11]. The spectral masking was done using linear prediction for spectral analysis of the speech signal and then shaping the watermark signal with those parameters, after so called bandwidth expansion [12]. A minimal watermark level was maintained to ensure a reliable transmission when no voice signal is present.



## 5.2 Watermark Decoder

At the decoder side the watermark data has to be extracted from the speech signal (Figure 6). First a whitening filter is employed to undo the frequency masking spectral shaping. The signal has to be synchronized to perform the watermark extraction. The error correction algorithm uses the added redundancy to correct errors made on the channel.

### 5.2.1 Whitening

At the receiver the incoming signal is first of all whitened to undo the spectral shaping process of the transmitter side. Analysis of the signal is again done using linear prediction analysis. This time the inverse of the linear prediction-filter is used to calculate the error signal. The bandwidth expansion is performed for the whitening filter, too.

### 5.2.2 Synchronisation

At the receiver side, a major issue is the synchronization of the watermarking signal. Since it is not known when the watermark signal starts it is crucial that the watermark synchronisation signal is detected in the received voice. Therefore in the encoder a special synchronisation sequence is added to the payload data. This known sequence has to be detected in the decoder by a matched filter. If this fails no valid data can be decoded.

### 5.2.3 De-spreading

After synchronisation, the exact location of the watermark signal is known, it can be de-spreaded, by multiplying the signal with the sequence used for spreading at the encoder. This works only when the watermark signal and the spreading sequence are exactly synchronized, multiplication with only one sample difference does not de-spread the signal.

### 5.2.4 Decoding and Error Correction

Decoding is done using a simple integrator. Since the decoder knows the positions and the length of the data bits, integrating over the period of one data bit gives the result. At this stage the down sampling from the signal rate to the binary symbol rate is performed as well. Based on these results error correction is done as far as needed and can be performed with the BCH decoder.

## 5.3 Performance Analysis and Conclusion

To evaluate the performance of the proposed watermarking algorithm a series of simulations were carried out. For the evaluation of the algorithm, different error rate results, characterising different levels of system performance and the perceptual quality were evaluated:

- Watermark channel Bit Error Rate (BER), before doing error correction
- Message Error Rate (MER): concerning whole message blocks (e.g. 12, 24, 36 bits).
- MER of confident results: MER taking only confident results (this is where the BCH-decoder decides that the number of bit errors is within the correction capability of the code).
- Occurrence of confident results: The percentage of occurrence of confident results is also evaluated. This is important, since even error-free confident results, which only occur for 10% of the transmitted data-blocks are not usable in practice.
- Perceptual quality: This was not evaluated systematically, but informally only. Extensive Mean Opinion Score (MOS) tests would have been necessary to substantiate the perceptual results. Such a study should be carried out with the intended user group.

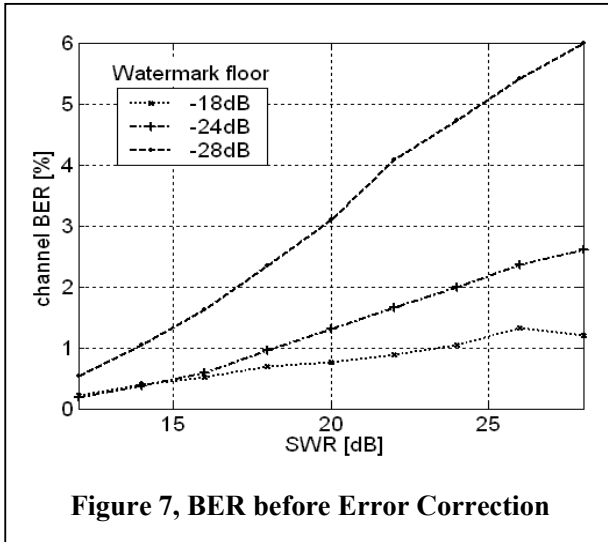
The following parameters, which are of interest for an implemented system were varied for the performance tests (1000 transmitted messages):

- Data rate: The watermark channel data rate (80, 100 bits/second).
- Message size: The AIT payload size (12, 24, 36 bits).
- Watermark floor: The effective minimum watermark energy in the signal (-16 to -32dB).
- Signal to Watermark Ratio (SWR): The ratio by which the watermark is attenuated relative to the speech signal (12dB to 28dB).

The results for the performance analysis, presented in this paper, are based on a channel data rate of 80 bits/second with 12 bits payload data. Results for other payload sizes (24, 36 bits) are

available. The tests have been performed as well with a channel data rate of 100 bits/second.

The channel BER is clearly dependent on both the watermark floor and the SWR. The error rate was under 10 % (see Figure. 7).



**Figure 7, BER before Error Correction**

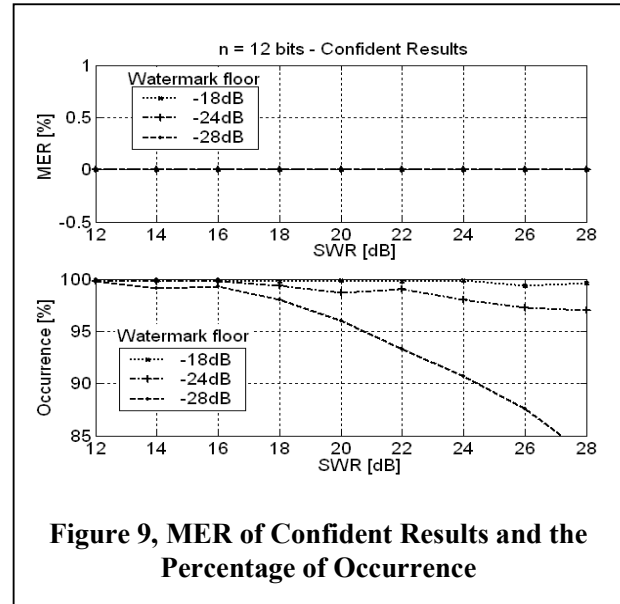
Extensive performance tests showed that the task is feasible within the given specification for the system. For the transmission of 12 payload bits in less than 1 second error rates were very low (< 0.1%) at a distortion level which is lower than the channel noise. Therefore, the desired AIT could be implemented as a simple add-on to the existing voice communication system without interfering with speech transmission quality. Figures 8 and 9 show the MER for the received blocks and their confidence values. In particular, for a watermark floor at or above -28dB, the receiver was always confident in at least 80% of the received messages. For those messages, not a single message error occurred in 1000 trials.

#### 5.4 Outlook

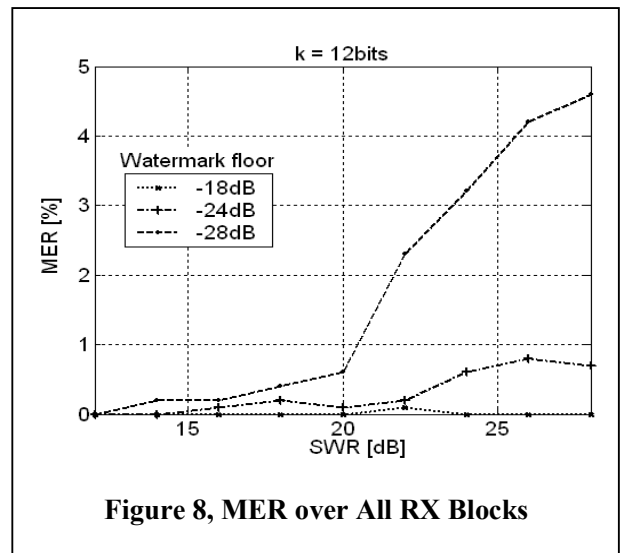
Further detailed information on the proposed system is available in the study report [14] With some more research effort higher data rates with lower error probability and lower perceptual distortion are clearly within reach, given the rapid development in the field. Additional research effort can be put into further improvements of the watermarking algorithm. Careful channel equalization at the receiver will also improve the data rates, since inter-symbol interference can be

reduced. This was not done in this research, since no dynamic aeronautic channel model was available.

Since watermarking applications usually cover security aspects (copyright enforcement), there is also a lot of potential of making data hiding algorithms resistant to misuse by a third party.



**Figure 9, MER of Confident Results and the Percentage of Occurrence**



**Figure 8, MER over All RX Blocks**

## 6 Conclusion

The prototype demonstrator based on a Digital Signal Processor (TMS320VC33) from Texas Instruments shows very high performance and robustness. Simulated difficult VHF transport layers

conditions (additional noise, fading, overlaying tone) have few influence on the results. Due to existing VHF channel noise the watermark is not noticeable by humans. The presented feasibility study is based on state of the art knowledge – further adapted research and optimisation will increase performance of such an AIT system significantly.

The AIT idea is based on standard mass market techniques. The AIT concept works automatically without any pilot intervention and no modification for existing certified AC equipment is required. The watermarked speech is not noticeable on the ATC party-line communication.

The future ATC communication strategy, for the upper airspace, is in favour of the digital Controller/ Pilot Data Link Communication (CPDLC) for dense traffic areas. This may reduce/limit in time the benefit from an AIT system implementation. But CPDLC will still include analogue VHF voice communication channels for special purpose (e.g. emergencies, ...). For continental low traffic areas VHF voice communication will probably persist for long time and therewith the benefit from AIT. Similar estimations are valid for approach and airport tower control.

With the results obtained so far we assume that:

- The safety lag of omitted, misspelled, misunderstood and so on AC call-signs could be eliminated for the ATC controller.
- The AIT system add supplementary redundancy to the audio channel and therewith increase the controller's situation awareness by facilitating the mental relation between audio VHF communication and visual AC position (e.g. simultaneous AC speech and highlighting of related AC track on the Radar display).
- AIT support the controller in the mental correlation between the heard spoken call-sign and this one displayed on the Radar (e.g. spoken: 'Airkibis one tree four'; Radar display 'KYV134'). This

will be from special interest for seldom companies.

- Small airfield structures (without Radar) may benefit from AIT for a long time as there the VHF voice communication will persist.
- AIT could also be used for HF voice communication of oceanic areas.
- AIT gives the possibility to address the security aspect of conventional pilot controller voice communication (fake communication of jokers, terrorists,...).

## 6.1 References

- [1] D. Kopitz/B. Marks. *RDS: The Radio Data System*, Artech House Publisher, UK, 1999.
- [2] M. Sajatovic. *White Paper TCAS over DiV*, Frequentis, Vienna, 2002.
- [3] Ch. Neubauer et al. A Compatible Family of Bitstream Watermarking Schemes for MPEG-Audio, *110<sup>th</sup> Convention Audio Engineering Society*, Amsterdam, The Netherlands, 2001.
- [4] Qiang Cheng and Jeffrey Sorensen. Spread spectrum signalling for speech watermarking. *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1337-1340, Salt Lake City, UT, USA, May 2001.
- [5] H. Hering. *Technical Analysis of ATC Controller to Pilot Voice Communication with Regard to Automatic Speech Recognition Systems*, Eurocontrol Research Centre, Bretigny, France, 2001.
- [6] Bernhard Sklar. *Digital Communications*. Prentice Hall PTR, Upper Saddle River, New Jersey, USA, 2001.
- [7] M. Litao Gang Akansu, A.N. Ramkumar. Security and synchronization in watermark sequence. *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 3736-3739, Orlando, Florida, May 2002.

[8] Ricardo A. Garcia. Digital watermarking of audio signals using a psychoacoustics auditory model and spread spectrum theory. *In Preprints of AES 107th Convention*, New York, USA, September 1999.

[9] N. Cvejic, A. Keskinarkaus, and T. Seppanen. Audio watermarking using m-sequences and temporal masking. *In IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 227-230, New Platz, NY, USA, October 2001.

[10] Petar Horvatic, Jian Zhao, and Niels J. Thorwirth. Robust audio watermarking: based on secure spread spectrum and auditory perception model, *In Proceedings of the Sixteenth Annual Working Conference on Information Security for Global Information Infrastructures*. pp. 181-190. Beijing, China, Kluwer Academic Publishers, Norwell, MA, USA, August 2000.

[11] D. Kirovski and H. Malvar. Spread-spectrum audio watermarking: requirements, applications, and limitations. *In Proc. IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 219-224, Cannes, France, October 2001.

[12] P. Vary, U. Heute and W. Hess. *Digitale Sprachverarbeitung*. B.G. Teubner Stuttgart, 1998.

[13] M. Hagmüller and G. Kubin. *Speech Watermarking for Air Traffic Control*. Technical Report TUG-SPSC-2003-02, Graz University of Technology, 2003.

[14] Stefan Derhaschnig. *Echtzeitsimulation eines VHF-Flugfunkekanals auf einem digitalen Signalprozessor TMS320C67*, Frequentis/ University of Technology Graz, Graz, Austria, 2001.