1st International Workshop on Wireless Multimedia Sensor Networks (WMSN'12)

QoE-aware FEC Mechanism for Intrusion Detection in Multi-tier Wireless Multimedia Sensor Networks

Zhongliang Zhao and
Torsten Braun
University of Bern
Bern – Switzerland
Email: {zhao, braun}@iam.unibe.ch

Denis do Rosário and
Eduardo Cerqueira
Federal University of Para
Belem – PA – Brazil
Email: {denis, cerqueira}@ufpa.br

Roger Immich and Marilia Curado
Department of Informatics Engineering
University of Coimbra
Coimbra – Portugal
Email: {immich, marilia}@dei.uc.pt

***Abstract*—Wireless Multimedia Sensor Networks (WMSNs) play an important role in pervasive and ubiquitous systems. The multimedia content in such networks has the potential of enhancing the level of information collected, enlarging the range of coverage, and enabling multi-view support. For WMSN applications, the multi-tier network architecture has proven to be more beneficial than a single-tier in terms of energy-efficiency, scalability, functionality and reliability. In this context, a multimedia intrusion detection application appears as a promising application of multi-tier WMSNs, where the lower tier can detect the intruder using scalar sensors, and the higher tier camera nodes will be woken up to send real time video sequences from the detected area. The transmission of multimedia content requires a certain quality level from the user perspective, while energy consumption and network overhead should be minimized. Among the existing mechanisms for improving video transmissions, Forward Error Correction (FEC) can be regarded as a suitable solution to improve video quality level from the user point-of-view. In this work, we propose a Quality of Experience (QoE)-aware FEC mechanism for WMSNs, which creates redundant packets based on impact of the frame on the user experience. According to the simulation results, our proposed mechanism achieved similar video quality level compared with standard FEC, while reducing the transmission of redundant packets, which will bring many benefits in a resource-constrained system.**

***Index Terms*—Wireless multimedia sensor networks, QoE, Forward error correction, Multi-tier architecture, Intrusion detection.**

I. INTRODUCTION

Wireless Multimedia Sensor Networks (WMSNs) [1] are new types of sensor networks gaining research interest due to the availability of low-cost and mature technologies in camera sensors and scalar sensors. As an extension of traditional scalar Wireless Sensor Networks (WSNs), WMSNs are composed of wirelessly interconnected sensor nodes equipped with multimedia devices, such as cameras and microphones, and capable to retrieve video and audio streams, still images, as well as scalar sensor data.

WMSNs promise a wide scope of potential applications in both civilian and military areas, which require visual and audio information such as multimedia surveillance, traffic monitoring and enforcement, personal health care, environmental and structural monitoring, industrial process control, etc. In these applications, multimedia content has the potential of enhancing the level of information collected, enlarging the range of coverage, and enabling multi-view support. Video content provides users with more precise information than simple scalar data.

Compared to WSNs, WMSNs have some additional characteristics and design challenges. The nature of the real-time multimedia data, which requires high bandwidth demand, real-time delivery, tolerable end-to-end delay, acceptable jitter and lower frame loss rate make the design of an efficient application for WMSNs a nontrivial task. Many different reference architectures have been proposed during past years to fulfill different application requirements. A multi-tier network architecture has proven to be more beneficial than a single-tier architecture in terms of less energy consumption, lower loss, higher functionality, better scalability and reliability [2].

In this work, we focus on the application of multi-tier WMSNs with video QoE-aware support for an intrusion detection scenario. In such an application, real-time video content has the potential to enhance the level of collected information, enable to detect and monitor the intruder. However, the video sequence should be delivered to the end user with an acceptable video quality level from the user perspective, enabling the intrusion detection process.

From the aspect of video characteristics, a compressed video is composed of three types of frames with different importance for the video quality level. The loss of more important frames can cause higher distortion in video quality than the loss of a less important frame [3].

In low-power communication, the constraints of sensor nodes will increase the effects of wireless channel errors. Thus, some error control schemes for multimedia communication over multi-hop WMSNs are needed. Application-level Forward Error Correction (FEC) can be applied as error control scheme for handling loss in real-time communication [4], denoted as FEC. FEC adds h redundant packets to n original source packets in the application layer to recover lost packets. The recovered data can be used to reconstruct a lost frame, and thus improve the video quality.

A multimedia intruder detection application requires high video quality from the user perspective, scalability, energy-efficiency and low network overhead. However, existing works do not take into account FEC mechanisms that consider the frame importance to create the redundant packets. The QoE-

978-1-4577-2014-7/12/$26.00

aware FEC can improve the video quality without increasing the overhead and energy consumption. Additionally, these proposals do not use multi-tier architectures to provide energy-efficiency and scalability.

In this context, this paper proposes a QoE-aware FEC mechanism for WMSNs for an intrusion detection application in a multi-tier architecture. The proposed FEC mechanism includes redundant packets based on frame importance from the user perspective to decrease the packet overhead, while achieving video sequences of intruders with high quality.

Simulations were carried out to show the impact and benefits of the proposed QoE-aware FEC mechanism for real-time video transmission in WMSNs for intrusion detection. This paper includes an analysis of energy-efficiency, overhead and video quality. Video quality was analyzed by means of well-known QoE objective metrics, which are Structural Similarity (SSIM) and Video Quality Metric (VQM).

The remainder of this paper is structured as follows. Section II explains the current technologies and architectures in WMSNs and outlines the related FEC mechanisms and their main drawbacks. The state of the art on intrusion detection/surveillance applications is also presented in this section. Section III describes the proposed multi-tier architecture and the QoE-aware FEC mechanism. Simulations were carried out and are described in Section IV. The paper concludes with Section V, which summarizes the main contributions and results of this paper.

II. RELATED WORK

Several efforts have been made to achieve promising results in various fields of WMSNs, both in the development of specific video camera hardware and the design of efficient algorithms and protocols for multimedia transmission. In the area of WMSNs, the main design objectives are to minimize energy consumption and to prolong network lifetime under certain video quality requirements. To achieve these goals, mainly two design approaches are applied: deploying a multi-tier network architecture and designing a scheduling algorithm for controlling the transmission of video packets in an efficient manner.

Most of the proposed works in traditional WSNs are based on a flat architecture of distributed homogeneous nodes, where low-power scalar sensors are in charge of performing simple tasks such as detecting scalar physical measurements, i.e., vibration or temperature. In WMSNs, with camera sensors that provide additional capabilities and functionalities, the amount of packets to be transmitted is much higher than in scalar sensor networks, and thus, this fact implies higher energy consumption and needs of larger buffers and memories. Therefore, traditional WSN architectures should be reconfigured in a way that the network can be more scalable and more efficient depending on different application requirements. An example of a single-tier video surveillance and monitoring system is presented in [5]. The main objective of the system is to use multiple, cooperative video sensors for continuous tracking and coverage. A framework for a single-tier multi-camera surveillance application [6] applies multi-source spatio-temporal data fusion for efficient tracking. Almalkawi et al. [1], summarize different network architectures, of which the multi-tier one is of great interest. A multi-tier architecture includes three tiers of sensors: in the lowest tier, scalar sensors perform simple tasks, e.g. motion detection. A second tier of camera sensors performs more complicated tasks such as object detection or recognition. At the top tier, high end video sensors are connected to wireless gateways. In this way, the high-tier nodes are only woken up by low-tier nodes when necessary. SensEye [2] demonstrates that in a surveillance application a multi-tier network can achieve an order of magnitude reduction in energy usage when compared to a single-tier network without sacrificing reliability.

Regarding the transmission of the video stream, [7] investigates issues associated with the transport of multimedia streams across WSNs, introducing a flow control algorithm based on pipelined transmission to increase network performance. Politis et al. [8] propose a scheduling algorithm for transmitting video packets over multiple paths according to their importance (high priority packets over high bandwidth paths), including a power-aware packet scheduling mechanism that selectively drops the least significant video packets prior to transmission in order to save energy. Guo [9] designs a QoS-enabled dynamic path formation algorithm to produce throughput-aware video delivery over WSNs by distributing a limited number of mobile sinks to the bottleneck location for each video stream. Chen et al. [10] propose a real-time video surveillance system composed of two IP cameras and sixteen low-cost wireless sensors in a multiple-tier architecture. The sensor network can detect and track an object and wake up the IP cameras to record these events. While in that work the communication between sensor nodes and IP cameras are via a sink controller, it is not a distributed solution and therefore can not scale well.

In the context of FEC, Jeong et al. [11] introduced a FEC mechanism into WSNs and remarked that FEC is more preferable than Automatic Repeat reQuest (ARQ) in a power scarce environment. Hurni [12] investigated the potential of FEC mechanisms and dynamic/run-time adaptive FEC variants in WSNs. Sarisaray et al. [13] present an error compensation technique, which uses FEC and multipath transmission. The FEC technique employs a modified version of the wavelets based error concealment algorithm. Yang el al. [14] propose a cross-layer FEC scheme for reliable block transfer of variable-length coded data in WMSNs. The proposal combines the iterative joint source channel fountain codes FEC at physical, transport and application layer. However, such proposal does not take into account the video content to create redundant information.

From the related work analysis it is evident that multi-tier and video-aware FEC mechanisms are required to enhance the video quality without increasing the network overhead, and thus minimizing the usage of network resources and saving energy.

III. QOE-AWARE FEC MECHANISM FOR MULTI-TIER WMSNS

This section introduces the network architecture used by the QoE-aware FEC mechanism for WMSNs, which creates redundant packets based on their impact of the user perception. The proposal keeps videos with high quality, and decreases the number of redundant packets. Thus, it minimizes the usage of scarce network resources and saves energy.

A. Network Architecture

It has been shown by many works, e.g., [2] [15], that a multi-tier architecture offers considerable advantages with respect to a single-tier architecture in terms of less energy consumption, better scalability, lower loss, higher functionality, and better reliability. Unlike the work in [1], we propose a multi-tier WMSN architecture of only two tiers, because for the intrusion detection application, a third tier of high resolution cameras is not really necessary to detect intruders. In the lower tier, scalar sensors perform simple tasks, such as detecting scalar physical measurements. Resource-rich camera sensors in the higher tier are responsible for complex tasks. The architecture is shown in Figure 1.

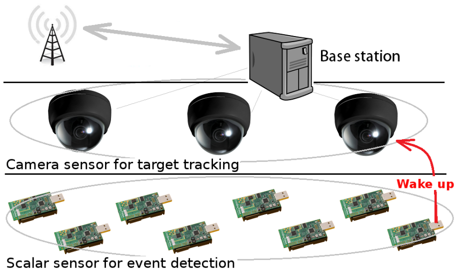


Fig. 1: Two-tier architecture

The scalar sensor nodes have a sensor range of a disk, which means that they can sense scalar physical measurements in an omnidirectional way. On the other hand, the sensing range of a camera node is called of Field of View (FoV). FoV is defined as a triangle, which depends of the direction (V), angle of view (α) and a depth of view (d), as shown in Figure 2. Thus, the sensing range of a camera node is limited, and depends on the direction of the camera and its features for angle and depth of view.

In this way, we design a multi-tier video intrusion detection system. The scalar sensor nodes perform intrusion detection, e.g. using vibration sensors. The camera sensors can be woken up on-demand to retrieve real-time video of the intruder that has been detected previously by the lower tier (scalar sensor), and send the video stream to the Base Station (BS).

When the camera node receives the wake-up message from a scalar sensor, it should change the direction of FoV to the location of the scalar sensor node, as shown in Figure 2. Then, it is possible to retrieve video from the target area. Video flows of an intruder provide users and authorities (e.g., police) with more precise information and allow them to decide a suitable action. The transmitted video will be useful to monitor and detect the intruder, and predict the intruder's moving direction.

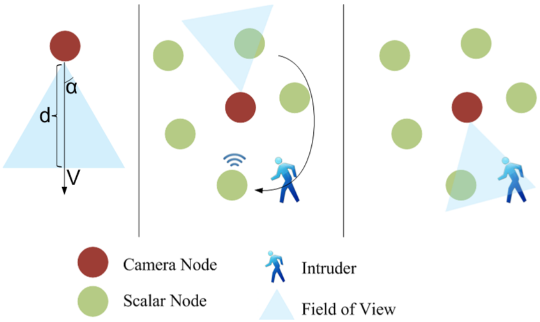


Fig. 2: Turnable FoV

In order to efficiently transmit video packets under certain application level QoE requirements, this paper extends the smart Multi-hop hierarchical routing protocol for Efficient VIdeo communication over WMSNs (MEVI) [16]. MEVI can provide a communication architecture suitable for the application scenario that we described above. MEVI proposes multi-hop communication with a cross-layer mechanism to select routes based on link quality, remaining energy and hop count.

As described above, the intrusion detection application requires scalar data transmission only when some events are detected. Then, the scalar node will wake up the camera node. However, MEVI was designed for continuous scalar data transmission and event multimedia data transmission. Therefore, MEVI was extended to provide the described data reports for intrusion detection applications. In the following, we will explain the extensions of MEVI.

Periodically, the setup phase will be performed during which camera nodes discover the routes to reach the BS. The paths are composed only of others camera nodes. The routes are scored based on the same cross-layer mechanism as proposed by MEVI.

Additionally, during the setup phase, each scalar sensor node selects its leader, a camera node. When an event has been detected, the scalar node has to wake up a camera node. Then, the camera node will change the direction of its FoV to retrieve and send multimedia content to the BS. The camera node is selected based on the higher value of the Link Quality Indicator of the Hello message, which is periodically sent from the camera node during the setup phase.

To enable the camera node to change the direction of its FoV to the target area, some initialization work is needed during the setup phase. In the first setup phase, the scalar sensor nodes should broadcast their location information to neighborhood nodes. The camera nodes receive and store this information. When the camera node receives a wake up message from a scalar sensor node, it is possible to change the direction of its FoV to the location of the scalar sensor node that reports the detection of an intruder.

B. Multimedia Content Characteristics

A compressed video is composed of three types of frames (I, P, B-frames), where: (i) Intra, or I-frames, are the reference for all the other frames that provide a reference point for decoding a received video stream; (ii) Predictive-coded, or P-frames, provide an increased rate of compression compared to I-frames; and (iii) Bi-directionally predictive-coded, or B-frames, use the previous and next I-frame or P-frame as their reference points for motion compensation [3].

The frame sequence that depends on an I-frame is called Group of Pictures (GOP). A GOP length of 10 frames means that the GOP starts with an I-frame, followed by 9 P or B frames.

A compressed video will apply spatial and temporal compression. The spatial compression is applied for I-, P- and B-frames. On the other hand, the temporal compression is applied only for P- and B-frames. The main consequence is that the loss of an I-frame will affect the other B- or P-frames of the same GOP. Thus, the errors will be propagated by other frames until a new I-frame reaches the receiver, i.e., the error will be propagated within the whole GOP.

For the case of loss of a P-frame, the error will be propagated by the remaining P- and B-frames in a GOP. Finally, if a loss of a B-frame happens, the error will not be propagated, since the B-frames are not used as a reference for other frames.

C. *QoE-aware FEC Mechanism for WMSNs*

The QoE-aware FEC mechanism considers the frame importance and its impact from the user point-of-view to create redundancy. The loss of I-frames causes more distortion than the loss of P- and B-frames from the user perspective. Additionally, the loss of P-frames at the beginning of a GOP causes more video distortion than P-frames at the end of a GOP.

In this context, I-frames need redundancy, since their loss will cause more video distortion. On the other hand, B-frames do not need redundant packets, since they are not used as a reference for other frames. Finally, the redundancy of P-frames can be applied based on their position within the GOP from a video application.

Figure 3 depicts how the proposed QoE-aware FEC mechanism is applied to a set of source multimedia packets that will be transmitted through a wireless channel to a destination node. Due to the low computational complexity as expected for WMSNs, Reed-Solomon (RS) coding was used to create redundant packets.

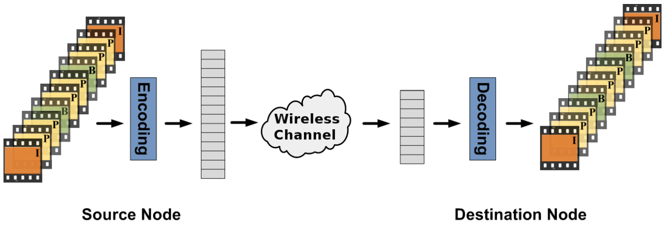


Fig. 3: QoE-aware FEC Mechanism

N given original packets are encoded into a group of *k* coded packets, and denoted as RS(k, n). Thus, k - n redundant packets were created, where this difference indicates the amount of redundancy (*r*). The destination node can reconstruct the *n* original packets by receiving any *n* out of *k* packets ($k > n$).

According to Algorithm 1, the proposed QoE-aware FEC mechanism should have a default value of the amount of redundancy (*r*), the GOP size (*s*) and the position (*p*) of each frame in the GOP (line 1 to 3).

Each video frame is divided into *n* original packets according to the fragment size (line 4). Then, redundant packets are created for the *n* original packets, depending on the type of the frame, and the location of the frame inside the GOP, as described in Algorithm 1.

Algorithm 1 QoE-aware FEC Mechanism

```
    Incoming Frame
 1: Let r = redundancy
 2: Let s = GOP size
 3: Let p = frame position inside a GOP
 4: Let n = number of packets after frame fragmentation
 5: if FrameType = I then
 6:     k ← ComputeRedundancy (n, r); //using RS coding, RS(k, n)
 7:     SendWithFEC(k, n)
 8: end if
 9: if FrameType = P and p < s/2 then
10:     k ← ComputeRedundancy (n, r); //using RS coding, RS(k, n)
11:     SendWithFEC(k, n)
12: end if
13: if FrameType = P and p > s/2 then
14:     SendWithoutFEC(n)
15: end if
16: if FrameType = B then
17:     SendWithoutFEC(n)
18: end if
```

I-frames and the first 50% of P-frames of each GOP will have their packets encoded using redundancy *r* (lines 5 to 12). These configuration values will be explained in Section IV-C. As described above, loss of these packets will cause high video distortion. Thus, their transmission needs redundancy to enable the recovery of the lost packets.

On the other hand, the last 50% of P-frames and B-frames are sent without redundancy (line 13 to 18), due to the fact that the loss of these packets are not sensitive from the user perspective.

IV. PERFORMANCE EVALUATION

A. *Simulation Scenario*

Simulation experiments were conducted to analyze the performance of FEC by using the Wireless Simulation Environment for Multimedia Networks (WiSE-MNet) [17]. WiSE-MNet is a simulator for WMSNs based on OMNeT++[18]/Castalia[19], a widely used network simulator for WSNs.

WiSE-MNet incorporates some functionalities from Castalia, and it provides a generic network-oriented simulation environment that addresses the need for co-design of network protocols and distributed algorithms for WMSNs.

Additionally, Wvsnmodel [20] was used, which proposes a simulation model for video sensor networks. Wvsnmodel efficiently defines a model to find subsets of nodes that cover the FoV area of a given node (denoted as cover set), and defines the sensing range by a FoV and not by a disk as usually done in sensor networks.

However, WiSE-MNet and Wvsnmodel do not enable multimedia transmission and evaluation, which are supported by Evalvid [21]. Due to the holistic property of WiSE-MNet, we extended it by integrating the functionalities of Wvsnmodel and Evalvid and take it as the basis of our work. Additionally, we implemented the change of the direction of the camera's FoV to make the camera able to retrieve multimedia content from the target area.

B. Parameters and Metrics

Simulations were carried out and repeated 20 times with different random seed numbers in order to have a confidence interval of 95%. Table I shows the simulation parameters used for a real-time video intrusion detection application.

TABLE I: Simulation parameters

Parameter	Value
Field Size	40x40
Location of Base Station	20, 0
Initial location of intruder	0, 0
Type of movement of intruder	Random mobility
Intruder velocity	1.5
Total number of Nodes	100
Number of Multimedia nodes	25
Location of Multimedia nodes	Grid
Location of Scalar sensor nodes	Uniform
Initial Energy for scalar nodes	14 J
Transmission Power	-15 dbm
Video sequence	Hall
Video Encoding	H.264
Format	QCIF (176 x 144)
Frame Rate	26 fps
Redundancy (r)	80 %, 100 %

For this paper, the Hall video sequence was chosen from the list of Video Trace Library. Hall is the video with similar motion and complexity as expected for the video intrusion detection application. Additionally, the video uses the QCIF format, since it is more suitable for WMSNs as shown in [22].

Scalar sensor nodes are assumed to be static, as expected for a real intrusion detection scenario. Furthermore, the camera sensor node parameters were set based on TelosB [23] equipped with a CMUcam3 [24]. Scalar sensor node parameters are based on TelosB.

The quality of the transmitted videos were evaluated using QoE metrics, which have an important role to measure the quality level of multimedia content based on the users' perspective [25]. Several objective QoE metrics have been formulated to estimate/predict (based on mathematical models) the quality level of multimedia content. The objective metrics used are: Structural Similarity (SSIM) and Video Quality Metric (VQM).

SSIM is a measurement of the structural distortion of the video, which tries to obtain a better correlation with the user's subjective impression. SSIM has values ranging from 0 to 1, a higher value means better video quality.

VQM measures the "perception damage" the video experienced, based on the human visual system characteristics, including distinct metric factors such as blurring, noise, color distortion and distortion blocks. A value close to 0 means a video with better quality.

C. Results

In order to evaluate the proposed multi-tier architecture and the QoE-aware FEC mechanism, the extended WiSE-MNet framework was used. However, we present only the results of video transmissions, since [2] demonstrated that multi-tier architecture can achieve energy reduction when compared to a single-tier.

The results in this section show the average value of network overhead, and the video quality according to the well-known QoE metrics. The SSIM and VQM values are obtained by using the MSU Video Quality Measurement Tool (VQMT). During the simulations, Hall video sequence was transmitted several times by each camera node and the video quality metrics presented here are the average of SSIM and VQM for all transmitted Hall video sequence of each camera node.

Simulations were carried out to find the best configuration values for the proposed FEC (QoE-aware FEC), as explained in the following. Then, we simulated a simple FEC approach, i.e. the RS coding with 80% and 100% of redundancy for all frames (standard FEC). An additional experiment was performed without any FEC mechanism (without FEC).

First, we analyze our approach (QoE-aware FEC mechanism) with different values for the redundancy of P-frames, which depend on the positions of P-frames inside the GOP. It is important to highlight that I-frames and B-frames are sent as proposed in Algorithm 1. The possible scenarios for sending P-frames are: (i) encode the first *x*% of P-frames using redundancy *r*, as defined in Section III-C, *x* is 60 or 50; (ii) encode the last *y*% of P-frames with *r/2*, where x + y = 100%; (iii) send the last *y*% of P-frames without redundancy. Table II shows these 8 possible scenarios and the number of redundant packets for r = 80% or r = 100%.

TABLE II: Possible scenarios for QoE-aware FEC mechanism

	Redundancy for the first 60 % of P-frames	Redundancy for the last 40 % of P-frames	Network Overhead
Scenario 1	80%	0%	96 packets
Scenario 2	80%	40%	152 packets
Scenario 3	100%	0%	159 packets
Scenario 4	100%	50%	215 packets
	Redundancy for the first 50 % of P-frames	Redundancy for the last 50 % of P-frames	Network Overhead
Scenario 5	80%	0%	84 packets
Scenario 6	80%	40%	152 packets
Scenario 7	100%	0%	145 packets
Scenario 8	100%	50%	213 packets

For these 8 scenarios, we analyzed the average value of SSIM for the distance between the source camera node and the BS, as shown in Figure 4. For space limitations, VQM results are not presented.

A further distant node suffers a higher packet loss, due to the fact that more hops are needed to reach the BS. Additionally, nodes with similar distances can select paths with different numbers of hops. These factors can cause more packet loss due to inference or buffer overflow, and explain why for some similar distances the values of SSIM can be below 0.05.

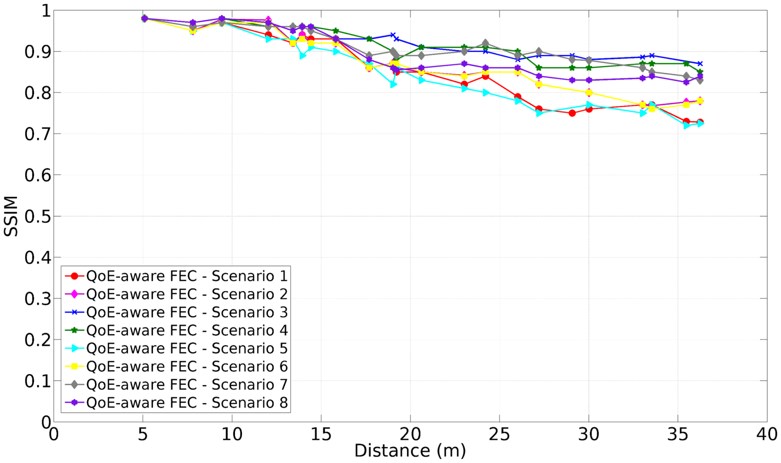


Fig. 4: SSIM according to distance for different scenarios of QoE-aware FEC mechanism

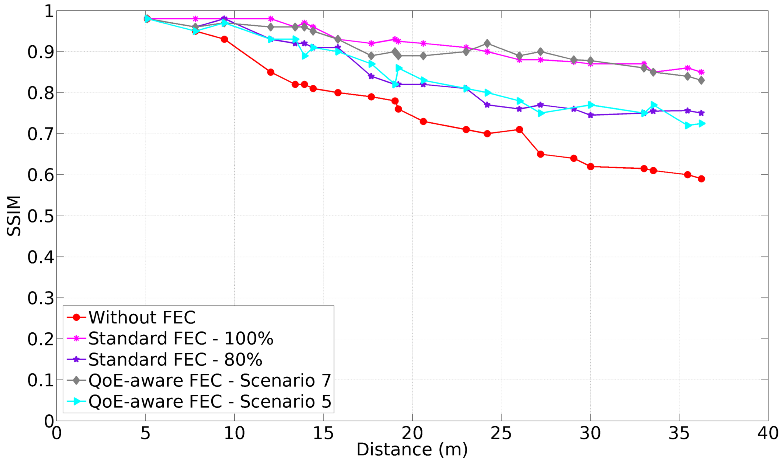


Fig. 5: SSIM according to distance from the camera node to BS

Figure 4 shows the SSIM results of different scenarios for the QoE-aware FEC mechanism with redundancy of 80% (scenarios 1, 2, 5 and 6). They can be divided into two groups. The first group sends the last P-frames without redundancy, as in scenarios 1 and 5. Scenario 1 improves video quality (in terms of SSIM) by less than 0.06 for some distances, compared with scenario 5. This is because it sent the first 60% of P-frames with redundancy. In contrast, in scenario 5 only the first 50% of P-frames are sent with redundancy.

The second group applies a redundancy of *r/2* for the last P-frames (scenarios 2 and 6). Scenarios 2 and 6 have almost the same video quality, because they create the same number of redundant packets.

Finally, by comparing these 4 scenarios, it is possible to conclude that scenarios 2 and 6 improve video quality (in terms of SSIM value) by less than 0.09 for some distances. The main difference between scenarios 2 and 6 against 1 and 5 is the fact that they (2 and 6) sent the last P-frames with redundancy of *r/2*, while in scenarios 1 and 5 there is no redundancy for P-frames. Therefore, with more redundant packets, which are used to reconstruct the lost packets, scenarios 2 and 6 achieve better results.

We now analyze the QoE-aware FEC mechanism using 100% of redundancy (scenarios 3, 4, 7 and 8). It is possible to group them in a similar way. The first group (scenarios 3 and 7) sent the last P-frames without redundancy. The second group (scenarios 4 and 8) sent the last P-frames with redundancy *r/2*. Scenarios 3 and 4 improve the video quality compared with scenarios 7 and 8, respectively. The same explanation as before applies here.

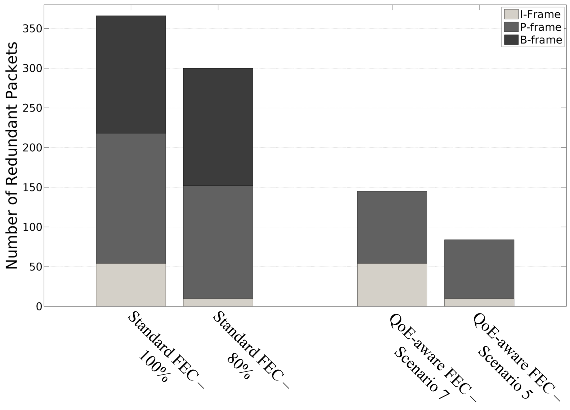


Fig. 6: Network overhead

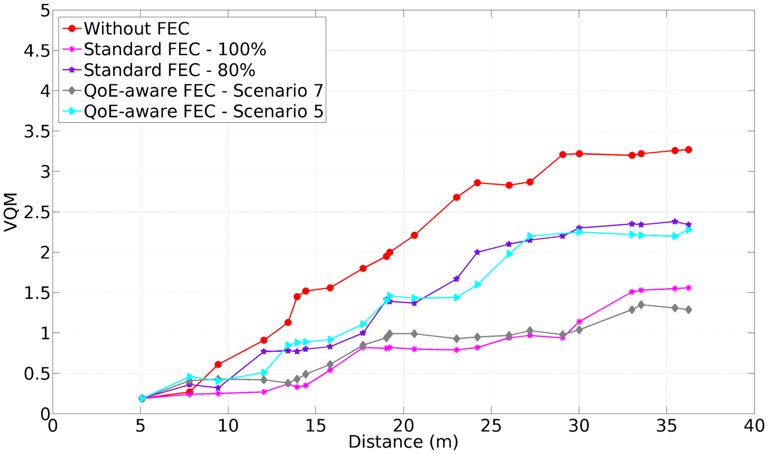
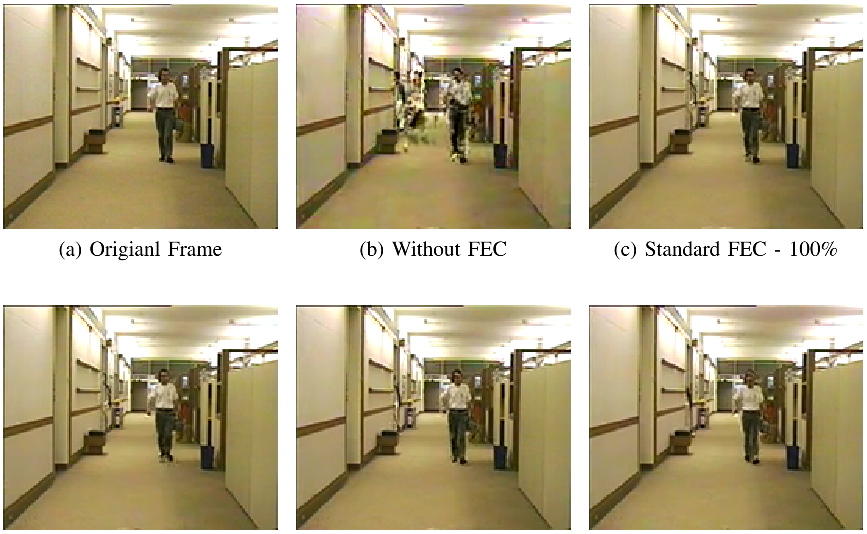


Fig. 7: VQM according to distance from the camera node to BS



(a) Origianl Frame (b) Without FEC (c) Standard FEC - 100%

(d) QoE-aware FEC - Scenario 7 (e) Standard FEC - 80% (f) QoE-aware FEC - Scenario 5

Fig. 8: Frame 258 of transmitted video

For the intrusion detection application that requires video with low mobility and complexity, we believe that the best trade-off between video quality, network overhead and energy-efficiency is to use either scenarios 5 or 7, with redundancy of 80% and 100% respectively. The video quality provided by other scenarios does not have a significant difference. Due to the fact that they create more redundant packets, which can causes more interference, the buffer overflow increases and more energy is consumed.

The selected scenarios 5 and 7 sent the last 50% of P-frames without redundancy, since loss of these frames causes lower video distortion from the user perspective. The first 50% of P-frames are sent using redundancy *r*, because the loss of these frames will lead to higher video distortion.

Figure 5 shows the SSIM value for the distance between the source node and the BS for video transmission without FEC, with standard FEC, and with the proposed QoE-aware FEC mechanism. For the case of 80% redundancy, the standard and QoE-aware FEC mechanisms improve the SSIM by around 10% for distances below 25m. For distances above 25m, they improve the SSIM at least by 20%.

For almost all distances between source node and BS, the standard and QoE-aware FEC mechanisms have similar video quality. However, standard FEC includes a higher network overhead. Instead, the proposed QoE-aware FEC approach achieves a lower overhead, as shown in Figure 6. Since less transmission means less energy consumption, we can conclude that our proposal can provide good energy-efficiency, while keeping the transmitted video with a good quality from the user perspective. The QoE-aware FEC mechanism creates redundant packets based on frame importance and user experience while reducing the network overhead compared to standard FEC.

Using a redundancy of 100%, it is possible to increase the video quality by at least 12% for distances above 25m. However, it will include 20% of extra overhead compared to the case of 80% of redundancy, as shown in Figure 6.

Figure 7 presents the video quality level, using the VQM of the transmitted videos for distances between the source camera node and the BS. The VQM results demonstrate the benefits of using FEC, for both the standard and QoE-aware approaches. Using a redundancy of 80%, the FEC mechanisms kept the VQM values below 1 for distances of less than 30m, and around 1.5 for distances above. For redundancy of 100%, it is possible to improve the video quality by around 20% for distances below 20m. To show the impact of transmitting video streams using FEC from the standpoint of the end-user, a frame was randomly selected (frame 258) from the transmitted video, as displayed in Figure 8. The benefits of the FEC mechanisms are visible in the frames. Frame 258 is the moment when a man (intruder in our application) was walking in a corridor. Therefore, this frame can be used to predict the moving direction and detect the intruder.

By comparing each transmitted frame with the original frame, it is possible to see a higher distortion when the video is transmitted without using any FEC, as shown in Figure 8b. Transmitting the frame using FEC with redundancy of 80% has the result of lower distortion, as shown in Figures 8f and 8e. Finally, sending the frame with redundancy of 100%, as shown in Figures 8d and 8c, the frame has almost the same quality as the original one. From the user perspective, the QoE-aware FEC mechanism keeps the video with good quality with reduced overhead, and therefore saves network resources and energy.

V. CONCLUSIONS

In this paper, we proposed a QoE-aware FEC mechanism for multi-tier WMSNs and simulated it using the OMNeT++ simulator. The proposal has been designed for intrusion detection in the area of interest, where the on-line video transmission has the potential of enhancing the level of information collected. Video flows provide more precise information than simple scalar data. The proposed multi-tier energy-saving architecture consists of static scalar sensors and camera sensors. The camera sensors are only woken up by scalar sensors on demand. The QoE-aware FEC mechanism targets on reducing redundant packet transmission while keeping videos with an acceptable quality level. Simulation results showed that our approach could achieve a good video quality experience from the end users' perspective, with an energy-efficient performance and saving network resources.

VI. ACKNOWLEDGMENT

This work is supported by M3WSN, a joint research project of Nano-Tera.ch and the Sino Swiss Science and Technology Cooperation (SSSTC).

REFERENCES

[1] I. Almalkawi, M. Guerrero Zapata, J. Al-Karaki, and J. Morillo-Pozo, "Wireless Multimedia Sensor Networks: Current Trends and Future Directions," *Sensors*, vol. 10, no. 7, pp. 6662–6717, 2010.

[2] P. Kulkarni, D. Ganesan, P. Shenoy, and Q. Lu, "Senseye: a multi-tier camera sensor network," in *13th annual ACM international conference on Multimedia*. ACM, 2005, pp. 229–238.

[3] J. Greengrass, J. Evans, and A. Begen, "Not all packets are equal, part i: Streaming video coding and sla requirements," *IEEE Internet Computing*, vol. 13, no. 1, pp. 70–75, 2009.

[4] M. Y. Naderi, H. R. Rabiee, M. Khansari, and M. Salehi, "Error control for multimedia communications in wireless sensor networks: A comparative performance analysis," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1028 – 1042, 2012.

[5] R. Collins, A. Lipton, and T. Kanadeand, "A system for video surveillance and monitoring," in *8th International Topical Meeting on Robotics and Remote Systems American Nuclear Society (ANS)*, 1999.

[6] L. Jiao, Y. Wu, E. Chang, and Y. Wang, "The anatomy of a multi-camera security surveillance system," *ACM Multimedia System Journal Special Issue*, pp. 144–163, 2004.

[7] J. Wang, M. Masilela, and J. Liu, "Suporting video data in wireless sensor networks," in *9 th IEEE International Symposium on Multimedia (ISM)*, 2007, p. 310.

[8] I. Politis, M. Tsagkaropoulos, and S. Kotsopoulos, "Optimizing video transmission over wireless multimedia sensor networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2008, p. 1.

[9] S. Guo, , and T. Little, "Qos-enabled video streaming in wireless sensor networks," in *9th IEEE International Symposium on Network Computing and Applications (NCA)*, 2010, p. 214.

[10] T. Chen, P. Chen, W. Lee, and C. Huang, "Design and implementation of a real time video surveillance system with wireless sensor nwetworks," in *Vehicular Technology Conference. VTC Spring 2008*, 2008, p. 218.

[11] J. Jeong and C. Ee, "Forward error correction in sensor networks," *University of California at Berkeley*, 2003.

[12] P. Hurni and T. Braun, "Link-quality aware run-time adaptive forward error correction strategies in wireless sensor networks," IAM, University of Bern, IAM-11-003, Tech. Rep., 2011.

[13] P. Sarisaray, G. Gur, S. Baydere, and E. Harmanc, "Performance comparison of error compensation techniques with multipath transmission in wireless multimedia sensor networks," in *15th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*. IEEE, 2007, pp. 73–86.

[14] Y. Yang, Y. Chen, and W. Yi, "Cross-layer fec for reliable transfer of variable-length coded data in wmsn," in *IET International Conference on Wireless Sensor Network*. IET, 2010, pp. 380–384.

[15] P. Kulkarni, D. Ganesan, and P. Shenoy, "The case for multi–tier camera sensor networks," in *International workshop on Network and operating systems support for digital audio and video*. ACM, 2005, pp. 141–146.

[16] D. Rosário, R. M. Costa, H. Paraense, K. Machado, E. Cerqueira, and T. Braun, "A smart multi-hop hierarchical routing protocol for efficient video communication over wireless multimedia sensor network," in *2nd IEEE International Workshop on Smart Communication Protocols and Algorithms (ICC'12 WS - SCPA)*, Ottawa, Canada, Jun. 2012.

[17] C. Nastasi and A. Cavallaro, "Wise-mnet: an experimental environment for wireless multimedia sensor networks," *Proceedings of Sensor Signal Processing for Defence (SSPD)*, 2011.

[18] A. Varga, "The OMNeT++ discrete event simulation system," in *Proceedings of the European Simulation Multiconference (ESM 2001)*, 2001.

[19] A. Boulis, "Castalia, a simulator for wireless sensor networks and body area networks, version 2.2," August 2009.

[20] C. Pham and A. Makhoul, "Performance study of multiple cover-set strategies for mission-critical video surveillance with wireless video sensors," *IEEE WIMOB*, 2010.

[21] Evalvid, "Evalvid - a video quality evaluation tool-set," available at: http://www2.tkn.tu-berlin.de/research/evalvid/fw.html. Accessed at Jun 2012.

[22] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[23] Memsic, "Telosb datasheet," available at: http://memsic.com/support/documentation/wireless-sensor-networks/category/7-datasheets.html?download=152%3Atelosb. Accessed at Jun 2012.

[24] A. Rowe, C. Rosenberg, and I. Nourbakhsh, "A low cost embedded color vision system," in *International Conference on Intelligent Robots and Systems (IEEE/RSJ 2002)*, vol. 1. IEEE, 2002, pp. 208–213.

[25] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin, "An overview of quality of experience measurement challenges for video applications in ip networks," *Wired/Wireless Internet Communications*, pp. 252–263, 2010.