Journal of Progress in Safety & Security

Research Article

TU Delft OPEN Publishing

An Extended CIA-Based Multi-Level Model for AHP-Driven Safety and Security Decision-Making in Last-Mile Robotic Systems

Christina Kolb [1,*] and Lin Xie [2]

[1]BMS-IEBIS, The University of Twente, The Netherlands

[2]Information Systems and Business Analytics, Brandenburg University of Technology Cottbus - Senftenberg, Germany

*Corresponding author: c.kolb@utwente.nl

Abstract: The rapid growth of e-commerce and the increasing demand for efficient last-mile logistics have led to the rising use of last-mile robots. While these robots promise faster and cheaper deliveries, their operation in complex and dynamic urban environments introduces significant safety and security challenges. Sensor failures, communication disruptions, and cyber-physical attacks may affect the behaviour of the robots and influence human safety.

This work models and analyses these challenges using the Extended Multi-Level Model to represent the different components of last-mile robotic systems and their influence on the environment. We apply Multi-Criteria Decision Making (MCDM) for parallel safety and security risk assessment, focusing on the confidentiality, integrity, and availability (CIA) of the last-mile systems. Considering these three properties together allows priorities to be set within the CIA triad, which is essential for financial and economic decision-making when only limited resources for countermeasures are available.

We extend the model to an Extended CIA Multi-Level Model that enables detailed evaluation of safety and security risks across all system levels. A case study involving robots transporting critical parcel contents demonstrates how confidentiality, integrity, and availability concerns arise throughout the model and how their violation may affect human safety. The approach supports structured decision-making and contributes to improving the safe and secure deployment of last-mile robots.

One sentence summary: This work models the safety and security challenges in last-mile robotic systems (LMRS) using the Extended Multi-Level Model and a Multi-Criteria Decision Making (MCDM) for risk assessment within the CIA triad.

Keywords: Autonomous robots, Multi-level model, Human safety, Safety and security co-engineering, Safety-security-dependability, Sensors, Confidentiality, Integrity, Availability.

Publishing history: Submitted: 20 December 2024; Revised: 09 May 2025; Accepted: 02 September 2025; Published: 29 December 2025

ISSN: 3050-4570
Vol. 1, 2025
DOI: 10.59490/pss.1.2025.8019

Cite as: Kolb, C., & Xie, L. (2025). An Extended CIA-Based Multi-Level Model for AHP-Driven Safety and Security Decision-Making in Last-Mile Robotic Systems. *Journal of Progress in Safety and Security, 1.* https://doi.org/10.59490/pss.1.2025.8019

1 Introduction

Given the rapid growth of e-commerce, the number of parcels delivered worldwide is expected to rise to 200 billion per year by 2025 - a dramatic increase from less than 90 billion in 2018 (von Szczepanski et al, 2021). This rise will increase the demand for last-mile logistics, which in turn will lead to an increase in traffic in cities and congestion. With the increase in delivery volumes, the costs associated with last-mile delivery have also risen significantly. In 2023, it is estimated that last-mile logistics will account for 53% of total delivery costs, up from 41% in 2018, mainly due to labour costs, which have increased due to the ongoing labour shortage (Pohowalla et al., 2024).

To overcome these challenges, the use of last-mile robots in logistics and urban environments is seen as a promising option to provide faster, cheaper and more efficient delivery options. As urban areas become more densely populated and customer demand for fast deliveries continues to rise, these robots are likely to play an even greater role in transforming logistics.

The rapid growth of last-mile delivery robots makes it clear that the safety and security of these systems is critical, especially given their frequent interaction with complex, dynamic environments. Unlike traditional robotic systems that operate in controlled or industrial environments, last-mile robots must navigate busy pavements, interact with pedestrians and are exposed to various risks. For example, sensor failures can prevent the detection of obstacles, or robots may be forced to make emergency stops in crowded areas, leading to safety concerns.

These challenges are exacerbated by environmental factors. Urban environments are inherently dynamic and subject to sudden changes, such as road closures, construction sites, temporary obstacles and unfavourable weather conditions. If a robot cannot adapt quickly to these disruptions - especially if it cannot connect to its control system due to network communication failures - there is a significant safety risk. For example, a robot may encounter a newly erected construction site barrier that is not marked on its map, or it may have difficulty navigating through heavy rain or snow, which could interfere with its sensors.

As these robots collect and process data to optimise their operation, data protection is becoming increasingly important. Protecting personal data and complying with data protection regulations is crucial, especially when robots are used in public spaces. One possible solution is to process the data before storing it to filter out sensitive information such as faces, although this can be computationally intensive.

Ultimately, there is a need to ensure that robots can safely and securely navigate unpredictable urban environments over the last mile - not only to protect the robots and the goods they transport, but also to maintain public confidence in these systems in urban spaces.

1.1 Overview of Related Work

For an overview of the recent development of last-mile delivery robots, covering operations, infrastructure, regulations and customer acceptance, see (Alverhed et al., 2024). In terms of safety, the importance of including safety modules in optimisation and navigation strategies has been highlighted in recent studies, such as (Li et al., 2020) and (Shaklab et al., 2023), who are paying increasing attention to the topic. These works emphasise the need to consider safety aspects in the navigation algorithms of autonomous last-mile robots to ensure their reliable operation in dynamic urban environments. The review paper on models for safety and security together (Nicoletti et al., 2021) points out that there is no formalism that assesses the risk for safety and security together in detail and that there is no safety model for robotics in general. It refers to safety as the absence of unintentional errors and to security as the absence of malicious attacks. We also relate safety to human welfare.

The publication (Kolb & Xie, 2024) contains a human layer to a model for safety and security of (Quamara et al., 2024) to analyse the impact of attacks on a human in the environment of robots. The paper (Kolb & Xie, 2024) presents a first idea on how to assess safety and security risks together, where safety refers to capturing a positive effect on human well-being, i.e., such that no last-mile robot collides with a human. We group the vocabulary for safety and security and investigate confidentiality, integrity and availability for last-mile robots in the multi-level model. In this work, we extend this model to an Extended CIA Multi-Level Model that allows risk assessment for safety and security decision making to improve the safety of a human while adding countermeasures to the last-mile robot system when there is limited budget available. Since last-mile robot attacks are mainly attacks on the communication network within a last-mile robot 2, i.e. sensor attacks, we focus on the confidentiality, integrity and availability of the data sent and the packages carried by the last-mile robots. The work on bow-ties (Abdo et al., 2018) identifies vulnerabilities via MCSs. The authors assign two probability levels to these cut sets: one for safety and one for security. This approach allows for trade-offs between safety and security, but the authors leave open a concrete method for decision making. For more detailed related work, we refer to section 2.

1.2 Our Contribution

This paper presents an original exploration of the security challenges unique to last-mile delivery robots. As the communication network for sending data between sensors and the computer within these robots plays an important role in the last-mile robot environment, we examine the CIA triad of confidentiality, integrity and availability for these robots. By defining these terms in the specific context of last-mile deployment, this study provides a basis for understanding the unique challenges faced by these systems compared to other robotic applications. It also provides a comparative analysis with existing robotic systems to highlight where these challenges diverge, and proposes models and approaches tailored to address security gaps in real-world last-mile delivery scenarios. This targeted focus on the unique operational context of last-mile delivery is an important contribution to the field, providing practical insights for improving the safe and reliable deployment of these systems.

Our key contributions are summarized as follows:

1. Development of an integrated model for confidentiality, integrity, and availability (CIA) assessment;
2. A fine-grained approach to security risk evaluation;
3. A structured framework for decision-making regarding safety and security countermeasures;
4. A method for refining safety and security concerns across multiple abstraction levels within the model; and

5. Demonstration of the proposed approach through a case study involving last-mile delivery robots.

1.3 Structure of this Work

The structure of this paper is organized as follows. In Section 2, we present related work, beginning with the case study of last-mile delivery robots (2.1) and an overview of potential attacks targeting them (2.2). We then introduce the Extended Multi-Level Model (2.3) and the CIA triad - confidentiality, integrity, and availability - as foundational security concepts (2.4). The background section concludes with an introduction to Multi-Criteria Decision Making (MCDM) (2.5), which will later be applied in the security evaluation of the CIA triad.

In Section 3, we propose the Extended CIA Multi-Level Model as an enhancement of the Extended Multi-Level Model. We detail its associated risk assessment methodology (3.1) using the last-mile delivery robots as a running example (3.2) and outline the decision-making process for selecting appropriate countermeasures (3.3).

Finally, Section 4 presents our conclusions and outlines directions for future work.

2 Background

2.1 Last-Mile Robots

An example of a last-mile delivery robot is illustrated in Figure 1. This robot features a battery-powered, four-wheeled design with a box-like structure and a cargo compartment situated on top. Its dimensions are 718 × 598 × 710 mm, allowing it to carry up to 15 kg of cargo at a maximum speed of 6 km/h.

To navigate its environment, the robot is equipped with three types of sensors: a LiDAR (Light Detection and Ranging) sensor, three RGB-D (Red, Green, Blue plus Depth) cameras, and eight ultrasonic sensors. The LiDAR sensor provides high-resolution 3D mapping and precise obstacle detection by emitting laser pulses and measuring their reflection time. The RGB-D cameras capture both color images and depth information, enabling object recognition, environmental mapping, and path planning. The ultrasonic sensors emit high-frequency sound waves to detect nearby objects, making them ideal for close-range obstacle avoidance, particularly in tight or low-visibility environments.

Additionally, the robot includes several other components, such as a microphone, a tablet with a speaker, and an antenna. The microphone allows for voice recognition or remote communication. The tablet with a speaker serves as an interactive interface for users, displaying delivery details or providing auditory feedback. The antenna facilitates wireless communication, enabling GPS connectivity, remote control, and real-time data transmission.

There are different types of last-mile robots in operation, such as Starship, which shares a similar technical framework. Another example is the FedEx SameDay Bot, designed to navigate sidewalks and roads to deliver parcels. Equipped with similar sensors, including LiDAR and cameras, it is built for safe and efficient delivery in urban and suburban areas. For a more detailed technical overview of the last-mile robot, please refer to (Kolb & Xie, 2024).

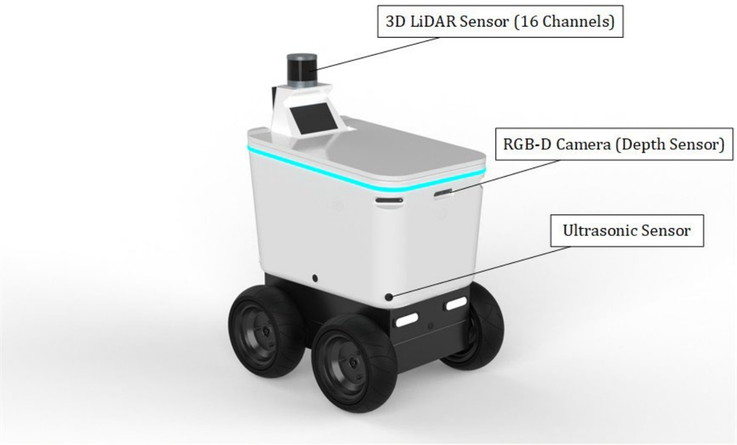

Fig.1: One example of the last-mile robot.

2.2 Potential Attacks on Last-Mile Robots

Last-mile robots share challenges with various robotic systems operating in dynamic environments. Autonomous vehicles (AVs) face similar issues with navigation, sensor failure, and cyberattacks (Cui et al., 2019). The safety frameworks from the automotive industry, such as SAE J3061 for cybersecurity and ISO 26262 for functional safety, can provide valuable insights into enhancing the security and safety of robotic systems. Drones are vulnerable to their dependence on communication networks (Tu et al., 2024); their regulations (e.g., FAA guidelines) can enhance network resilience for last-mile robots. Healthcare robots prioritize human-robot interaction and safety (Fosch-Villaronga et al., 2021), offering valuable insights for pedestrian interactions. Collaborative robots (cobots), focused on safety in human environments (Arents et al., 2021), provide lessons on real-time monitoring and collision avoidance. An overview of cybersecurity aspects about robotics is given in (Tanimu & Abada, 2024).

Last-mile robots face diverse security and safety risks, classified as *risks from signals, codes, and objects*, as outlined in (Tu & Piramuthu, 2023) for last-mile drones. A critical vulnerability lies in their dependence on communication networks such as 5G and Wi-Fi. This reliance exposes them to *network attacks* (risks from signals), including Denial of Service (DoS), where overloaded communication channels disrupt service; GPS spoofing, which manipulates navigation signals; eavesdropping, enabling attackers to intercept sensitive data; malware injection via exploited network vulnerabilities and ransomware, locking critical operational data, or compromising software supply chains. It can demand ransom payments while rendering robots inoperative. For example, compromised robots stuck in hazardous locations like busy streets or train tracks can disrupt both service and public safety.

Man-in-the-Middle (MITM) attacks, categorized as risks from codes, further compromise security. Attackers intercept communication between robots and control systems to alter navigation, manipulate sensor readings, or redirect operations. Such attacks can result in misdeliveries, theft, or misuse, threatening operational and public safety.

Physical attacks (risks from objects) are another pressing concern. Exposed sensors like LiDAR, cameras, and ultrasonic systems are vulnerable to tampering or damage. For instance, obstructing or disabling cameras can cause the robot to misinterpret its environment, increasing the risk of accidents. Damage to antennas, a common physical attack, can impair 5G connectivity, indirectly causing communication failures.

2.2 Extended Multi-Level Model

(Kolb & Xie, 2024) introduce the Extended Multi-Level Model. The Extended Multi-Level Model allows modelling not only the different component levels of complex systems, but also their effect on a human being. The authors use the Extended Multi-Level Model to classify cyber attacks, physical attacks, and cyber-physical attacks in order to simplify the analysis of attacks on last-mile robots. They find that all three types of attacks have an impact on human safety. Human safety is important in this context because the parcels transported by the last-mile robot may contain poisonous elements that are harmful to humans if touched, or certain types of medicine that should only be consumed under medical supervision, such as morphine.

The **extended multi-level model** consists of the following levels:

Level 4, *Human level,* represents the human, represents the needs or consequences for a human to be safe.

Level 3, is the *System level,* and it represents the system as an atomic unit and concentrates on its high-level strategic concerns, like goals and objectives. In our case it is the robot.

Level 2, *Sub-system level,* represents the functional decomposition of the system, wherein different sub-systems have dedicated functionalities that must be integrated to make the whole system. In our case, this could be the industrial computer of the robot.

Level 1, *Information level,* captures the information that facilitates interaction within or between the sub-systems. In our case, these could be the messages and signals that are sent between computer and sensors.

Level 0, *Component level,* focuses on the self-contained computational, communication, and physical elements that make up a sub-system or a system. In our case, this could be a sensor.

Each level has two numbers (*h,s*), where *h* denotes the human safety category and *s* denotes the security category.

2.4 CIA Triad

The CIA triad, introduced in (Anderson, 1972) and repeated in (Saltzer & Schroeder, 1975), consists of *confidentiality* (the property of preventing disclosure of information to unauthorized individuals or systems), *integrity* (means that data cannot be modified without authorization) and *availability* (any information system to serve its purpose, the information must be available when it is needed).

The article (Lacava et al. 2021) analyses potential and actual cyber-attacks and provides a classification according to the CIA triad concept, dividing them into categories of threats. The result of this analysis is that robotics faces prominent security challenges in the areas of collaborative robotics, autonomous vehicles, autonomous robotic platforms, and regulation and legal frameworks. However, to the best of our knowledge, there is no mention of a security assessment for confidentiality, integrity or availability, either separately or together that considers the safety too.

Among the CIA properties, *confidentiality* is generally the easiest for an attacker to compromise, as it only involves intercepting data without modifying it. *Availability* is somewhat more difficult to breach, requiring the attacker to block or delay message delivery. *Integrity*, however, is typically the hardest to compromise, as it involves altering message content, which is more complex and often detectable.

In the context of last-mile delivery robots, this results in the following order of attack difficulty: confidentiality, availability, and integrity.

However, when human safety becomes a factor—such as when the robot delivers critical items like medication—the order of vulnerability shifts in terms of impact. In such safety-critical scenarios, *integrity* becomes the most critical, followed by *availability*, and finally *confidentiality*. This shift guides the assignment of risk values used in our application (see Section 3.3).

2.5 Multiple Criteria Decision Making

Multi-Criteria Decision Making (MCDM), also known as Multi-Criteria Decision Analysis, is a process that involves evaluating and comparing multiple alternatives based on a set of criteria (Triantaphyllou, 2000). The goal is to help decision-makers select the best alternative that optimally balances various conflicting objectives. The simplest and most widely used method is called the weighed sum model. In such a method, each alternative is calculated by the weighted sum of the criteria.

$$\text{Weighted Score for Alternative} = \sum_{i=1}^{m} w_i \cdot x_{ij}$$

Where:

- w_i is the weight of criterion i,
- x_{ij} is the score of alternative j on criterion i,
- m is the number of criteria.

The recent comprehensive review by Sahoo and Goswami (2023) provides an in-depth analysis of the various methods used to address MCDM problems. It highlights their strengths, limitations, and applications across diverse domains, such as healthcare, business, and engineering. Among these methods, the Analytic Hierarchy Process (AHP), introduced by Saaty (1987), stands out as one of the classical and widely adopted approaches. Despite being introduced decades ago, AHP remains a popular method, with numerous modifications and advancements over the years (see Triantaphyllou, 2000, for a detailed discussion).

AHP and the term CIA has also seen recent applications in other works, such as (Lee & Geum, 2017). But in this field, the term CIA has a different meaning and is not related to security. In the field of security, AHP particularly is mentioned in areas such as banking (Syamsuddin & Hwang, 2009), egovernment (Syamsuddin, 2011), and economics (Lee & Geum, 2017). However, none of these fields investigates the safety too.

To demonstrate the workings of the Analytic Hierarchy Process (AHP), we consider an example involving three objectives: C (Criterion 1), I (Criterion 2), and A (Criterion 3). The goal is to determine the priority weights for these objectives through the following structured steps:

Step 1: Pairwise Comparison

Construct a pairwise comparison matrix to evaluate the relative importance of the objectives C, I, and A. For n objectives, the comparison matrix M will be of size $n \times n$, as shown below:

$$M = \begin{bmatrix} 1 & w_{CI} & w_{CA} \\ \frac{1}{w_{CI}} & 1 & w_{IA} \\ \frac{1}{w_{CA}} & \frac{1}{w_{IA}} & 1 \end{bmatrix}$$

Here, w_{CI},w_{CA}, and w_{IA} represent the relative importance of the objectives based on expert judgment. These values are assigned using a scale ranging from 1 to 5:

- 1: Equal importance,
- 2: Moderate importance,
- 3: Strong importance,
- 4: Very strong importance,
- 5: Extreme importance,
- Reciprocal values (e.g., $1/3$,$1/5$) are used for inverse importance.

Step 2: Normalize the Matrix

To normalize the matrix, divide each element in a column by the sum of its column. This produces a normalized pairwise comparison matrix. Then, compute the average of the normalized values in each row to derive the *priority vector*, which represents the weights of each objective.

Step 3: Check for Consistency

Assess the consistency of the judgments by calculating the Consistency Index (CI):

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the comparison matrix, and n is the number of objectives.

Next, compute the Consistency Ratio (CR):

$$CR = \frac{CI}{RI}$$

Here, RI is the Random Index, a constant based on the matrix size n. For $n = 3$, $RI = 0.58$. If $CR < 0.1$, the judgments are considered consistent. Otherwise, the matrix should be revised.

Step 4: Evaluate Alternatives

Once the priority weights of the objectives are determined, evaluate the alternatives for each objective using a similar pairwise comparison process. The weighted score for each alternative can then be calculated as:

$$\text{Weighted Score} = (\text{Score for } C \times \text{Weight of } C) + (\text{Score for } I \times \text{Weight of } I) + (\text{Score for } A \times \text{Weight of } A)$$

Repeat this process for all alternatives and rank them based on their weighted scores to identify the best option.

3 Extended CIA Multi-Level Model

Building upon the foundational elements introduced in the previous chapter, we now extend these concepts. Specifically, we enhance the Extended Multi-Level Model by introducing the *Extended CIA Multi-Level Model*. This extended model enables the parallel assessment of confidentiality, integrity, and availability (CIA) for last-mile delivery robots. Using the Analytic Hierarchy Process (AHP), we subsequently aggregate these individual assessments to compute an overall security score.
We begin by formally defining the Extended CIA Multi-Level Model $M_{multi}(safe_{multi}, sec_{multi})$.

3.1 Definition (Extended CIA Multi-Level Model)

The *Extended CIA* ***Multi-Level Model*** **is defined as the tuple**

$$M_{multi}(safe_{multi}, sec_{multi}),$$

where M_{multi} is a set of three specialized sub-models extending the Extended Multi-Level Model from Section 2.3, each corresponding to one of the CIA components:

- $M_C(safe_C, sec_C)$ models *confidentiality*,
- $M_I(safe_I, sec_I)$ models *integrity*, and
- $M_A(safe_A, sec_A)$ models *availability*.

Each sub-model $M_X(safe_X, sec_X)$, where $X \in \{C, I, A\}$, provides safety and security values at level 4 of the corresponding model.

The tuple $(safe_{multi}, sec_{multi})$ represents the aggregated safety and security values derived from the three sub-models M_C, M_I, and M_A.

A method for calculating this overall value will be discussed in Section 3.3.

3.2 Risk Assessment

In this subsection, we provide the *Risk Assessment Method for the Extended CIA Multi-Level Model.* Therfor, we investigate which effect to the human level an attack can have when the confidentiality, integrity, and availability of the communication network are violated, i.e., the influence of an attack for all levels for $\mathrm{M}_C(safe_C,sec_C)$ to model confidentiality, $\mathrm{M}_I(safe_I,sec_I)$ to model integrity, and $\mathrm{M}_A(safe_A,sec_A)$ to model availability in parallel as

$$(\mathrm{M}_C(safe_C,sec_C)||\mathrm{M}_I(safe_I,sec_I)||\mathrm{M}_A(safe_A,sec_A)).$$

We calculate the safety $safe_k i$ and security value $sec_k i$ for each of the CIA models $k \in \{C,I,A\}$ per each level i, $i \in \{0,1,2,3,4\}$ by summing up the corresponding values as

$$((level_k0(safe_k0,sec_k0)),(level_k1(safe_k1,sec_k1)),(level_k2(safe_k2,sec_k2)),$$

$$(level_k3(safe_k3,sec_k3)),(level_k4(safe_k4,sec_k4))).$$

We adopt a value range from 0 to 4 for both safety and security assessments, following the approach used by the authors in (Abdo et al., 2018) for pure security evaluations without incorporating safety aspects. In our context, a value of 0 indicates no safety or security concern. Higher values from 1 to 4 indicate increasing levels of risk or threat severity:

1 - Represents a minor safety risk or a negligible security issue that poses minimal danger and may be safely ignored.

2 - Indicates a moderate risk or threat, more significant than level 1 but not yet critical.

3 - Corresponds to a high safety risk or security threat that requires active mitigation.

4 - Denotes a critical, maximum-level danger, whether safety- or security related, that must be addressed with the highest priority.

The following section provides a concrete example of how these values are applied in the context of last-mile delivery robots.

3.2 Example for Risk Assessment for Last-Mile Robots with Critical Parcel Content

We consider the example of last-mile delivery robots transporting critical parcels. Such parcels may contain essential medications intended for direct delivery to a patient's home - for instance, insulin for a person with diabetes or morphine for severe pain.
In this scenario, we examine the potential impact of a sensor-based attack on the human level, specifically focusing on how breaches in *confidentiality*, *integrity*, or *availability* of the robot's communication network may affect safety and security.

We model the influence of these attacks across all levels using the submodels for confidentiality, integrity, and availability: $\mathrm{M}_C(safe_C,sec_C)$ (see Table 1), $\mathrm{M}_I(safe_I,sec_I)$ (see Table 2), and $\mathrm{M}_A(safe_A,sec_A)$ (see Table 3). These are evaluated in parallel as:

$$(\mathrm{M}_C(safe_C,sec_C) \parallel \mathrm{M}_I(safe_I,sec_I) \parallel \mathrm{M}_A(safe_A,sec_A)).$$

$\mathrm{M}_C(safe_C,sec_C)$: (see Table 1) To calculate the overall safety value for confidentiality $safe_C$ and the overall security value sec_C for confidentiality under a sensor attack, we first set a 0 for the safety, because it is not influenced so far in the sensor level 1, and we set the security value as a 1, because the security is clearly harmed when a sensor is hacked but in this level not yet really problematic. Thus, we achieve $(level_C0(safe_C0,sec_C0)) = (level_C0(0,1))$.

In the data level 1, where the attacker is able to read the messages, for example, this could be the message of the current position of the last-mile robot or the address of the user, the security problem increases clearly and we update the security value for this level as a 2. The safety value stays the same, because the attacker read the message but did not act accordingly to harm a person. Thus, we obtain $(level_C1(safe_C1,sec_C1)) = (level_C1(0,2))$.

In the sub-system level 2, the correct message is sent to the computer of the last-mile robot, because the attacker did not change the message and both the safety and the security value stay the same as $(level_C2(safe_C2,sec_C2)) = (level_C2(0,2))$.

Because there is no attacker action or change in the system level, and the robots act normal with going the correct way, the safety and security values stay the same for this level as $(level_C3(safe_C3,sec_C3)) = (level_C3(0,2))$.

Finally, in the human level 4, both the safety and security values are increased by 1. In this level, the attacker uses his knowledge of the confidentiality breach by knowing the person's address, and might use this knowledge to harm this person. We obtain $(level_C4(safe_C4,sec_C4)) = (level_C4(1,3))$.

$M_I(safe_I,sec_I)$: (see Table 2) As an example for the integrity for last-mile robots, we can consider the change of the parcel content which might harm a person. Therefore, in the sensor level 1 where an attacker attacks the sensor, we obtain the following with the same reasons as above: $(level_I0(safe_I0,sec_I0)) = (level_I0(0,1))$.

In the data level 1, we obtain the same values as above but for a different reason. The attacker changes the message and we obtain $(level_I1(safe_I1,sec_I1)) = (level_I1(0,2))$.

The forged message reaches the computer, in the sub-system level 3. Thus, we set the security value one up and obtain $(level_I2(safe_I2,sec_C2)) = (level_I2(0,3))$.

Since the computer received a forged message which might consist of the allowance to change the parcel content of the last-mile robot, the security value for integrity increases by one, and we have $(level_I3(safe_I3,sec_I3)) = (level_I3(0,4))$.

Finally, in the human level 4, the parcel reaches a person and can potentially explode due to the changed content. We consider this as very harmful and increase the safety value up to 4 as $(level_I4(safe_I4,sec_I4)) = (level_I4(4,4))$.

$M_A(safe_A,sec_A)$: (see Table 3) For the availability, we have the same safety and security values for the first level as above, thus $(level_A0(safe_A0,sec_A0)) = (level_A0(0,1))$.

For the data level 1, we increase the security value to 2, because the attacker cancels a message. The safety value still stays the same, and we have $(level_A1(safe_A1,sec_A1)) = (level_A1(0,2))$.

Because there is no message received by the computer, in the sub-system level 2, we increase the security value to 3. Nothing happened to the safety, thus we stay with 0 as the safety value, and obtain $(level_A2(safe_A2,sec_A2)) = (level_A2(0,3))$.

In the system level 3, the robot stops because it did not receive any message, which is a serious security problem and we increase the security value again by 1. Since a robot that does not move also does not harm a person, the safety value stays the same, and we have $(level_A3(safe_A3,sec_A3)) = (level_A3(0,4))$.

Finally, in the human level 4, the parcel is not delivered because the robot could not move forward from the previous level. This might be a harmful situation for a human being, because he does not get his medication. Thus, we increase the safety value up to 2. There is no effect on the further security. So the security value stays the same. In total, we obtain $(level_A4(safe_A4,sec_A4)) = (level_A4(2,4))$.

It is important to examine confidentiality, integrity and availability together to understand the interdependencies and common properties between the three.

For example, while the first-level safety and security values may appear to be the same for a similar or identical attack, the impact on human safety may be different. Therefore, in the next subsection, we use the AHP to quantify and prioritize the relative importance of confidentiality, integrity and availability at multiple levels.

Table 1: Example of confidentiality for last-mile robots with critical medications as parcel contents

Human Level 4	attacker knows address (1,3)
System Level 3	robot goes the correct way (0,2)
Sub-System Level 2	correct message to computer (0,2)
Data Level 1	read message (0,2)
Sensor Level 0	hack sensor (0,1)
	Confidentiality

Table 2: Example of integrity for last-mile robots with critical medications as parcel contents

Human Level 4	robots explodes (4,4)
System Level 3	changed package content (0,4)
Sub-System Level 2	forged message to computer (0,3)
Data Level 1	forge message (0,2)
Sensor Level 0	hack sensor (0,1)
	Integrity

Table 3: Example of Availability for last-mile robots with critical medications as parcel contents

Human Level 4	parcel not delivered (2,4)
System Level 3	robot stops (0,4)
Sub-System Level 2	no message to the computer (0,3)
Data Level 1	cancel message (0,2)
Sensor Level 0	hack sensor (0,1)
	Availability

3.3 Decision Making

To calculate the overall safety-security value, we make use of the AHP hierarchy.

The goal is defined as the identification of the optimal CIA configuration for last-mile robotic systems.

$$\text{Goal : Optimal CIA Configuration}$$

The hierarchy includes the three CIA criteria, each divided into two subcriteria:

$$\text{Criteria : } C, I, A$$

Where:

- **C** represents Confidentiality, subdivided into c_1 (Confidentiality of safety) and c_2 (Confidentiality of security),
- **I** represents Integrity, subdivided into i_1 (Safety integrity) and i_2 (Security integrity),
- **A** represents Availability, subdivided into a_1 (Safety availability) and a_2 (Security availability).

Thus, the hierarchy can be expressed as:

Goal → (C, I, A) → (Sub-Criteria of C, Sub-Criteria of I, Sub-Criteria of A)

Step 1: Conduct Pairwise Comparisons.

We construct the pairwise comparison matrix for the main criteria C, I, and A:

$$M = \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{3} \\ 5 & 1 & 3 \\ 3 & \frac{1}{3} & 1 \end{bmatrix}$$

Where: - The element M_{ij} represents the relative importance of criterion i over criterion j. When the safety of a human comes into play, for example, the application of last-mile robots with critical parcel content where this parcel content could be a medication, this order of vulnerability with respect to the safety within the CIA triad changes to the integrity, availability, and confidentiality. In the previous example, we assume that the delivery of the wrong medication (Integrity) is more critical than a failed delivery (Availability), while the exposure of a known address (Confidentiality) is comparatively less critical. Consequently, we assign the relative importance values as follows: $I/C = 5$ and $I/A = 3$.

Step 2.1: Normalize the Matrix and Compute the Priority Vector.

To normalize the matrix, divide each element by the sum of its column:

$$\text{Normalized} \quad M = \begin{bmatrix} 0.1111 & 0.1304 & 0.0769 \\ 0.5556 & 0.6522 & 0.6923 \\ 0.3333 & 0.2174 & 0.2308 \end{bmatrix}$$

Compute the average of each row to derive the priority vector:

$$\text{Priority Vector} = \begin{bmatrix} 0.1062 \\ 0.6333 \\ 0.2605 \end{bmatrix}$$

Step 2.2: Repeat Comparison and Weight Calculation for Sub-Criteria.

We determine the relative weights of *Safety* and *Security* based on the values presented in the table of the previous subsection, specifically at the human level (Level 4). To avoid a weight of zero, we apply a proportional upscaling of all values. For example, in the case of *Confidentiality*, the original values at Level 4 are adjusted proportionally (see Table 1), resulting in a relative comparison of $c_1/c_2 = 2/4 = 1/2$.

$$M_C = \begin{bmatrix} 1 & \frac{1}{2} \\ 2 & 1 \end{bmatrix}$$

Normalize M_C by dividing each element by the sum of its column:

$$\text{Normalized} \quad M_C = \begin{bmatrix} 0.3333 & 0.3333 \\ 0.6667 & 0.6667 \end{bmatrix}$$

Priority Vector for C (the average of each row):

$$[0.3333, 0.6667]$$

Similarly, we compute the priority vectors for the sub-criteria of *Integrity (I)* and *Availability (A)* using their respective pairwise comparison matrices:

$$M_I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad M_A = \begin{bmatrix} 1 & \frac{3}{5} \\ \frac{5}{3} & 1 \end{bmatrix}$$

For *Integrity*, the matrix M_I reflects equal importance between *Safety* and *Security*, resulting in the priority vector:

$$[0.5, 0.5]$$

For *Availability*, the matrix M_A expresses a moderate preference for *Security* over *Safety*, yielding the priority vector:

$$[0.375, 0.625]$$

Step 2.3: Calculate the Overall Weighted Score

Based on the previously computed priority vectors for the main criteria and their sub-criteria, we now calculate the overall weighted risk score. We assume the following risk values (on a scale from 0 to 4) for each sub-criterion:

- **Confidentiality (C):** Safety = 1, Security = 3
- **Integrity (I):** Safety = 4, Security = 4
- **Availability (A):** Safety = 2, Security = 4

Using the sub-criteria priority vectors:

$$\text{C: } [0.3333, 0.6667], \quad \text{I: } [0.5, 0.5], \quad \text{A: } [0.375, 0.625]$$

We compute the weighted risk scores for each main criterion:

$$\text{Confidentiality Score} = 0.3333 \times 1 + 0.6667 \times 3 = 2.3334$$

$$\text{Integrity Score} = 0.5 \times 4 + 0.5 \times 4 = 4.0$$

$$\text{Availability Score} = 0.375 \times 2 + 0.625 \times 4 = 3.25$$

Using the main priority vector:

$$\text{Main Priority Vector} = [0.1062, 0.6333, 0.2605]$$

We compute the final weighted score:

$$\text{Overall Score} = 0.1062 \times 2.3334 + 0.6333 \times 4.0 + 0.2605 \times 3.25$$

$$= 0.2478 + 2.5332 + 0.8466 = \boxed{3.6276}$$

Step 3: Interpret the result

A total risk score of 3.6276 (on a scale from 0 to 4) indicates a very high level of overall risk. This result is primarily driven by the criticality of *Integrity*, which carries both high risk values and the highest weight in the priority vector.

Step 4: Perform Sensitivity Analysis

Due to the inherent uncertainty in pairwise comparisons, alternative comparison matrices can be constructed. This enables us to examine how variations in the input affect the final result.
In this sense, we also test different pairwise comparisons as follows:

$$M' = \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{4} \\ 5 & 1 & \frac{4}{5} \\ 4 & \frac{4}{5} & 1 \end{bmatrix}, \qquad M'' = \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{5} \\ 5 & 1 & 1 \\ 5 & 1 & 1 \end{bmatrix}$$

The priority vectors derived from these matrices are:

$$\text{Priority Vector}, \quad M' = \begin{bmatrix} 0.1012 \\ 0.4328 \\ 0.4660 \end{bmatrix} \quad \text{Priority Vector} \quad M'' = \begin{bmatrix} 0.0909 \\ 0.4545 \\ 0.4545 \end{bmatrix}$$

Using the same sub-criteria weights and risk values from previous steps, we compute the final weighted scores:

Final Risk Score using $M' = 0.1012 \cdot 2.3334 + 0.4328 \cdot 4.0 + 0.4660 \cdot 3.25 = \boxed{3.4818}$

Final Risk Score using $M'' = 0.0909 \cdot 2.3334 + 0.4545 \cdot 4.0 + 0.4545 \cdot 3.25 = \boxed{3.5072}$

Despite varying the input comparison matrices, the final risk scores remain consistently high (above 3.48), indicating the robustness of the initial conclusion. Integrity and Availability continue to contribute most significantly to the overall risk in the context of last-mile delivery robots.

Interpretation: This approach provides a single safety-security score to support decision-making, especially under budget constraints where prioritization of countermeasures is necessary. In our case, the overall risk value of 3.6276 is close to the maximum of 4, indicating a critical situation. Such a high value suggests a need for deeper analysis using the detailed CIA submodels (see Tables 1, 2, and 3) to identify which dimension—confidentiality, integrity, or availability—poses the greatest risk. If the overall safety-security score remains within an acceptable range (e.g., 0–2), no further action may be required. However, values above this threshold call for inspecting the submodels to locate the highest risk area. In this example, Tables 2 and 3 indicate that integrity and availability require targeted countermeasures.

4 Conclusion and Future Directions

Conclusion: The example in the tables 1, 2 and 3 shows that confidentiality, integrity and availability are present at all levels in the multi-level model for the practical implementation of last-mile robots with critical content such as medication and creates an awareness that all levels need to be defended against attackers. Furthermore, the grouping of safety and security issues shows that there are overlaps between safety and security. Thus, a countermeasure for one of these issues will avoid both. Using multi-criteria decision making (MCDM) for risk assessment for both safety and security in parallel, economic and financial decisions can be made as to which defences to add.

Future work: Investigating AI and machine learning for LMRS would be interesting as a future approach to discuss how AI can improve safety and security, as well as the risks involved. In addition, to determine the tailored regulatory standards that are needed for last-mile robots. So far, the paper investigates the influence starting from the security part which affects the safety of people. As future work, it will also be crucial to understand how the human level influences the safety.

Funding or Grant

Our last-mile robot is funded by the German Federal Ministry for Digital and Transport as part of the “Usage Scenarios for Innovation Networks in 5G” (USIN5G) project.

CRediT authorship contribution statement

Christina Kolb: Conceptualization, Methodology, Formal analysis, Writing, Visualization, Review, Editing

Lin Xie: Conceptualization, Methodology, Formal analysis, Writing, Visualization, Writing, Review, Editing

Use of AI

During the preparation of this work, the author(s) used ChatGPT 4 in order to improve the grammar. After using this tool/service, the author(s) reviewed, edited, made the content their own and validated the outcome as needed, and take(s) full responsibility for the content of the publication.

Declaration of competing interests

There is no conflict of interest.

References

Abdo, H., Kaouk, M., Flaus, J.-M., & Masse, F. (2018). A safety/security risk analysis approach of industrial control systems: A cyber bowtie - combining new version of attack tree with bowtie analysis. *Computers & Security, 72*, 175–195. https://doi.org/10.1016/j.cose.2017.09.004

Alverhed, E., Hellgren, S., Isaksson, H., Olsson, L., Palmqvist, H., & Flodén, J. (2024). Autonomous last-mile delivery robots: A literature review. *European Transport Research Review, 16*. https://doi.org/10.1186/s12544-023-00629-7

Anderson, J. P. (1972). *Computer security planning study: Technical report ESD-TR-73-51* (Tech. Rep.). Air Force Electronic Systems Division. https://csrc.nist.rip/publications/history/ande72.pdf

Arents, J., Abolins, V., Judvaitis, J., Vismanis, O., Oraby, A., & Ozols, K. (2021). Human–robot collaboration trends and safety aspects: A systematic review. *Journal of Sensor and Actuator Networks, 10*(3), 48. https://doi.org/10.3390/jsan10030048

Cui, J., Sabaliauskaite, G., Liew, L. S., Zhou, F., & Zhang, B. (2019). Collaborative analysis framework of safety and security for autonomous vehicles. *IEEE Access, 7*, 148672–148683. https://doi.org/10.1109/ACCESS.2019.2946632

Fosch-Villaronga, E., & Mahler, T. (2021). Cybersecurity, safety and robots: Strengthening the link between cybersecurity and safety in the context of care robots. *Computer Law & Security Review, 41*, 105528. https://doi.org/10.1016/j.clsr.2021.105528

Kolb, C., & Xie, L. (2024). Security and safety in urban environments: Evaluating threats and risks of autonomous last-mile delivery robots. In A. Ceccarelli, M. Trapp, A. Bondavalli, E. Schoitsch, B. Gallina, & F. Bitsch (Eds.), *Computer safety, reliability, and security. SAFECOMP 2024 workshops* (pp. 36–49). Springer. https://doi.org/10.1007/978-3-031-68738-9_3

Lacava, G., Marotta, A., Martinelli, F., Saracino, A., La Marra, A., Gil-Uriarte, E., Mayoral-Vilches, V., et al. (2021). Cybersecurity issues in robotics. *Journal of Wireless Mobile Networks, Ubiquituous Computing and Dependable Applications, 12*(3), 1-28. https://doi.org/10.22667/JOWUA.2021.09.30.001

Lee, H., & Geum, Y. (2017). Development of the scenario-based technology roadmap considering layer heterogeneity: An approach using CIA and AHP. *Technological Forecasting and Social Change, 117*, 12–24. https://doi.org/10.1016/j.techfore.2017.01.016

Li, B., Liu, S., Tang, J., Gaudiot, J.-L., Zhang, L., & Kong, Q. (2020). Autonomous last-mile delivery vehicles in complex traffic environments. *Computer, 53*(11), 26–35. https://doi.org/10.1109/mc.2020.2970924

Nicoletti, S., Peppelman, M., Kolb, C., & Stoelinga, M. (2021). Model-based joint analysis of safety and security: Survey and identification of gaps. *Computer Science Review, 50*. https://doi.org/10.1016/j.cosrev.2023.100597

Pohowalla, F., Collins, T., & Chang, J. (2021). *Supply chain technology market update* [PowerPoint slides]. Cascadia. https://www.cascadiacapital.com/wp-content/uploads/Supply-Chain-Technology-Winter-Spring-2024.pdf

Quamara, M., Kolb, C., & Lohachab, A. (2024). Where do safety and security mutually reinforce? A multi-level model-based approach for a consistent interplay. In A. Ceccarelli, A. Bondavalli, M. Trapp, E. Schoitsch, B. Gallina,

& F. Bitsch (Eds.), *Computer Safety, Reliability, and Security. SAFECOMP 2024 Workshops – DECSoS, SASSUR, TOASTS, and WAISE* (Lecture Notes in Computer Science, pp. 316–328). Springer. https://doi.org/10.1007/978-3-031-68738-9_25

Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling, 9*(3–5), 161–176. https://doi.org/10.1016/0270-0255(87)90473-8

Sahoo, S. K., & Goswami, S. S. (2023). A comprehensive review of multiple criteria decision-making (MCDM) methods: Advancements, applications, and future directions. *Decision Making Advances, 1*(1), 25–48. https://doi.org/10.31181/dma1120237

Saltzer, J., & Schroeder, M. (1975). The protection of information in computer systems. *Proceedings of the IEEE, 63*(9), 1278–1308. https://doi.org/10.1109/PROC.1975.9939

Shaklab, E., Karapetyan, A., Sharma, A., Mebrahtu, M., Basri, M., Nagy, M., Khonji, M., & Dias, J. (2023). *Towards autonomous and safe last-mile deliveries with AI-augmented self-driving delivery robots* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2305.17705

Syamsuddin, I. (2011). Strategic information security decision making with analytic hierarchy process. *International Research Journal of Applied and Basic Sciences, 2*(11), 426-432. https://repository.poliupg.ac.id/id/eprint/196/1/J022_ORIGINAL%20PAPER_%20Irfan%20ITPOSMO%20AHP%20Sec.pdf

Syamsuddin, I., & Hwang, J. (2009). The application of AHP model to guide decision makers: A case study of e-banking security. In *Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology* (pp. 1469–1473). IEEE. https://doi.org/10.1109/ICCIT.2009.251

Tanimu, J. A., & Abada, W. (2024). Addressing cybersecurity challenges in robotics: A comprehensive overview. *Cyber Security and Applications, 3,* 100074. https://doi.org/10.1016/j.csa.2024.100074

Triantaphyllou, E. (2000). Multi-criteria decision making methods. In *Multi-criteria decision making methods: A comparative study* (pp. 21–43). https://doi.org/10.1007/978-1-4757-3157-6_2

Tu, Y.-J., & Piramuthu, S. (2023). Security and privacy risks in drone-based last mile delivery. *European Journal of Information Systems, 33*(5), 617–630. https://doi.org/10.1080/0960085X.2023.2214744

von Szczepanski, K., Wagener, C., Mooney, T., McDaniel, L., Mathias, O., & Sharp, L. (2021). *Only an ecosystem can solve last-mile gridlock in package delivery.* Boston Consulting Group. https://www.bcg.com/publications/2021/solving-the-package-delivery-system-problems-with-a-new-ecosystem