Abstract—SafeCOP (Safe Cooperating Cyber-Physical Systems using Wireless Communication) is an European project that targets cyber-physical systems-of-systems whose safe cooperation relies on wireless communication. In particular, SafeCOP will provide an approach to the safety assurance of such systems in the healthcare, maritime, vehicle-to-vehicle and vehicle-to-infrastructure sectors. The vehicle-to-infrastructure (V2I) cooperation for the traffic management scenario is a key use cases of SafeCOP, where the cooperation between different cyber-physical systems can also include a significant interaction with remote servers. In this paper, we outline the V2I traffic management scenario and assess the research goals that arise from it, with special focus on the IoT characteristics.

1. Introduction

According to the U.S. Department of Transportation, up to 80% of automobile accidents can be prevented with improved vehicle connectivity, using Vehicle-to-Anything (V2X) communications [1]. Additionally, it is projected that over 10% of the driving time is wasted in traffic jam, 12% of urban traffic is created by vehicles trying to park, and up to 17% of urban fuel is wasted at traffic lights when there is no cross-traffic. Cooperative-Intelligent Transportation Systems (C-ITS) [2] can be very useful when situations such as construction site warnings and traffic congestion in highways caused by an accident or road damage are encountered. With vehicles connected together and connected with roadside infrastructure, all of these problems can be mitigated. It is worth noting that C-ITS require more stringent safety requirements and standardization procedures to enable the cooperation among systems developed by different producers. However, V2X communications and driver assistance platform large scale diffusion can be negatively impacted by security concerns. This consideration arises from a 2014 UMTRI poll [3]: Some 30% of those questioned said they are ‘very concerned’ about security breaches from hackers, and about data privacy in tracking speed and location. Another 37% are ‘moderately concerned’. In addition, most expressed concern about system failure and performance, and about drivers relying too much on the technology or being distracted by it. Moreover, it is important to highlight that people percepts the usefulness of V2X and driver assistance services. According to the same poll “…75% of respondents believe that connected vehicles will reduce the number and severity of crashes, improve emergency response times and result in better fuel economy”. Thus, it clear that the acceptance of V2X technology is not limited by its perceived value, but by the potential security and safety issues.

Safety assurances for vehicles relying on V2X communications require reliable cooperation among the components of the whole system, as the reliability of the whole system allows both to prevent accidents and to detect malfunctioning subsystems. The problem is twofold: on one hand it is necessary to develop innovative and efficient solutions to increase the security and safety of these systems; on the other hand it is necessary to standardize the security assurance procedures, and to use prototypes and demonstrator to reassure people about security concerns and foster their wide utilization. A key point is that a fast widespread adoption of V2X technology is crucial for it to provide significant benefits. Indeed, in a context where V2X-equipped vehicles are the norm, any vehicle on the road not V2X-equipped has the potential to increase the risk of accidents, and therefore to decrease the overall impact of V2X technology.

Moreover, V2X technology acts as a support tool for autonomous vehicles, which are under development and foreseen to reach maturity in the next decade. Google, Nissan, and other companies expect that by 2020 cars on the market will be completely driverless in many situations. These cars will employ sensors to detect other vehicles, pedestrians, and objects around them, but this information
can be integrated with additional data coming from roadside monitoring systems, using V2I communication, improving the quality and speed of the detection, as well as providing information that cannot be detected by sensors at all (e.g., the timing of traffic lights ahead).

For more efficient drive assistances and future autonomous vehicles a large amount of information needs to be communicated from roadside monitoring units to the vehicles in transit. At the same time, an efficient monitoring roadside system needs to acquire information from the vehicles in transit. Such information can be obtained in many ways, including Adaptive Traffic Light System (A-TLS) Co-operative Awareness Message (CAM), Green Light Optimal Speed (GLOSA) Cameras, radar, lidar (light detection and ranging) and laser technologies. There are already proposals for combining CAM with automated traffic lights [4].

1.1. Contributions

We propose an holistic approach to integrate the advantages offered by technologies such as CAM, GLOSA and A-TLS, in a Traffic Management Application to enhance road safety via functions such as traffic condition warning (rapid traffic evolution), stationary vehicle warning (disabled vehicles), and wrong way driving warning. In the SafeCOP project we explore the integration, on the V2I network, of the traffic management application and a Video Content Analysis (VCA) platform (acquisition from video cameras and elaboration algorithms) for detecting possible dangerous road events/situations (such as vehicles slowing down, vehicles queue, motionless objects) and contributing to the active road safety, to alert drivers of such traffic anomalies.

Specific industrial and research challenges in the above framework deal with the need for a robust approach using sensor fusion to ensure reliability for drive assistance (and, in the future, for autonomous vehicles) Sensor fusion allows to rely on multiple, heterogeneous data sources to extract relevant knowledge to take decisions. In the SafeCOP Traffic Management scenario, sources are the results of locally performed analyses on individual sensors, including Video Content Analysis (VCA) on roadside camera feeds, as well as navigation and driver behaviour data from in-vehicle sensors. Machine learning is crucial for performance discovery of autonomous vehicles and driver assistance systems due to the large amount of data and its heterogeneity. The application of machine learning theory to the field of C-ITS represents an innovative contribution.

Moreover, SafeCOP will need to guarantee security of communication. While confidentiality of the individual communication link can be guaranteed via encryption, the authenticity and trustworthiness of the endpoints needs to be established [1]. In SafeCOP, we will investigate the engineering trade-offs between security security level and performance of the crypticographic subsystem included in the RSU and OBU platforms.

2. Organization of the paper

In Section 3 we introduce the V2I traffic management use case scenario. In Section 4 we provide an overview of the main technologies involved in the solution. In Section 5 we focus on Video Content Analysis and its computational and communication requirements, which provide the main constraints for the components of the system residing on the remote server. Finally, in Section 6, we draw our conclusions.

3. Industrial Use Case Scenario

Intelligent Transportation Systems (ITS) optimize the efficiency and improve the safety of transportations, exploiting the possibilities offered by the state-of-the-art of Information and Communication Technology (ICT).

- **Active road safety** ITS applications class decreases the probability of traffic accidents by providing assistance to drivers (e.g. in order to avoid collisions with other vehicles).
- **Efficiency and management** ITS applications class optimizes the traffic flow, by coordinating the vehicles kinematics— speed and route—with respect to the traffic flow.\(^1\)

Both classes need a solid (in terms of reliability and response time) communication channel connecting all transportation actors, the vehicles and the infrastructure: wireless vehicular networks are the most important components of ITS enabling technologies. Vehicle to Infrastructure (V2I) communications are part of this use case.

Another part of the use case is the Adaptive Traffic Light System (A-TLS), which is an efficiency and management application. A-TLS changes the traffic lights signaling plan (the duration of red, yellow and green phases) at least according the time and the day. A better A-TLS optimizes the signaling plan according to the changing traffic conditions, usually by extending the green phase when vehicles are closely spaced. Currently, the time interval between two consecutive passing vehicles is measured, by inductive loops sensors.\(^2\)

Whilst A-TLS adapts the signaling plan to traffic conditions, another applications considered by the use case, the Green Light Optimal Speed Advisory (GLOSA), tries to adapt the traffic flow to the signaling plan. GLOSA computes the optimal vehicles speeds that minimizes the average (of all vehicles) travel cost (e.g. stop time at traffic lights and total travel time), for a given traffic lights signaling plan. Then GLOSA informs vehicles drivers about the optimal speed they should maintain, using some communications mean that, currently, consists of auxiliary roadside signs.

Figure 1 shows the use case block diagram.

\(^1\) There is a third class of miscellaneous applications (e.g. infotainment), which are not considered here.

\(^2\) An inductive loop is a wire, of square, circle or rectangle shape, that is installed into or under the surface of the road.
It is worth noting that, other than to improve driver comfort, efficiency and management ITS applications bring environmental benefits, because a smooth driving, with limited (de)accelerations and shorter travel time, reduces fuel consumption and CO₂ emissions.

Besides common sensor techniques, such as the use of inductive loop sensors, every vehicle will have V2I communication enabled. Also, by applying Video Content Analysis (VCA) techniques, on the installed video infrastructure, it is possible to detect passing vehicles, and in this manner to contribute to the traffic management. Moreover, VCA is capable of extracting information about dangerous situations, such as vehicle queues or vehicles moving along forbidden directions. Other examples include traffic condition warning (rapid traffic evolution), stationary vehicle warning (disabled vehicles), and wrong way driving warning (forbidden heading). V2I based efficiency and management applications can be extended in order to inform vehicles of such traffic anomalies. In this manner, vehicles cooperate through the infrastructure.

4. SafeCOP Technology Overview

This section explains the technical advances of the SafeCOP V2I experiment.

4.1. On Board and Road Side Units

The SafeCop V2I traffic management system employs a combination of Road Side Units (RSU) and On Board Units (OBU) to collect information about the ongoing traffic at a crossroad. The RSU has a single sensor, a camera, which is used to provide a video feed to the Control Center. The OBU constitutes the core system for collecting individual vehicle data, namely dynamic data (accelerations, angular speeds and magnetic field), position data (latitude/longitude, speed, heading) and — when available — vehicle data (brake, engine rpm, gear, etc.). To this purpose the system is equipped with the following main components: 1) A 9-DOF inertial measurement unit for acceleration, angular speed and magnetic field 2) A high-precision GPS receiver 3) A CAN bus interface to be connected to the OBD vehicle diagnostic interface. In addition to these sensing units, the OBU will host a microcontroller and a 3G/4G module to provide connectivity to the Control Center. The same communication interface is used also for configuration and diagnostic of the device itself. The processing of dynamic, position and diagnostic data will be partitioned into an online portion, performed by the OBU itself, and an offline part performed either in real-time or in batch mode in the Control Center. The on-line processing is structured according to a synchronous data streaming model and will consist of the following main phases.

**Acquisition.** Data from all the sensors is acquired at fixed but configurable sampling rates. The accelerometer and the gyroscope will be sampled at a frequency between 1.5kHz and 5kHz, while the magnetometer will be sampled at a much lower rate, typically around 20Hz to 40Hz.

**Pre-filtering.** Previous experience in the specific domain shows that acceleration and angular speed data have a significant bandwidth between 40Hz and 100Hz, while magnetometer data has a bandwidth around 10Hz. The pre-filtering phase is constituted by a low-pass anti-alias digital filter followed by suitable decimation.

**Synchronization.** The synchronization phase has the purpose of guaranteeing that the ratio between sampling frequencies of the different source — inertial unit, GPS and CAN bus — are constant over time, which is needed by the subsequent sensor fusion phase.

**Sensor fusion.** Raw dynamic and position data are the combined with a filtering scheme based on techniques such as Kalman filters or AHRS (Attitude and Heading Reference System).

**On-line processing.** Using the data produced by sensor fusion, this phase executes a set of algorithms that need to process the data in real-time. The main purpose of such algorithms is to compute a set of features with a much slower rate of change. As an example, maximum positive/negative acceleration, average acceleration and average speed will typically be computed over a 1 second period.

Data produced by the OBU is then transmitted via the 3G/4G module to the Control Center for further processing.

4.2. Hazard Detection Functions

The understanding of the security risks involved in connected vehicle streams through V2V/V2I communication is an hot topic of research [5]. Hazard detection presents a scientific challenge as the inference of anomalies may be derived in the absence of a-priori knowledge about how such abnormalities can be realized and how they can be measured through the available sensors. The concept of anomalies not only includes cyber-attacks, but also fault events (e.g., string instability of the vehicle platoon) or unexpected network congestion to be prevented. The proposed approach involves...
the use of innovative machine learning models [6], [7]. Even the feature extraction phase (i.e., the processing of the raw data before the application of machine learning algorithms) may include advanced statistical methods, such as spectral analysis and advanced statistics, such as the mutual information metric [8]. The most difficult task is the lack of a-priori knowledge of the anomalies (known in the literature as unsupervised learning) [9]. In order to identify anomalies, a “clustering” processing of data [10], [11] is performed. Roughly speaking, the process “brings together” all the data which can be considered as a regular, except for the few samples that could be linked to abnormalities. Clustering can be achieved through traditional algorithms such as $k$-means or be flanked by newer methods, such as one-class classification [12]. If, on the other hand, data can be a-priori characterized (labelled) as regular or anomalous, the process can be empowered by supervised learning that exploits the a-priori knowledge of the data. The aggregation of data through clustering is then processed to classify the aggregations (clusters) according intelligible rules. To do so, the Switching Neural Network (SNN) is used [7], [13]. It creates intelligible rules with an accuracy comparable to traditional black box techniques, such as neural networks. Intelligible rules mean a sequence of Boolean functions that link the data collected, their values and the fulfillment of certain conditions. The interested reader finds in [14] a detailed comparison between the SNN and traditional classification methods. Models based on intelligible rules and high accuracy are of main interest in this context because they drive effective actions by human operators, for example, by setting certain operational conditions in order to maintain a desired security level. The rules and the inherent actions may be easily integrated within the cyber-physical system as well. Since intelligible rules are expressed in terms of boolean functions, their application requires a very small amount of computational resources and therefore allows its efficient implementation even on simple hardware devices, such as an FPGA or an 8-bit microcontroller. The study is carried out through the Rulex platform [15]. Rulex contains an implementation of the SNN, with focus on both reliability (test extended on numerous case studies) and computational efficiency (speed calculation, memory optimization). The following figure shows all the elements that make Rulex a machine learning platform over the state of the art. Once the model has been derived from historical data, one can update it with new data (re-training). Depending on the impact on quality of the processing steps from feature extraction to the derivation of rules, the re-training stage may involve the intervention of an analyst who oversees the accuracy of the updated model or be fully automated.

4.3. Safety functions

Wireless cooperative systems in vehicular application scenarios enable a set of traffic management applications, including the above mentioned A-TLS, for the adaptation of signal plan to the instantaneous traffic conditions, and the GLOSA, which shapes the vehicles traffic by suggesting the optimal speed to reduce (de)acceleration and stopping time, aiming to reach the green light on time. Traditionally, this information is derived from sensors mounted under or above the road surface (referred to as the Infrastructure), and auxiliary roadside signs may be used to inform the drivers. These systems are currently developed independently; however, changes in the dynamic signal plan according to the instantaneous traffic conditions impose the selection of a different speed optimal value, i.e. the integration of these two systems will strongly improve the effectiveness of this traffic management function [16]. In this context, SafeCOP will design and demonstrate an integrated A-TLS and GLOSA system based on Cooperative Awareness Messages (CAM) that a vehicle periodically transmits on the V2I network, to inform the infrastructure of its position and speed (as well as other status information).

Road safety applications aim at providing drivers with information about critical situations in order to prevent accidents; these include Cooperative Forward Collision Warning, Approaching Emergency Vehicle, Emergency Electronic Brake Lights, Pre-crash Sensing Warning. All road safety functions rely on Vehicle-to-Infrastructure (V2I) communications, i.e. vehicles are equipped with short range wireless communication capabilities which collaborate to form a temporary distributed network enabling communications with road infrastructure nodes (or other vehicles) located in line of sight or even out of radio range (if a multi-hop network is built among vehicles). Possible dangerous road events/situations, such as vehicles slowing down, vehicles queue, motionless objects, etc., can be detected through a specific platform for the acquisition from video cameras and elaboration algorithms, which is called the Video Content Analysis (VCA) platform [17]. SafeCOP will contribute to the active road safety by exploring the integration, on the V2I network, of the traffic management application and a VCA platform, in order to alert drivers of emerging traffic anomalies. SafeCOP V2I networks will support communications for both management and safety information.

Furthermore, starting from the available standards from IEEE 802.11p, ITS-G5 and CEN, SafeCOP will develop new mechanisms for safe wireless communications between cooperative embedded systems in vehicles and for V2I communications, in order to enhance the automotive functional safety. This include a revision of the messages that enable the V2x applications, e.g. the CAMs (Cooperative Awareness Messages) and DENMs (Decentralized Environmental Notification Messages), transmitted by vehicles and infrastructure periodically.

4.4. Security

Cryptographic primitives are the foundational building-blocks to provide security and privacy assurances in complex computational and communication systems, and V2X communications are no exception. When the device is exposed to physical threats, cryptographic primitives may be subject to applied cryptanalysis techniques, i.e. implementation attacks. The largest class of implementation attacks
is represented by side-channel attacks (SCAs), where the attacker exploits the information leakage happening on an unintended channel, typically an environmental parameter of the computation which is dependent on the computed data. Instances of such side-channels include energy consumption, execution timing or electro-magnetic (EM) emanations: all these environmental parameters provide enough information to infer the value of secret data intended to be stored within the device in an otherwise un-accessible way.

Designing efficient and effective countermeasures against side-channel attacks is a topic which has received warm attention by the research community. Typically, countermeasures against the aforementioned threats involve modifying the cipher at either the algorithmic or the implementation level [18], [19], [20], [21], [22], [23], or changing the underlying hardware architecture so to suppress the side-channel leakage. In SafeCop, we will investigate the engineering trade-offs between security level and performance of the cryptographic subsystem included in the RSU and OBU platforms [24].

5. Video Content Analysis

The road-side monitoring system is equipped with video cameras and communicates the acquired information to the vehicles in transit. This system is mainly composed of a Video Content Analysis (VCA) platform running on the Control Center server in charge of monitoring the roads, and elaboration algorithms which generate warning messages. In more detail, this system is able to extract the following information about potentially dangerous situations, analysing the images acquired by video cameras opportunely installed along the roads: 1) presence of objects moving inside the reference area; 2) presence of motionless objects in the reference area for longer than a minimum time threshold; 3) detection of vehicles slowing down inside the scene; 4) presence of vehicle queues in the video camera scene; 5) detection of objects moving along a reference direction (e.g. vehicles moving in forbidden directions); 6) presence of people inside sensitive areas; 7) detection of smoke or fog in sensitive areas.

5.1. Adopted metrics

The requirements of the Video Content Analysis (VCA) systems are reported in this paper using the following three metrics: bandwidth, computational load, and storage capacity.

Bandwidth. The amount of bandwidth in bps necessary for the transmission of the video flow. This quantity is determined by the resolution, the number of frames per second and the compression level of the video.

Computational load. It is quite difficult to estimate the computational load because it is extremely conditioned by the particular application case. Consequently, in this work the analysis is proposed starting from real implementations of VCA systems. The adopted configuration is used as a measure of the computational load required. In more details, we have considered some parameters of the machine which perform the analysis: 1) characteristics of the processor and the number of core units used; 2) available RAM; 3) percentage of CPU used.

Storage capacity. It identifies the amount of memory required to store the video stream acquired.

Two different systems configuration can be adopted: the VCA algorithm can be implemented in a Network Video Recorder (NVR) or directly on the camera that acquire the video. Consequently, the video flow can be elaborated on the camera or coded and transmitted by the camera to a remote network element in charge of recording and elaborating it. The onboard elaboration it is possible only with some cameras (e.g. AXIS, HIKVISION, and partially BOSCH).

Further consideration are necessary: the requirements reported in this chapter are deeply affected by several other factors such as size of the monitored area, presence of moving object and dynamic background, presence of obstacles and not homogeneous lights. In general: i) the larger the monitored area the more stringent the requirements, ii) the more complex the monitored area and the more stringent the requirements.

5.2. Estimation of the monitoring area dimensions

Before starting the discussion of the requirements, expressed using the aforementioned metrics, it is necessary to estimate how large is the area which can be monitored by a single camera. Again it is quite difficult because the size of the area is affected by several heterogeneous factors. In any case a general idea can be the following: the camera needs to be installed at a certain height and not too close to the scene that is monitored and the higher the camera, the larger the area but the higher the resolution required. The minimum requirements are shown in Figure 2.

If these conditions are not respected, the monitoring process could be ineffective, because some objects in the foreground can be too large, covering other objects that cannot be seen.

In general, we have experimented in several applicative cases that with a camera at a height of at least 8 or 10 meters, with a resolution of 4 CIF PAL [704 x 576] it is possible to monitor an area of up to 80 x 60 meters, in open field, with no obstacles and obstructions.

Some examples regarding the monitoring area dimensions are shown below. In Figure 3 and Figure 4 are represented the images acquired by two cameras (C1 and
C2) installed on the Aitek building located in Genoa. The cameras are installed respectively at 10 and 5 meters from the ground.

Figure 3. Image from camera C1 in the Aitek building

Figure 4. Image from camera C2 in the Aitek building

5.3. Bandwidth

As previously said the bandwidth required is determined by the quality of the images, as consequence of a different resolution and frame rate. So the quantities reported in this chapter are referred to two different resolutions, adopted in two different use cases where the VCA algorithms are implemented and used in a real monitoring system. The two values reported below can be considered as the minimum and maximum amount of bandwidth required.

• **For a low resolution video** (i.e. CIF 352x288) the bandwidth in transmission required is equal to 1 - 2 Mbps.

• **For high resolution video** (i.e. 1920x1080): the bandwidth in transmission required is equal to 5-6 Mbps.

5.4. Computational load

Also for this metric, the higher the resolution of images and the frame rate of the video and the more stringent the computational requirements for the analysis. Moreover, also for this metric are considered two applicative cases to give two examples of measured computational loads of real monitoring systems. As previously said the requirements are expressed in terms of processor, CPU and the amount of RAM of the machine (NVR or camera) which performs the processing.

**Elaboration on NVR.** In the **Use Case 1** images in 896x504 resolution with 10 fps are elaborated by a NVR equipped with a processor xeon 2670 @ 2.5GHz with a single core for each of the 8 cameras which compose the system. In this case, we have measured an average use of about 70-80% of the CPU using the tool `top`. The overall amount of RAM available is equal to 16 GB of RAM.

In the **Use Case 2** are used 40 cameras to capture 720p video streams at 30 fps which are compressed, sent to a remote NVR and finally elaborate. The NVR received a video stream at 10 fps in CIF format. In this case the video analysis is performed by a machine with 16 GB of RAM. Moreover the machine is also equipped with two Xeon E5-2630 processors v2@2.4GHz octacore, for a total of 16 cores. Information concerning the percentage of CPU utilization will be available in the future, with a more precise measurement campaign.

**Elaboration on camera.** The computational load of the analysis executed on cameras depend on the model considered. The following results are obtained during preliminary tests on two different cameras:

- **HIKVISION:** this type of camera requires at least one ARM dual core processor to analyze a portion of a CIF image with a resolution of 5 fps. The amount of RAM required is approximately equal to 18-20 MB. In general, this type of camera provides better VCA performance than Axis models tested by the authors.

- **AXIS:** these cameras require at least Etrax4 processor to analyze a portion of a CIF image with a frame rate of 5fps. For these cameras the minimum RAM requested is equal to 27 MB.

Some cameras, such as those from Bosch, proprietary software is used, so that it is not possible to extract useful metrics from their elaboration. Nevertheless, the output of these proprietary preprocessing software can be used as input for the Video Content Analysis algorithms.

5.5. Storage capacity

This metric identifies the memory necessary to store the captured video. It can be a stringent requirement in the case of video processing executed on the camera, if the video has to be stored locally. On contrary, it could be less relevant in the case of processing on a centralized NVR, where storage is a less stringent requirements.

However, it is necessary to define a criterion for the quantification of the storage capacity required. A possible solution is to use this formula: \( b \times t / 8 \), where \( b \) is the video flow bandwidth (in [bps]) and \( t \) is the duration of the video that it is necessary to store (in [s]). For example to store a...
video of 1 Mbps for a whole day it is necessary a storage capacity approximately equal to 11 GByte.

6. Conclusions

SafeCOP is an ECSEL project, funded by the European Commission and several EU national governments, aiming at the definition of a framework for cooperating cyber-physical systems with safety and security concerns. Within SafeCOP, a relevant goal is to support V2X applications, which impose significant safety and security constraints. In this paper, we have provided an overview of the V2I cooperation for traffic management industrial use case, which will drive the development of the SafeCOP framework and serve as a benchmark for the validation and verification of several of its components. In particular, we have focused on the Video Content Analysis component, which provides the main driver for the server-side part of the V2I system-of-systems.

In the course of the SafeCOP project, a medium fidelity demonstrator system operated in a laboratory simulated environment will be developed. The purpose of the demonstrator will be to act as a proof-of-concept for SafeCOP. The demonstrator will be developed incrementally, starting from an initial version of the Communication Generator and the Traffic Management Application, which will be initially developed separately to provide two simulation environments to integrate in the early demonstrator.

Acknowledgements

This work has been funded in part by the European Commission and the Italian government under grant 692529-2 of the ECSEL Joint Undertaking.

References


