Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2007, Article ID 31976, 13 pages
doi:10.1155/2007/31976

Research Article

Wireless Mesh Networks to Support Video Surveillance: Architecture, Protocol, and Implementation Issues

Francesco Licandro and Giovanni Schembra

Dipartimento di Ingegneria Informatica e delle Telecomunicazioni, University of Catania, Viale A. Doria 6, 95125 Catania, Italy

Received 29 June 2006; Revised 21 December 2006; Accepted 30 January 2007

Recommended by Marco Conti

Current video-surveillance systems typically consist of many video sources distributed over a wide area, transmitting live video streams to a central location for processing and monitoring. The target of this paper is to present an experience of implementation of a large-scale video-surveillance system based on a wireless mesh network infrastructure, discussing architecture, protocol, and implementation issues. More specifically, the paper proposes an architecture for a video-surveillance system, and mainly centers its focus on the routing protocol to be used in the wireless mesh network, evaluating its impact on performance at the receiver side. A wireless mesh network was chosen to support a video-surveillance application in order to reduce the overall system costs and increase scalability and performance. The paper analyzes the performance of the network in order to choose design parameters that will achieve the best trade-off between video encoding quality and the network traffic generated.

1. INTRODUCTION

Video-surveillance systems are very important in our daily lives due to the number of applications they make possible. The reasons for interest in such systems are diverse, ranging from security demands and military applications to scientific purposes. Video-surveillance systems are currently undergoing a transition from traditional analog solutions to digital ones. This paradigm shift has been triggered by technological advances as well as increased awareness of the need for heightened security in particular vertical markets such as government and transportation. Compared with traditional analog video-surveillance systems, digital video surveillance offers much greater flexibility in video content processing and transmission. At the same time, it can also easily implement advanced features such as motion detection, facial recognition, and object tracking. Many commercial companies now offer IP-based surveillance solutions. For example, Texas Instruments DSPs can be used to design various video-surveillance systems from low-end to high-end and from a portable implementation to plug-in implementation. The TMS320C64x DSP provides users with a high-resolution video-surveillance system over the Internet protocol, thanks to its architecture and peripherals such as video ports and on-chip ethernet media access controller (EMAC) [1].

The paper starts from an experience of deployment of a prototype of a large-scale distributed video-surveillance system that the authors' research group is realizing as a common testbed for many research projects. It consists of sixty video cameras distributed over the campus of the University of Catania, transmitting live video streams to a central location for processing and monitoring.

Deployment and maintenance of large-scale distributed video-surveillance systems are often very expensive, mainly due to the installation and maintenance of physical wires. The solution is chosen in order to significantly reduce the overall system costs, while increasing deployability, scalability, and performance is the use of wireless interconnections [2, 3].

With this in mind, the idea at the basis of this work is to apply multihop wireless mesh networks (WMN) [4–9] as the interconnection backbone of a wireless video-surveillance network (WVSN). The proposed architecture is shown in Figure 1. As we will see in Section 2, it fits in well with the structure of a WMN [10], where traffic sources are networked digital video cameras, while the nodes of the WVSN are fixed and wirelessly interconnected to provide video

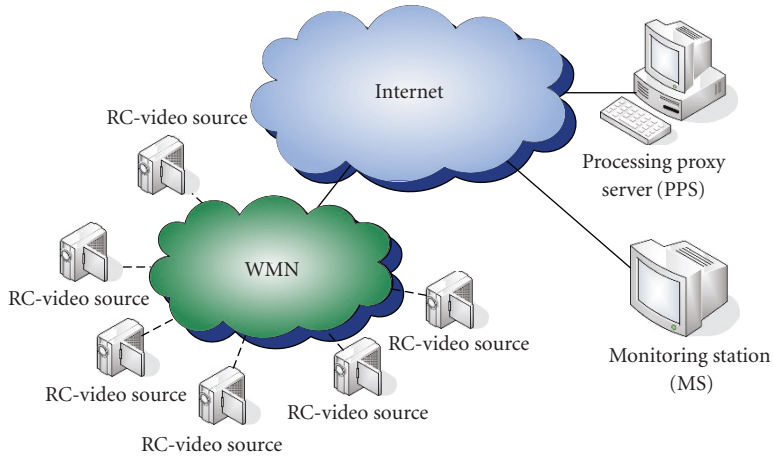


Figure 1: WVSN architecture.

sources with connections towards a video proxy with processing and filtering capabilities. Video proxies are typically located in the wired network.

Nevertheless, implementing an intelligent, scalable, and massively distributed video-surveillance system over wireless networks remains a research problem and leads to at least the following important issues, some of which have been raised by Feng et al. [11]:

(i) transmission bandwidth and transmission power are scarce resources in wireless environments;
(ii) wireless links are more vulnerable to interceptions and external attacks;
(iii) the high loss percentage of wireless links requires sophisticated techniques for channel encoding that often increase transmission delays.

On the contrary, a video-surveillance system presents the following features:

(a) commonly used computer vision algorithms in a video-surveillance environment perform better when video is encoded with high PSNR and temporal quality. However, increasing video quality causes an increase in both the required transmission bandwidth and transmission power;
(b) interceptions and external attacks are a serious problem in video-surveillance applications;
(c) delay and delay jitter are very harmful, and therefore must be kept below a given acceptable threshold.

So, the target of this paper is twofold: describing a real experience based on a new WVSN architecture, defined by the authors, which is based on a wireless mesh network as the interconnection backbone; analyzing its performance in order to evaluate some protocol and implementation issues, and provide some insights into the choice of design parameters that will optimize the quality of video received at destination by the processing proxy server. This can be achieved by trying to obtain the best trade-off between video encoding quality and the network traffic generated at the source side, and using suitable routing algorithms in the wireless mesh network. The video quality at destination is evaluated through an objective quality parameter which is able to simultaneously account for packet losses in the network (impacting the received frame rate) and encoding quality (impacting the PSNR of the decoded frames).

The paper is structured as follows. Section 2 introduces the related work. Section 3 describes the proposed architecture. Section 4 discusses the achieved performance. Finally, Section 5 concludes the paper.

2. RELATED WORK

Analog video-surveillance systems (e.g., CCTV) are increasingly being replaced by more advanced digital video-surveillance (DVS) solutions, often utilizing IP technologies and networked architectures. Besides the ever-increasing demand for security, the low cost of cameras and networking devices has contributed to the spread of *digital distributed multimedia surveillance systems*. This now constitutes an emerging field that includes signal and image processing, computer vision, communications, and hardware.

The automated analysis and processing of video surveillance is a central area of study for the computer vision and pattern recognition research community. IBM Research's PeopleVision [12] project, for example, has focused on the concept of *Smart Surveillance* [13], or the application of automated analysis of surveillance video to reduce the tedious, time-consuming task of viewing video feeds from a large number of security cameras. There have been a number of famous visual surveillance systems. The real-time visual surveillance system W4 [14] employs a combination of shape analysis and tracking and constructs models of people's appearances in order to detect and track groups of people as well as monitor their behaviors even in the presence of occlusion and in outdoor environments. This system uses a single camera and grayscale sensor. The VIEWS system [15]

is a 3D-model-based vehicle tracking system. The Pfinder system [16] is used to recover a 3D description of a person in a large room. It tracks a single nonoccluded person in complex scenes, and has been used in many applications. The system at CMU [17] can monitor activities over a large area using multiple cameras that are connected into a network.

As far as hardware for video surveillance is concerned, companies like Sony and Intel have designed equipments suitable for visual surveillance, for example, active cameras, smart cameras [18], omnidirectional cameras [19, 20], and so on. Networking devices for video surveillance are the Intelligent Wireless Video Systems proposed by Cisco with the 3200 Series Wireless and Mobile Routers. Cisco Systems offer, for example, an outdoor and mobile wireless router with intelligent video functions, addressing public safety and transportation customer needs for highly secure, cost-efficient, and standards-based video-surveillance applications [21].

Another important focus of research into video-surveillance systems is on communications between networked cameras and video processing servers. This is the field of this paper.

The classical approach to digital video-surveillance systems is based on wired connections with existing Ethernet and ATM dedicated-medium networks [22]. Another wired-based approach is proposed in [23], where IEEE 1394b FireWire is investigated as a shared medium protocol for ad hoc, economical installation of video cameras in wireless sensor networks (WSNs). However, they are the cost and performance bottleneck to further deployment of large-scale video-surveillance systems with highly intelligent cameras [11]. A hybrid routing protocol for future arbitrary topology WSNs is presented. It uses distributed location servers which maintain the route-attribute-location knowledge for routing in WSNs.

The latest step in the evolution of video-surveillance systems, aimed at increasing the scalability of large video-surveillance systems, is the migration to wireless interconnection networks. Many solutions have been proposed in this context, by both industries and research institutions. Firetide Inc., a developer of wireless multiservice mesh technology, and Axis Communications, a company working on network video solutions, have announced a strategic partnership to deliver high-quality video over wireless mesh networks, which are being used by a number of cities to provide wireless video surveillance. In Massachusetts, for example, the Haverhill Police Department selected these technologies for its own video-surveillance system [24]. Initially installed in a small, high-crime area downtown, the solution consists of Firetide HotPort outdoor and indoor wireless mesh nodes and AXIS 214 PTZ (pan-tilt-zoom) and AXIS 211 fixed cameras.

A great amount of work has been done to reduce power consumption in wireless video-surveillance networks. Reference [25] defines some QoS-parameters in video surveillance, like video data quality and its distortions in network transmission (jitter). Further parameters include quality metrics such as image size, data rate, or the number of frames per second (fps). The work in [3] investigates the trade-off between image quality and power consumption in wireless video-surveillance networks. However, existing implementations lack comprehensive handling of these three correlating parameters. In [26], an adaptive checkpointing algorithm is proposed that also minimizes energy consumption.

Another important issue to be considered from the communications point of view is routing. A very large amount of research has been carried out regarding routing in ad hoc wireless networks. Now we have to take into account that the network environment we are considering in this paper is a wireless mesh network, which is a particular case of wireless ad hoc networks. In addition, as we will illustrate in the following section, we will apply multipath routing, given that multiple paths can provide load balancing, fault-tolerance, and higher aggregate bandwidth [27]. Load balancing can be achieved by spreading the traffic along multiple routes. This can alleviate congestion and bottlenecks. From a fault tolerance perspective, multipath routing can provide route resilience. Since bandwidth may be limited in a wireless network, routing along a single path may not provide enough bandwidth for a connection. However, if multiple paths are used simultaneously to route data, the aggregate bandwidth of the paths can satisfy the bandwidth requirement of the application. Also, since there is more bandwidth available, a smaller end-to-end delay can be achieved.

Many multipath routing protocols have been defined in the past literature for ad hoc wireless networks. The multipath on-demand routing (MOR) protocol [28] was defined to connect nodes in wireless sensor networks. Other important routing protocols for ad hoc networks are DSR [29], TORA [30], and AODV [31]. DSR is an on-demand routing protocol which works on a source routing basis. Each transmitted packet is routed carrying the complete route in its header. TORA is an adaptive on-demand routing protocol designed to provide multiple loop-free routes to a destination, thus minimizing reaction to topological changes. The protocol belongs to the link reversal algorithm family. AODV is an on-demand distance-vector routing protocol, based on hop-by-hop routing. It is a modified DSR protocol incorporating some features presented in the DSDV protocol, such as the use of hop-by-hop routing, sequence numbers, and periodic beacon messages.

However, all the above protocols are reactive, or on-demand, meaning that they establish routes as needed. The advantage of this approach is obvious if only a few routes are required, since the routing overhead is less than in the proactive approach of establishing routes whether or not they are needed. The disadvantage of on-demand establishment of routes is that connections take more time if the route needs to be established. However, given that the wireless mesh networks considered in this paper have stable topologies because nodes are fixed and powered, the proactive approach works better. For this reason we propose to use the distance-vector multipath network-balancing routing algorithm [32], which is a proactive routing algorithm.

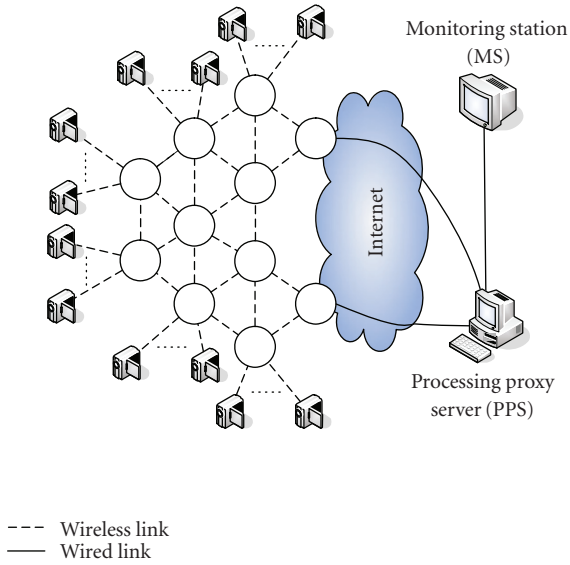


Figure 2: WVSN topology.

(a) Edge node (b) Core node

Figure 3: Mesh network routers.

3. DESCRIPTION OF THE WVSN SYSTEM

In this section, we will describe the video-surveillance platform considered in the rest of the paper. The system topology is shown in Figure 2. It is made of an access network and a core network. In order to monitor six different areas of the campus, the access network comprises six edge nodes. Edge and core nodes are sketched in Figure 3. Each edge node is equipped with one omnidirectional antenna to allow wireless access to video cameras. Both edge and core nodes are routers wirelessly connected to the other nodes by high gain directional antennas to minimize interferences, and so to avoid network capacity degradation. All the links of the mesh network are IEEE 802.11b wireless connections at 11 Mbps. More specifically, the following antennas have been used:

(a) omnidirectional antenna installed in each edge node for connection of wireless cameras: pacific wireless 2.4 GHz PAWOD24-12, with a gain of 12 dBi, a frequency range of 2400–2485 MHz, and a vertical beam width of 7 degrees;
(b) unidirectional antenna installed in each node (both edge and core) for point-to-point connection with the other nodes: pacific wireless 2.4 GHz Yagi PAWVA24-16, with a gain of 16 dBi, a frequency range of 2400–2485 MHz, and a beam width of 25 degrees.

Radio frequencies have been designed in such a way that different radio interfaces on the same node use different radio channels.

The wireless mesh network is connected to the Internet through the gateway nodes. We have chosen a number of two gateway nodes to distribute network load and to guarantee path diversity towards the proxy server.

Video sources are networked digital video cameras connected to the edge nodes through IEEE 802.11b wireless connections at 11 Mbps. Specifically, ten wireless video cameras are connected to each edge node. Video cameras are set to encode video with a 352×288 CIF format, using a standard MPEG-4 encoder at a bit-rate settable in the range between 100 kbps and 600 kbit/s, and a frame rate of 12 fps. Each frame is encoded as an I-frame.

4. WVSN ARCHITECTURE

The distributed architecture defined for the video-surveillance system is sketched in Figure 1. It consists of a number of wireless networked rate-controlled video cameras (*RC-video sources*) which, thanks to the WMN, access the Internet and continuously transmit their video flows to a *processing proxy server* (PPS) for processing and filtering. The PPS is directly, or again through the Internet, connected to one or more *monitoring stations* (MS). Not every video stream that is sent to the PPS for processing is shown to the end user at the MS. In fact, the PPS analyzes all the received video flows, and alerts the MS only if a suspicious event is detected. The focus of our paper is concentrated on the RC-video sources (and video stream destination at the PPS) and the wireless mesh network. They will be described in Sections 4.1 and 4.3. The Processing Proxy Server will be briefly described in Section 4.2, even though the internal algorithms are beyond the scope of this paper.

4.1. RC-video source

The logical architecture of the RC video system is sketched in Figure 4. It is an adaptive-rate MPEG video source over a UDP/IP protocol suite. The video stream generated by the video source is encoded by the *MPEG encoder* according to

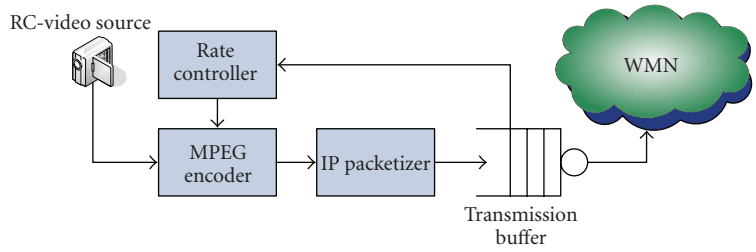


Figure 4: RC-video source architecture.

the MPEG-4 video standard [33, 34]. In the MPEG encoding standard, each frame, corresponding to a single picture in a video sequence, is encoded according to one of three possible encoding modes: intraframes (I), predictive frames (P), and interpolative frames (B). Typically, I-frames require more bits than P-frames, while B-frames have the lowest bandwidth requirement. For this reason the output rate of MPEG video sources needs to be controlled, especially if the generated flow is transmitted on the network. Thus, as usual, a rate controller combined with the transmission buffer has been introduced in the video encoding system. It works according to a feedback law by appropriately choosing the so-called *quantizer scale parameter* (QSP) in such a way that the output rate of the MPEG encoder results in as much constant as possible. The *MPEG encoder* output is packetized in the *packetizer* according to the UDP/IP protocol suite and sent to the *transmission buffer* for transmission.

The QSP value can range within the following set $[1, 31]$: 1 being the value giving the best encoding quality but requiring the maximum number of bits to encode the frame, and 31 the value giving the worst encoding quality, but requiring the minimum number of bits. However, let us note that it is not possible to encode all the frames with the same number of bits at least for the following three reasons: (1) quantizer scale is chosen a priori before encoding, and this choice is only based on long-term video statistics, and not on the particular frame to be encoded; (2) quantizer scale parameter can assume 31 values only, and therefore granularity is not so high to obtain any value desired for the number of bits of the encoded frame; (3) sometimes, for example, when scene activity is too high or too low, the desired number of bits cannot be obtained for none of the 31 quantizer scale parameter values. Taking into account this, the transmission buffer role is necessary to eliminate residual output rate variability. In fact, the transmission buffer is served with a constant rate, and therefore its output is perfectly constant, except for the cases when it empties. Of course, the transmission buffer queue should not saturate because high delays and losses should be avoided, and therefore the rate controller presence is fundamental to maintain the queue around a given threshold, avoiding both empty queue and saturation states.

So the rate controller is necessary to make the output rate of the MPEG video source constant, avoiding losses in the transmission buffer, and maximizing encoding quality and stability. As said so far, it works according to a given feedback law. This law depends on the activity of the frame being encoded and the current number of packets in the transmission buffer. More specifically, in order to keep the output rate as constant as possible, a frame-based feedback law is used [35]. According to this law, the target is to maintain the queue of the transmission buffer very close to a given threshold, θ_F. This is based on the statistics of the video flow, expressed in terms of rate and distortion curves [36, 37].

The rate curves, $R_{a,j}(q)$, give the expected number of bits which will be emitted when the jth frame in the GoP has to be encoded, if its activity value is a, and is encoded with a QSP value q. The distortion curves, $F^{(j)}(q)$, give the expected encoding PSNR for each value of the QSP [38]. The rate and distortion curves for the implemented video-surveillance system are shown in Figure 5.

As said so far, the aim of the rate controller is to maintain the transmission buffer queue length lower than and very close to θ_F at the end of each frame encoding interval. Indicating the frame to be encoded as j, its activity as a, and the number of data units in the transmission buffer queue before encoding as s_Q, the expected number of packets to encode can be calculated from the rate curve, $R_{a,j}(q)$. So, the frame-based feedback law works by choosing the QSP as follows:

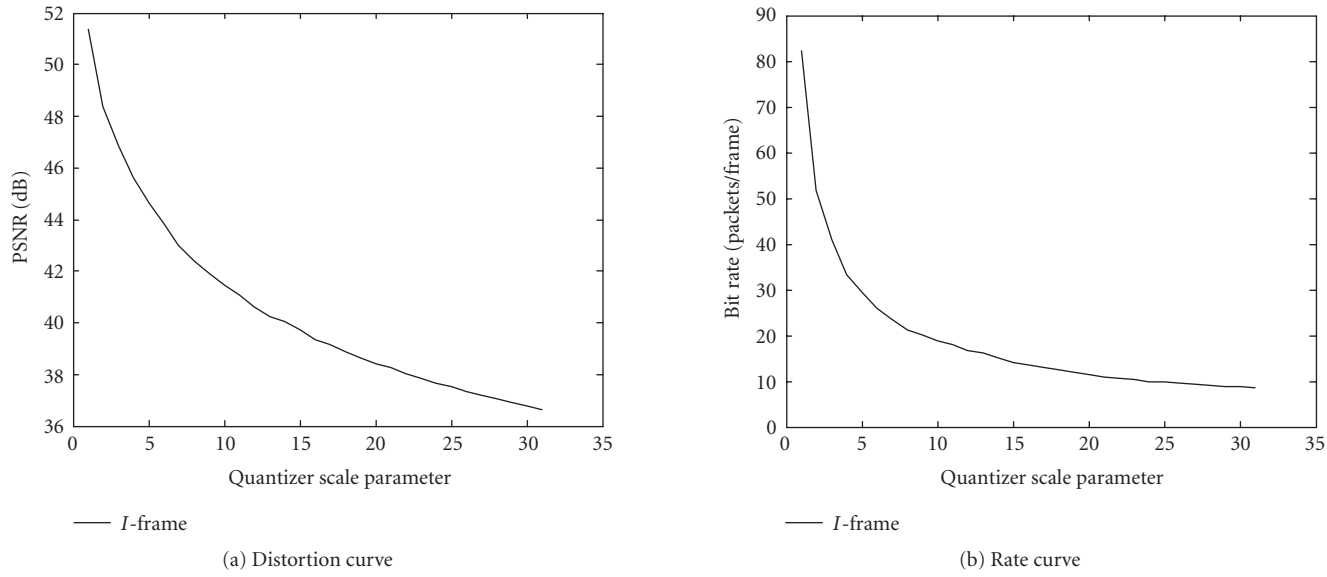
$$q = \Phi(s_Q, a, j) = \min_{\overline{q}\in[1,31]} (\overline{q} : s_Q + R_{a,j}(\overline{q}) \leq \theta_F). \quad (1)$$

4.2. Processing proxy server

The logical architecture of the PPS is sketched in Figure 6. Its main task is to process the video signals in order to detect an intrusion in the controlled area and to send the relative video to the MS.

The RC-video receiver block receives the video flows from the distributed video-surveillance network through the Internet. It is made up by three fundamental blocks: a *packet reordering buffer* and a *jitter compensator buffer*, with the aim of eliminating loss of packet order and delay variations introduced by the network, and an *MPEG decoder block* to decode the received video flow.

The decoded video streams are processed by the *video processor and the alarm trigger* block. When an intrusion is



(a) Distortion curve (b) Rate curve

Figure 5: Rate-distortion curves.

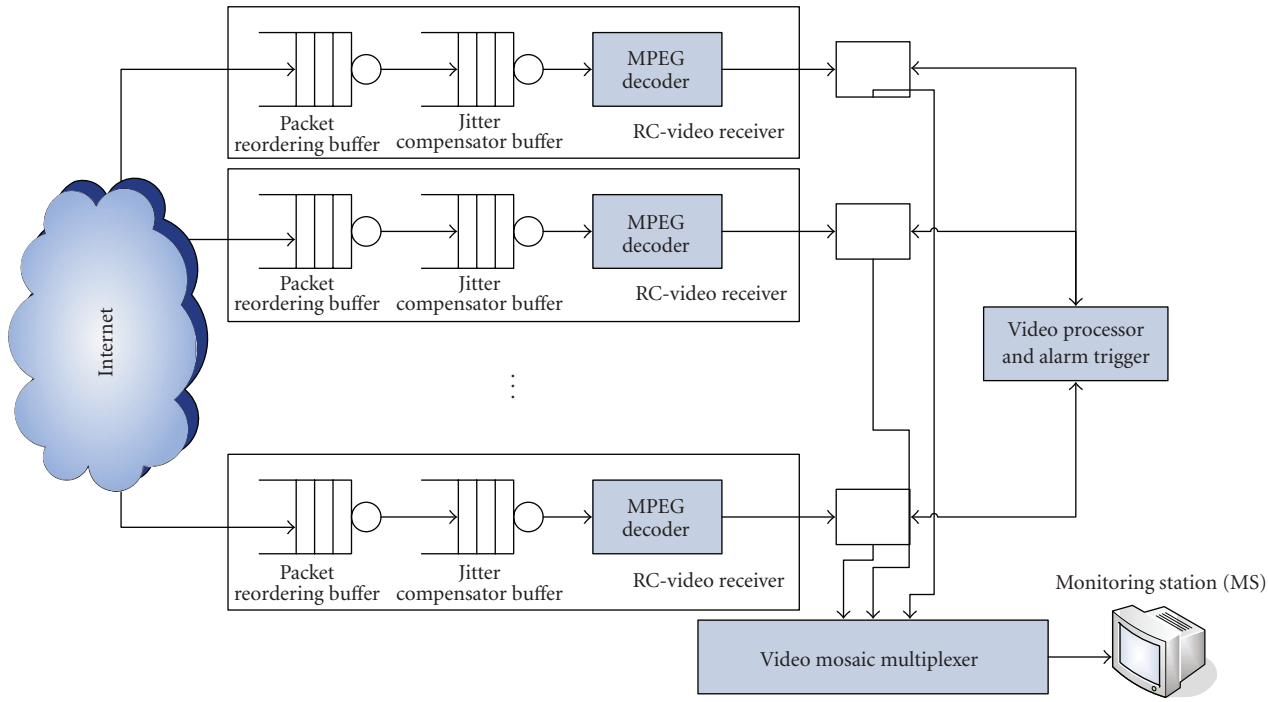


Figure 6: Processing proxy server architecture.

detected by the video processor, the trigger system sends the relative video images to the *video mosaic multiplexer* block which makes a spatial composition of the videos. Finally, the multiplexer output video is sent to the monitoring station (MS) for visualization by the final user.

4.3. Wireless mesh network

The WMN constitutes the infrastructure interconnection network for the wireless video-surveillance system. It comprises a number of edge nodes, a number of core nodes, and

a number of gateway nodes, all interconnected through wireless links. It is a multihop wireless network which, unlike mobile ad hoc networks (MANET), is constituted by fixed nodes. RC-video sources are connected to edge nodes, while the WMN is connected to the Internet through the gateway nodes. The number and location of the edge nodes have to be chosen in such a way as to allow the connection of all the networked wireless cameras.

An important role in this architecture is played by the routing algorithm. Given that the WMN is stable in time because nodes are powered and fixed, a proactive discovery of paths is the best solution since it provides reduced packet delays (deleterious for video-surveillance applications) [39]. On the other hand, additional packet latency due to on-demand route discovery, typical in reactive routing strategies, is not acceptable.

Bearing in mind the above-mentioned issues, we have used a distance-vector multipath network-balancing routing algorithm [32]. According to this algorithm each node, thanks to a distance-vector algorithm, knows the distance from the Internet through each path in the mesh network, and forwards packets, in a round-robin fashion, through all the paths having the same minimum cost to reach the Internet, whatever the destination gateway node is. The distance-vector multipath network-balancing routing algorithm is used for two reasons: first it is able to reduce delay [32, 40, 41]; second, thanks to its multipath peculiarity, it increases the robustness of the architecture to external attacks and interceptions. In fact, if a path is (maliciously or not) shielded, or its quality is temporally degraded, all the packets flowing through it are lost; however, the application of the multipath network-balancing routing algorithm guarantees that a high percentage of packets are able to reach the video decoder block, and therefore frames can be decoded, by applying an error concealment video decoding algorithm [42–44].

Mesh nodes are implemented as software routers running on low-cost computers with the Click modular router [45, 46] on a Linux platform. Hardware of each node is realized by using the Soekris Engineering net4801 single board computer, shown in Figure 7(a), chosen as a good trade-off between costs and performance.

Click is a software architecture for building flexible and configurable routers. A Click router is assembled from packet processing modules called elements. Individual elements implement simple router functions like packet classification, queueing, scheduling, and interfacing with network devices. A router configuration is a directed graph with elements at the vertices; packets flow along the edges of the graph. A standards-compliant Click IP router has sixteen elements on its forwarding path. Click configurations are modular and easy to extend. The Click modular router configuration we have designed and implemented for mesh nodes is shown in Figure 7(b). The *AOMDV* element implements the multipath routing algorithm by communicating with the other network nodes through the network interfaces, represented as eth0 and eth1 in Figure 7(b). Then it elaborate information and manages the IP routing table, which is read by *the LookupIPRoute* element.

5. NUMERICAL RESULTS

In this section, we will analyze the performance of the wireless video-surveillance system described so far, and the quality of service (QoS) perceived at the PPS video processor block, which is crucial for the detection of suspicious events.

More specifically, we will discuss the following two main issues:

(i) delay analysis for jitter compensation buffer dimensioning;
(ii) quality of service (QoS) perceived at destination by the PPS, and in particular by its video processor block, which is crucial for the detection of suspicious events.

Both analyses are carried out by comparing the distance-vector multipath network-balancing routing algorithm proposed for this system with classic single-path minimum hop count routing, in order to evaluate the advantages and disadvantages of the proposed approach.

The analysis has been carried out versus the encoding rate imposed by the rate controller to each video source. This rate was changed in the range $[200, 600]$ kbps, given that greater rates cannot be supported by the four bottleneck links connecting the mesh network to the gateway nodes, because each link has a maximum transmission rate of 11 Mbps.

As regards the delay analysis, we considered both the end-to-end average delay and the delay jitter, represented by the standard deviation of the delay distribution [21].

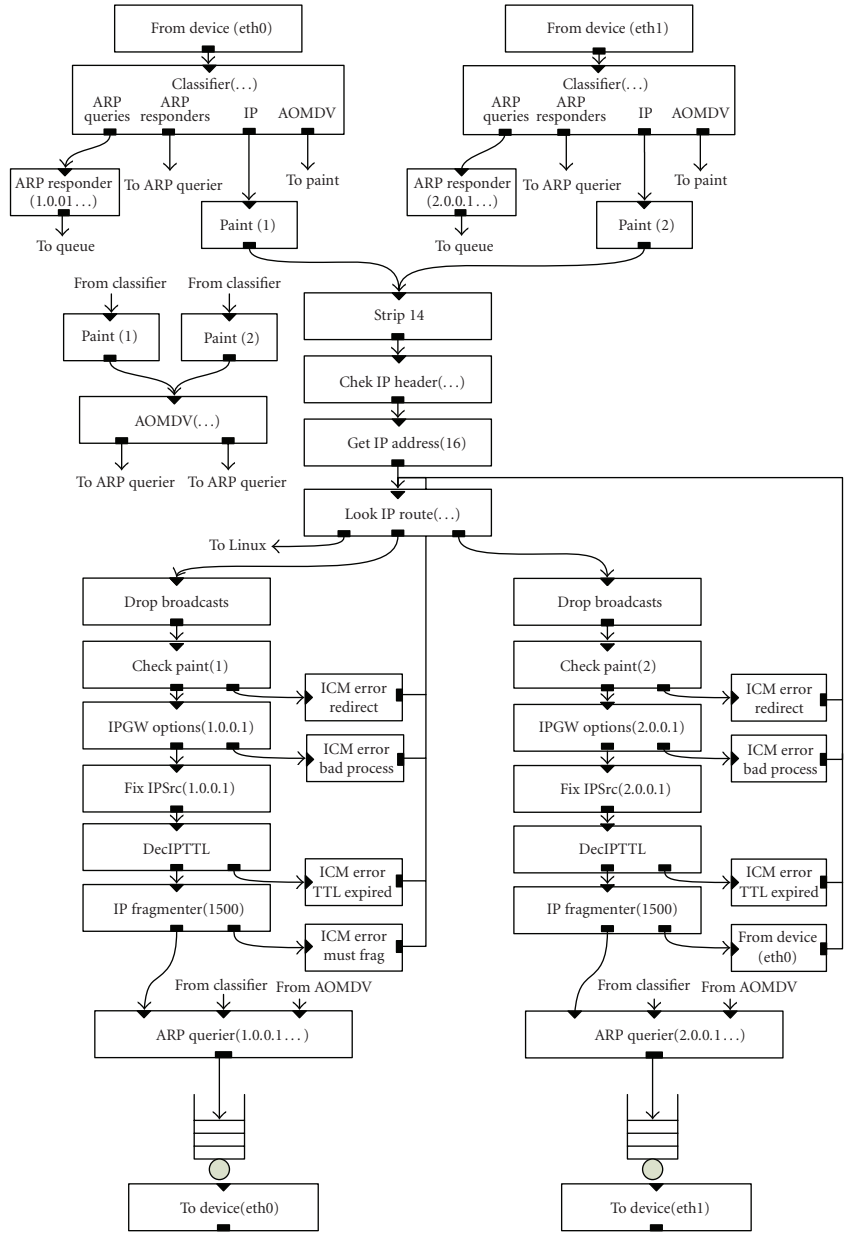
The quality of service (QoS) perceived at destination by the PPS video processor block depends on both the encoding quality at the source and losses occurring in the network and the jitter compensation buffer. More specifically, the encoding quality is decided by setting the quantizer scale parameter, q, as described in Section 4.1. Losses in both the network and the jitter compensation buffer cause an additional degradation of the quality of the decoded frames at destination, given that some frames will never arrive at destination, while other frames will arrive corrupted because not all their packets are available to the decoder at the right time. In this case an error concealment technique is used at destination to efficiently reconstruct corrupted and missing frames and thus improve the quality of the decoded video.

Given that the concealment technique used is beyond the scope of our paper, in order to achieve results independently of it, we assumed that all frames which have registered a loss percentage greater than a given threshold, here set to $\tau = 20\%$, are not decodable; instead, frames with fewer lost packets are reconstructed, with a quality depending on the percentage of arrived packets.

To summarize, losses in the network and the jitter compensation buffer cause both a reduction in the quality of decoded frames, and a frame rate reduction due to nondecodable and nonarrived frames.

The quality of decoded frames at destination is described by the peak signal-to-noise ratio (PSNR), defined as the ratio

(a) Soekris Engineering net4801 board for hardware implementation



(b) Click modular router configuration for software implementation

Figure 7: Edge and core node implementation.

between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. The PSNR is most easily defined via the mean squared error (MSE). For two $m \times n$ monochrome images I and K, where I is the original image before encoding and K is the reconstructed image at destination, MSE is defined as

$$\mathrm{MSE} = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left| I(i,j) - K(i,j) \right|^2. \tag{2}$$

Then the PSNR is defined as

$$\mathrm{PSNR} = 20 \log_{10} \left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}} \right), \tag{3}$$

where MAX_I is the maximum pixel value of the image. Since pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with B bits per sample, MAX_I is $2^B - 1$. PSNR is usually expressed in terms of the logarithmic decibel scale because many signals have a very wide dynamic range.

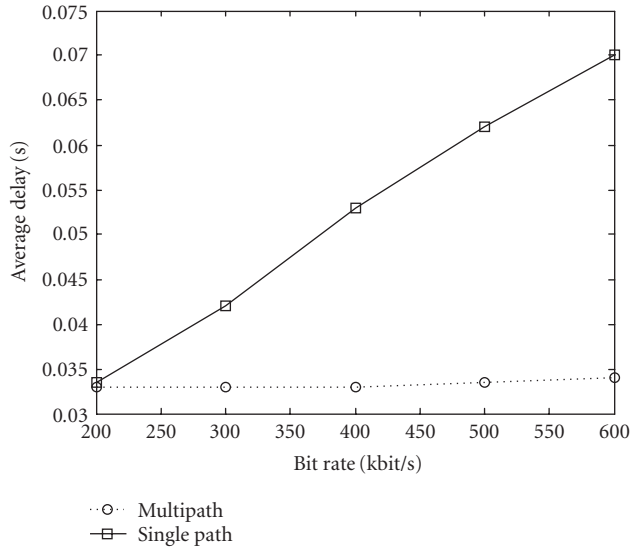


Figure 8: End-to-end average delay.

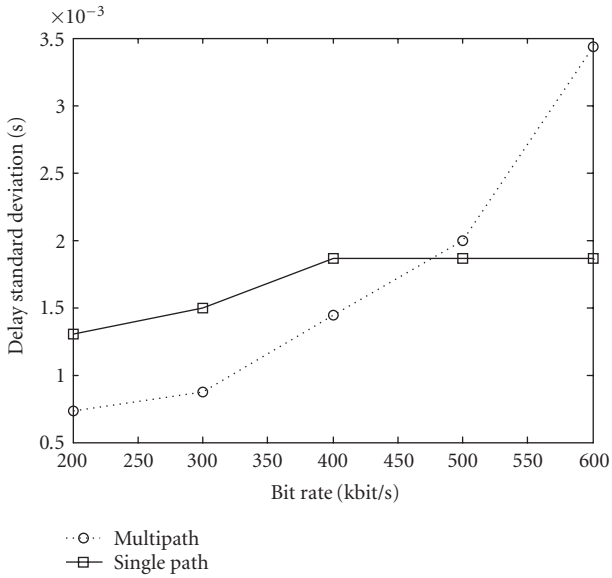
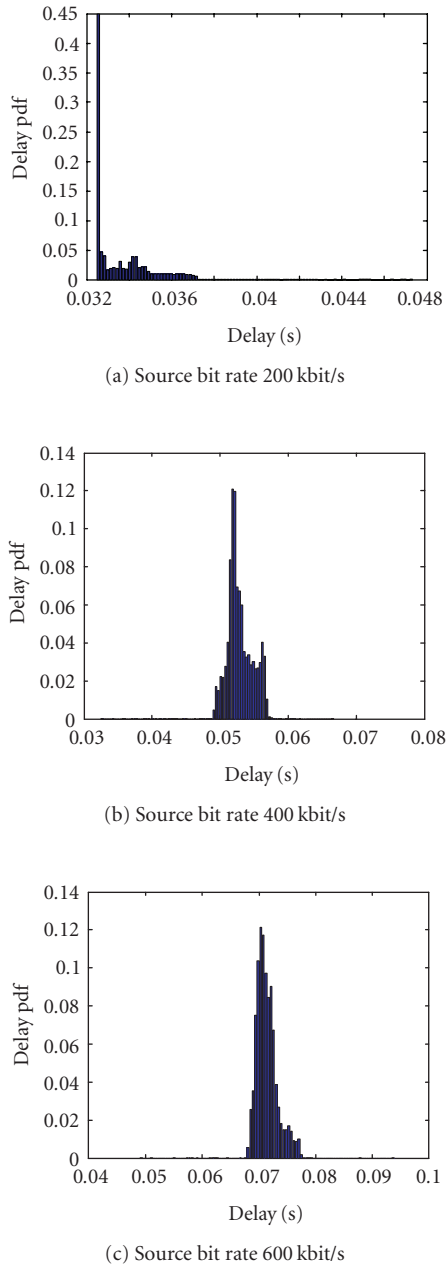


Figure 9: End-to-end delay standard deviation.



(a) Source bit rate 200 kbit/s

(b) Source bit rate 400 kbit/s

(c) Source bit rate 600 kbit/s

Figure 10: End-to-end delay distribution for single-path routing.

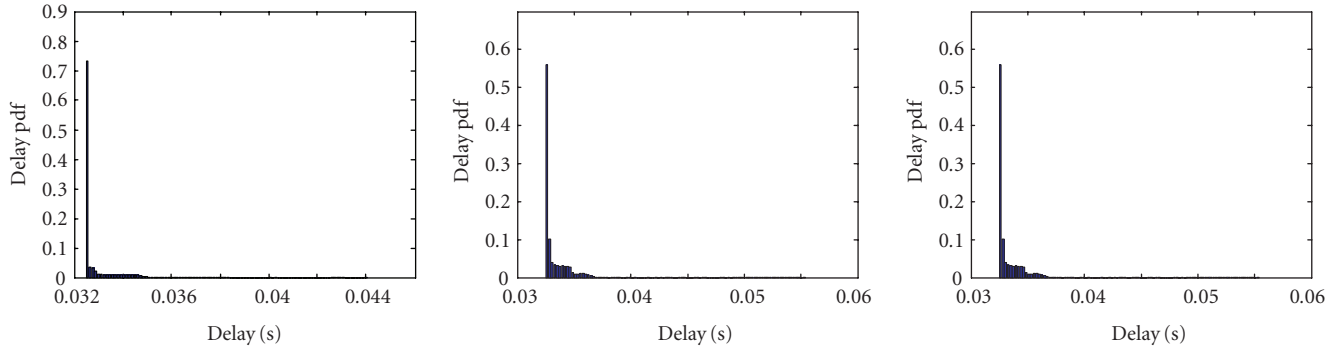
In order to account for the frame rate reduction as well, we used the objective quality parameter Q proposed in [22], defined as

$$Q = 0.45 \cdot psnr + \frac{(fr - 5)}{10} - 17.9, \tag{4}$$

where $psnr$ is the PSNR value measured at the destination, after error concealment processing, while fr is the frame rate of the video sequence perceived at destination, counting decoded frames only. The constant coefficients in (4) were calculated in [22] by evaluating the data set obtained in a survey, and assuming a minimum acceptable frame rate of 5 frame/s. According to the above definition, the greater the PSNR and the frame rate at destination are, the greater the Q parameter is.

Given that the WMN is made up of wireless lossy links, usually constituting bottlenecks due to their low transmission capacity, the Internet is considered as lossless, jitter and losses being introduced by the WMN only.

Figures 8 and 9 show the average value and the measured standard deviation of the end-to-end delay, respectively. We can see that multipath routing allows a lower average delay to be achieved, as compared to single-path routing; however, it introduces a larger delay jitter, due to the fact that packets



(a) Multipath 200 kbit/s (b) Multipath 400 kbit/s (c) Multipath 600 kbit/s

Figure 11: End-to-end delay distribution for multipath routing.

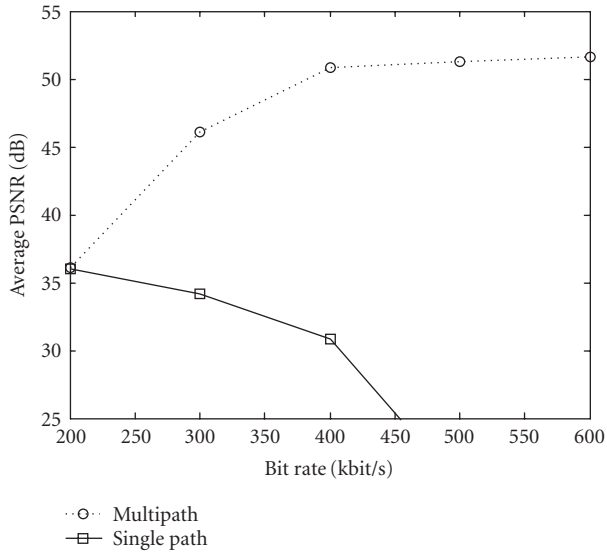


Figure 12: Average PSNR.

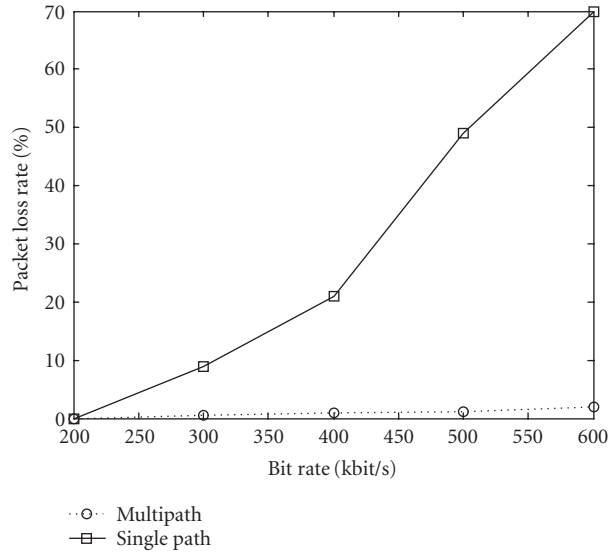


Figure 13: Packet loss rate.

follow different paths, and therefore may experience different delays. In order to highlight this phenomenon better, Figures 10 and 11 present the end-to-end delay probability distributions for both the single-path and multipath routing techniques, respectively.

By comparing all the figures from 8 to 11 we can deduce that multipath routing causes higher jitter values, which have to be compensated for by the jitter compensator buffer at the PPS. To this end, delay distributions are used to choose the value of the threshold σ_J leaving on its right a negligible portion of probability, representing the percentage of packets that are lost if the jitter compensator buffer equalizes delays to the chosen threshold σ_J. Of course, the greater the value of σ_J is, the less the loss percentage introduced by the jitter compensation buffer is, but the higher the equalization delay is. In our system we chose σ_J such that 0.1% of packets suffer a delay greater than σ_J, and are therefore discarded.

In order to evaluate the QoS perceived at destination, we first calculated the following:

(i) the average PSNR, measured at the destination side as specified in (2) and (3) on the frames fully or partially arrived and decoded (Figure 12);
(ii) the packet loss rate in the WMN network (Figure 13);
(iii) the video frame corruption percentage (Figure 14), and the consequent effective frame rate, *fr*, measured at destination (Figure 15), obtained as the ratio of the number of frames that have been decoded (also thanks to the application of the error concealment decoding technique) over the measurement period.

Figure 12 shows the *psnr* term, defined in (4) as the PSNR calculated at the destination, after error concealment processing. We can observe that the *psnr* obtained with multipath routing is higher than that obtained with single-path

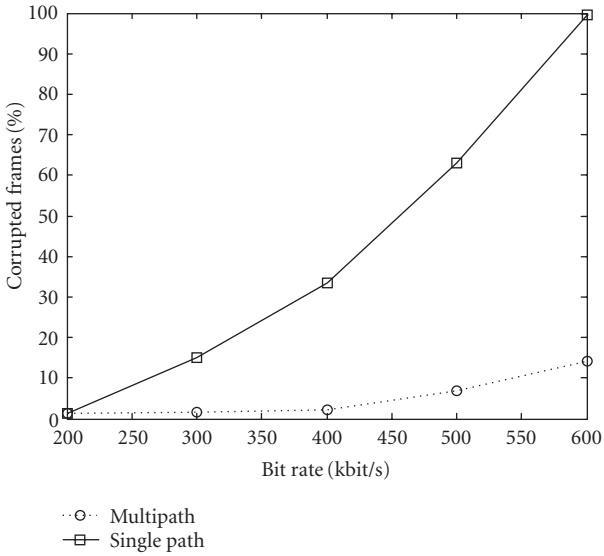


Figure 14: Video frame corruption percentage.

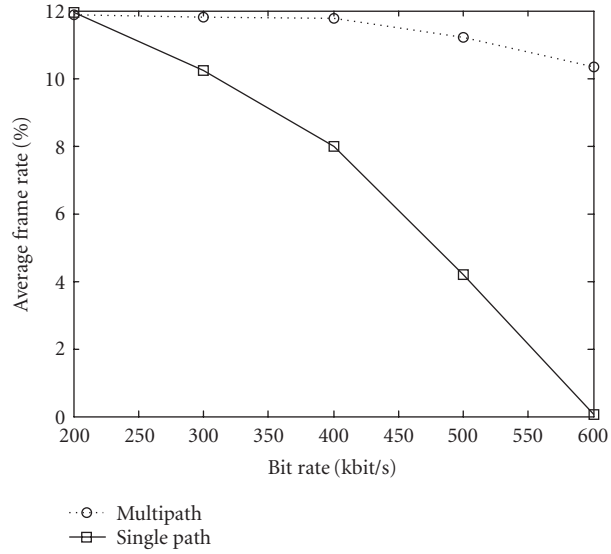


Figure 15: Average video frame rate, *fr*.

routing. In this case, in fact, the reduced packet loss rate in the network allows the error concealment algorithm run at destination to work better, therefore providing frames with a better quality, more similar to the original ones. However, when the encoding bit rate is too high (over 400 kbit/s), the PSNR increase at the source side corresponds to PSNR degradation due to network losses, and the PSNR therefore exhibits a flat trend. Of course, with encoding bit rate values higher than 600 kbit/s, not shown here because of being unrealistic due to the enormous loss rate, the curve would have exhibited a decreasing trend. On the other hand, the huge number of losses encountered with single-path routing, which increases with the encoding bit rate, causes a decreasing PSNR trend, although the PSNR at the source increases.

Figures 13 and 14 present the packet loss rate in the WMN network, and the consequent video frame corruption percentage. From these figures we can notice that when the output bit rate increases, the destination frame rate achieved with single-path routing soon becomes too low, while multipath routing allows the source to encode at a high rate while maintaining a high destination frame rate: losses remain low up to 400 kbps.

As shown in Figure 14, with low bit rate values, we can reduce packet losses by decreasing the video source transmission bit rate. In fact, by decreasing it, the probability of a packet being discarded decreases, and the received video quality grows.

Finally, Figure 16 summarizes the QoS perceived at destination by showing the overall objective quality parameter Q defined in (4), and demonstrates the power of multipath routing in guaranteeing a perceived QoS greater than that achieved by single-path routing with any video source output rate. The behavior of this parameter is determined by both the *psnr* parameter shown in Figure 12, and the *fr* parameter shown in Figure 15. We can observe that when single-path routing is used the overall quality decreases with increasing

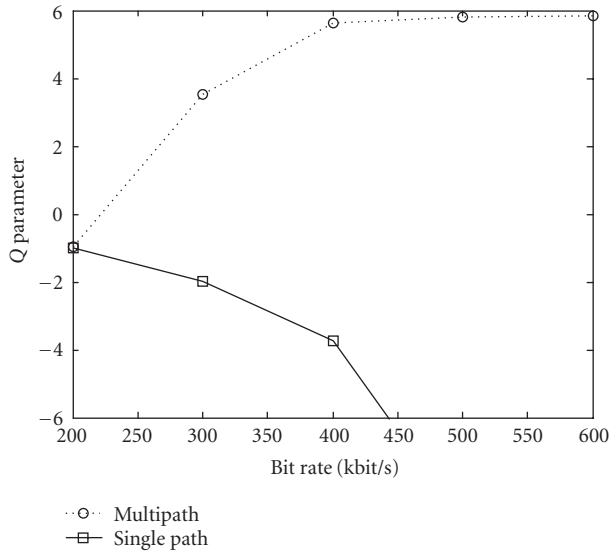


Figure 16: Objective parameter Q.

encoding bit rates, and the best quality is achieved with the minimum considered encoding bit rate, equal to 200 kbit/s. On the contrary, using multipath routing allows us to encode at a higher bit rate the best being between 500 and 600 kbit/s.

To summarize, taking into account that multipath routing, besides robustness to external attacks and interceptions, provides a higher decoding quality and less delay than single-path routing, it is the best solution for the proposed video-surveillance system. The only problem of multipath routing is that delay jitter is higher, but this can be compensated for by a compensation buffer at destination.

6. CONCLUSIONS

This paper describes a real experience of a wireless video-surveillance system, illustrating the overall architecture and the structure of each component block. Specifically, video sources use rate control to emit a constant bit-rate flow, while the access network is a WMN implementing a multipath routing algorithm to minimize delay and intrusions. However, this causes jitter, which is not acceptable for video-surveillance applications but can be compensated at destination if delay statistics are known. Analysis is carried out against the emission bit rate, and quality perceived at destination is evaluated with an objective parameter. Numerical results have demonstrated that multipath routing guarantees less delay and the best quality at destination. So it is the best solution for the proposed video-surveillance system with any encoding bit rate.

ACKNOWLEDGMENTS

The authors wish to thank the anonymous reviewers for their detailed reviews and many constructive suggestions which have improved the paper significantly. The work was partially supported by the Italian Ministry for University and Scientific Research (MIUR) through the BORA-BORA Prin project under Grant 2005097340.

REFERENCES

[1] C. Peng, "Introduction to Video-Surveillance Systems over the Internet Protocol," Real World Video & Imaging: Texas Instruments White Papers, October 2003.

[2] R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, "Algorithms for cooperative multisensor surveillance," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456–1477, 2001.

[3] C.-F. Chiasserini and E. Magli, "Energy consumption and image quality in wireless video-surveillance networks," in *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '02)*, vol. 5, pp. 2357–2361, Lisbon, Portugal, September 2002.

[4] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, 2005.

[5] Mesh Networks, http://www.meshnetworks.com/.

[6] Radiant Networks, http://www.radiantnetworks.com/.

[7] R. Karrer, A. Sabharwal, and E. Knightly, "Enabling large-scale wireless broadband: the case for TAPs," in *Proceedings of the 2nd Workshop on Hot Topics in Networks (HotNets-II '03)*, Cambridge, Mass, USA, November 2004.

[8] P. Bhagwat, B. Raman, and D. Sanghi, "Turning 802.11 inside-out," in *Proceedings of the 2nd Workshop on Hot Topics in Networks (HotNets-II '03)*, Cambridge, Mass, USA, November 2003.

[9] Tropos Networks, http://www.tropos.com/.

[10] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MOBICOM '04)*, pp. 114–128, Philadelphia, Pa, USA, September-October 2004.

[11] W. Feng, J. Walpole, W. Feng, and C. Pu, "Moving towards massively scalable video-based sensor networks," in *Proceedings of the Workshop on New Visions for Large-Scale Networks: Research and Applications*, pp. 12–14, Washington, DC, USA, March 2001.

[12] PeopleVision Project, IBM Research, http://www.research.ibm.com/peoplevision/.

[13] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Senior, and Y. Tian, "Smart surveillance: applications, technologies and implications," in *Proceedings of the Joint Conference of the 4th International Conference on Information, Communications and Signal Processing, and the 4th Pacific Rim Conference on Multimedia*, vol. 2, pp. 1133–1138, Singapore, December 2003.

[14] I. Haritaoglu, D. Harwood, and L. S. Davis, "W^4: real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809–830, 2000.

[15] T. N. Tan, G. D. Sullivan, and K. D. Baker, "Model-based localisation and recognition of road vehicles," *International Journal of Computer Vision*, vol. 27, no. 1, pp. 5–25, 1998.

[16] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.

[17] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proceedings of the 4th IEEE Workshop on Applications of Computer Vision (WACV '98)*, pp. 8–14, Princeton, NJ, USA, October 1998.

[18] S. E. Kemeny, R. Panicacci, B. Pain, L. Matthies, and E. R. Fossum, "Multiresolution image sensor," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 4, pp. 575–583, 1997.

[19] T. Boult, "Frame-rate multi-body tracking for surveillance," in *Proceedings of DARPA Image Understanding Workshop*, pp. 305–308, Monterey, Calif, USA, November 1998.

[20] A. Basu and D. Southwell, "Omni-directional sensors for pipe inspection," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3107–3112, Vancouver, BC, Canada, October 1995.

[21] Cisco Systems® documentation, "Intelligent Wireless Video Surveillance Solutions," http://www.cisco.com/en/US/products/hw/routers/ps272/prod_brochure0900aecd804a8cad.html.

[22] "People-Mover Project Brings 21st Century Surveillance System to Dallas Airport," Telindus, 2002 http://www.telindus.com/resources/ref_case_dallas-screen.pdf.

[23] V. Chandramohan and K. Christensen, "A first look at wired sensor networks for video surveillance systems," in *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks (LCN '02)*, pp. 728–729, Tampa, Fla, USA, November 2002.

[24] "Firetide, Axis partner to deliver wireless video surveillance," *Telematics Journal*, September 2006, http://www.telematicsjournal.com/content/newsfeed/8344.html.

[25] A. Doblander, A. Maier, and B. Rinner, "Increasing service availability in intelligent video surveillance systems by fault detection and dynamic reconfiguration," in *Proceedings of the Telecommunications and Mobile Computing Workshop on Wearable and Pervasive Computing (TCMC '05)*, Graz, Austria, March 2005.

[26] Y. Zhang and K. Chakrabarty, "Energy-aware adaptive checkpointing in embedded real-time systems," in *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition (DATE '03)*, pp. 918–923, Paris, France, March 2003.

[27] S. Mueller, R. P. Tsang, and D. Ghosal, "Multipath routing in mobile ad hoc networks: issues and challenges," in

Performance Tools and Applications to Networked Systems, M. C. Calzarossa and E. Gelenbe, Eds., vol. 2965 of *Lecture Notes in Computer Science*, pp. 209–234, Springer, Berlin, Germany, 2004.

[28] E. Biagioni and S. H. Chen, "A reliability layer for ad-hoc wireless sensor network routing," in *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS '04)*, pp. 4799–4806, Big Island, Hawaii, USA, January 2004.

[29] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks," in *Ad Hoc Networking*, chapter 5, pp. 139–172, Addison-Wesley, New York, NY, USA, 2001.

[30] V. D. Park and M. S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proceedings of the 16th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '97)*, vol. 3, pp. 1405–1413, Kobe, Japan, April 1997.

[31] C. E. Perkings, E. Belding-Royer, and S. R. Das, "Ad hoc on-demand distance vector routing," RFC 3561, 2003.

[32] S. Vutukury and J. J. Garcia-Luna-Aceves, "MDVA: a distance-vector multipath routing protocol," in *Proceedings of the 20th Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '01)*, vol. 1, pp. 557–564, Anchorage, Alaska, USA, April 2001.

[33] "Coded Representation of Picture and Audio Information," International Standard ISOIEC/JTC1/ Sc29/WG11, MPEG Test Model 2. July 1992.

[34] "Coding of Moving Pictures and Associated Audio for Digital Storage Media up to 1.5 Mbit/s—Part 2," Video. International Standard ISO-IEC/JTC1/SC29/WG11, DIS11172-1. March 1992.

[35] A. Lombardo and G. Schembra, "Performance evaluation of an adaptive-rate MPEG encoder matching intserv traffic constraints," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 47–65, 2003.

[36] C.-F. Chang and J.-S. Wang, "A stable buffer control strategy for MPEG coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 6, pp. 920–924, 1997.

[37] A. Cernuto, F. Cocimano, A. Lombardo, and G. Schembra, "A queueing system model for the design of feedback laws in rate-controlled MPEG video encoders," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 4, pp. 238–255, 2002.

[38] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 1, pp. 12–20, 1996.

[39] X. Yuan, Z. Sun, Y. Varol, and G. Bebis, "A distributed visual surveillance system," in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '03)*, pp. 199–204, Miami, Fla, USA, July 2003.

[40] S. Vutukury and J. J. Garcia-Luna-Aceves, "A simple approximation to minimum-delay routing," in *Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM '99)*, pp. 227–238, Cambridge, Mass, USA, August-September 1999.

[41] S. Vutukury and J. J. Garcia-Luna-Aceves, "A practical framework for minimum-delay routing in computer networks," *Journal of High Speed Networks*, vol. 8, no. 4, pp. 241–263, 1999.

[42] Y.-C. Lee, Y. Altunbasak, and R. M. Mersereau, "Multiframe error concealment for MPEG-coded video delivery over error-prone networks," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1314–1331, 2002.

[43] S.-C. Pei and Y.-Z. Chou, "Novel error concealment method with adaptive prediction to the abrupt and gradual scene changes," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 158–173, 2004.

[44] S. Tsekeridou and I. Pitas, "MPEG-2 error concealment based on block-matching principles," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 4, pp. 646–658, 2000.

[45] "The Click Modular Router Project," http://pdos.csail.mit.edu/click/.

[46] R. Morris, E. Kohler, J. Jannotti, and M. F. Kaashoek, "The click modular router," in *Proceedings of the 17th ACM Symposium on Operating System Principles (SOSP '99)*, pp. 217–231, Kiawah Island Resort, SC, USA, December 1999.