# Chapter 1.

# Introduction

Digitalization of vertical industries is inevitable considering the growing demands for flexibility, agility, and cost-effectiveness. The usage of wireless communication in vertical industries has increased with the launch of Fifth Generation (5G) technology. 5G facilitates the Ultra Reliable Low Latency Communication (URLLC), reducing the end-to-end latency to 1 ms and increasing the reliability to $10^{-9}$, hence, promoting new applications and use cases. Applications and services like telesurgery, connected vehicles, Human-Machine Communication (H2M), smart factories for production and automation are necessities of future society. Most of these applications are driven by closed-loop feedback controls and are dependent on wireless technology. In recent years, advancements in wireless technology have led to major transformations in the envisioned automation industry, leading to Fourth Industrial Revolution (Industry 4.0).

Industry 4.0 will enable flexible wireless factory infrastructure, smart production and manufacturing, wireless control of mobile robots, and a high density of connected sensor devices and actuators. Industry 4.0 applications can be realized by numerous Industrial Internet of Things (IIoT) and sensor devices communicating, collaborating, and coordinating with each other over a wireless channel to process the information and execute a given task. These IIoT devices have limited computational and power capacity to process the acquired sensor data within latency constraints and provide the decision to actuators. Therefore, the complex computations need to be processed on a centralized server as shown in Fig. 1.1. In conventional networks (Long Term Evolution (LTE), Third Generation (3G), Second Generation (2G)) data is stored and processed on the centralized cloud server, causing high end-to-end latency, hence failing to fulfill the stringent URLLC requirements. Therefore, cloud computation near the edge is necessary for faster and more reliable processing. An edge cloud is the promising solution for faster computation and ultra-low latency requirements in Industry 4.0.

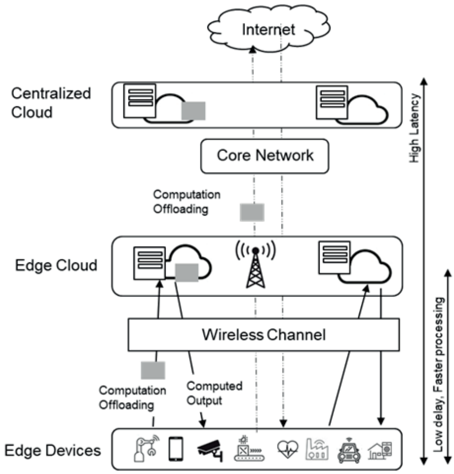


**Figure 1.1.:** Edge cloud offloading

The edge clouds are additional low power-consuming servers with computing and storage capabilities that are placed near the network edge. The offloaded computation can be executed on the edge cloud instead of the centralized cloud, as shown in Fig. 1.1. The servers at the network edge reduce latency by allowing user-plane processing/connectivity at the edge cloud, hence only requiring control plane connectivity via the centralized cloud. The edge devices are connected devices capable of offloading tasks to an edge cloud, unlike traditional Internet of Things (IoT) devices that can only process data in-device. An Automated Guided Vehicle (AGV) used in a factory for logistics, sensor devices from measurement and production plants, health monitoring devices, connected vehicles driving in a platoon, etc., are all edge devices that offload tasks to the edge cloud. The edge cloud is connected to several edge devices over a wireless channel and can process the computation from the diverse edge devices simultaneously, as shown in Fig. 1.1.

Industry 4.0 will rely to a large extent on intelligence at the edge to instantly process the massive information generated by IIoT devices. The communication requirements, such as data rate, latency, payload size, and frequency of information exchange in industrial control, vary with the applications. For example, a motion control application with closed-loop feedback requires a latency of 1 ms, whereas process automation can sustain a latency of up to 50 ms. The edge cloud placement may vary according to the application requirements in terms of latency, reliability, security, and privacy constraints. The edge cloud can be placed on the premises of an enterprise to protect technology and ensure data privacy or near the core network for latency-tolerant applications.

Although 5G technology can satisfy the stringent communication requirements by adapting Massive Multiple Input Multiple Output (mMIMO), edge cloud, and Millimeter Wave (mmWave) technologies, the integration of cellular technology with industrial controls still needs to be researched. The Wireless Networked Controlled System (WNCS) is a well-researched topic among control engineers in the last few decades. Many robust and adaptive controllers are designed to mitigate the effects of network errors. However, the potential gains of communication resources by designing control-aware strategies still remain unexplored. Therefore, a cross-domain approach for control-communication co-design is needed to optimally use the communication resources.

## 1.1. Background

The Programmable Logic Controllers (PLCs) were launched at the end of the 1960s to accelerate production and manufacturing processes. Since then, industrial communication standards and protocols have been developed to facilitate communication among sensor devices, controllers, and actuators by the International Electrotechnical Commission (IEC). In the last decade, the automation industry has evolved faster to accommodate time-critical applications using communication protocols and standards like Time Sensitive Networking (TSN) developed by the Institute of Electrical and Electronics Engineers (IEEE) 802.1 working group. TSN networks provide low latency and reliable data transmission over Ethernet, which is crucial for real-time applications. Moreover, TSN can precisely synchronize time between actuators, sensor devices, controllers, and network devices using the IEEE Precision Time Protocol (PTP). In recent years, upcoming applications and use cases require real-time wireless communication. The Third Generation Partnership Project (3GPP) has focused on developing IoT standards that can function over existing Fourth Generation (4G) and 5G networks. IoT standards like Long Term Evolution for Machines (LTE-M) and Narrow Band (NB)-IoT provide low power and low latency communication by connecting massive devices. Furthermore, standalone 5G-A and upcoming Sixth Generation (6G) technology will further enhance industrial wireless communication by providing low power, ultra-reliable low latency communication.

Additionally, concepts like network slicing can efficiently provide network resources by identifying the communication requirements of the application. Cellular network providers may offer network slices based on private networks to industrial customers. These cellular standards are developed with the intent to provide URLLC communication for the automation industry by analyzing its control requirements. A lot of research has been done to satisfy URLLC requirements that can be achieved by mMIMO technology, allocating higher bandwidth, or transmitting over sub-6GHz bands. However, few research studies are conducted to adapt and optimize the communication system by utilizing industrial control information. Therefore, the focus of my thesis is to develop control-aware, power-efficient communication strategies and algorithms to optimally use communication resources for industrial applications. Moreover, as energy efficiency is an essential aspect of IIoT devices due to their small battery power, wireless channel-aware task offloading strategies to minimize energy consumption are also investigated in the thesis.

## 1.2. Motivation and research questions

The IIoT devices are battery-limited and will require offloading complex data computations to the cloud for faster processing and to reduce in-device energy consumption. In the case of time-critical applications, computation must be offloaded, processed, and feedback provided within specified latency constraints. As the data will be offloaded to the cloud over an uncertain wireless channel, communication delay is introduced, and additional energy is consumed to transmit the data to the cloud. In the case of a bad channel condition, the data cannot be successfully transmitted and would require multiple transmissions, leading to higher energy consumption at the device. Since the wireless channel is time-varying, an optimal offloading strategy must be adopted considering the energy consumption of the device and the latency constraints. As there exists a trade-off between energy consumption for in-device computation and transmit energy, the following questions need to be addressed:

- Can offloading the computation to an edge cloud save energy for the sensor device?
- How to minimize the energy consumption of the device by optimally offloading over a wireless channel?
- How much data to offload over a wireless channel under given latency constraints?

Moreover, in industrial control systems, communication between sensors, controllers, and actuators might need latency around 1 $\mu$s and have high data dependency within the control entities. For example, in a closed-loop feedback process, the controller's output must be sent in real-time and applied as input to the actuator. Actuators provide feedback on the current state to the controller instantaneously to generate new control commands. These feedback control systems in an Industry 4.0 are mostly driven over a wireless channel as controllers and actuators can be located at a distant location. For example, an AGV in the warehouse or in a mining area can be controlled remotely from the building premises. Therefore, while offloading functionality to the cloud, data dependency and the impact of the wireless channel must also be considered. Therefore, in the thesis, we answer the following questions:

- Can a real-time feedback control system be partially offloaded?
- What is the impact of the wireless channel on the performance of a real-time control system?
- How are the communication and control aspects interdependent?

Furthermore, many practical control systems are nonlinear and time-varying systems. Most research papers consider a Linear Time Invariant (LTI) system to study the impact of the wireless channel on the control system. Therefore, we analyzed the impact of the wireless channel on a pragmatic nonlinear time-varying real-time system. Moreover, communication requirements vary with the control application; a one-size-fits-all system will lead to inefficient usage of communication resources. The communication system for Industry 4.0 needs to be further optimized by considering the control aspect. Therefore, a cross-domain approach is required to identify dependencies between the control and communication domain. The key characteristics of industrial communication are short data packets, URLLC requirements, and processing computation in an edge cloud over a wireless channel. The transmission of short packet size will have higher error probability due to finite block length coding and hence less redundancy. In case of high reliability requirements, the redundancy should be higher, whereas with lenient reliability requirements, the redundancy can be lower. Therefore, an optimal channel coding is required to adapt as per the requirement of the control system. Channel resources can be used more efficiently if the communication requirements of the control system are known. Therefore, control-aware communication strategies like channel coding, resource allocation are needed to efficiently utilize the communication resources. That instigates to find answers to the following questions:

- How to integrate control system aspects to efficiently use communication resources?
- Can the control-aware channel coding strategy be designed?
- Can control-aware resource allocation and scheduling schemes increase resource efficiency?

## 1.3. Research contributions and dissemination

### 1.3.1. Device energy optimization for an edge cloud offloading

To investigate whether offloading computation can minimize energy consumption, an analytical framework to optimally offload computation over a wireless channel is developed. Moreover, a channel-aware optimal offloading strategy is evaluated by solving an energy optimization problem that determines how much data to offload and under what channel conditions. Furthermore, the energy optimization problem is extended to process the offloaded computation within the latency constraints.

- S. Tayade, P. Rost, A. Maeder, and H. D. Schotten, Device-centric energy optimization for edge cloud offloading in 2017 IEEE Global Communications Conference (Globecom), 1-7.
- S. Tayade, P. Rost, A. Maeder and H. D. Schotten, "Delay constrained energy optimization for edge cloud offloading," 2018 IEEE International Conference on Communications (ICC) Workshops.

### 1.3.2. Edge cloud controlled AGV system over a wireless channel

To study the impact of an edge cloud-controlled real-time system in an Industry 4.0 setting, we consider an AGV use-case. The AGV is controlled by an edge cloud-based controller over a wireless channel. We evaluate the impact of channel outages and delays on the stability and performance of the control system. Furthermore, analytical results are evaluated to describe the interdependencies between the control and communication channels, which can be used for cross-domain optimization.

- S. Tayade, P. Rost, A. Maeder and H. D. Schotten, "Cloud Control AGV over Rayleigh Fading Channel - The Faster The Better," SCC 2019; 12th International ITG Conference on Systems, Communications and Coding, Rostock, Germany, 2019, pp. 1-6.
- S. Tayade, P. Rost, A. Maeder, "Error Convergence Analysis and Stability of a Cloud Control AGV" NGNA 2021 Kaiserslautern.

### 1.3.3. Control aware communication strategies

We also proposed cross-domain strategies to optimize communication resources for a cloud-based AGV system. The thesis focuses on evaluating the optimal coding rate and

Modulation and Coding Scheme (MCS) scheme for a stable control system. Moreover, we proposed control-aware communication Key Performance Indicators (KPIs) for resource allocation and scheduling, which increase resource efficiency.

- S. Tayade, P. Rost, A. Maeder and H. D. Schotten, "Impact of Short Blocklength Coding on Stability of an AGV Control System in Industry 4.0," 2020 IEEE International Conference on Communications Workshops (ICC Workshops), Dublin, Ireland, 2020, pp. 1-6.
- S. Tayade, P. Rost, A. Maeder and H. D. Schotten, "Resource Allocation with Stability Constraints of an Edge-cloud controlled AGV," European Wireless 2022 VDE ITG society Dresden Germany.

## 1.4. Structure of work

In Chapter 2, formal problem statements and an evaluation methodology are described in detail. Chapter 3 presents device-centric energy optimization solutions for edge cloud offloading. The analytical solutions, theorems, and algorithms for optimal offloading over a wireless channel are also presented. In Chapter 4, a real-time use case of a cloud-based AGV controller is considered to study the impact of the wireless channel on closed-loop feedback systems. The impact of channel outages and delay tolerance on the stability of the control system is evaluated. Moreover, control-aware communication strategies for channel coding short packets in industrial communication are presented in Chapter 5. The control-aware resource allocation and scheduling algorithm to increase resource efficiency is presented in Chapter 6.

# Chapter 2.

# Formal Problem Statements and Evaluation Methodology

The Chapter presents the detailed problem description and challenges addressed in the thesis. In every section a specific problem is formally elaborated and formulated. Along with the problem formulation, we further describe in brief the existing as well as the proposed solution approaches and evaluation strategies.

## 2.1. Challenges of edge cloud offloading

According to [Eri19], 25 % of Industry 4.0 usecases will require edge computing by 2023. Moreover, the power consumption of the edge devices can be substantially reduced by skillfully offloading the computation to the edge cloud. As edge computing may become the driving technology for future applications, a deeper investigation and analysis on the performance gains is necessary due to its dependency on dynamic and uncertain factors as shown in Fig. 2.1. Offloading the computation depends upon the factors like latency constraints, the amount of offloading data and the dependencies, as the data dependent algorithms and codes are required to be partitioned before offloading. Furthermore, other factors like the processing and power capabilities of a device and an edge cloud, as well

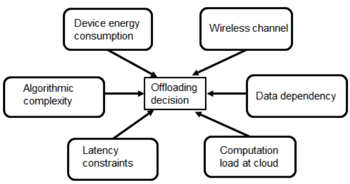


**Figure 2.1.:** Factors for edge cloud offloading

as load conditions at the edge cloud should be considered to avoid longer waiting times in the queue before processing.

One of the major challenges for offloading data is the time varying wireless channel. Many frameworks have been developed with the focus on saving energy by optimally offloading computations [CBDk+10, SSB14]. However the focus is on partitioning the source code during run-time, Virtual Machine (VM) migration assuming constant network specifications like bandwidth, Packet Error Rate (PER) etc. Due to changing network conditions it is crucial to adapt offloading decisions with respect to wireless channel conditions. Therefore, we investigate device centric performance gains achieved by optimal offloading decision considering the varying channel and factors as shown in Fig. 2.1.

### 2.1.1. Energy consumption trade-off

The IIoT consists of numerous small devices with limited battery and processing capabilities. The main challenge to overcome is the faster draining batteries that would make these devices non-functional, hence, increasing the maintenance cost. Offloading the computation to an edge cloud might reduce the energy consumption at these devices. However an additional energy is consumed for offloading the data over a wireless channel. As shown in Fig. 2.2, $E_{tr,i}$ is the energy consumption of the $i^{th}$ edge device for transmitting the computational data $D_i$ to an edge cloud and $E_{u,i}$ is energy consumed for in-device execution. The $E_{tr,i}$ increases for transmitting the data if the edge device is at farther distance from the edge cloud due to path-loss and fading. The device needs higher transmit power, lower MCS and several re-transmissions if the channel is in deep fade causing higher usage of device energy and communication resources (bandwidth) to transmit the data to the cloud.

Moreover, $E_{tr,i}$ can exceed the $E_{u,i}$ if the communication load is higher than computational load. The communication load is the amount of data required to transmit to an edge cloud for processing while computational load depends upon the complexity of the data processing. For example a multiplication of any given matrix with itself requires less computations for processing whereas the inverse of matrix require complex computations. The data that is required to be communicated to an edge cloud is the matrix itself, hence the multiplication operation can be executed in the edge-device, as only single operation is require, whereas matrix inverse that have higher computational load will need an edge cloud processing as the $E_{tr,i} < E_{u,i}$. Therefore, to reduce the energy consumption at the device, it is necessary to make an optimal offloading decision considering the Energy trade-off for transmitting the data and for in-device execution.