An OMNeT++ Framework to Evaluate Video Transmission in Mobile Wireless Multimedia Sensor Networks

Denis Rosário
Institute of Computer Science
and Applied Mathematics
University of Bern
Bern, Switzerland
rosario@iam.unibe.ch

Zhongliang Zhao
Institute of Computer Science
and Applied Mathematics
University of Bern
Bern, Switzerland
zhao@iam.unibe.ch

Claudio Silva
Federal University of Pará
Belém, PA, Brazil
claudio@ufpa.br

Eduardo Cerqueira
Federal University of Pará
Belém, PA, Brazil
cerqueira@ufpa.br

Torsten Braun
Institute of Computer Science
and Applied Mathematics
University of Bern
Bern, Switzerland
braun@iam.unibe.ch

ABSTRACT

The development and evaluation of new algorithms and protocols for Wireless Multimedia Sensor Networks (WMSNs) are usually supported by means of a discrete event network simulator, where OMNeT++ is one of the most important ones. However, experiments involving multimedia transmission, video flows with different characteristics, genres, group of pictures lengths, and coding techniques must be evaluated based also on Quality of Experience (QoE) metrics to reflect the user's perception. Such experiments require the evaluation of video-related information, i.e., frame type, received/lost, delay, jitter, decoding errors, as well as inter and intra-frame dependency of received/distorted videos. However, existing OMNeT++ frameworks for WMSNs do not support video transmissions with QoE-awareness, neither a large set of mobility traces to enable evaluations under different multimedia/mobile situations. In this paper, we propose a Mobile MultiMedia Wireless Sensor Network OMNeT++ framework (M3WSN) to support transmission, control and evaluation of real video sequences in mobile WMSNs.

Categories and Subject Descriptors

I.5 [**Simulation of communication networks**]: Miscellaneous; I.7 [**Integration with other simulation tools**]: Miscellaneous

General Terms

Design, Simulations

Keywords

Mobile Multimedia Simulation, Node Mobility, Wireless Multimedia Sensor Networks

1. INTRODUCTION

With the rapid development of low-cost technologies involving camera sensors and scalar sensors, Wireless Multimedia Sensor Networks (WMSNs) [2] have attracted the interest of researchers, citizens and industries due to their wide scope of potential applications in both military and civilian areas. As an extension of traditional scalar Wireless Sensor Networks (WSNs), WMSNs are composed of wireless interconnected camera sensor nodes equipped with multimedia devices, e.g., a camera and/or microphone. The camera nodes are capable of retrieving video/audio streams, still images, as well as scalar sensor data.

WMSNs promise a wide scope of Internet of Things (IoT) or smart cities applications that require visual and audio information, such as multimedia surveillance, traffic monitoring, personal health care, and environmental monitoring. The multimedia data in such applications enable the end-user (or system) to visually determine the real impact of an event and being aware of what is happening in the environment. Users can also detect the objects or intruders, and/or analyse scenes based on visual information [12].

Additionally, the advances in mobile communications are enhancing WMSN scenarios with mobility support for objects, sensor nodes or both. The objects that have to be monitored (e.g., cars, people or animals) are naturally mobile. These objects might have sensors, and, thus, the sensors might become mobile. For example, the sensors might be integrated into mobile personal devices, such as smartphones or cars. Furthermore, the objects or special mobile devices, such as robots or unmanned aerial vehicles, might have Base Stations (BS) responsible for receiving multimedia data.

In the context of IoT/smart cities applications, several algorithms and protocols have been developed for fixed and mobile WMSNs [2, 12]. However, due to the mobility and multimedia-awareness of future WMSNs, it is necessary to provide an in-depth analysis of protocols, algorithms and applications that can lead to optimizations that can also be based on video-related characteristics and impairments. Thus, an event-driven simulation is required to allow the evaluation of different parameters before the real deployment in a more comprehensive manner (reducing cost, time and human resources).

In terms of performance evaluation, solutions involving multimedia transmission and management must evaluate the video content from the user's perspective, by using Quality of Experience (QoE) metrics, such as Structural Similarity (SSIM) and Mean Opinion Score (MOS). This is due to the fact that Quality of Service (QoS) metrics do not reflect the user's perception in the same way as the QoE metrics do, and, hence, the QoS metrics fail to reflect the subjective factors associated with human's experience [13, 16]. However, the current network simulators and frameworks for WMSNs do not support the QoE and video-aware transmission natively. Additionally, existing simulator frameworks do not provide a large set of mobility models to enable more complex mobile simulations as expected for many smart cities applications, such as car-based video surveillance schemes.

In this paper, we propose an OMNeT++ framework, called Mobile Multi-Media Wireless Sensor Network (M3WSN) [14]. The proposed framework is based on OMNeT++ and Castalia [4] to support multimedia transmission for fixed and mobile scenarios with QoE-awareness. There are initial works to enable WMSN simulations using OMNeT++, such as WiSE-MNet [10] and WVSN [11], which are both specific frameworks based on Castalia. Castalia is a well-known framework for WSN simulation and provides detailed implementation of realistic wireless channel and radio models. Thus, we decide to choose Castalia as our development platform, and integrate the functionalities of WiSE-MNet and WVSN due to their importance for WMSN experiments. Moreover, the M3WSN framework implements full support for delivering, control, and evaluating real video sequences in fixed and mobile scenarios with the aid of mobile traces. We believe that a new set of mobile multimedia-based solutions can be evaluated and optimized with the aim of M3WSN.

The remainder of this paper is structured as follows. In Section 2 provides the state of the art analysis for the network simulators/frameworks, and also discuss their main problems with regard to multimedia transmission and mobility models. Section 3 outlines the proposed M3WSN framework. Simulations were carried out to show the benefits of using the M3WSN framework, and are described in Section 4. The paper concludes with Section 5, which summarizes the main contributions and results of this paper.

2. RELATED WORK

OMNeT++ [17] is a standard and general-purpose tool employed to study protocols in wired and wireless networks. Additionally, OMNeT++ has many frameworks to support mobility, wired/wireless, standard/non-standard protocols, and energy models. The main frameworks for OMNeT++ include INET [6], INETMANET [7], MiXiM [9], and Castalia.

INET is an open source communication network simulation package for OMNeT++ environment. It contains models for several wired and wireless network protocols, including UDP, TCP, IP, and many others. On the other hand, INETMANET includes additional protocols and components to simulate ad hoc wireless networks. It contains multiple radio wave propagation models, simple battery models, and a good support for node mobility. Moreover, INETMANET has many protocols for the MAC layer and network layer that are already implemented.

Castalia includes advanced wireless channel and radio models, a physical process (e.g., for event detection). A sensing model, CPU drift clock and power consumption model, as well as MAC and routing protocols, including IEEE 802.15.4 are also implemented. However, Castalia does not provide any advanced functionality for video transmission, control and evaluation as expected for emerging multimedia applications, due to the fact that it was originally designed for WSNs simulation. Furthermore, Castalia was designed for scalar sensor network simulation and it only includes a very basic model for mobility, i.e., linear mobility.

Wireless Simulation Environment for Multimedia Networks (WiSE-MNet) [10] incorporates some of Castalia's functionalities/features to provide a generic network-oriented simulation environment for WMSNs. WiSE-MNet addresses the need for co-designing network protocols and distributed algorithms for WMSNs. Even though designed for WMSNs, WiSE-MNet does not provide video control and QoE support, which is a key characteristic to enable multimedia evaluation from the user's perspective. Additionally, it considers an idealistic communication mechanism to test algorithms without take into account the unreliable nature of wireless medium. Moreover, WiSE-MNet does not support node mobility with complex traces as expected in many smart cities applications.

The Wireless Video Sensor Network (WVSN) model proposes a simulation model for video sensor networks. It defines the sensing range of camera nodes by a Field of View (FoV), which is more realistic for WMSNs. Additionally, depending on the number of nodes, the model determines the cover-sets for each sensor node and computes the percentage of coverage for each cover-set. Then, this information is used to increase the frame capture rate, e.g., a node has a higher capture rate when it has more covers. However, this work also fails in providing an accuracy video transmission and evaluation approach, and no mobility is supported.

In this context, it is clear that existing OMNeT++ frameworks have no support for transmission, control and evaluation of real video sequences as required for many WMSNs and smart cities scenarios. Hence, it is required a QoE-aware and video-related framework to manage video flows with different characteristics, types, GoP lengths, and coding, while collecting information on the type of each frame received/lost, frame delay, jitter and decoding errors, as well as inter and intra-frame dependency of the received/distorted videos are required for emerging IoT/WMSNs. Thus, a set of mobile multimedia-based solutions can be evaluated and improved with the aim of a new framework.

3. THE M3WSN FRAMEWORK

This section details the proposed Mobile MultiMedia Wireless Sensor Network OMNeT++ framework (M3WSN). The proposed framework integrates functionalities of WiSE-MNet and WVSN model, and also implements new functionalities to provide mobile multimedia-aware management. M3WSN enables to analyse the impact and benefits of novel video-aware algorithms and protocols for fixed or mobile MWNS scenarios. The M3WSN framework is publicly available as open-source at [14], together with all relevant information and supporting documentation (e.g. installation and user guides).

3.1 Castalia and WiSE-Mnet frameworks

Castalia is a framework designed to model distributed algorithms for classic WSNs under realistic communication conditions. The overall architecture of Castalia is composed of a wireless channel, physical process, and nodes, as shown in Figure 1. The wireless channel simulates the behaviour of wireless link, and interconnects different nodes. The physical process also interconnects nodes, and models a sensed-data generation to feed the sensor manager of different nodes with data, which corresponds to a real process with spatial correlation of data and variability over time. Different modules compose a node, i.e., communication and application, sensor, resource, and mobility manager.

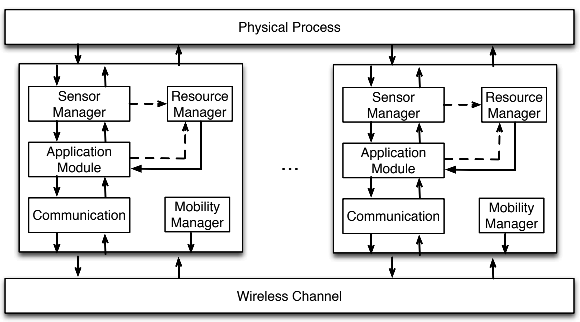


Figure 1: Node architecture of Castalia

From a network layer's standpoint, the application module is equivalent to the application layer of the Internet architecture. The communication module uses a simplified network stack composed of three layers, namely radio (physical layer definition), MAC and routing. The sensor manager is responsible for providing the application module with new samples from physical processes. The mobility manager is responsible to enable and control mobility within the simulation area. The resource manager models the local usage of resource, such as, energy consumption, memory usage, and CPU states.

In contrast to Castalia, WiSE-MNet supports complex data types, from scalar-based to any data type. At the physical process module, WiSE-MNet proposes the use of moving objects on a ground plane with different types of motion, e.g. linear, circular, and random. Unlike existing works that represent object as point, the moving object in WiSE-MNet is represented as a box to enable object detection. In other words, the moving object has a defined area (bounding box) that can be used by the camera node to detect if the object is within its sensing range. At the sensor manager module, the cameras are assumed to be top-down facing to obtain a simplified projection model. A distributed tracking algorithm is implemented at application module. Finally, WiSE-MNet also includes a simple GUI for 2D world representation.

3.2 The Wireless Video Sensor Network model

The WVSN model [11] defines the sensing range of a camera node by a FoV. Unlike to that, wireless sensor nodes have a sensing range defined by a disk (right side of Figure 2), and, thus, they can sense scalar physical measurements in an omni-directional way. On the other hand, FoV is defined as triangle, as shown on the left side of Figure 2. The FoV depends on the direction of the camera (V), angle of view (α) and depth of view (d). Thus, the sensing range of a camera node is limited, and depends on the direction of the camera and its features of angle and depth of view.

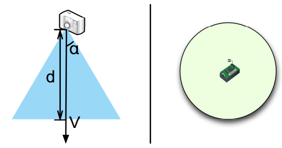


Figure 2: Sensing range for camera/scalar nodes

The WVSN model introduces the notion of cover-set. Where, depending on the number of nodes, the model defines the cover-sets. For example, a node x has a sensing range, and the cover-set defines a set of nodes that can cover (sensing) the same area of the node x. The idea is that when a node has several nodes covering its sensing range, the node can increase its frame capture rate. This is due to if it runs out of energy, it can be replaced by one of its covers. Additionally, the model defines the cover-sets for sensing range defined by a FoV and not by a circle to improve the accuracy for WMSNs scenarios. The frame captures rate increases depending on the application criticality level. Moreover, the WVSN model does not aim to improve the coverage by rotating the camera to the area of interest or including node mobility.

3.3 M3WSN Architecture

The M3WSN framework integrates the functionalities of the WiSE-MNet and WVSN models due to their initial efforts to model WMSNs. The functionalities include moving objects, object detection, FoV, cover-set and application criticality. Figure 3 depicts the M3WSN framework architecture, including where we integrate the WiSE-MNet and WVSN functionalities, and also the new functionalities.

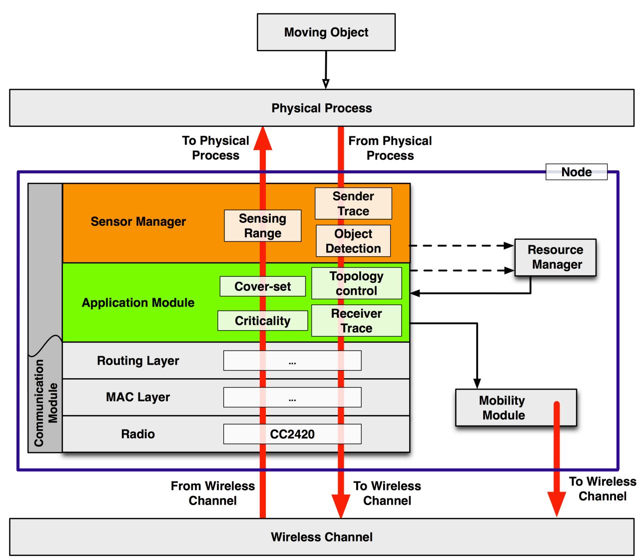


Figure 3: The M3WSN Framework Architecture

In many WMSN applications, such as intrusion detection, there is a need to simulate the behaviour of moving objects (intruder). To this end, we implemented the moving object with bound box (area used to detect the object) at the physical process to simulate a realistic mobile intruder/object and to enable intrusion detection. Additionally, there are available different types of motion for mobile objects, e.g. linear, circular, random, and others.

The M3WSN framework implements the sensing range models at the sensor manager module. For scalar sensor nodes, we include a disk model to enable the nodes to detect the intruder in an omni-directional way. For example, scalar node detects the presence of mobile objects by detecting the object bound box within its sensing range. On the other hand, we implemented the FoV model to represent a camera node with a realistic sensing range model. At the application module, we included cover-set to improve the simulation performance for some specific applications.

3.4 Multimedia management

Applications involving multimedia transmission must evaluate the video quality level from the user's perspective, and also collecting video-related characteristics. In specific terms, frames with different priorities (I, P and B) compose a compressed video, and the loss of high priority frames causes severe video distortion from human's experience. For the loss of an I-frame, the errors propagate through the rest of the Group of Picture (GoP), because the decoder uses the I-frame as the reference frame for all other frames within a GoP. When this occurs, the video quality recovers only when the decoder receives an unimpaired I-frame. Second, for the loss of a P-frame, the impairments extend through the remaining of the GoP. Third, the loss of a B-frame affects the video quality only of that particular frame.

In this context, multimedia flows enable the end-users (systems) to visually determine the real impact of a detected event, perform object/intruder detection, and analyse the scenes based on collected visual information. However, Castalia and its extensions (both WiSE-MNet and WVSN models) do not enable the transmission, control and evaluation of real video sequences. Therefore, we ported Evalvid [5] to M3WSN, because Evalvid provides video-related information, such as frame type, received/lost, delay, jitter, and decoding errors, as well as inter and intra-frame dependency of the received/distorted videos. This video-related information enables the creation of new assessment and optimization solutions for fixed and mobile WMSN scenarios.

Evalvid is a framework for video transmission and quality evaluation. Thus, before transmitting a real video sequence, we need a video source, for example from a video library [8] or the user can create a new one. Once the video has been encoded, trace files have to be produced. The trace files contain all relevant information for transmission, and the evaluation tools provide routines to read and write these trace files for multimedia evaluation.

Specifically, there are three kinds of trace files. Two of them are created at the source side, namely video and sender traces. On the other hand, the destination node creates the receiver traces. Figure 4 illustrates the steps to generate the trace files, and reconstruct the transmitted video based on these trace files and original video. More information on how to create these trace files can be found in [5].

The video trace is created once, and contains all relevant information about every frame that composes the video. This information includes frame number, type and size; number of segments in case of (optional) frame segmentation; and the time to transmit each frame. The source node has to create a sender trace file for every video transmission, based on information from the video trace file. The sender trace is generated with the help of the input routines and data structures provided by Evalvid, which contains information about every packet generated before the transmission. The information consists of time stamp, the packet id and the packet size. These two trace files together represent a complete video transmission (at the sender side) and contain all the information needed for further evaluations. The destination node creates a receiver trace file for every received video, which is created with the help of the output routines of Evalvid. The receiver trace also contains the time stamp, packet id, and payload size.

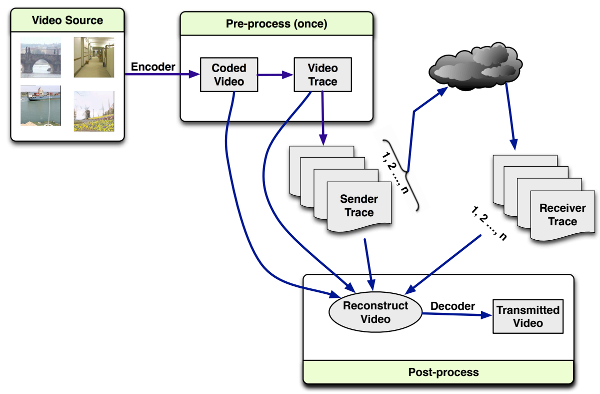


Figure 4: Video sequence transmission

We implemented the creation of sender traces at the sensor manager module, because it supports a camera retrieving a video. On the other hand, the receiver trace is created at the application layer module, due to it represents the application layer receiving multimedia packets and reconstructing the video. Moreover, the user can define the energy consumption rate for retrieving each frame, and this value could be chosen for a real camera.

Regarding multimedia evaluation from a user's perspective, there are basically two QoE approaches: objective and subjective (hybrid methods will not be detailed in this paper). Objective approaches estimate/predict the video quality level by means of mathematical models or signal processing algorithms. The main objective metrics are the following: Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM), and Video Quality Metric (VQM). Despite objective metrics easily evaluate the video quality level, they fail in capturing all the details that might affect user's experience. In this context, subjective evaluations are required. The Mean Opinion Score (MOS) is one of the frequently used approaches for subjective video evaluation, which consists of human observers rating the overall video quality. For MOS evaluation, the viewers watch only once a video and then rate the video quality using the following scale: Bad; Poor; Fair; Good; and Excellent [1].

In this context, Evalvid provides evaluation traces, which contain packet and frame losses, as well as delay/jitter. In specific terms, Evalvid needs the sender, receiver, and video source traces to generate the evaluation traces. On the other hand, the video quality assessment is performed frame by frame. Thus, we need to reconstruct the transmitted video, which is achieved with the aid of sender, receiver, and video source traces, as well as the original encoded video. Thus, it is possible to measure the PSNR by using Evalvid. Additionally, we can measure SSIM or VQM by using external tools. As soon as we have the original and transmitted videos, we can perform subjective evaluations following the International Telecommunication Union (ITU-T) recommendations. Detailed information on how to produce these results can be found in [5].

3.5 Mobility traces

Mobility is one of the most challenging issues in WMSNs. To understand how the network behaves under different mobile situations, the node mobility has to be simulated in a reasonable way. In this context, BonnMotion [3] is a simulator-independent tool to generate mobility traces for different mobility models. BonnMotion provides several mobility models, such as the Random Waypoint model, the Gauss-Markov model, the Reference Point Group Mobility model, and others. The generated traces can be exported to NS-2, QualNet, MiXiM, and other compatible simulator. However, the current implementation of BonnMotion is unable to generate traces that could be used by Castalia, WiSE-MNet and WVSN frameworks. However, node mobility is not the main focus of the current Castalia framework, which has only support for static and linear mobility models.

According to M3WSN architecture, the mobility manager is responsible to enable and control node mobility within the simulation area. In this context, we implemented BonnMotion at the mobility manager module to fully support different mobility models. Additionally, the M3WSN framework enables the user to configure the energy consumption for a node to move to a certain distance to make the simulation more realistic.

Mobile WMSNs consist of geographically dispersed nodes with a lack of a central infrastructure, which implies that the network does not have a fixed topology. In this context, topology control is important for enabling nodes to be aware of topology changes and, thus, determining an appropriate topology to keep the connectivity. To this end, the topology control enables the mobile nodes to move with a certain formation to keep their connectivity with each other [19].

In contrast to Castalia architecture that does not allow the mobility manager and application module to communicate with each other, the M3WSN framework extended the node architecture, as outlined in Figure 3. In specific terms, M3WSN enables the application module send messages to mobility model. Then, the node can change the movement/direction based on topology control information.

4. EVALUATION

In this section, we introduce a use case that makes use of M3WSN framework to obtain key video-related information, such as frame type and GoP length for creating new assessment and optimization solutions. Additionally, the proposed use case shows the importance to evaluate the transmitted video sequences from the user's perspective. This use case scenario is expected for MWSNs/smart cities applications.

4.1 Scenario description

We implemented a QoE-aware FEC (Forward Error Correction) mechanism for intrusion detection in multi-tier WMSNs, which was proposed by [20]. The multi-tier architecture is used due to its considerable advantages in terms of reduced energy consumption, better scalability and reliability, lower loss, and higher functionality compared to single-tier architecture. The constraints of sensor nodes increase wireless channel errors, and application-level FEC can be employed as an error control scheme for handling losses in WMSN communications. FEC schemes achieve robust video transmission by sending redundant packets. In case of packet loss, the original frame can be recovered from the redundant information.

The QoE-aware FEC mechanism creates redundant packets based on frame importance from user's experience, and thus reduces the packet overhead, while keeping the video with a good quality. The QoE-aware FEC mechanism creates redundant packets for I-frames and the first 50% of P-frames. On the other hand, B-frames and the last 50% of the P-frames are transmitted without create redundant packets. We defined the packet redundancy of 80%.

On the basis of the multi-tier intrusion detection scenario described above, simulations were carried out to evaluate the transmitted video from the user's perspective by using QoE-aware FEC. Following this, we simulated a simple FEC approach, i.e. creating redundancy for all the frames (simple-FEC), and also without any FEC mechanism (no-FEC). The simulations were carried out and repeated 20 times with different seeds for random number generation to provide a confidence interval of 95% (vertical bars in graphics). Table 1 shows the simulation parameters for these solutions.

Table 1: Simulation parameters

Parameter	Value
Field Size	80x80
Base Station Location	40, 0
Initial location of intruder	0, 0
Intruder movement type	Random mobility
Intruder velocity	1.5
Total number of Nodes	100
Number of high-tier nodes	25
High-tier nodes topology	Grid
Low-tier nodes topology	Uniform
Initial Energy for low-tier nodes	20 J
Transmission Power	-10 dBm
Path loss model	Lognormal shadowing model
Radio model	CC2420
Video sequence	Hall
Video Encoding	H.264
Video Format	QCIF (176 x 144)
Frame Rate	26 fps

The simulations consist of a multi-tier architecture, composed of two tiers, as shown in Figure 5. In the lower tier, scalar sensor nodes perform intrusion detection, e.g. using vibration sensors, to wake up the higher tier camera nodes. On the other hand, the high-tier nodes, i.e. camera sensors nodes, will only 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 BS. As mentioned before, M3WSN enables to include the energy consumption for retrieving each frame that composes a video, as expected to model realistic simulation scenario. In this context, we set this value according to the values from CMUcam3 [15]. Additionally, it is also possible to define the energy consumption for transmitting packets, and we configured this value based on TelosB.

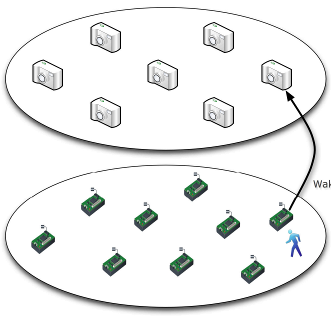


Figure 5: Multi-tier architecture for intrusion detection

The simulation starts with an intruder located at (0,0), and during the simulation the intruder moves in a random way. As soon as the low-tier detects the intruder, it must wake up the high-tier to send the video flows. Video sequences provide more precise information for users and authorities (e.g. the police) about the intruder, and enable them to monitor, detect, and predict the intruder's moving direction. Additionally, they allow the authorities to take precise decisions based on valuable visual information.

Existing works on multimedia area classify the videos into three categories, according to their motion and complexity levels, i.e. low, median and high. For example, Aguiar et al. classify the Hall video sequence (taken from the Video Trace Library) as low movement, which means that there is a small moving region on a static background, i.e. men walking in a hall [1].

We evaluated the transmitted videos by means of two well-known objective QoE metrics, i.e. SSIM and VQM, obtained by using the MSU Video Quality Measurement Tool (VQMT) [18]. SSIM measures the structural distortion of the video, and attempts to obtain a better correlation with the user's subjective impression. SSIM has values ranging from 0 to 1, a higher value meaning a better video quality. On the other hand, VQM measures the "perception damage" of video experienced, based on features of the human visual system, including distinct metric factors such as blurring, noise, colour distortion and distortion blocks. For VQM, a value closer to 0 means a video with a better quality.

4.2 Simulations results

First, we analyse the performance of no-FEC, simple-FEC and QoE-aware FEC solutions using Packet Delivery Ratio (PDR) with respect to the number of hops. By analysing the results, we can see that the solution without creating redundant packets has a higher PDR compared to simple-FEC and QoE-aware FEC. Additionally, the QoE-aware FEC mechanism has intermediate PDR compared to the simple-FEC and no-FEC solutions. This is due to simple-FEC and QoE-aware FEC solutions create redundant packets, which cause more interference and thus more packet losses.

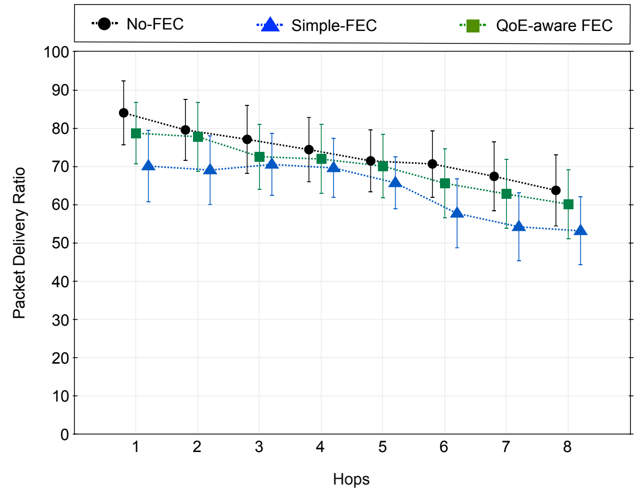


Figure 6: PDR with respect to number of hops

Moreover, the reason for QoE-aware FEC outperforms the simple-FEC solution in terms of PDR is related to the amount of redundant packets created by simple-FEC, as shown in Figure 7. Instead of QoE-aware FEC that creates redundant packets based on the frame importance, the simple-FEC generates redundant packets for all frames that compose a video (in a black-box way). Hence, QoE-aware FEC has a reduced packet overhead, creates less interference, and thus lower packet loss rate.

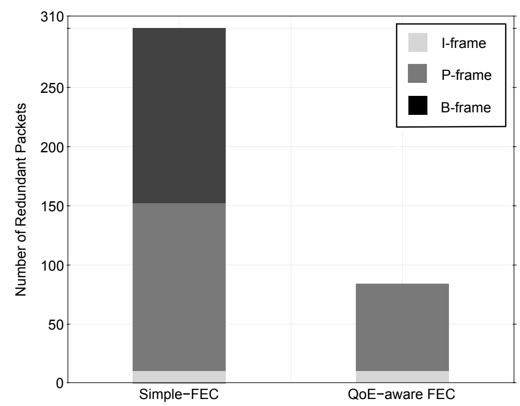


Figure 7: Overhead

However, multimedia solutions should evaluate the transmitted video from the user's perspective, by using QoE metrics. In this context, in our experiments, we measure the SSIM and VQM for transmitted videos with respect to the length of the transmission route (number of hops), as shown in Figures 8 and 9 respectively. The reason for conducting this analysis was that more hops have more interference and packet drops due to restricted buffers, and thus these nodes suffer higher packet loss.

Figure 8 shows the SSIM value according to the number of hops for video transmission without FEC, with simple FEC and with QoE-aware FEC solutions. Analysing the results, the solutions that create redundant packets improve the SSIM by around 25% compared to the solution without

FEC. This is due to the fact that application-level FEC is applied as error control scheme for handling packet losses in real-time communication. Hence, the redundant packets can be used to reconstruct a lost frame, and thus improve the video quality from a user perspective.

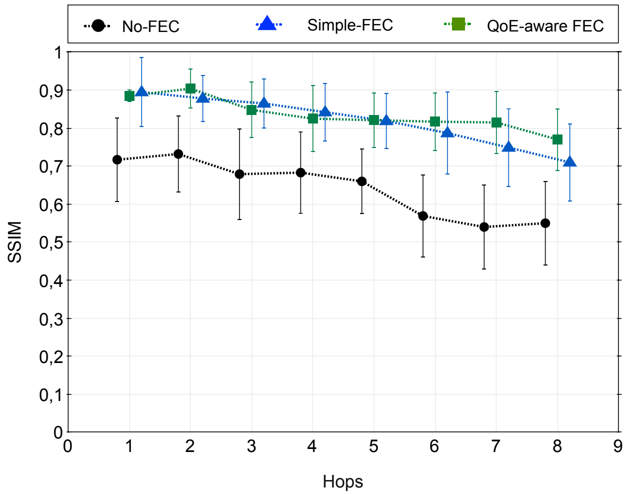


Figure 8: SSIM with respect to number of hops

Moreover, the simple and QoE-aware FEC solutions have similar video quality. However, the simple FEC solution includes a higher overhead, as shown 7. In contrast to that, the QoE-aware FEC approach achieves a lower overhead, while keep the video with the same video quality, which will bring many benefits in a resource-constrained system.

Due to less transmission means less energy consumption, we can conclude that QoE-aware FEC can provide energy-efficiency, while keeping the transmitted video with a good quality from the user perspective. The reason for this is that the QoE-aware FEC mechanism creates redundant packets based on frame importance and user experience to reduce the network overhead compared to simple FEC.

Figure 9 presents the video quality level by using VQM, and it is important to highlight that low VQM value means higher video quality level. The VQM results demonstrate the benefits of using FEC and confirm the SSIM values. For both the simple and QoE-aware approaches kept the VQM values below the solution without FEC, and thus improve the video quality level. However, the QoE-aware FEC mechanism reduces the amount of generated redundant packet while keeping videos with an acceptable quality level.

Finally, to show the impact of transmitting video streams from the standpoint of the end-user, a frame was randomly selected (i.e., Frame 258) from the transmitted video, as displayed in Figure 10. Frame 258 is the moment when a man (the intruder in our application) was walking along a corridor. For intruder detection application, this is an important frame to provide users and authorities (e.g., police) with more precise information and allow them to decide a suitable action. Additionally, it will be useful to monitor/detect the intruder, and predict the intruder's moving direction.

The benefits of the FEC mechanisms are visible by analysing the frames of Figure 10. By comparing each transmitted frame with the original frame (Figure 10(a)), it is possible to see a higher distortion for the frame transmitted without

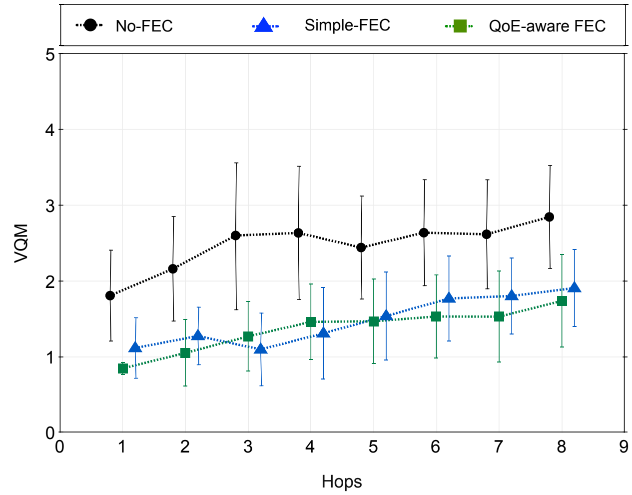
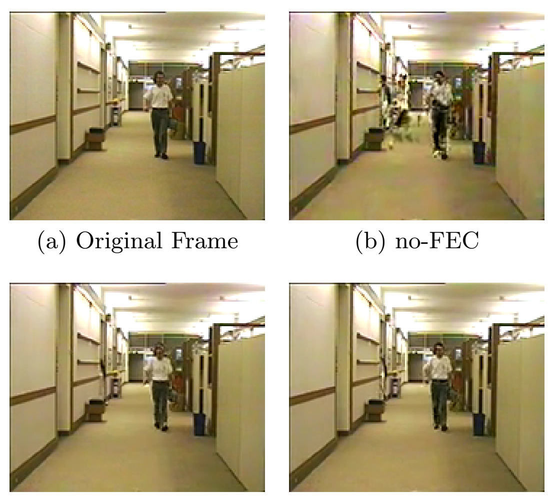


Figure 9: VQM with respect to number of hops

using any FEC, as shown in Figure 10(b). The frames transmitted using FEC mechanism achieves lower distortion, as shown in Figures 10(c) and 10(d). From the user's perspective, the QoE-aware FEC mechanism keeps the video with acceptable quality level, while the network overhead is significantly reduced. The visual evaluation is only possible due to M3WSN framework enables the transmission and control of real video sequences.



(a) Original Frame (b) no-FEC

(c) QoE-aware FEC (d) Simple FEC

Figure 10: Frame 258 of Hall video sequence

After analysing the QoS and QoE results, we can conclude that QoS metrics do not reflect the user's perception and, consequently, fail in capturing subjective aspects associated with human experience. In this context, QoE metrics or approaches overcome the limitations of current QoS schemes regarding to human perception and subjective aspects related to multimedia content. Nevertheless, QoE evaluation and optimization are only possible by transmitting real video sequences. In contrast to the M3WSN framework, the existing OMNeT ++ frameworks for WMSNs do not support multimedia transmission, control and evaluation.

5. CONCLUSIONS

The paper contributes to the OMNeT++ community by introducing the Mobile MultiMedia Wireless Sensor Network (M3WSN), which is available as open-source at [14]. M3WSN extends Castalia with functionalities of both WiSE-MNet and WVSN models, because they have realistic models for WMSNs. Additionally, M3WSN enables the multimedia transmission of real video sequence. Thus, a set of multimedia algorithms, protocols and services can be now evaluated by using QoE metrics, such as cross-layer QoE-routing and load-balance. Moreover, key video-related information, such as frame type, GoP length and intra-frame dependency can be used for creating new assessment and optimization solutions for WMSNs as expected for IoT/smart cities applications. Node mobility is becoming more important and is one of the most challenging issues in WMSNs. In this context, the M3WSN framework also supports several mobility traces to enable the understanding of how the network behaves under different mobile situations. Our experiments have proven the importance to obtain the key video information, i.e. frame type and GoP length, for creating packet redundancy based on video characteristics. Moreover, the transmissions of real video sequences enable multimedia evaluation from user's perspective by means of QoE metrics.

6. ACKNOWLEDGMENTS

This work is supported by M3WSN, a joint research project of Nano-Tera.ch and the Sino Swiss Science and Technology Cooperation (SSSTC), and also the National Council for Scientific and Technological Development (CNPq) - Brazil.

7. REFERENCES

[1] E. Aguiar, A. Riker, A. Abelém, E. Cerqueira, and M. Mu. Video Quality Estimator for Wireless Mesh Networks. In *20th International Workshop on Quality of Service (IWQoS)*, pages 1–9. IEEE, 2012.

[2] I. Almalkawi, M. Guerrero Zapata, J. Al-Karaki, and J. Morillo-Pozo. Wireless Multimedia Sensor Networks: Current Trends and Future Directions. *Sensors*, 10(7):6662–6717, 2010.

[3] N. Aschenbruck, R. Ernst, E. Gerhards-Padilla, and M. Schwamborn. Bonnmotion: a mobility scenario generation and analysis tool. In *Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques*, SIMUTools '10, pages 51:1–51:10, ICST, Brussels, Belgium, Belgium, 2010.

[4] A. Boulis. Castalia, a Simulator for Wireless Sensor Networks and Body Area Networks, Version 2.2. *User manual, NICTA*, Aug 2009.

[5] Evalvid. EvalVid - A Video Quality Evaluation Tool-set. [OnLine] Available: http://www2.tkn.tu-berlin.de/research/evalvid/fw.html. (Access date: Sep 2012).

[6] INET. INET framework. [OnLine] Available: http://www.inet.omnetpp.org/. (Access date: Sep 2012).

[7] INETMANET. INETMANET framework. [OnLine] Available: http://www.github.com/inetmanet/inetmanet/. (Access date: Sep 2012).

[8] V. T. Library. Yuv video sequences. [OnLine] Available: http://trace.eas.asu.edu/yuv/. (Access date: Sep 2012).

[9] MiXiM. MiXiM framework. [OnLine] Available: http://www.mixim.sourceforge.net/. (Access date: Sep 2012).

[10] C. Nastasi and A. Cavallaro. WiSE-MNet: an Experimental Environment for Wireless Multimedia Sensor Networks. *Proceedings of Sensor Signal Processing for Defence (SSPD)*, pages 1–5, 2011.

[11] C. Pham and A. Makhoul. Performance study of multiple cover-set strategies for mission-critical video surveillance with wireless video sensors. In *IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2010)*, pages 208–216. IEEE, Oct 2010.

[12] D. Rosário, R. Costa, H. Paraense, K. Machado, E. Cerqueira, T. Braun, and Z. Zhao. A Hierarchical Multi-hop Multimedia Routing Protocol for Wireless Multimedia Sensor Networks. *Network Protocols and Algorithms*, 4(2):44–64, 2012.

[13] D. Rosário, K. Machado, A. Abelém, D. Monteiro, and E. Cerqueira. *Recent Advances and Challenges in Wireless Multimedia Sensor Networks*, pages 74–96. Mobile Multimedia - User and Technology Perspectives. InTech, Dian Tjondronegoro (Ed.), Jan 2012.

[14] D. Rosário, Z. Zhao, C. Silva, E. Cerqueira, and T. Braun. A Mobile MultiMedia Wireless Sensor Network OMNeT++ framework (M3WSN). [OnLine] Available: http://cds.unibe.ch/research/M3WSN/. (Access date: Sep 2012).

[15] 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)*, volume 1, pages 208–213. IEEE, 2002.

[16] 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*, pages 252–263, 2010.

[17] A. Varga. The OMNeT++ Discrete Event Simulation System. In *Proceedings of the European Simulation Multiconference (ESM 2001)*, 2001.

[18] D. Vatolin, A. Moskvin, O. Petrov, and N. Trunichkin. Msu video quality measurement tool. [OnLine] Available: http://www.download3k.com/Install-MSU-Video-Quality-Measurement-Tool.html. (Access date: Sep 2012).

[19] Z. Zhao and T. Braun. Topology control and mobility strategy for uav ad-hoc networks: A survey. In *Joint ERCIM eMobility and MobiSense Workshop*, pages 27–32, Santorini, Greece, Jun 2012.

[20] Z. Zhao, T. Braun, D. Rosário, E. Cerqueira, R. Immich, and M. Curado. QoE-aware FEC Mechanism for Intrusion Detection in Multi-tier Wireless Multimedia Sensor Networks. In *1st International Workshop on Wireless Multimedia Sensor Networks (WMSN')*, pages 689–696, Oct. 2012.