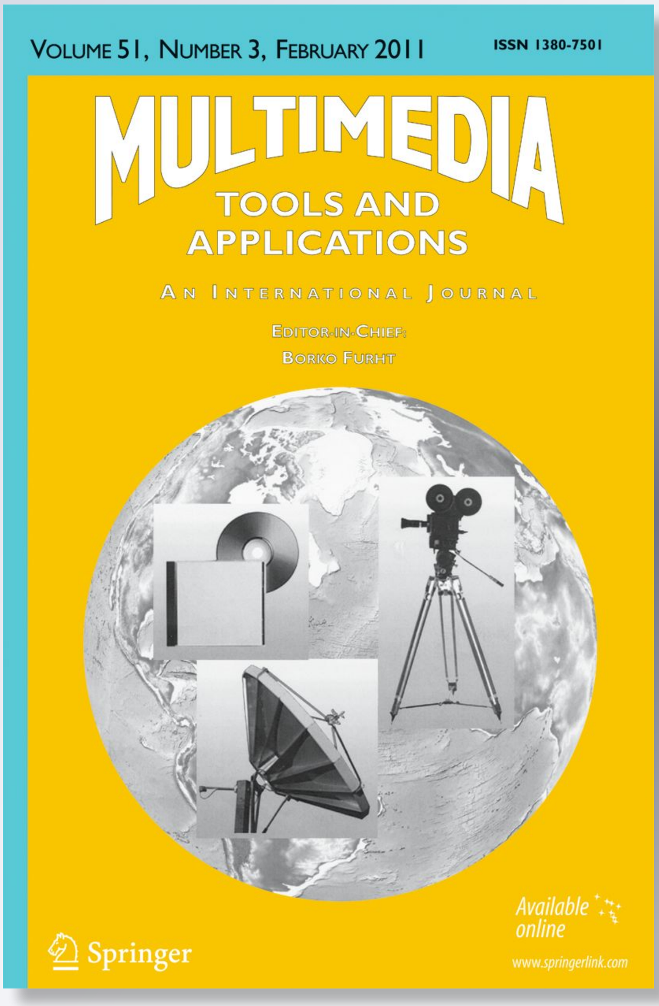
Quality-driven secure audio transmissions in wireless multimedia sensor networks

Honggang Wang, Wei Wang, Min Chen & Xingmiao Yao

Multimedia Tools and Applications
An International Journal

ISSN 1380-7501

Multimed Tools Appl
DOI 10.1007/s11042-011-0928-5



Multimed Tools Appl
DOI 10.1007/s11042-011-0928-5

Quality-driven secure audio transmissions in wireless multimedia sensor networks

**Honggang Wang · Wei Wang ·
Min Chen · Xingmiao Yao**

Abstract Many audio applications such as audio surveillance and human acoustic health monitoring require security protections for audio streaming over WSNs. The process of watermarking which embeds small amounts of data (i.e., the watermark) into the original audio is an effective technique to ensure the integrity of received audio data at the receiver in energy-constrained WSNs. However, the selection of positions to embed watermark into audio streams is critical to both received audio quality and watermarking authentication performance in error-prone wireless transmission environments. In this paper we propose an approach that dynamically determines the range of middle sub-band components for embedding the watermark with minimum quality distortions, based on psycho-acoustic models and adaptive sub-band thresholds. In addition, through unequal network resource allocation schemes the proposed approach protects both middle sub-bands and high sub-bands, which include the important audio components. Our theoretical analysis and simulation results demonstrate that the proposed quality-driven energy-efficient watermarking approach for audio transmissions can achieve considerable performance gains in WSNs.

H. Wang
University of Massachusetts, Dartmouth, MA, USA
e-mail: hwang1@umassd.edum

W. Wang
South Dakota State University, Brookings, SD, USA
e-mail: wei.wang@sdstate.edu

M. Chen (✉)
School of Computer Science and Technology, Huazhong University
of Science and Technology, Wuhan, 430074, China
e-mail: minchen@ieee.org

X. Yao
University of Electronic Science and Technology of China, Chengdu, China
e-mail: yxm@uestc.edu.cn

Keywords Multimedia sensor networks · Audio · Watermarking · Security

1 Introduction

A number of sensor-based audio applications such as acoustic surveillance and health monitoring require security protection for audio delivery in WSNs. In these applications, malicious intruders may access and modify the content of audio, destroying the integrity of audio content. As shown in Fig. 1, malicious users in wireless environments manipulate or eliminate the key audio elements so that the users at the sink node receive forged audio with altered meanings. Therefore, an audio authentication scheme is needed. Traditionally, authentication schemes [7, 21, 22] at stream level require that the integrity of the data must be strictly maintained. Any bit flip-flop can fail the authentication. Since Hash-chain or signature-based authentication at the packet level introduces comparatively large data overhead that consumes extra energy and bandwidth [6, 8, 10, 15], it is challenging to directly apply these authentication techniques to sensor networks for audio authentication.

Audio watermarking, on the other hand, is an effective technique to offer security protection for audio transmissions through the embedding of a small watermark into the original audio data. There are many existing research approaches [16–18, 20] for audio watermarking, which are primarily designed for copyright protection. However, to the best of our knowledge, no significant research work has been conducted for secure audio transmissions in wireless sensor networks based on watermarking techniques.

In a traditional watermarking system, robustness and transparency are two major goals. In sub-band codec-based watermarking schemes, e.g. watermarking MP3-compressed audio, the main factor affecting these two goals is the appropriate band selection for embedding watermarks. As reported in [11], in the sub-band domain, low sub-bands are the most perceptually significant coefficients and contain most of the audio signal's energy. Therefore, only minor band distortions can be

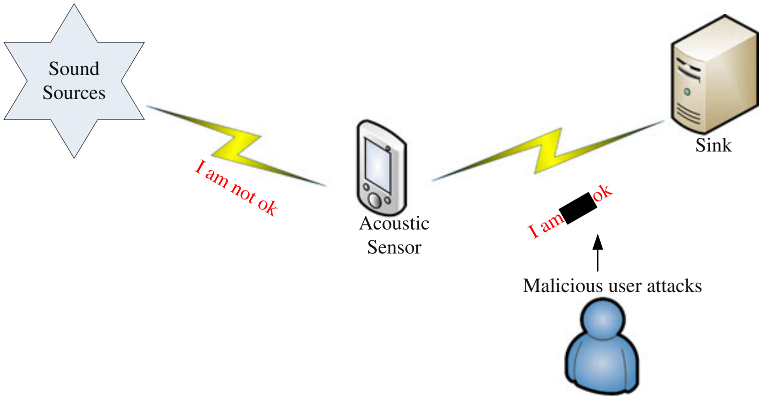


Fig. 1 A motivational example for audio authentication in WSNs

allowed for this region. High sub-bands contain the least perceptually significant coefficients but are also most sensitive to noise. Therefore, only middle sub-bands are appropriate places for embedding watermark data. However, determining the positions of middle sub-bands using strict thresholds limits the flexibility and capacity of the watermarking schemes.

In addition, in resource-limited WSNs the audio watermarking scheme must be carefully designed to be energy-efficient and tolerable to error-prone wireless transmissions. To conserve energy in audio transmissions, the immediate solution is to compress the audio data. One of the most efficient audio codec is the MPEG/Audio standard. In this codec, the uncompressed PCM input audio signal is organized into a sequence of 32 blocks of 12 time-domain samples each. Each set of 32 blocks is then transformed into the frequency domain through filter banks. At the same time, a psycho-acoustic model is utilized to process the PCM audio to find the masking thresholds to be applied to the frequency components for sub-band coding. In [11], the authors proposed a sub-band coding based watermarking scheme. In order to achieve the robustness and transparency required for watermarking, the scheme embeds the watermark into the middle sub-bands. As reported in [11], the PCM audio is transformed into 32 sub-bands. Only mid-band coefficients are suitable for embedding watermark information. The watermarking function for embedding a spread-spectrum watermark is shown as follows:

$$F_w = \begin{cases} S_i^{'} = S_i + a_i \times W_i, & \text{Embedding } '1' \\ S_i^{'} = S_i - a_i \times W_i, & \text{Embedding } '0' \end{cases} \tag{1}$$

where a_i is a scaling factor derived from the psycho-acoustic model, and W_i represents the specified signature. A detailed analysis of this approach is provided in [11]. The basic idea of this scheme is to embed "1" or "0" into middle sub-bands of each frame. Each frame uses the specified signature (e.g. PN sequence) to accommodate one bit information. The statistic testing is employed for the watermark detection procedure.

However, the thresholds for separating low, middle, and high sub-bands are not easily determined for different types of audio, which increased the difficulty in determining appropriate watermarking positions. Their watermarking scheme also requires a large number of sub-bands in each frame for embedding the watermark effectively, and thus necessitates statistical verification later at the receiver. A single audio frame in sensor networks typically contains only a few coefficients, making the watermarking verification procedure in [11] at the receiver site impractical.

In this study, we first propose a watermarking scheme to embed the watermark into two consecutive frames, which can accommodate more signature information in the middle band with less distortion. In this proposed approach, if any of the two frames is attacked, the authentication will fail due to the incompletely extracted watermark information. Thus the dual-frame security protection improves the authentication performance in malicious wireless environments. Secondly, the thresholds for the separation of low and high sub-bands are adaptively selected based on network conditions, audio quality and authentication performance requirements. The proposed watermarking scheme can offer the watermarking flexibility and the gains of optimizing both the quality and the authentication performance in wireless environments. In WSNs, the error prone wireless channel can cause packet loss, resulting in degraded audio quality and reduced watermark reliability in WSNs. If some of the data components containing watermark information are degraded or lost during wireless transmissions, the watermark information may not be extracted or be statistically verifiable. The watermarked components' quality becomes critical and extra network resources should be allocated to protect their transmissions over WSNs. As shown in our previous studies [25, 26], there is a tradeoff between the quality and resource allocation strategy under resource constraints (e.g. energy, bandwidth, computation capability). When more resources are allocated to protect the transmissions by decreasing the packet loss ratio in wireless environments, the audio quality gain can be achieved. In addition, efficient resource allocation strategies should consider the unequal importance of audio data. In the proposed watermarking scheme, more resources should be allocated to protect the transmissions of high-importance sub-bands rather than less important sub-bands. There is a tradeoff between the overall audio quality and watermark quality under resource constraints. For example, if more energy is budgeted for protecting the low sub-bands, less energy will be available for the middle sub-bands. The optimal tradeoff point between them should be determined in order to improve both the watermark component quality and overall audio quality.

In this paper, we propose a quality-driven and energy-efficient watermarking system for audio transmissions in WSNs. The watermark is embedded into the middle sub-bands, while the low sub-bands containing most of the audio signal's energy are highly protected. The least amount of resources is invested into the high sub-bands that represent the least important information (e.g. noises). The significance and original contributions of our proposed approach are described as follows:

1) It enhances the watermarking algorithm in [11] for audio authentication in WSNs. In our scheme the watermark is embedded into two consecutive audio frames, which improves the signature capacity of watermarking and authentication performance.
2) It provides Unequal Error Protection (UEP) on the low sub-bands.
3) It also provides UEP on the middle sub-bands.
4) It adapts both the low and high sub-band thresholds collectively for two consecutive frames to assure audio quality and security performance.
5) It provides a solution to optimize both the watermark components and overall audio quality under the resource constraints in WSNs.

The proposed scheme can achieved both robustness and transparency. Especially, it can tolerate the distortions caused by both the audio compression (source coding) and wireless channel errors. Compared with traditional authentication methods such as MAC (message authentication code), the proposed watermarking approach has less overheads when it is aware of the multimedia content.

The remainder of this paper is organized as follows: Section 2 provides a literature review of existing techniques for both audio watermarking and multimedia transmission quality protection. In Section 3, a quality-driven watermarking optimization problem for audio transmissions is formulated. The detailed algorithm for the proposed scheme is described in Section 4. Simulation and experimental results are shown in Section 5. The conclusions are provided in Section 6. The following presents the notations of the symbols used in the paper.

Symbols	Notations
W_i	Specified *i*th signature
α_i	Scaling factor
TL_n	Low subband threshold for *n*th frame
TH_n	High subband threshold for *n*th frame
$\mu\left[Q_m^n\right]$	The expected audio distortion reduction of the frame *n* when the unequal resource allocation (URA) is employed
d_j	The distortion reduction resulting from the source data in the *j*th subbands
$\kappa(j)$	The packet error probability under channel conditions when the appropriate network resources are allocated to protect the *i*th packet.
e	Energy consumption
$RTY_{\max}$	Maximum retransmission retry limit

2 Literature review

The existing audio watermark techniques can be divided into three major classes: a) time domain watermarking [2, 3, 5], b) frequency domain watermarking [11] and c) compressed domain watermarking [1, 4, 29]. Spread spectrum-based coding, echo hiding and phase coding are three major types of time domain watermarking techniques. In [3], the authors proposed a pseudo noise (PN) sequence-based watermarking scheme, which makes embedded watermark almost inaudible. In this scheme, the original signal and the PN-sequence must be available for the verification and detection. The scheme has the advantages of high detection rates. In [11], authors studied audio data hiding based on spread spectrum (SS) technique in the sub-band domain. It employs psycho-acoustic models to limit the introduced distortion. The novelty of this approach is to embed the watermark into the middle sub-bands rather than high or low sub-bands with less distortion to the audio quality. It calculates SMR (Signal-to-Masking Ratio) the corresponding scaling factors that control introduced distortion due to watermarking. In [1], the authors present a phase modulation-based audio watermarking. The inaudibility of the watermark is achieved by using small absolute phase shift. Their approach transforms the audio into multiple segments using overlapping windows and FFT (Fast Fourier transform), and the watermark is embedded into every other segment. The watermarking approach in [4] embeds watermarks into MPEG-2 AAC bit streams such that the watermark can be detected in the decompressed audio data. The advantage of this approach is to support efficient real time watermarking with less computation overheads. In [29], audio authentication is provided for speech protection through CELP speech coder and fragile watermarking. However, all these existing watermarking schemes are not easily applied to the resource-constrained WSNs to enable audio data authentication. These techniques lack the following features for successful audio watermarking for WSNs.

Traditional plain authentication techniques are to assure data integrity when malicious intruders intend to access and modify content delivered over wireless networks. However, in these schemes, a single bit failure due to the channel error could fail the authentication, which is not appropriate for multimedia whose semantic meaning of the content may not be changed due to this failure. In error-prone wireless sensor networks, bit errors and packet losses are inevitable due to ambient interferences. Therefore, there is a strong need for designing robust channel aware authentication schemes for multimedia transmissions. Preliminary researches [12–14, 31, 32] have been developed to provide robust authentication based on the invariant features extracted from the multimedia content. However, these schemes are conducted at the frame level using MAC (Message Authentication Code), which leads to extra authentication bits embedded in the frame. These extra bits increase communication overheads, and thus degrade the energy efficiency in WSNs. The feature of multimedia content are not considered in these schemes.

- Any audio watermarking techniques in WSNs must be energy-efficient and computation-efficient. The watermarking should not introduce significant data overhead into the original audio stream. Any extra data overhead consumes extra energy resources for both communication and computation.
- Any audio watermarking in WSN must tolerate the transmission distortions in error-prone wireless environments, and the watermark information should be extracted and statistically verifiable at the receiver site.

The watermarking technique proposed in [11] has the advantage of embedding a small watermark into the middle sub-bands of the original audio and is robust against compression distortion. It can potentially be applied in WSNs to achieve energy, computational and security efficiency. To resist the transmission distortions due to the wireless channel error in WSNs, unequal error protection (UEP) approaches have been studied by the research community. The simplest method of UEP is to allocate different channels for different types of data. In [9], a perceptually controlled UEP scheme is proposed for transmitting audio over IP networks. In [28] Reed Solomon FEC (Forward Error Correction) codes are employed to provide UEP for audio streams. In [19], the authors consider a Media-Specific FEC for audio transmission quality protections. In our previous work [24], the unequal importance of audio stream is identified and unequal resource allocation is employed to achieve high-quality audio transmission. Since digital audio streams typically have non-uniform perceptual importance, UEP can be used for multimedia transmission protection over wireless channels. In addition, the transmission distortion in wireless environment could also degrade the watermark part quality and make watermark undetectable at the receiver site. The current published research lack the interaction between watermarking and the UEP communications schemes to resist distortion and provide the audio authentication and quality transmission in resource-constrained WSNs. In this study, our contributions are two-fold:

1) The watermarking technique based on sub-band coding is explored for secure audio transmissions over WSNs. The watermark is embedded into two consecutive frames to accommodate a larger signature and thus make the verification procedure more accurate and robust.
2) The transmissions of low sub-bands that contain important information and middle sub-bands with the embedded watermark are highly protected through efficient UEP with resource allocations. The thresholds of separation among high sub-bands, middle sub-bands and low sub-bands are dynamically determined based on the network conditions and performance requirements. Through them, both audio quality and watermark component quality are improved.

3 Problem statement

To ensure high quality and secure audio transmissions over WSNs, a quality-driven energy-efficient audio transmission scheme based on watermarking is required. The watermarking technique in [11] has great potential to be applied to WSNs due to its robustness, computational efficiency, reduced watermarking overhead, and a sub-band coding based compression-domain technique that is applicable to the popular MPEG codec. Also, it can tolerate the distortion that results from audio compression, strongly required for audio transmissions in WSNs.

However, there are several challenges for applying that scheme to energy-constrained WSNs:

- The watermarking scheme requires a significant number of embedding coefficients in each audio frame. Nevertheless, a single short frame in sensor networks is not enough to accommodate the signature for accurate watermark verification.
- Wireless transmission distortion degrades the audio quality as well as the quality of the frequency components that include the watermark information. The distorted watermark components make it difficult to verify the trusted audio at the receiver site. Most current watermarking schemes do not consider the distortion caused by the error-prone wireless channel.
- When the watermark is embedded in the middle sub-bands, ensuring the received quality of these sub-bands is vital to watermark retrieval and verification. Therefore, both low and middle sub-bands need to be protected. Given a limited energy consumption budget, less invested energy for middle sub-band protections lead to more invested energy left for low sub-bands. Thus, there is a tradeoff between the protection on the middle bands and low bands.

In response to these challenges, we have studied a quality-driven energy-efficient audio watermarking scheme. As shown in Fig. 2, in our proposed scheme, the watermark information is carried by the generated specified signature ($W_n + W_{n+1}$), which is embedded into two consecutive audio frames. The middle band positions where the watermark is embedded are based on network conditions and dynamically determined by the required watermark and audio quality. In order to assure the required audio quality, the low sub-bands that contain the most important audio

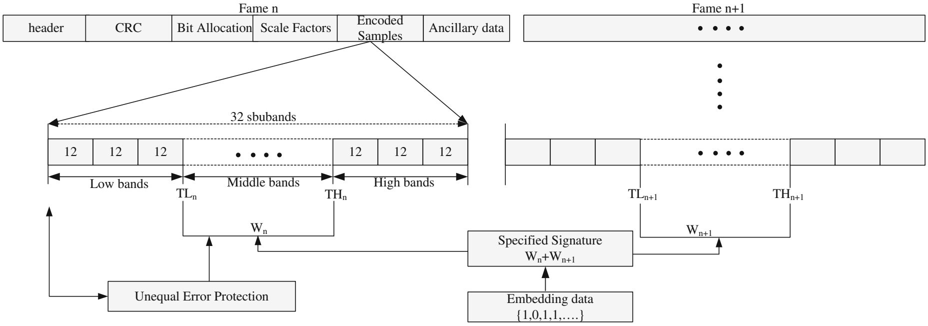


Fig. 2 A MPEG audio codec based secure audio transmission scheme

information are highly protected through effective resource allocation. The middle sub-bands are also protected to assure the watermark component quality and to improve the authentication performance in WSNs. In this paper, a cross-layer optimization approach for secure audio transmission is formulated, combining resource allocation, adaptive sub-band threshold determination, as well as watermarking control to maximize the audio quality. The detailed optimization framework is presented in chapter IV.

4 Optimization framework and adaptive watermarking algorithm

An energy-efficient quality-driven adaptive watermarking scheme is proposed as shown in Fig. 1. Two groups of adaptive thresholds $\{TL_n, TH_n\}$ $\{TL_{n+1}, TH_{n+1}\}$ are defined for two frames. TL_n and TL_{n+1} are the low sub-band thresholds for n and $n+1$ frame separately, and TH_n and TH_{n+1} are the high sub-band thresholds, respectively. An audio quality $\mu(Q_A)$ maximization problem can be formalized for each of the two consecutive audio frames as shown in the following equation:

$$\left\{(TL_n, TH_n)^{opt},\ (TL_{n+1}, TH_{n+1})^{opt}\right\} = \underset{\{(TL_n, TH_n),\ (TL_{n+1}, TH_{n+1})\}}{\arg\max} \left[\mu(Q_A)\right] \tag{2}$$

$$\text{s.t. } Q_m < \Gamma_m,\ e < \Gamma_E$$

The middle sub-band quality at the nth frame can be expressed as shown in the following equation:

$$\mu\left[Q_m^n\right] = \sum_{i=1}^{N_s} \left(\sum_{j=1}^{i} d_j \left(\prod_{j=1}^{i} \left[1 - \kappa(j)\right] \right) (\kappa(j+1)) \right) \tag{3}$$

where $\mu[Q_m^n]$ is the expected audio distortion reduction of the frame n when the unequal resource allocation (URA) is employed and is affected by the packet error probability, d_j is the distortion reduction resulting from the source data in the jth sub bands. $\kappa(j)$ denotes the packet error probability under channel conditions when the appropriate network resources are allocated to protect the ith packet. The middle sub-band quality $\mu[Q_m^{n+1}]$ at frame $n+1$ can be similarly calculated. The related detail study is refereed to [27, 30].

Due to the additive nature of the distortion reduction, the total quality of the middle sub-bands for two frames can be derived in (4).

$$\mu\left[Q_m\right] = \sum_{i=1}^{2} \left(\sum_{j=1}^{i} Q_m^{n-j+1} \right) \tag{4}$$

The total audio quality for two frames can be described in (5):

$$\mu\left[Q_A\right] = \sum_{i=1}^{6} \left(\sum_{j=1}^{i} Q_j \right) \tag{5}$$

where $\{Q_1, Q_4\}$ represents the high sub-band distortion reduction, $\{Q_2, Q_5\}$ is the middle sub-band distortion reduction, and $\{Q_3, Q_6\}$ is the low sub-band distortion reduction. The energy consumption is calculated as follows:

$$e = \left\lceil \frac{\mu}{S} \right\rceil (\xi_t + \xi_r) \frac{(S + \rho_o)(1 - \Delta^{RTY_{MAX}+1})}{R_s b (1 - \Delta)} \tag{6}$$

where e is the resulting energy consumption, μ is the packet size, ξ_t is the power consumed by the transmitter, ξ_r is the power consumed by the receiver, S is the frame size, ρ_0 is the header size, R_s the symbol rate, b is the constellation size, Δ is the packet error rate, and $RTY_{\max}$ is the maximum retransmission retry limit. For more details on this energy model please refer to our related previous work on WSN communication energy conservation in [26].

Figure 3 shows the cross-layer architecture unequal resource allocation (URA) for audio watermarking over WMSNs. At the upper layer, the MDCT coding (Modified Discrete Cosine Transform) of MPEG standard is applied. In this framework, the audio data are divided into the M (middle) bands and HL (high and low) bands after two adaptive low subband (TL) and high subband (TH) thresholds are applied. M bands contain the watermark information, while L bands do not. It is critical to protect the transmission of these packets in the error-prone wireless channel. In this architecture, these blocks are further packetized into WM (frame with watermark information) and MW (frame without watermark information) frames. The unequal resource allocation (URA) (e.g. power and transmission rates) are allocated to protect WM frame transmissions. At the receiver site, the received frames are assembled and decoded. The left bar shows the cross-layer solution based on multiple parameters from multiple equivalent layers. When the receiver obtains these frames, it decodes them. This process is called de-packetization. At the application layer, these packets are reassembled to decode the whole audio. The

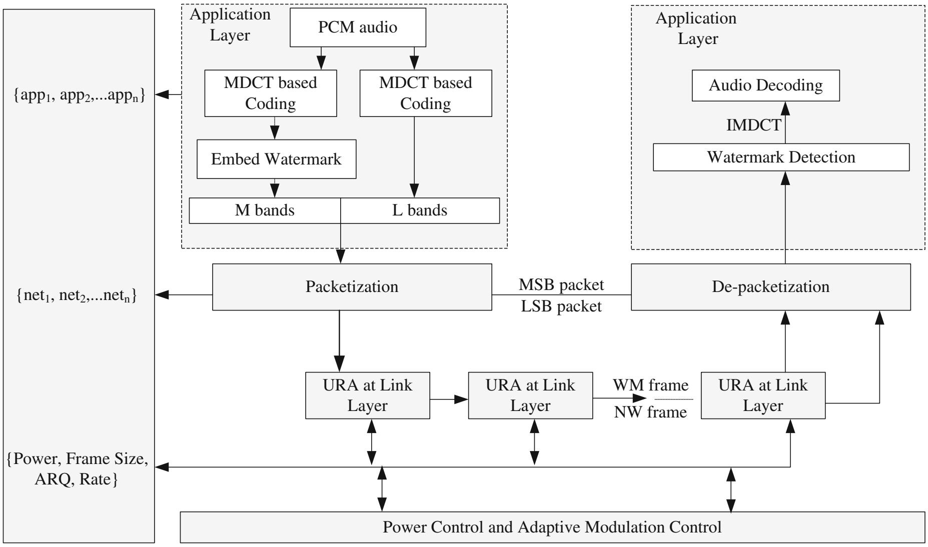


Fig. 3 Energy efficient adaptive audio watermarking framework

Table 1 The proposed quality-driven watermarking optimization algorithm

```
Input: The distortion reductions of each sub-band of two audio frame $(n, n+1)$
Output: Optimized $\{(TH_n, TL_n, TH_{n+1}, TL_{n+1}\}$, optimized resource allocation parameters $r$
Step 1: Reorder the sub-bands in order from low to high frequencies for each frame
Step 2: The threshold set$\{(TH_n, TL_n, TH_{n+1}, TL_{n+1}\}$and resource allocation vector $r$ are
  represented as each chromosome, and each element in the chromosome is coded as gene by binary
  coding and decoding for each chromosome.
Step 3: Re-organize the audio frame into multiple smaller frames (middle band frame, low sub-band
  frame and high sub-band frame).
Step 4: Initialize the size $S_{pop}$ of population and the maximal evolution iterations.
Step 5: Calculate the fitness of each chromosome in the current generation $g$ :
For $i = 0;\ i \leq S_{pop} - 1;\ i++$
{
  Calculate expected audio distortion reduction $\mu\left[Q_A\right]$ using the current threshold set
    $\{(TH_n, TL_n, TH_{n+1}, TL_{n+1}\}$.
  The fitness $f$ of each chromosome can be evaluated using $f(i) = \mu\left[Q_A\right]$.
  Calculate the expected energy consumption $e$ and distortion reduction for middle sub-bands $Q_m$.
  If $e > \Gamma_E$ or $Q_m < \Gamma_m$
    remove the unqualified chrome due to energy consumption and middle sub-bands constraints
    f(i)=0;
  end
}
Step 6: In the current generation, sort the chromosome in descending order based on the fitness
  value $f$. Look for the best chromosome in the current population with the largest fitness value
  satisfying energy consumption budget $\Gamma_E$ and $\Gamma_m$.
Step 7: Loop step 5 and 6 for the evolution.
Step 8: Based on the optimized sub-bands threshold $\{(TH_n, TL_n, TH_{n+1}, TL_{n+1}\}^{opt}$, the middle
  sub-bands are determined, the watermarking is embedded in to the middle sub-bands.
```

watermark is also extracted through the inverse-MDCT process, and then the audio is then authenticated.

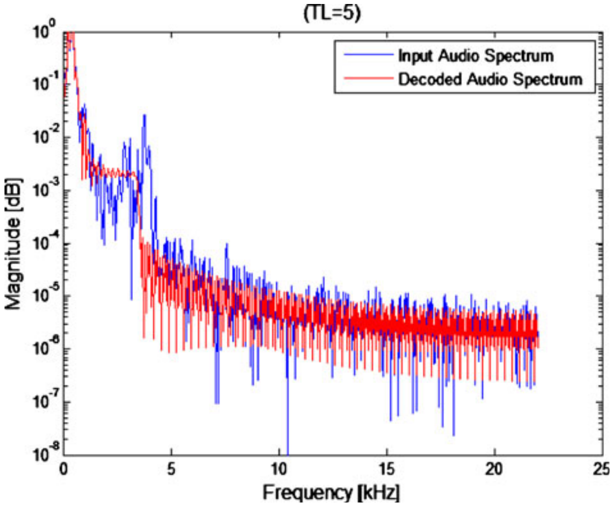
The quality-driven adaptive watermarking algorithm for audio transmissions is described in Table 1.

5 Simulation and experiment

The Matlab simulation is based on the continuous speech frames from audio recordings. The sample frequency is 44100 Hz. 1024-point Fast Fourier Transform (FFT) is applied only for the audio waveform representation in frequency domain shown in the figures below, whereas the codec follows the specifications for MPEG-Audio compression. For calculating the energy consumption and to analyze wireless transmissions we apply sensor network protocol parameters for T-MAC [23]. In T-MAC, data packets for TinyOS have a MAC header of 11 bytes. For the control packets such as RTS and ACK, the length is 13 bytes. The preamble length is 18 bytes. A CTS packet is 15 bytes. The transmission of the control packets uses the basic modulation scheme, while the transmission of DATA packets utilizes the scaled modulation schemes (e.g., BPSK, QAM).

Figure 4 shows the decoding quality of the watermarked audio when the watermark is embedded in the first five low sub-bands. As shown in Fig. 4, this obviously

Fig. 4 The audio quality when watermark is embedded in low sub-band (TL = 5)



impacts in particular the low frequencies that contain significant information. This graph clearly demonstrates that the lower sub-bands are in no way suitable for embedding the watermark. It also suggests the possibility of adapting the low sub-band thresholds to avoid embedding watermark in these sub-bands.

Figure 5 shows the degraded audio quality at the receiver site when the high sub-bands are lost during transmission. The high sub-bands are defined from 20th sub-band to 32nd sub-band in this scenario. It is observed that the decoded audio spectrum is similar to the input audio spectrum. It demonstrates that high sub-band losses during transmissions do not significantly degrade the audio quality when

Fig. 5 The received audio quality when the high sub-band is lost (TH = 20)

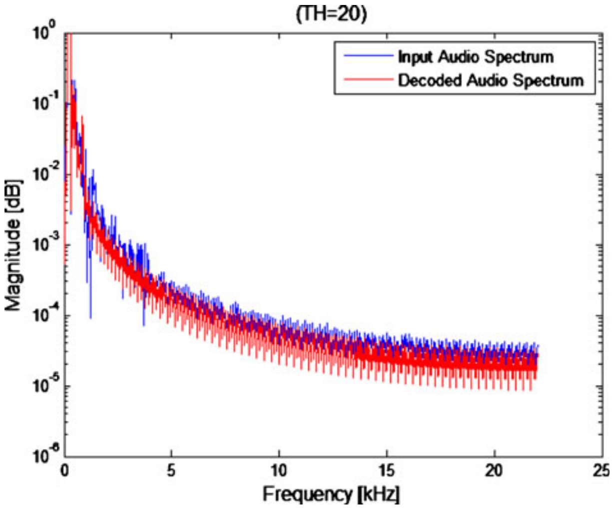
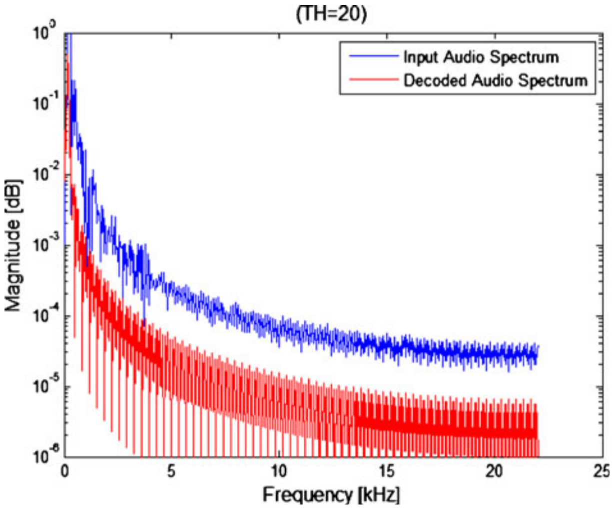


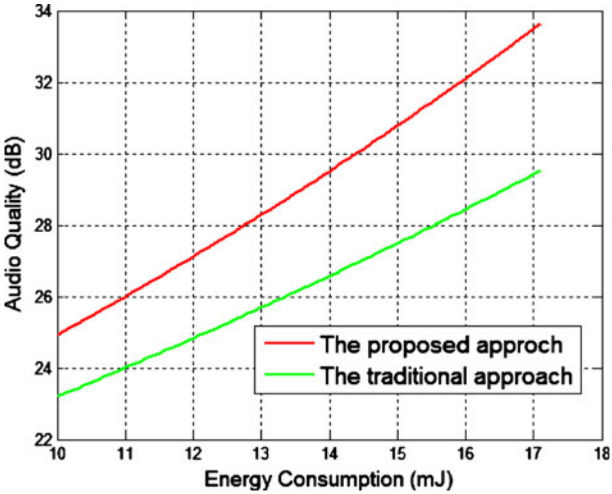
Fig. 6 The received audio quality when the (2–5) sub-band is lost (TL = 5)



these sub-bands contain the least significant information. Therefore, some amount of packet losses of high sub-bands can be tolerated by the system in resource-limited sensor network environments.

In Fig. 6 we show the decoded audio quality when the sub-bands two to five are lost. Compared with the original audio spectrum, the received audio quality has been significantly distorted. It is clear that the quality of these middle sub-bands is critical to the whole audio quality. The loss of only four sub-bands can significantly degrade the whole audio quality. This demonstrates that pursuing a high quality of middle sub-bands should have higher priority than protecting the lower sub-bands in resource constrained WSNs. As shown in Fig. 7, the proposed approach can achieve better audio quality than the traditional approach under the same energy consumption budget.

Fig. 7 Energy consumption vs. audio quality with unequal resource allocation



5.1 Transparency and robustness

In this section, we demonstrate that the transparency goal has been achieved in the proposed watermark scheme. We conducted the experiment and compare the SNR (Signal-to-Noise Ratio) result for the watermarked audio and the decoded audio for 12 types of audio test samples at a general 64 kbps bit rate. As we can see the second column from Table 2, the proposed adaptive watermarking scheme can still achieve very good SNR performance (26.79 dB at average) when the watermark is regarded as the the noise. In this situation, no transmission distortion is considered and only the watermark is considered as noises. In addition, wireless channel errors could also distort the audio transmission quality, and thus the transmission robustness is affected. In this experiment, the distortion caused by transmission errors is also considered as noises like the watermark to the original audio data. The experiment is conducted under different channel condition represented by Bit Error Ratio (BER) such as 1e-3, 1e-4 and 1e-5. It is clear that the lower BER leads to high SNR generally. It is because the worse channel quality leads to more packet loss during the wireless transmission. Even with BER 1e-3, the average 21.3 dB(SNR) still can be achieved, which has very close SNR performance to normal MP3 compression at 64 kbps reported in [11]. These results demonstrated that the proposed watermarking scheme can achieve the robust transmission in WSNs.

5.2 Energy efficiency

We implemented the proposed algorithm described in Section 4 and compared it with the traditional approach in [11]. Two fixed thresholds (TL = 4, TH = 16) as suggested in [11] are applied for the traditional watermarking scheme. This traditional scheme is defined without unequal resource allocation for audio transmissions. Figure 6 shows the comparison of our proposed approach combining audio watermarking and unequal resource allocation against the traditional approach in terms of audio quality. Under the same energy consumption budget and middle band quality requirements, the proposed approach can achieve up to 3.4 dB improvement in the achieved audio quality.

Table 2 The transparency and robustness study for the proposed adaptive audio watermarking

Audio type	SNR (dB) (no transmission noises)	SNR (dB) (BER = 1e-5)	SNR (dB) (BER = 1e-4)	SNR (dB) (BER = 1e-3)
Airplane	23.15	23.15	20.90	22.03
Electrical drill	29.22	23.53	21.70	22.52
Elephant	27.53	23.50	21.71	22.07
Motor	29.22	23.84	22.36	22.12
Thunder	28.69	23.93	22.52	21.91
Ocean wave	26.64	23.91	22.50	21.39
Bubbles	30.93	24.14	22.93	21.58
Laugh	28.41	24.17	22.99	21.40
Frog	23.65	24.14	22.99	20.88
Computer	23.76	24.10	22.97	20.66
Machine	25.02	24.15	23.06	20.70
Train_horn	25.26	24.20	23.15	20.76
Average	26.79	23.89	22.48	21.51

6 Conclusions

In this paper, an energy-efficient and quality-driven watermarking system is proposed for secure audio transmissions in WSNs. It simultaneously controls the resource allocation at the MAC-PHY layers and the watermark embedding at the application layer to optimize audio quality and assure the authentication performance. The proposed watermarking scheme does not only achieve the robustness and transparency desired in general compression approaches (i.e., MPEG-1, MEPG-2 and MP3), but also it is shown to be resistant to the distortion due to the wireless channel. It improves both the watermark component quality and overall audio transmission quality in WSNs. Our study showed that significant gains in terms of energy efficiency, authentication performance, and audio transmission quality have been achieved with our proposed adaptive watermarking scheme in WSNs.

Acknowledgements The work of Min Chen was supported in part by the Program for New Century Excellent Talents in University (NCET), and through the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2011-0009454).

References

1. Arnold M (2000) Audio watermarking: features, applications and algorithms. In: Proc. IEEE international conference on multimedia and Expo (ICME '00). New York, NY, USA, pp 1013–1016
2. Bassia P, Pitas I, Nikolaidis N (2001) Robust audio watermarking in the time domain. IEEE Trans Multimedia 3(2):232–241
3. Boney L, Tewfik AH, Hamdy KN (1996) Digital watermarks for audio signals. In: Proc. 3rd IEEE int conf multimedia computing and systems, pp 473–480
4. Boney L, Tewfik AH, Hamdy KN (1996) Digital watermarks for audio signals. In: Proc. VIII European signal proc. conf. (EUSIPCO '96). Trieste, Italy
5. Daniel G, Anthony L, Walter B (1996) Echo hiding, vol 1174. Lecture Notes in Computer Science, Information Hiding. Springer, Berlin, Germany, pp 295–315
6. Ge X, Huang K, Wang C, Hong X, Yang X (2011) Capacity analysis of a multi-cell multi-antenna cooperative cellular network with co-channel interference. IEEE Trans Wirel Commun (TWC) 10(10):3298–3330
7. Haitsma J, Kalker T, Oostveen J (2001) Robust audio hashing for content identification. In: Proc international workshop content-based multimedia indexing. Brescia, Italy
8. Humar I, Ge X, Xiang L, Jo M, Chen M, Zhang J (2011) Rethinking energy-efficiency models of cellular networks with embodied energy. IEEE Netw 25(3):40–49
9. Khalifeh CA, Yousefi'zadeh H (2008) An optimal UEP scheme of audio transmission over MIMO wireless links. In: Proc. wireless communications and networking conference, 31 March–3 April 2008, pp 3191–3196
10. Lai CF, Huang Y, Chao HC (2010) DLNA-based multimedia sharing system for OSGI framework with extension to P2P network. IEEE Systems Journal 4(2):262–270
11. Li X, Yu H (2000) Transparent and robust audio data hiding in sub-band domain. In: Proc. IEEE int conf information technology: coding and computing, pp 74–79
12. Li Z, Sun Q, Lian Y, Chang WC (2007) Joint source-channel-authentication resource allocation and unequal authenticity protection for multimedia over wireless networks. IEEE Trans Multimedia 9(4):837–850
13. Lin C-Y, Chang S-F (1998) A robust image authentication method surviving JPEG lossy compression. SPIE Storage and Retrieval of Image/Video Database 3312:296–307
14. Lin C-Y, Chang S-F (2001) A robust image authentication method distinguishing JPEG compression from malicious manipulation. IEEE Transaction on Circuits System and Video Technology 11(2):153–168
15. Lin K, Lai CF, Liu X, Guan X (2011) Energy efficiency routing with node compromised resistance in wireless sensor networks. Mobile Networks and Applications. doi:10.1007/s11036-010-0287-x

16. Neubauer C, Herre J (1998) Digital watermarking and its influence on audio quality. In: Proc. 105th AES convention, audio engineering society. San Francisco (preprint 4823)
17. Neubauer C, Herre J (2000) Audio watermarking MPEG-2 AAC bitstream. In: 108th AES convention, audio engineering society. Paris (preprint 5101)
18. Neubauer C, Herre J (2001) Advanced audio watermarking and applications. In: 109th AES convention, audio engineering society. Los Angeles (preprint 5176)
19. Perkins C, Hodso O, Hardman V (1998) A survey of packet loss recovery techniques for streaming audio. IEEE Network, pp 40–48
20. Qiao L, Nahrstedt K (1999) Non-invertible watermarking methods for MPEG encoded audio. In: Wong PW, Delp EJ (eds) Proc. SPIE, security and watermarking of multimedia contents, vol 3657, pp 194–202
21. Skraparlis D (2003) Design of an efficient authentication method for modern image and video. IEEE Trans Consum Electron 49(2):417–426
22. Sun Q, Chang S (2005) A secure and robust digital signature scheme for JPEG2000 image authentication. IEEE Trans. Multimedia 7(3):480–494
23. van Dam T, Langendoen K (2003) An adaptive energy-efficient MAC protocol for wireless sensor networks. In: Proc. Sensys‘03
24. Wang H, Hempel M, Peng D, Wang W, Sharif H, Chen HH (2010) Index-based selective audio encryption for wireless multimedia sensor networks. IEEE Trans Multimedia 12(3):215–223. doi:10.1109/TMM.2010.2041102
25. Wang H, Peng D, Wang W, Sharif H, Chen HH (2008) Energy-aware adaptive watermarking for real-time image delivery in wireless sensor networks. In: Proc. IEEE international conference on communications (ICC)
26. Wang W, Peng D, Wang H, Sharif H, Chen HH (2008) Energy-constrained distortion reduction optimization for wavelet-based coded image transmission in wireless sensor networks. IEEE Trans Multimedia (TMM) 10(6):1169–1180
27. Wang H, Peng D, Wang W, Sharif H, Chen HH (2009) Image transmission with security enhancement based on region and path diversity in wireless sensor networks. IEEE Trans Wirel Commun (TWC) 8(2):757–765
28. Wang Y, Ahmaniemi A, Isherwood D, Huang W (2003) Content-based UEP: a new scheme for packet loss recovery in music streaming. In: Proc. the eleventh ACM international conference on multimedia. ACM Press, New York, NY, USA, pp 412–421
29. Wu C-P, Kuo C-CJ (2002) Comparison of two speech content authentication approaches. In: Photonics West, vol 4675. San Jose, Calif, USA, pp 158–169
30. Wu Z, Bilgin A, Marcellin M (2005) Joint source/channel coding for multiple images. IEEE Trans Commun 53(10):1648–1654
31. Zhou L, Chao H (2011) Multimedia traffic security architecture for internet of things. IEEE Netw 25(3):35–40
32. Zhou L, Chao H, Vasilakos A (2011) Joint forensics-scheduling strategy for delay-sensitive multimedia applications over heterogeneous networks. IEEE J Sel Areas Commun 29(7):1358–1367

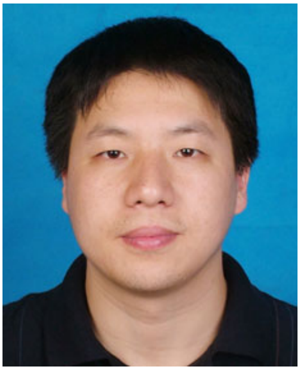
Honggang Wang (M’10) received the B.S. and M.S. degrees from Southwest Jiaotong University, China, in 1996 and 2001, respectively, and the Ph.D. degree from University of Nebraska-Lincoln in

2009. He worked for Lucent Technologies China from 2001 to 2004 as a Senior Software Developer. He is currently an Assistant Professor at the University of Massachusetts, Dartmouth. His research interests include networking, wireless sensor networks, multimedia communication, network and information security, embedded sensory system, biomedical computing, and pattern recognition.

Wei Wang (M'10) received the B.S. and M.S. degrees from Xian Jiaotong University, Xian, China, in 2002 and 2005 respectively, and the Ph.D. degree from the University of Nebraska-Lincoln in 2009. He is currently an Assistant Professor in the Department of Electrical Engineering and Computer Science, South Dakota State University, Brookings. His research interests include computer networks, wireless communications, embedded systems, and multimedia computing. Dr.Wang is one of the recipients of the Best Paper Award in IEEE WCNC 2008. He also served as a TPC member in IEEE Globecom 2010, WiMob 2008 and 2009, and IEEE WCNC 2009, as well as a reviewer for several journals/transactions.

Min Chen (SM'09) was an assistant professor in School of Computer Science and Engineering at Seoul National University (SNU). He was a research associate in Dept. of Computer Science at University of British Columbia (UBC) for half year. He has worked as a Post-Doctoral Fellow in Dept. of Electrical and Computer Engineering at UBC for three years since Mar. 2009. Before joining UBC, he was a Post-Doctoral Fellow at SNU for one and half years. He has published more than 140 technical papers. He is the sole author of a textbook OPNET Network Simulation (Tsinghua Univ. Press, 2004). Dr. Chen received the Best Paper Runner-up Award from The Fifth International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine) 2008. He serves as editor or associate editor for Wiley I. J. of Security and Communication Networks, Journal of Internet Technology, KSII Transactions on Internet and

Information Systems, IET Communications, International Journal of Sensor Networks (IJSNet). He is a managing editor for International Journal of Autonomous and Adaptive Communications Systems. He has worked as session chairs in several conferences, such as VTC'08, QShine'08, ICACT'09, Trientcom'09, and ICC'09. He is a TPC co-chair of BodyNets 2010. He is a symposia co-chair and workshop chair of CHINACOM 2010. He is the co-chair of MMASN-09, UBSN-10, and NCAS-11. He was the TPC chair of ASIT-09, ASIT-10, TPC co-chair of PCSI-09, PCSI-10, and UMES-11 publicity co-chair of PICom-09. He serves as the corresponding guest editors for several international journals, such as ACM/Springer Mobile Networks and Applications (MONET), International Journal of Communications System (IJCS). He is an IEEE senior member.



Xingmiao Yao is an associate professor of the Communication Engineering Department, University of Electronic, Science and Technology of China. His research interests include networking and multimedia processing.