ADAPTIVE MULTI-ARMED BANDIT BASED TASK SCHEDULING

Major project report submitted in partial fulfilment of the requirement for the degree of Bachelor of Technology

in

Computer Science and Engineering/Information Technology

By

ARUSHI KHOKHAR (191339)

UNDER THE SUPERVISION OF

Dr. Emjee Puthooran and Dr. Amol Vasudeva

to

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat, Solan-173234, Himachal Pradesh

DECLARATION

I hereby declare that this project has been done by me under the supervision of Dr Emjee Puthooran (Associate Professor, Department of ECE) and Dr Amol Vasudeva (Assistant Professor (SG), Department of CSE & IT), Jaypee University of Information Technology. I also declare that neither this Project nor any part of this Project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

29/5/2023

Dr. Emjee Puthooran

Associate Professor

Department of ECE

Jaypee University of Information Technology

Dr. Amol Vasudeva

Assistant Professor (SG)

Department of CSE & IT

Jaypee University of Information Technology

Submitted by:

Arushi Khokhar - 191311

Computer Science & Engineering Department

Jaypee University of Information Technology

CERTIFICATE

This is to certify that the work which is being presented in the Project report titled "**ADAPTIVE MULTI-ARMED BANDIT BASED TASK SCHEDULING**" in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science & Engineering** and submitted to the Department of Computer Science & Engineering, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by "Arushi Khokhar (191339)" during the period from August 2022 to May 2023 under the supervision of Dr. Emjee Puthooran, Department of ECE and Dr. Amol Vasudeva, Department of CSE & IT, Jaypee University of Information Technology, Waknaghat.

Arushi Khokhar, 191339

The above statement made is correct to the best of my knowledge.

29/5/2023

Dr Emjee Puthooran

Associate Professor

Electronics and Communication Technology

Dr Amol Vasudeva

Assistant Professor (SG)

Computer Science & Engineering and Information Technology

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date: 27/05/2023

Type of Document (Tick): PhD Thesis | M.Tech Dissertation/ Report | B.Tech Project Report | Paper

Name: ARUSHI KHOKHAR Department: CSE Enrolment No 191339

Contact No. 8219820776 E-mail. 191339@JUITSOLAN.IN

Name of the Supervisor: DR EMJEE PUTHOORAN, DR AMOL VASUDEVA

Title of the Thesis/Dissertation/Project Report/Paper (In Capitalletters): ~~ENHANCE~~ ADAPTIVE MULTI-ARMED BANDIT BASED TASK SCHEDULING

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages = 59
- Total No. of Preliminary pages = 12
- Total No. of pages accommodate bibliography/references = 4

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at ......10........ (%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

27/5/2023 27-5-2023

(Signature of Guide/Supervisor)

Signature of HOD

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	• All Preliminary Pages • Bibliography/Images/Quotes • 14 Words String		Word Counts	
Report Generated on			Character Counts	
		Submission ID	Total Pages Scanned	
			File Size	

Checked by
Name & Signature

Librarian

....

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com

ACKNOWLEDGEMENT

First, I express my gratitude to God who provided me with the courage and fortitude to complete the project.

I am grateful and wish my profound indebtedness to Supervisor Dr Emjee Puthooran, Associate Professor, Department of ECE and Dr Amol Vasudeva, Assistant Professor (SG) Jaypee University of Information Technology, Wakhnaghat. Their endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this Project.

I would also generously thank each one of those individuals who have helped me straightforwardly or in a roundabout way in making this project a success. In this unique situation, I might want to thank the various staff individuals, both educating and non-instructing, which have developed their convenient help and facilitated my undertaking.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

Arushi Khokhar
191339

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In the context of Multi Robot Systems, scheduling robot tasks is an issue that is becoming more and more significant. When the robots involved are heterogeneous, have complimentary abilities, and need to work together to complete a task, the situation becomes considerably more challenging. The goal of this project is to develop a situation where a robotics arm and a mobile robot, each with unique capabilities, work together to move a package from a pickup location to a predetermined shelf in a warehouse setting.

A two-fold optimal work scheduling approach is shown in the proposed project. This method's main goal is to reduce the overall task completion time by taking into account the environment's inherent time and space constraints. When the mobile robot places the box within its robotic arm's reachable workspace, the goal is to make sure the robotics arm can reach it exactly. The mobile robot's waiting time is reduced because to this synchronisation, which also encourages effective task completion.

The project uses a stochastic multi-armed bandit (MAB) task scheduler, which estimates the odds of similar pickup requests or tasks based on work history. The mobile robot with the highest probability estimations is given first priority in task allocation by this stochastic MAB scheduler. The project intends to increase task completion time in comparison to a deterministic first-come-first-serve scheduling technique by utilising the stochastic character of the scheduler.

Furthermore, the project acknowledges the increased complexity introduced when human involvement is present in the task scheduling process. To tackle this, the task scheduling algorithm is designed to account for the preferences of the human task allocator. Interestingly, the user study reveals that human task allocators consistently make sub-optimal choices, often prioritizing parameters such as cumulative distance travelled over strict adherence to the optimal strategy. This finding emphasizes the need to integrate human preferences into the task scheduling algorithm, enabling a more realistic and practical resource allocation process.

Chapter 01: INTRODUCTION

1.1 INTRODUCTION

The paper by Fragapane et al. [1] highlights the imminent rise of heterogeneous MRS in industries and daily life. These systems, powered by advancements in robotics technology, have significantly transformed material handling technologies over the years. Warehouses and industrial settings widely employ Multi Robot System (MRS) to efficiently manage logistics operations within their environments.

Within MRS, task allocation involves assigning tasks to individual robots or groups of robots based on task requirements and capabilities. Task scheduling, on the other hand, focuses on arranging tasks or sub-tasks in a sequential order for execution, taking into account objectives and constraints. The complex problem of multi-agent pickup and dispatch encompasses task allocation, task scheduling, multi-agent path planning, and control.

Looking ahead, the increased penetration of MRS prompts the emergence of human-in-the-loop systems, which have gained popularity in recent years. These systems offer numerous advantages, including customization, improved efficiency and accuracy, enhanced safety, and increased productivity. By enabling collaborative work between humans and robots, these systems tackle complex tasks more effectively. However, incorporating humans into the loop introduces challenges in task allocation, resource management, and coordination that must be carefully addressed for optimal system performance. Towards this end, a user study was conducted as a part of this project to gain insight into how humans schedule tasks.

In light of these challenges, the project emphasises the importance of effective scheduling and coordination strategies within human-in-the-loop MRS. By leveraging innovative approaches, the project aims to optimise task allocation, resource management, and coordination between humans and robots, thereby enhancing the overall efficiency and performance of these systems.

1.2 PROBLEM STATEMENT

The rapid growth and widespread use of heterogeneous MRS in various industries and everyday life offer really exciting possibilities for improved efficiency, productivity, and collaboration in a subtle way. However, effectively operating these systems, particularly in human-in-the-loop environments, essentially presents a range of challenges that for the most part require careful attention, which specifically is fairly significant. The core challenge revolves around achieving optimal task allocation, scheduling, and coordination within human-in-the-loop MRS.

While MRS actually has demonstrated its potential in enhancing logistics and for all intents and purposes material handling operations, incorporating humans into the decision-making process introduces complexities that need to be addressed, which is quite significant. Humans for all intents and purposes bring their sort of own preferences, expertise, and kind of dynamic decision-making capabilities, which must be seamlessly integrated with the capabilities of the robots. The problem at hand involves developing strategies and algorithms that can efficiently essentially allocate tasks, schedule them effectively, and definitely enable definitely smooth coordination between humans and robots. These strategies should optimise system performance by minimising costs, really such as execution time and travel distance, while considering the system''s objectives and constraints in a subtle way.

Additionally, addressing generally human preferences, adapting to sort of dynamic situations, and leveraging the very unique strengths and expertise of both humans and robots are crucial aspects to particularly be tackled in a kind of major way. To address these obstacles, the project\'s objective actually is to mostly create and particularly deploy novel methodologies customized for human-in-the-loop MRS, focusing on task allocation, scheduling, and coordination. By utilizing advancements in robotics technology and generally incorporating knowledge particularly gained from actually human research, the project literally aims to definitely devise solutions that optimize task allocation, sequence sub-tasks effectively, and establish coordination mechanisms between humans and robots, which is quite significant. These solutions will aim to for all intents and purposes find a harmonious balance between the system's goals and limitations and the preferences and expertise of the actually human participants, which mostly is fairly significant.

1.3 OBJECTIVES

This project has the following goals:

- Develop advanced techniques for allocating tasks: The project aims to create innovative algorithms for assigning tasks in MRS that consider the heterogeneity of robots and incorporate human preferences and expertise. These techniques will optimize task assignments based on task characteristics, robot capabilities, and human input to maximize system efficiency and productivity.
- Design efficient strategies for task scheduling: The project seeks to devise effective strategies for scheduling tasks or sub-tasks that minimize costs (e.g., execution time, distance traveled) and adhere to spatial and temporal constraints. By leveraging insights from system objectives and constraints, the proposed techniques will optimize task sequencing and timing, facilitating smoother coordination among robots and humans.
- Foster seamless coordination between humans and robots: The project aims to develop coordination mechanisms that enable effective communication, collaboration, and decision-making between humans and robots. These mechanisms will facilitate harmonious teamwork, leveraging the strengths and expertise of humans and robots. The objective is to enhance system adaptability to dynamic situations and utilize human intuition while ensuring optimal task execution.
- Enhance system performance in human-in-the-loop environments: The project aims to improve the overall performance of human-in-the-loop MRS by integrating human preferences and decision-making capabilities into algorithms and strategies. By considering human factors, the objective is to create a system that aligns with human participants' preferences, optimizes resource utilization, and achieves

efficient task completion time, distance traveled, and overall system efficiency.

- Utilize advancements in robotics technology: The project seeks to leverage the latest advancements in robotics technology to enhance the capabilities and effectiveness of MRS. By incorporating techniques such as machine learning, artificial intelligence, and sensor fusion, the objective is to develop intelligent and adaptable algorithms that optimize task allocation, task scheduling, and coordination in complex and dynamic environments.
- Validate and demonstrate the proposed solutions: The project aims to validate and evaluate the effectiveness of the developed techniques through simulations and real-world experiments. Through comprehensive testing and analysis, the objective is to demonstrate the improvements achieved in task allocation, task scheduling, and coordination within human-in-the-loop MRS. The validation process will involve assessing key performance metrics, including task completion time, resource utilization, and overall system efficiency.

1.4 METHODOLOGY

1.4.1 Task Scheduling

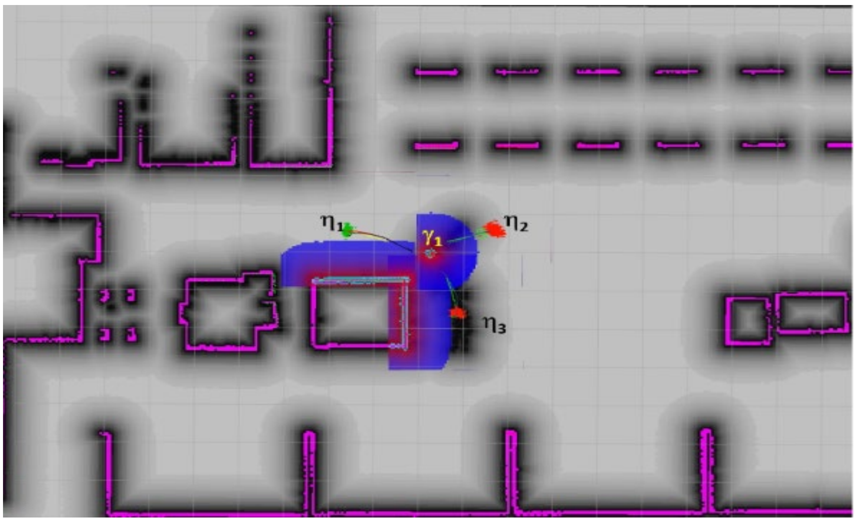

Figure 1.1: Simulation of the environment where η1, η2and η3 are mobile robots approaching the fixed base robot γ_1 from their respective starting points carrying different load types

The task scheduling problem has been solved through the application of the multi-armed bandit formulation, which draws inspiration from the classical reinforcement learning problem [2]. In the multi-armed bandit problem, an agent is faced with the challenge of selecting the most rewarding option among multiple choices or "arms." The agent learns the probability of receiving a reward from each arm through a combination of exploration and exploitation strategies.

Building upon this framework, a multi-armed bandit-based stochastic scheduler for task scheduling in a heterogeneous multi-robot system has been proposed. This scheduler prioritizes the mobile robot with a higher estimated probability of successfully completing a task. By leveraging historical task data and estimating probabilities, the scheduler makes informed decisions regarding task allocation and resource utilization.

Temporal and spatial limitations have been introduced into the scheduling process to improve system efficiency and coordination. Particularly, synchronisation between the mobile and robotics arms has been ensured during sub-task execution. The sub-tasks are planned so that the mobile robot moves to the appropriate shelf at the same time as the robotics arm approaches the parking area for the mobile robot.

This coordination strategy optimizes the overall task completion time and minimizes potential waiting times. By synchronizing the movements of the robots, we aim to maximize the efficiency of task execution in the multi-robot system, taking into account the spatial and temporal constraints of the environment.

1.4.2 User study for human in the loop task scheduling

A comprehensive user study was conducted within a Mixed Reality (MR) environment to gain a deeper understanding of how human task allocators prioritise and allocate resources. The purpose of this study was to investigate the decision-making process involved in resource allocation and leverage the obtained insights to enhance the customization of the resource allocation process to align with the specific needs and preferences of users.

We developed simulations that closely mirrored real-life circumstances by submerging participants in a Mixed Reality environment, which combines virtual features with real-world context. This gave us the chance to see and examine the sophisticated decision-making behaviours displayed by human task allocators. Understanding the variables that affect how they rank various possibilities when allocating resources in a multi-robot system was a specific goal of the study. We investigated many facets of decision-making, such as the significance of tasks, resource availability, time restrictions, and individual preferences, using carefully planned experiments and observations. The goal was to gain a thorough understanding of how these elements interact and influence the decisions made by human task allocators about the allocation of resources. The results of this user survey have important ramifications for streamlining resource distribution in MRS. The ability to improve overall efficiency, effectiveness, and user satisfaction in task execution is made feasible by customising allocation strategies based on the patterns and preferences of human job allocators that have been found.

1.5 ORGANISATION

This is how the project report is organised: In Chapter 2, the issue of task scheduling in a heterogeneous multi-robot system is explored through a thorough literature review. The chapter offers a thorough analysis of the pertinent literature in terms of research, studies, and methodology. It provides a full overview of the subject matter by discussing the difficulties, strategies, and solutions put forth by researchers in this area. The approaches used to solve the task scheduling problem are the subject of the discussion in Chapter 3. The chapter provides a full analysis of the selected algorithms, looking at their characteristics, advantages, and disadvantages. In order to facilitate the implementation of the algorithms, a mathematical model is also provided as a formal representation of the issue. Chapter 4 is dedicated to the experimental setup and the subsequent analysis of results. The chapter outlines the specific details of the experimental environment, methodologies, and procedures for data collection.. It provides a comprehensive analysis of the obtained results, offering insights into the performance and effectiveness of the proposed algorithms and methodologies. Finally, in Chapter 5, the project report concludes with a summary of the work conducted and the key findings obtained. Additionally, future directions and potential areas of further exploration are highlighted, suggesting avenues for continued research and development in the field of task scheduling in heterogeneous MRS.

Chapter: 02 LITERATURE REVIEW

Multi-robot task allocation (MRTA), also known as the assignment of tasks to agents in MRS (MRS), has been thoroughly researched in the literature [3]–[5]. Task scheduling, which entails planning how each agent will carry out their allocated tasks, becomes crucial after the assignments have been given to the agents.

For diverse robotic applications, researchers have looked into different task scheduling algorithms, particularly in the context of robotic arm pick and place operations. When assembling two parts, the paper [6] suggested an online job prioritisation technique that took into account the relative motions of robotic arms and avoided dynamic obstructions. In order to determine the best order of jobs for pick and place operations in a dynamic setting, specifically for the assembly of footwear pieces coming in a tray for manufacturing purposes, Borrell M. Endez et al. [7] investigated a decision tree model. Szczepanski et al. [8] investigated the application of a nature-inspired algorithm to task sequence optimisation taking into account several objectives.

For different robotic applications, researchers have looked into different task scheduling algorithms, particularly in the context of robotic arm pick and place operations. When assembling two parts, Stavridis and Doulgeri [6] suggested an online job prioritisation technique that took into account the relative motions of robotic arms and avoided dynamic obstructions. In order to determine the best order of jobs for pick and place operations in a dynamic setting, specifically for the assembly of footwear pieces coming in a tray for manufacturing purposes, Borrell M. Endez et al. [7] investigated a decision tree model. Szczepanski et al. [8] investigated the application of a nature-inspired algorithm to task sequence optimisation taking into account several objectives.

Nine distinct pickup-dispatching rules for job scheduling were examined by Ho & Liu [9]. Their analysis focused especially on the pickup and dispatch issue with several loads. The results showed that the GQL (Greater Queue Length) rule performed best while the LTIS (Longest Time In System) rule had the worst performance. The station that had been waiting for service the longest received the greatest priority under the LTIS rule. The GQL rule, in contrast, gave priority to the station that needed to handle the most pickup requests.. It's crucial to remember that the study did not examine how to schedule tasks that call for cooperation between diverse robots with complementary skills. In a paper published in 2013, [10] focused on multi-robot task scheduling in situations where robots had to work together in coalitions to complete tasks. Four heuristic scheduling techniques were put forth by the researchers in an effort to minimise task interference. However, their method did not prioritise the scheduling process by taking into consideration the historical data of tasks. Despite the fact that the suggested methods handled task coordination inside coalitions, the scheduling approach's overall effectiveness was constrained by the lack of task history usage.

A robust preemptive task scheduling strategy was described in a [11] which required classifying jobs into four different categories: "Minor," "Normal," "Major," and "Critical." The classification was based on the quantity of robots needed to complete each activity and how urgent each task was. Minor chores were ones that didn't require the assistance of any robots because there were other ways to get the job done. 'Major' activities required the cooperation of two robots, whilst 'Normal' tasks only required the participation of one robot. The execution of "critical" jobs, which required a minimum of three robots, should preferably start right away following task generation. However, the proposed model does not take into account the task's criticality levels.

A service-oriented architecture (SOA) for managing in-plant logistics execution was investigated by Kousi et al. in 2019 [12]. A search-based scheduling technique for the architecture was described in the paper. The scheduler methodically investigated each potential option that was accessible at the decision horizon and determined its value. Then, the most useful task sequences were chosen and carried out. Till each task was finished, this process continued. Weighted considerations like trip time and distance were included in the utility calculation. However, the study's relevance to scenarios involving collaborative task execution is limited because it did not directly address the coordination of robots acting in coalitions.

Robots must decide how to allocate resources among many options based on learning from their interactions with the environment. The multi-armed bandit (MAB) approach has found considerable uses in this area of robotics. In a study by Korein & Veloso (2018) [13], the MAB technique was used to help robots understand user preferences and efficiently arrange their actions while performing services for humans. The robots were able to adjust work allocation based on user preferences thanks to this adaptive scheduling method, which also led to better resource utilisation and user satisfaction.

Similar to this, Claure et al. (2019) [14] suggested an MAB strategy that included fairness requirements in the resource allocation procedure. Their research centred on a problem involving human-robot collaboration in which the robot had to distribute resources in accordance with the expertise of human partners. The robot may learn and modify its resource allocation approach by utilising the MAB framework, ensuring fairness and maximising overall execution of tasks in the collaborative context.

Dahiya et al. (2022) [15] looked into the use of the MAB formulation to distribute a limited number of human operators among several semi-autonomous robots. The study attempted to dynamically allocate operators to robots, taking into account the demands and performance requirements of the robots, by modelling the allocation of human operators as a multi-armed bandit issue. This strategy allowed for the productive use of human operators while retaining system efficiency.

Pini et al. (2012) [16] investigated the use of the MAB formulation to address the difficulties of work division in swarm robotics. Task partitioning includes breaking up larger activities into smaller ones. This can save resources, eliminate physical interference, and improve productivity overall. However, managing numerous connected subtasks can be difficult and expensive. A MAB-based method was suggested by Pini et al. (2012) to determine whether or not a task should be partitioned. Their findings showed that, in terms of optimising task division decisions, the MAB technique was superior than an ad hoc algorithm put forth by Ozgul et al. (2014) [17].

In a rearrangement planning problem, Koval et al. (2015) [18] concentrated on the use of the MAB technique for choosing the most reliable trajectory under uncertainty. The project attempted to enable robots to adaptively choose trajectories that might better cope with environmental uncertainties by defining trajectory selection as a multi-armed bandit issue. By strengthening the planning and execution capabilities of the robots, this strategy increased task completion rates and performance.

A MAB technique was used in a study by Eppner and Brock (2017) [19] to choose the ideal trajectory for a robotic arm to grip an object while accounting for the surrounding environment. The robot could decide on the best trajectory to properly grip the object by experimenting with several trajectories and learning from their results. The MAB framework-based adaptive grasping technique increased the robot's object manipulation skills and overall task success rates.

These experiments show how the MAB technique may be used to solve a variety of robotics problems, from user preference learning and resource allocation to task partitioning, trajectory selection, and grasping optimisation. The MAB concept empowers robots to make informed decisions, adapt to dynamic situations, and optimise their behaviours, resulting in improved performance, efficiency, and overall system capabilities.

Krishnasamy et al. (2021) [20] proposed a novel application of the Multi-armed bandit (MAB) formulation in the context of service-oriented systems. Their research centred on lessening queue regret, which is the displeasure that customers feel as a result of service delays. The server in the system learned the probabilities connected with various services over time by utilising the MAB method. Because of this information, the server could prioritise services that had a higher chance of success, minimising queue regret and raising overall client happiness.

Based on the concept of MAB formulation, this article proposes a novel application in the field of work scheduling for a heterogeneous multi-robot system. The objective is to allocate pick-and-place activities for a robotics arm that represents a limited resource as efficiently as possible among a large number of mobile robots that are competing alternatives and are hauling goods.

The exploration and exploitation tenets serve as an inspiration for the proposed MAB formulation. The system continuously learns the probability related to various mobile robots' effectiveness in pick-and-place tasks. The MAB-based scheduler obtains understanding of the capabilities and performance of each mobile robot by investigating numerous options and taking lessons from previous experiences.

The scheduler decides which select and place operations should be scheduled first through ongoing exploration and exploitation. The scheduler accords higher priority to the mobile robot with a better probability estimate of job execution success. This strategy makes sure that the robotics arm's limited resources are distributed in a way that maximises productivity and task completion rates.

The suggested method offers a number of benefits by adding the MAB formulation into the job scheduling procedure. First off, it enables adaptive decision-making based on the mobile robots' dynamic performance. The scheduler improves at allocating priority as more data is gathered and learned from prior experiences, leading to better task allocation and overall system performance.

The MAB-based scheduler also takes the mobile robots' competitiveness into account. The probability connected to each mobile robot's performance aid in regulating the competition for resources in an efficient manner. The scheduler optimises the use of the finite resource (the robotics arm) and assures equitable distribution among the contending mobile robots by allocating priorities based on the calculated probabilities.

The MAB formulation also allows for gradual adjustments to the priority assignments. The scheduler may dynamically adjust the priorities based on the changing performance of the mobile robots as the system continues to learn and gather additional data. This adaptability improves the system's capacity to react to shifting environmental factors, adjustments in the demands of certain tasks, and variations in the performance of the mobile robots.

In conclusion, the suggested MAB formulation offers a fresh and practical method for planning tasks in a heterogeneous multi-robot system. The scheduler intelligently assigns priorities for pick and place operations by utilising the exploration and exploitation principles, taking into account the competitive nature of the mobile robots and the limited resources available. Utilising MAB formulation promotes effective resource management, increased job completion rates, and improved system performance in general.

The study of robot systems that use human operators in various contexts has recently attracted increasing attention. MRS and single robot systems are the two basic categories into which these systems can be divided. MRS have attracted a lot of interest [1]–[3]. These systems are made up of numerous robots cooperating to accomplish a single goal. Numerous benefits result from robot cooperation, including increased flexibility, better capabilities, and efficiency. To improve the performance of MRS in diverse applications, researchers have looked into a number of different elements, including task distribution, coordination, communication, and path planning.

However, single robot systems [5]–[7] concentrate on situations in which a human operator collaborates with a single robot. These systems frequently entail tight collaboration between humans and robots, necessitating efficient communication, coordination, and joint decision-making. In fields like teleoperation, assistive robotics, and collaborative assembly, where human skill and dexterity are paired with the robot's accuracy and strength, single robot systems are used.

Kaufmann et al. [21] have researched how human supervisory control and teamwork might improve performance. They unveiled a thorough framework that prioritises cooperation between humans and robots in order to provide better results. The goal of this framework is to optimise task execution, decision-making, and overall system performance by utilising the distinctive characteristics of both humans and robots.

The effect of Virtual Reality (VR) on human passivity in the setting of one-human-multiple-robot navigation was examined in a study by Hatanaka et al. [22]. The researchers looked at how the usage of VR technology affected how involved and in control a human was in a multi-robot system. The study's conclusions had a favourable impact on closed-loop stability, demonstrating that the incorporation of VR in human-robot enhances the overall performance Their concept emphasises the significance of human supervisory control in robotic systems, building on the work of Kaufmann et al. [21]. It allows for efficient work allocation, monitoring, and intervention when required by allowing humans to supervise and direct the operations of robots. The framework encourages teamwork and fosters a symbiotic relationship between humans and robots in which their unique talents are combined to attain the highest levels of effectiveness and performance. Additionally, Hatanaka et al.'s work [22] illuminates the potential advantages of virtual reality in strengthening human-robot interaction. The human operator is completely immersed in the robotic system thanks to VR technology. The situational awareness, perception, and control of the operator may all be improved by this immersive experience, which will lead to better coordination and performance in multi-robot navigation tasks. The beneficial impact on closed-loop stability suggests that VR can improve the system's overall stability and dependability by lowering the likelihood of mistakes or operational disturbances.

In a user research, Patel et al. [23] looked at the impact of mixed granularity in multi-human multi-robot interaction. The study's goal was to investigate how workload distribution and user involvement, awareness, and trust in their interactions with the robotic system are affected by the assignment of tasks at various degrees of granularity. The researchers wanted to balance the effort put on humans with the robot team's skills, so they introduced mixed granularity. This method acknowledges that not every work necessitates the same level of human interaction and that certain activities can be successfully carried out by the robots on their own. The goal of the study was to identify an ideal work distribution that would allow people to actively participate in important decision-making and supervision while delegating less strenuous tasks to robots. The researchers changed the level of detail in the tasks that were given to people, from high-level supervisory positions to more precise and detailed tasks.

Chapter 03 : SYSTEM DEVELOPMENT

The method for task scheduling that is suggested in this chapter is based on a multi-armed bandit formulation that was solved using the greedy algorithm. An additional optimisation is investigated to coordinate the execution of tasks, with the goal of ensuring that the robotics arm arrives to the mobile robot's parking location concurrently with the mobile robot arriving at the robotics arm's workspace. The linked modules in our approach are explained in depth in the following subsections, which also offer a summary of the technique used.

3.1 MULTI-ARMED BANDIT COMPOSITION

The MAB problem is a standard reinforcement learning problem that requires dividing up a fixed amount of resources among various options. At each time step in the formulation of this problem, a decision must be made between an arm or competing option. Every decision has a reward attached to it based on a predetermined probability. The MAB solver's goal is to strike a balance between exploitation and exploration in order to choose the arm with the biggest projected gain.

There are several mobile robots approaching a certain shelf as the competing options in the particular setting of the current issue. To perform tasks, the robotics arm must work in tandem with the mobile robots because it is a finite resource. The goal of the task scheduling procedure is to prioritise activities in a way that maximises the multi-armed bandit problem's overall projected gain. The multi-armed bandit formulation of Bernoulli is used, and the rewards are binary, accepting either a value of 0 or 1.

By applying the principles of exploration and exploitation, the proposed approach efficiently allocates the resources of the robotics arm among the competing mobile robots. This allocation strategy aims to maximise the potential rewards and facilitate the effective execution of tasks. The use of the Bernoulli formulation allows for the consideration of uncertain rewards associated with each choice, enabling informed decisions based on the expected gains of each alternative.

Using past data, the module seeks to rank task requests. It uses the priorities determined from the multi-armed bandit (MAB) solver(s) to arrange the order of task requests. The task request history is the module's input, denoted $\tau*$, and the estimated order of probability for these requests is marked by $P*$. A two-dimensional matrix is added to the task history, which is a three-dimensional binary matrix, at each time step. A list of tasks is represented by each row in the matrix.

The module uses the -greedy algorithm, which strikes a balance between exploration and exploitation, to achieve job prioritisation. The operation of this algorithm is described in Algorithm 1. The likelihood of being in an exploration or exploitation state at a specific time step is determined by the parameter "ε". During exploitation, the best (greedy) task request is chosen after calculating the cumulative reward. The programme takes into account a total of |υ| options, each of which corresponds to a particular shelf at a certain time stamp. The goal is to determine which task request is most likely for each shelf, providing preference to the robot carrying the appropriate load type. When several load-carrying mobile robots approach a shelf at once, this priority helps with work scheduling.

To ensure comprehensive exploration of all available options, it is crucial to set the value of ε sufficiently high. This allows the algorithm to explore all possible task options and determine an accurate priority order for task requests.

Algorithm 1 $\varepsilon - greedy(Timesteps)$

$T_s = 0$
n = a randomly generated number at each time step
for $T_s < Timesteps$ **do**
    if $\varepsilon > n$ **then**
        Explore
        $T_s = T_s + 1$
    else
        Exploit
        $T_s = T_s + 1$
    end if
    Update reward
end for

The order of the arms, representing the task requests, is dynamic and changes as the algorithm learns from the updated task history. The estimation of bandits' probabilities is calculated using the following equation:

$$P^*(i,j) = \frac{1}{N_j(a_T=a(i,j))}\sum_{T=1}^{T=cur} \tau^*(i,j,T)\beta\,(a_T = a(i,j)) \qquad (1)$$

In Equation 1, $P*(i, j)$ denotes the estimated probability of a task request from the ith pickup point to the jth shelf. The variable T represents the time step when the tasks are generated. $N_j(a_T = a(i,j))$ represents the number of times the action $a(i,j)$ was chosen by the MAB solver corresponding to shelf j until the current time step T=*cur*.

The function β(X) is a binary function that returns 1 if the condition X is satisfied and 0 otherwise. The denominator $N_j(a_T = a(i,j))$= $\sum_{T=1}^{T=cur}(a_T = a(i,j))$ represents the total number of times the arm i was chosen by the MAB solver at shelf j until the current time step T=*cur*.

3.2 TIME SYNCHRONISATION BASED TASK SCHEDULING

The scheduling of job requests for a robotics arm stationed at a specific shelf is the main topic of this subsection. Based on the estimated probability ($P*$) received from the MAB solvers and the anticipated arrival time(s) of the mobile robot(s) involved, the order of these task requests is prioritised.

A joint effort between a mobile robot and a robotics arm is necessary to complete a task. The mobile robot moves into a parking spot within the robotics arm's reachable workspace while carrying the package. The missions are planned such that the robotics arm and mobile robot arrive at their destinations simultaneously. The goal of this cooperative strategy is to reduce the total amount of time needed to perform the work.

Based on the probabilities derived from $P*$, a priority order (p) is determined by selecting the element with the highest probability in each column. Sorting the indices of the rows based on the descending order of their corresponding highest estimates provides us with the prioritised order.

By following this approach, the task requests can be scheduled in an optimised manner, enabling efficient collaboration between the fixed-base and mobile robots and reducing the overall task completion time.

$$p = sort(i \leftarrow max-to-min(max(P*(i,:)))) \quad (2)$$

The estimated likelihood of task requests ($P*$), which is sorted in descending order based on the highest value among the rows of P, determines the priority order (p). The row numbers are sorted from highest to lowest, yielding the priority order p, by comparing the highest probability value in each row with the corresponding values in other rows. According to this priority order, robots carrying a particular load type that emerge at the start of p are more likely to accumulate duties than ones that appear later.

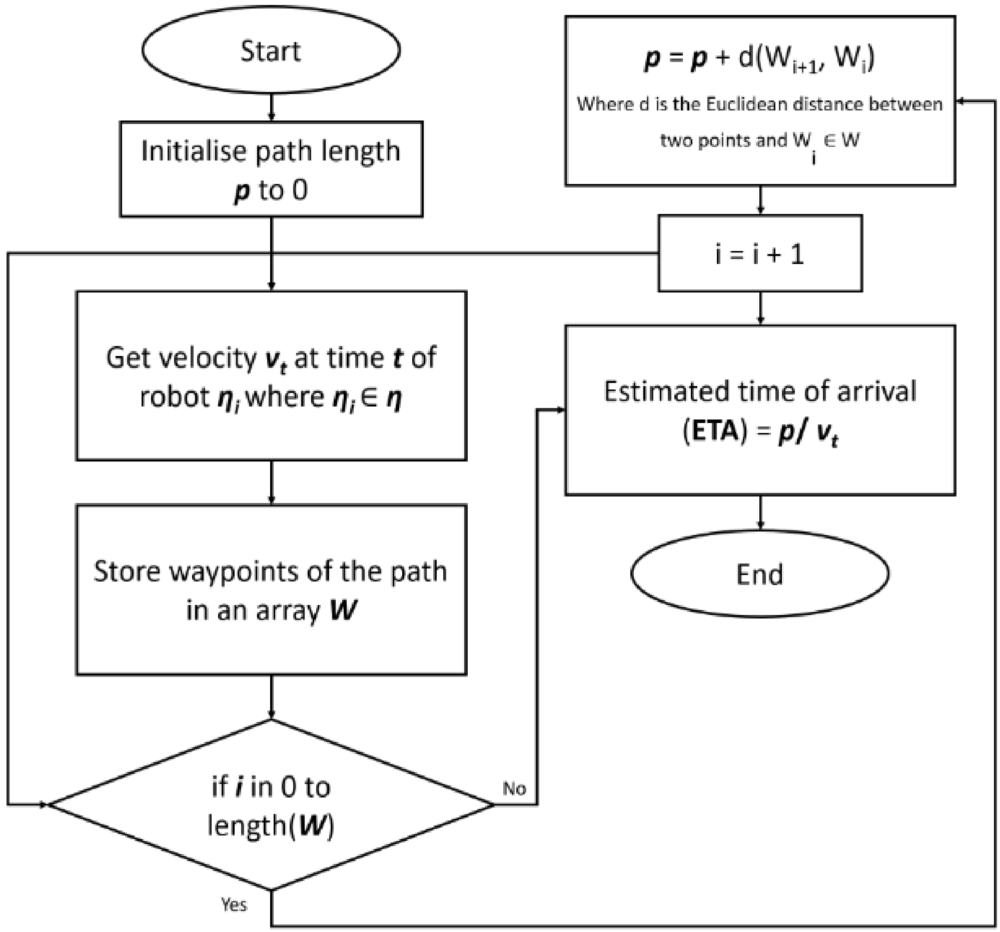

Figure 3.1: Calculation of ETA

As shown in Figure 3.1, the path length(s) and current velocity are both taken into account when estimating the time(s) at which the mobile robot(s) will arrive. RRT-connect (Rapidly-exploring Random Trees Connect), a motion planning algorithm created by Kuffner and LaValle in 2000, is used in the robotic arm. From the source and goal points, this method generates trees that steadily approach and eventually meet. The robotic arm's movement is then determined by choosing the shortest path among the set of nodes (tree).

When the robotics arm's movement time coincides with the mobile robot's anticipated arrival time, the robotics arm will begin to carry out the requested tasks that have been arranged. The mobile robot's expected arrival time is derived by dividing its present velocity by the distance it still needs to cover. Contrarily, the angular distance (argument) that the base joint must travel and the controller's velocity profile are what decide how long the robotics arm takes to complete the task.

The dynamic-window strategy was developed by Fox et al. [24] and is used by the mobile robot as both a local and global planner for navigation. The mobile robot's navigation while avoiding obstacles is made easier by the dynamic-window technique, an online collision avoidance system.

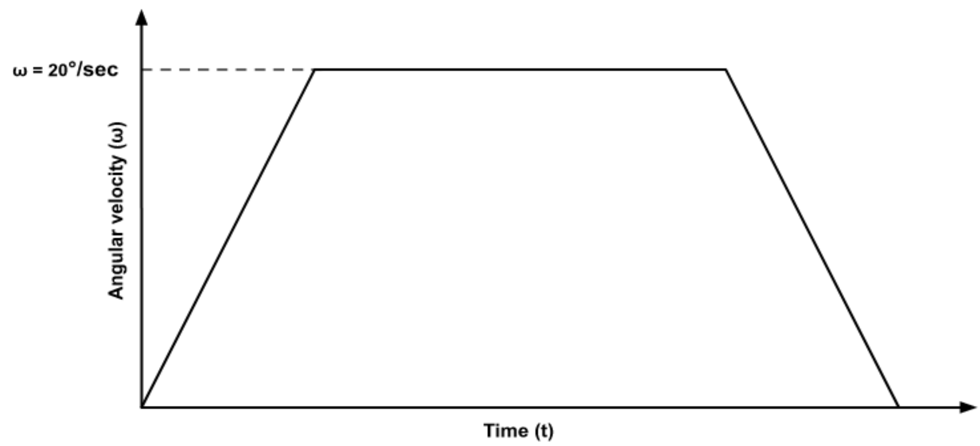

Figure 3.2: Velocity profile of the fixed base robot

The robotics arm's velocity profile, illustrated in Figure 3.2, follows a specific pattern. The angular velocity, denoted as ω, initially undergoes uniform angular acceleration for angular displacements θ less than 10 degrees. Once θ reaches 10 degrees, the angular velocity remains constant at 20 degrees per second.

However, during the last 10 degrees of the angular displacement, the angular velocity gradually decreases until it reaches zero.

In Algorithm 2, the time required to execute the movement of the robotics arm, denoted as γ_1, is calculated based on the angular displacement θ of the base joint. This calculation takes into consideration the angular velocity profile described earlier. By determining the specific value of θ, we can accurately compute the corresponding time required for the robotics arm to complete its movement.

Algorithm 2 Grid Approach(i,G_i,v_i,ω_i)

```
Robot Id ← i
Current location of robot ← G_i = [xg_i, yg_i]
Current linear velocity of robot ← v_i
Current angular velocity of robot ← ω_i
if GRID_LOCK(i,G_i) then
    POLICY_2(G_i)
    Grid Approach(i,G_i,v_i,ω_i)
end if
Forward simulation(i,v_i,ω_i) → Priority_i
if Conflict_Check(Priority_i) then
    POLICY_1(G_i, Priority_i)
end if
Continue moving to the goal using ROS-NAV1 architecture
```

Figure 3.3 provides a thorough overview of the suggested methodology, displaying the system's architecture and its associated parts. Two essential components make up the architecture: "Multi-armed bandit" and "Multi-agent coordination."

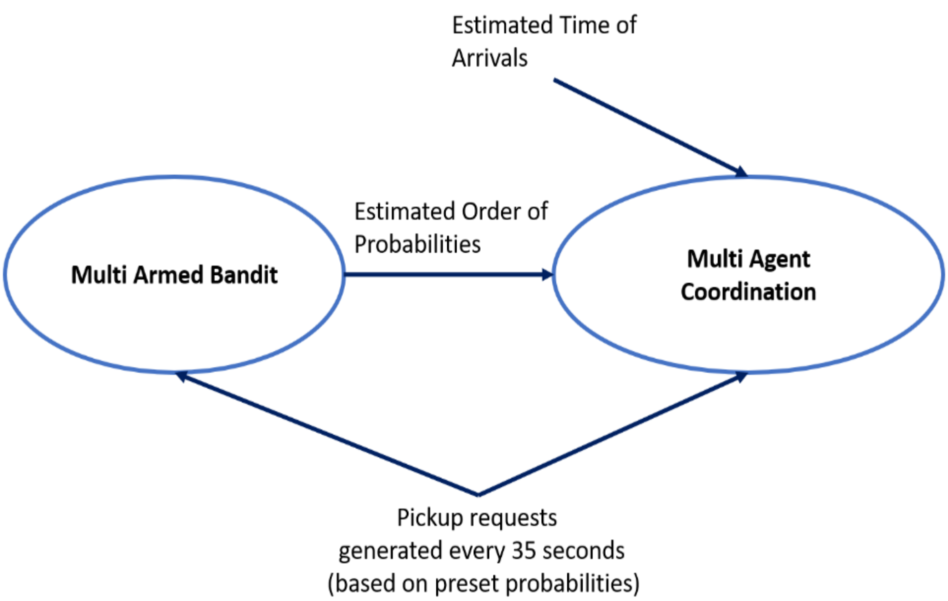

Figure 3.3: Data exchange within the proposed architecture's submodules

Estimated probability for the job requests are crucially provided by the "Multi-armed bandit" module. To produce these probabilities, this module makes use of task history. It assesses the likelihood of each task request by examining the past data, allowing for informed task scheduling decision-making.

On the other hand, the "Multi-agent coordination" module takes into account various factors such as movement time, estimated time of arrivals, and the probability estimates obtained from the "Multi-armed bandit" module. This module focuses on coordinating the actions of multiple agents involved in the system. It carefully plans the sequence of tasks, considering factors like the time taken for movement and the expected arrival times of the agents. By incorporating these variables, the module optimizes the coordination between the agents, ensuring efficient task execution and coordination.

The research primarily focuses on the task scheduling of robotics arms, specifically in a scenario involving a single shelf and three pickup points. The system consists of four agents, namely a robotics armic arm denoted as γ_1, and three distinct load-carrying mobile robots identified as η_1, η_2, *and* η_3. These mobile robots originate from three different pickup points.

In this context, the allocation of tasks is carried out by assigning each task to the mobile robot capable of carrying the specific load type associated with that task. After completing a task, the respective mobile robot returns to its designated pickup point from the shelf. This movement ensures that the mobile robots are positioned appropriately to execute future tasks, if any.

Figure 3.4 displays a flow chart outlining the sequential phases involved in the execution process of requests by a load-carrying robot. This flow chart shows the series of tasks carried out by a mobile robot, including the picking up of a load, moving to the shelf, carrying out the task, and then returning to the pickup place for additional task assignments.

This study dives into the complexities of coordinating and optimising the movement and work allocation of robotics arms and load-carrying mobile robots by investigating the task scheduling process in the context of one shelf and three pickup spots. The particular configuration taken into account in this work enables a targeted examination of the suggested approach's effectiveness in this specific environment.

The method presented in Algorithm 3 gives a thorough description of how the load-carrying mobile robots' task execution is carried out. A list of tasks, designated as *T*, is the algorithm's input. Each item in the list relates to the movement of a cargo from a particular pickup location to the specified shelf.

The algorithm runs based on four different states that the mobile robot can adopt while carrying out the task. The mobile robot is in the "start" state when it is at the pickup location and ready for the cargo to be placed on it. In the "to shelf" mode, the mobile robot moves with the load in the direction of the shelf, attempting to locate a parking space that the robotic arm can access.

When the mobile robot has arrived at the shelf, it switches to the "pick" state and waits for the robotics arm to arrive so that the pick and put operation can be completed. This synchronisation guarantees that the robotics arm and the mobile robot both arrive to the package at the same time, facilitating effective task performance.

Upon completing the task, the mobile robot transitions to the "to start" state, indicating that it has finished its assigned task and needs to return to its corresponding pickup point. The mobile robot follows a predefined path to navigate back to the pickup point, allowing it to be in position for future task assignments.

The algorithm continues to iterate through this loop as long as there are unfinished tasks remaining in the list. By managing the mobile robot's states and coordinating its movements with the robotics arm, the algorithm facilitates the smooth execution of tasks and ensures that the mobile robot is optimally positioned to carry out subsequent load transportation assignments.

The robotics arm stationed at a specific shelf employs the MAB scheduler algorithm, described in Algorithm 1, to prioritise the pickup requests it receives. The algorithm outputs a sequence of tasks denoted as E, which represents the order in which the robotics arm should execute the tasks. Using the multi-armed bandit formulation, the algorithm estimates the priority to be assigned among the mobile robots approaching the shelf at the current moment.

To ensure efficiency and timeliness, the robotics arm focuses on prioritising the requests from mobile robots that are within a predefined threshold δ, which corresponds to the time required for executing a pick and place task. By considering only the requests within this threshold, the scheduler can make informed decisions and reduce waiting times for the mobile robots.

The movement of the mobile robot and robotics arm are synchronised by precisely timing the robotic arm's package collection and delivery. Strong task scheduling is ensured by this synchronicity. When the mobile robot's anticipated arrival time coincides with the robotics arm's movement time, the robotics arm begins to move in the direction of the mobile robot's parking location.

Algorithm 3 MAB Scheduler(ETA,$\widehat{p}$)

```
Estimated Time of Arrivals ← ETA
for ∀x ∈ ETA do
    if x < δ then
        Append load type of x to list E
    end if
end for
count=0
for ∀y ∈ p̂ do                                        ▷ sort E subject to p̂
    if y ∈ E then
        Move y in E to position count
        Shift the requests after count by one position
        count = count+1
    end if
end for
```

With this method, the robots' ability to coordinate with one another is optimised, and the scheduler's overall performance is improved.

The priority order p derived by the MAB algorithm is used to schedule tasks. The objective is to reduce waiting times for mobile robots that have a higher likelihood of accruing additional jobs. A deterministic scheduler based on the first-come-first-serve (FCFS) technique is also taken into consideration while assessing the efficiency of the MAB task scheduler. According to the FCFS strategy, regardless of the task request history, the mobile robot that is projected to reach the shelf first based on its estimated time of arrival (ETA) is scheduled first for the pick and place operation.

It is possible to investigate and assess the efficiency of the MAB algorithm for job prioritising and scheduling by contrasting it with the FCFS scheduler.

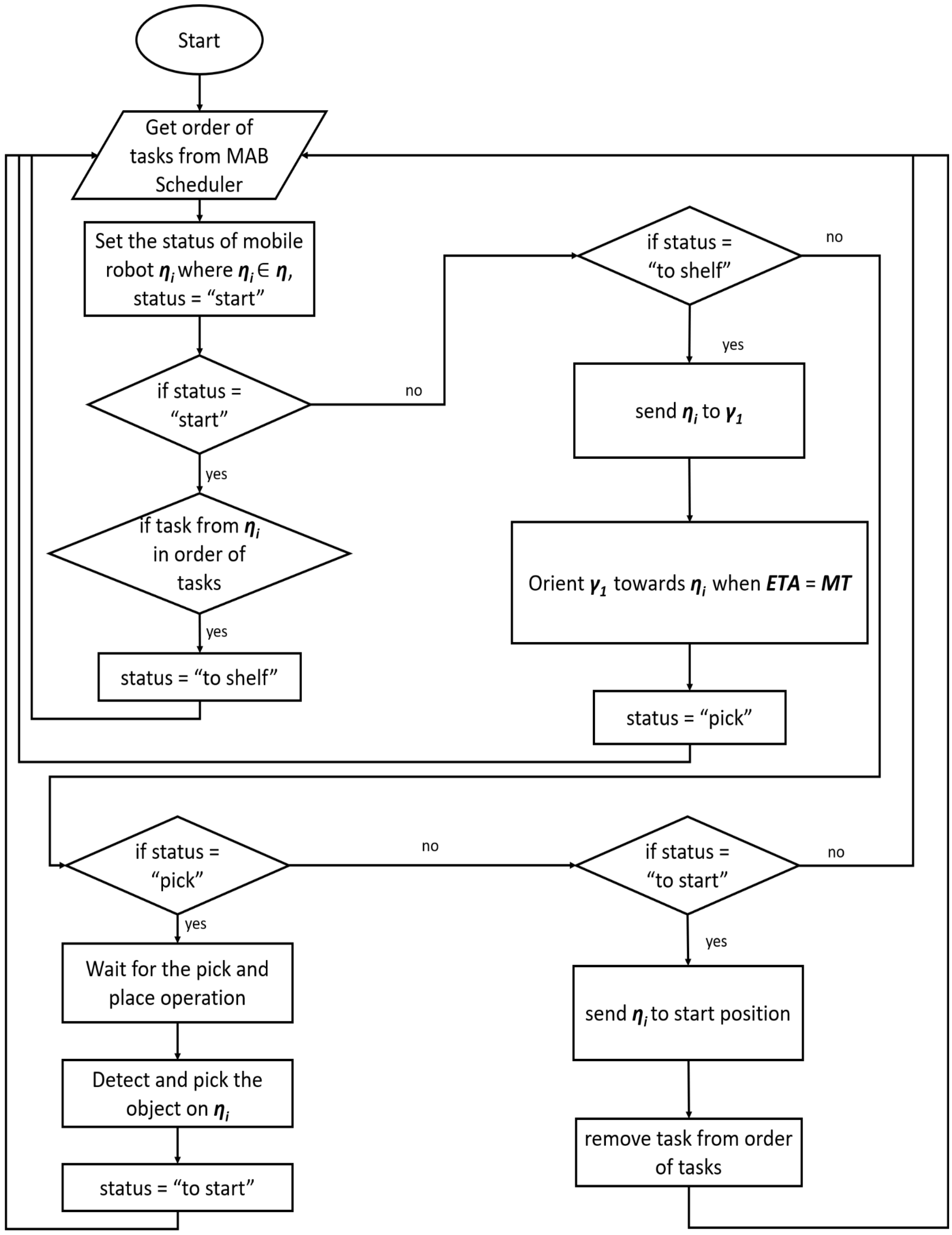

Figure 3.4: Mobile robot task execution

3.3 USER STUDY FOR HUMAN IN THE LOOP

3.3.1 Microsoft HoloLens 2

To conduct the user study, a cutting-edge device known as Microsoft HoloLens 2 was used. This advanced mixed reality headset, developed by Microsoft Corporation, merges the physical world with virtual holographic content, facilitating seamless interaction with digital objects and information. HoloLens 2 represents a significant leap forward from its predecessor, HoloLens, offering enhanced features and capabilities suitable for a wide range of applications and industries.

To precisely detect and interpret user movements and gestures, the headgear incorporates a number of sensors and tracking technologies. These include depth sensors and RGB cameras, which make it possible to map the environment and precisely track hand motions. This makes it possible for users to manipulate virtual objects by pinching, grabbing, and other motions while viewing holographic material. Additionally, voice commands are available, allowing for hands-free navigation and control.

Figure 3.5: Microsoft HoloLens2

The design of HoloLens 2 placed a high priority on comfort and prolonged wear. The user's head and neck experience the least amount of stress possible thanks to the headset's lightweight design and careful weight distribution. It has a headband that is adjustable and a padded visor to ensure a snug fit for people of different sizes. Additionally, the visor has a flip-up feature that enables seamless switching between mixed reality and the outside world without taking off the full headset.

Advanced spatial sound technology is used into HoloLens 2 to enhance the immersive experience. It creates a three-dimensional audio environment that precisely arranges sound cues in relation to the user's perception of space by utilising built-in speakers and powerful audio algorithms. As a result, authentic and engaging aural experiences are produced, which heightens the sense of presence in the mixed reality environment.

The gadget has a powerful onboard computer that makes it possible to process and render sophisticated holographic content in real time. Additionally, it provides a variety of cutting-edge networking options, such as Wi-Fi and Bluetooth, enabling easy integration with other tools and services. For effective data transfer and charging, it also supports USB-C connectivity.

For HoloLens 2, Microsoft offers a thorough development environment that enables programmers to produce a wide range of applications and experiences. Developers can create mixed reality applications with interactive holographic content by utilising resources like the Microsoft Mixed Reality Toolkit (MRTK) and well-known game engines like Unity or Unreal Engine. The Windows Mixed Reality platform provides a wide range of experiences and apps for use in many different sectors, including design and engineering, healthcare, entertainment, and more.

3.3.2 User Study

A user study was conducted in a Mixed Reality (MR) environment to gain insights into how a human task allocator prioritises tasks in a multi-armed bandit scenario. The study aimed to understand the decision-making process and priorities of human task allocators, which can then be utilised to customise resource allocation strategies.

The MR environment utilised in the study allows for teleoperation by a human operator. It consists of a warehouse model that closely matches the spatial layout of a real-world warehouse. The mobile robot within the environment can be controlled by the user, either from inside the warehouse or remotely, through a stereoscopic display that provides a visual representation of the warehouse layout. Microsoft HoloLens 2.0 was employed to create the mixed reality experience.

Figure 3.6: Immersive MR setup for the study

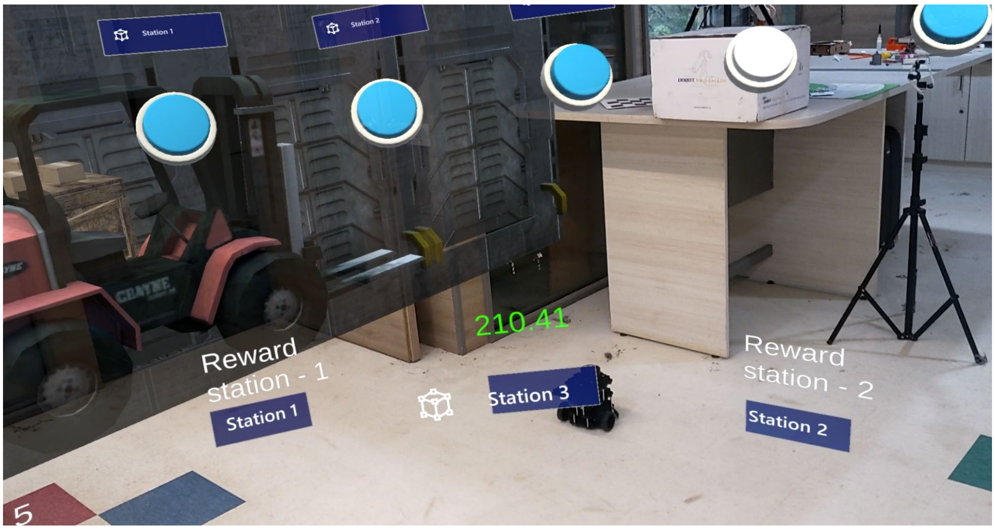

Figure 3.7: Mobile robot executing the task scheduled by the human operator.

Through the MR interface, users can observe holographic representations of virtual objects that are seamlessly superimposed onto the real world. These objects include forklifts, wooden boxes, shelves, and the boundaries of the warehouse. The addition of these objects enhances the realism of the warehouse environment, providing users with an immersive experience (refer to figure 3.6). Users have the ability to interact with a mobile robot situated in the real-world setting, controlling its movements and operations through the MR interface.

By conducting the user study in this mixed reality environment, researchers aimed to capture and analyze the decision-making patterns and priorities of human task allocators. The study offers valuable insights into the allocation of tasks in a multi-armed bandit scenario, thereby facilitating the development of customised resource allocation strategies based on human preferences and expertise.

The study included the design of five reward stations, represented as holographic objects, placed on a two-dimensional map within the simulation environment. Users interacted with the environment by selecting the position of the mobile robot using corresponding buttons associated with each reward station. Once a reward station was chosen, the mobile robot would navigate from the base location to the selected reward station.

At the base station, the mobile robot could perform various activities such as scanning codes, picking up loads, and delivering materials. Each reward station displayed a reward value generated based on a predetermined probability distribution. The objective for users was to maximise their cumulative reward by identifying the station with the highest potential reward at each iteration.

The simulation environment, as depicted in figures 3.6 and 3.7, provided users with a visual representation of the reward stations, the mobile robot, and the associated rewards. Figure 3.7 specifically illustrates the observed reward by the user after scheduling the execution of task at reward station three using the mobile robot.

The rewards assigned to the stations followed either a Bernoulli distribution (eq. 3a) or a Gaussian distribution (eq. 3b) within a specified range of minimum and maximum values. This variation in reward distributions aims to introduce different levels of uncertainty and diversity in potential rewards, challenging users to make informed decisions to maximise their overall reward accumulation.

$$P(X = x) = \begin{cases} p, & if\ x = 1 \\ 1 - p, & if\ x = 0 \\ 0, & otherwise \end{cases} \tag{3a}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3b}$$

Chapter 04: EXPERIMENTS RESULT ANALYSIS

4.1 EXPERIMENT SETUP

4.1.1 Task Scheduling

Using the ROS-Gazebo architecture, a warehouse scenario was simulated with a number of pickup sites and a robotic arm. The robotic arm was placed on a shelf at coordinates [0,0.7] and the pickup locations were situated at (3,2), (-2.5,1.5), and (2,-4). At regular intervals of 40 seconds, task requests were produced with probabilities of 0.5, 0.3, and 0.9 from each pickup location to the shelf.

After completing empirical investigation, it was found that the variable, which indicated the threshold for giving task requests approaching the robotics arm priority, was 4.0 seconds. In Simulations 1 and 2, a comparison of the deterministic and stochastic job scheduling methodologies was carried out. The value of the parameter was fixed to 0.3 in both simulations.

Figure 4.1 presents a comparison between the deterministic and stochastic task scheduling approaches in Simulation 1, while Figure 4.2 illustrates the comparison in Simulation 2. The stochastic approach demonstrated notable improvements in reducing the total time taken to complete tasks for Robot 1, Robot 2, and Robot 3 in Simulation 1, with reductions of 24.5%, 62.3%, and 40.2%, respectively. In Simulation 2, the stochastic approach resulted in reduced total time for Robot 2 and Robot 3 by 2.3% and 11.1%, respectively. However, the total time taken for Robot 1's tasks increased by 43.2% when employing the multi-armed bandit-based approach.

Table 4.1 shows how long it took each robot (Robot 1, Robot 2, and Robot 3) to finish tasks in Simulations 1 and 2 using both the deterministic first-come, first-serve technique and the stochastic multi-armed bandit approach. using a stochastic method demonstrated significant time reductions for Robots 2 and 3

Table 4.1: Time required (in hours) to finish 100 tasks in Simulations 1 and 2

Robot	Simulation 1		Simulation 2	
	FCFS (hours)	MAB (hours)	FCFS (hours)	MAB (hours)
1	7.2	5.4	7.2	10.4
2	1.2	0.4	1.2	1.2
3	34.9	20.3	34.9	30.7

in both simulations. However, for Robot 1, the stochatic approach increased the total time taken to complete tasks. Additionally, Table 4.2 displays the cumulative time taken by Robot 3 to complete consecutive sets of 20 tasks using both the deterministic and stochastic approaches.

Table 4.2: Time (in hours) required for Robot 3 to perform 20 tasks in succession in Simulations 1 and 2

Tasks	Simulation 1		Simulation 2	
	FCFS (hours)	MAB (hours)	FCFS (hours)	MAB (hours)
20	2.52	1.42	2.52	1.76
40	6.33	3.73	6.33	5.41
60	10	6.25	10	9.42
80	15.29	8.54	15.29	13.53

Overall, the results indicate the effectiveness of the stochastic multi-armed bandit approach in reducing task completion time for certain robots in the simulated warehouse scenario.

The time difference between completing the first 20 tasks using the suggested multi-armed bandit-based strategy and the first-come-first-serve (FCFS) technique for Robot 3 in Simulation 2 is 0.76 hours. The difference in Simulation 1 is also 1.1 hours. Due to the accumulated waiting time, this disparity increases exponentially as the number of subsequent tasks rises. Figures 4.3 and 4.4 provide a visual representation of the rising completion time discrepancy.

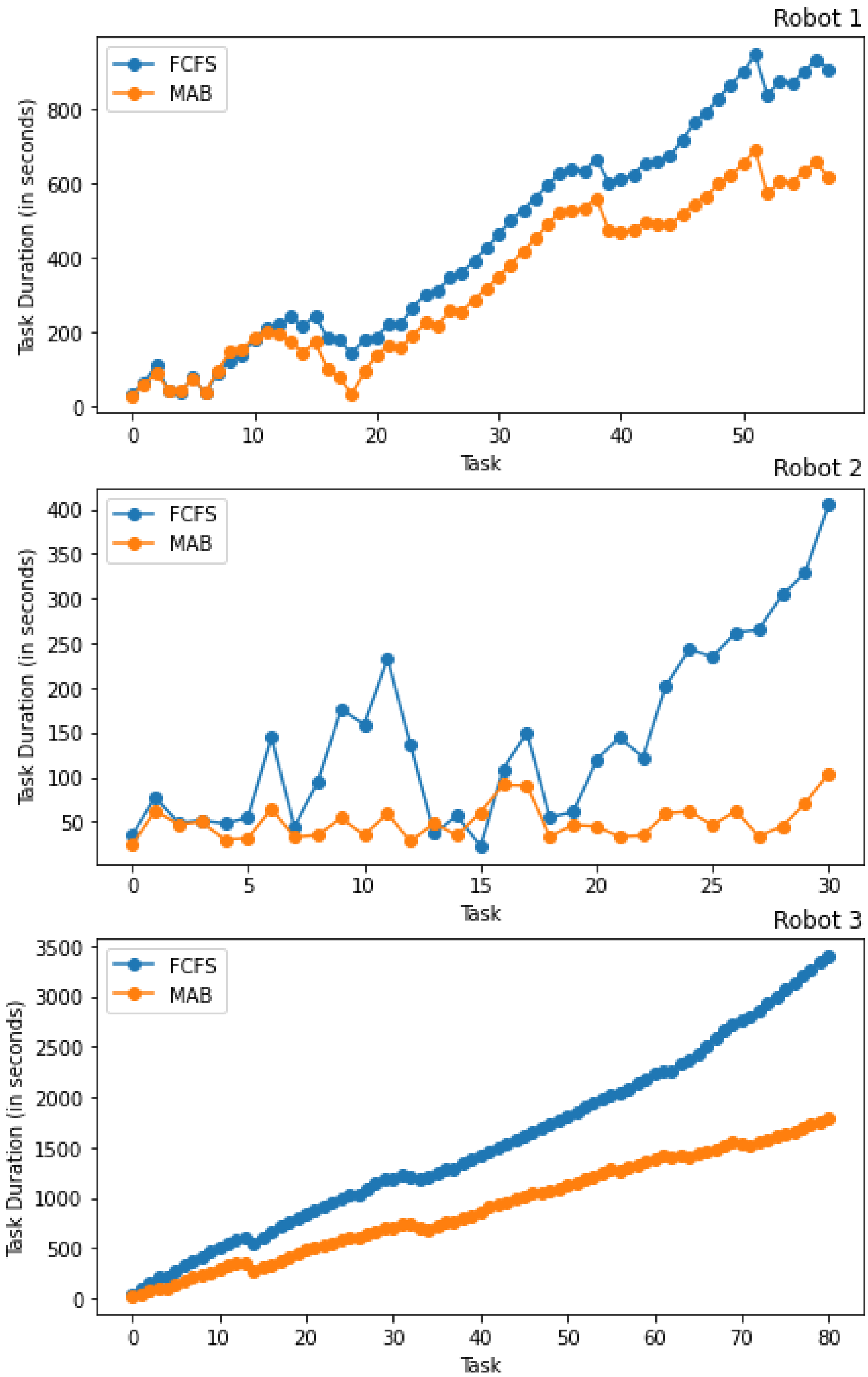

Figure 4.1 Task duration in Simulation 1

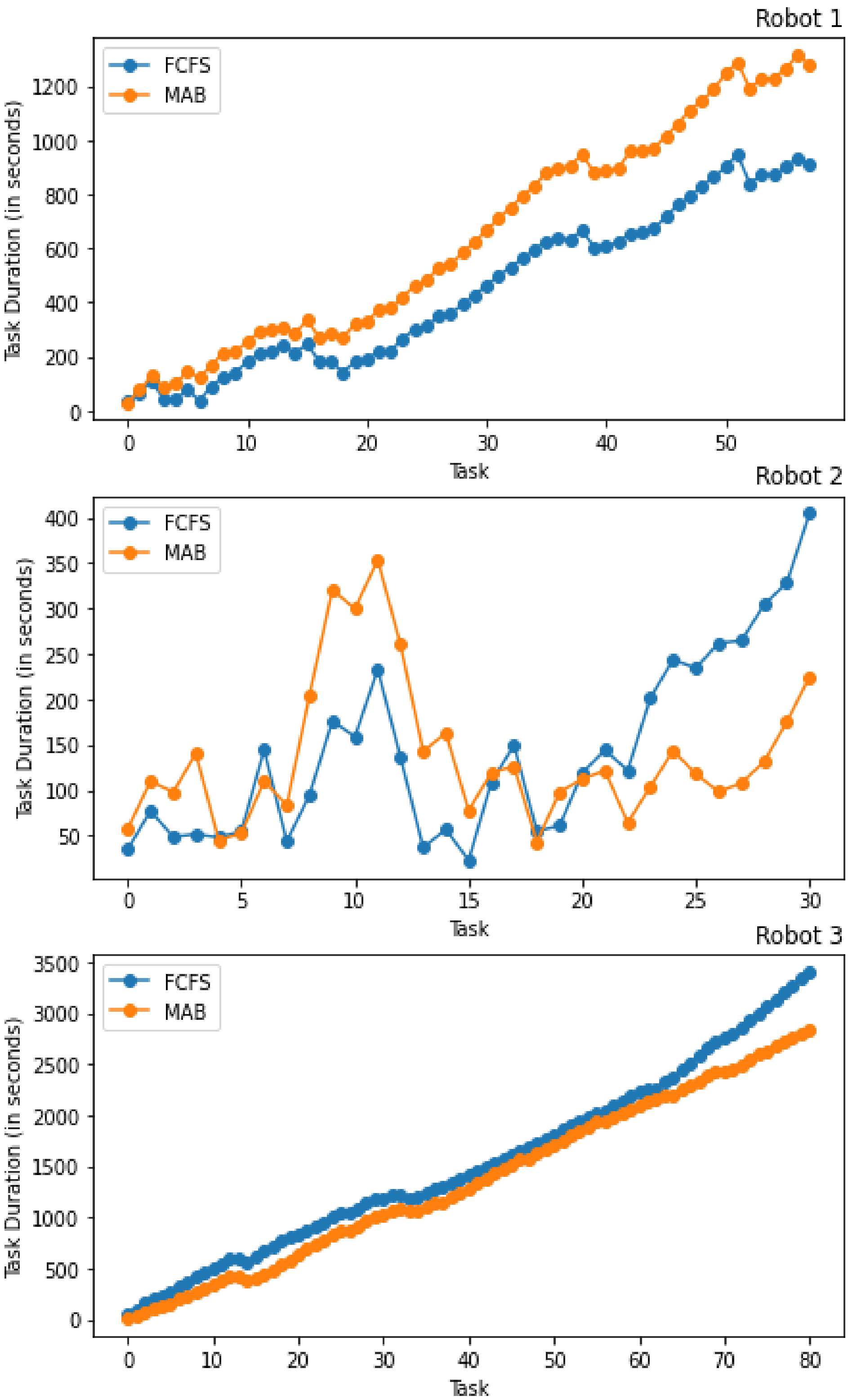

Figure 4.2: Task duration in Simulation 2

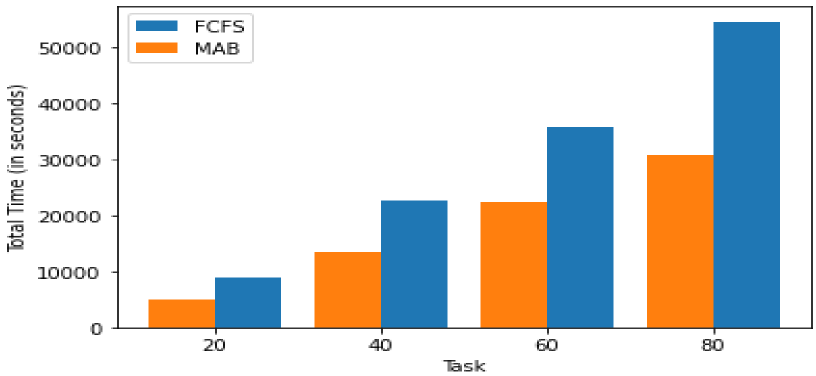

Figure 4.3: Cumulative task duration in Simulation 1 for Robot 3

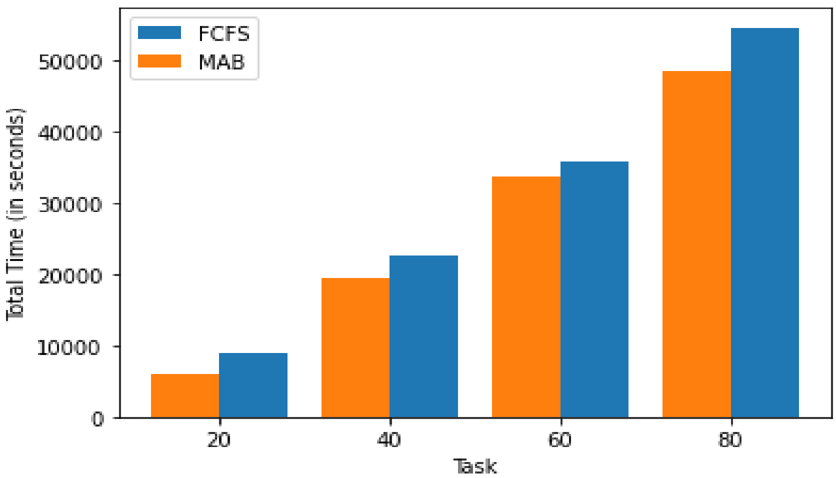

Figure 4.4: Cumulative task duration in Simulation 2 for Robot 3

Regarding Robot 2, in Simulation 1, the difference in time taken to complete the first six tasks between the proposed multi-armed bandit-based approach and the FCFS approach is 0.02 hours (as shown in Figure 4.5 and 4.6). However, in Simulation 2, the FCFS approach is faster by 0.05 hours. Notably, for the last set of six tasks, the multi-armed bandit-based approach outperforms the FCFS approach by 0.45 hours in Simulation 1 and by 0.23 hours in Simulation 2.

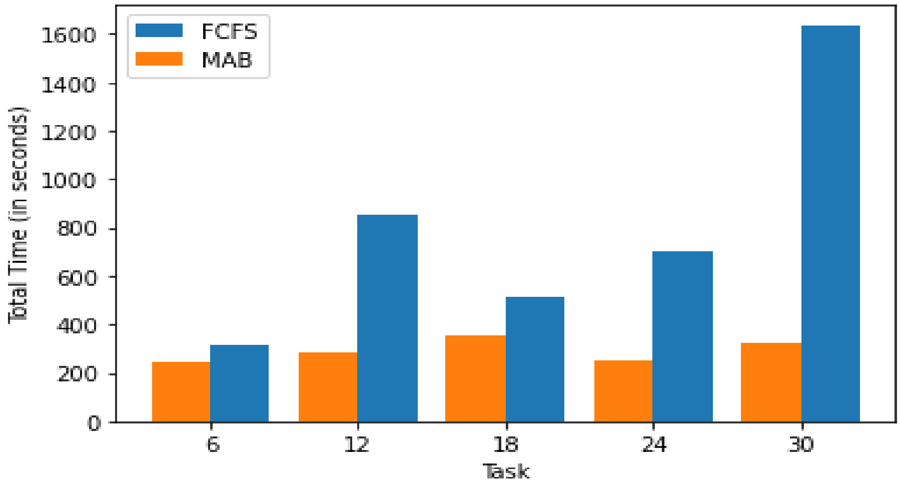

Figure 4.5: Cumulative task duration in Simulation 1 for Robot 2

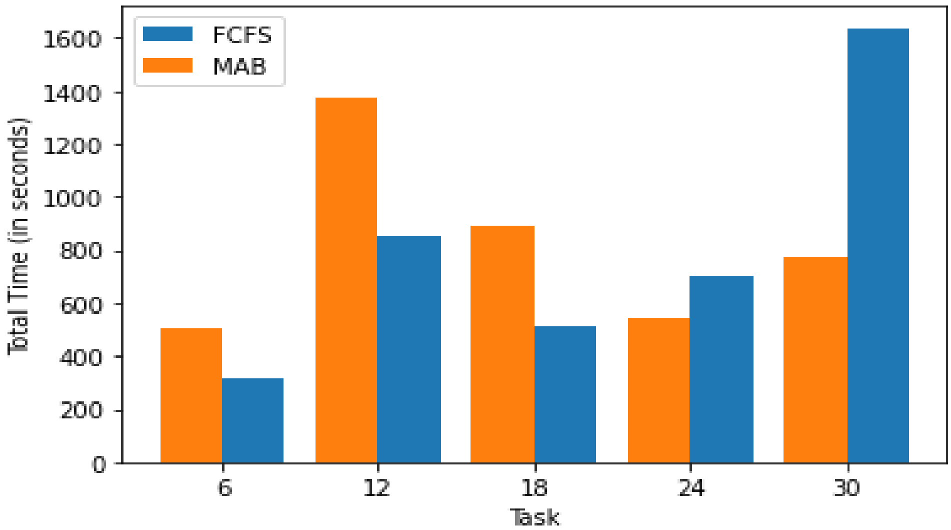

Figure 4.6: Cumulative task duration in Simulation 2 for Robot 2

For Robot 1 in Simulation 1, the difference in time taken to complete the first 11 tasks using the multi-armed bandit approach and the FCFS approach is 0.01 hours, with the FCFS approach being faster. In Simulation 2, the difference is 0.17 hours in favor of the deterministic FCFS approach.

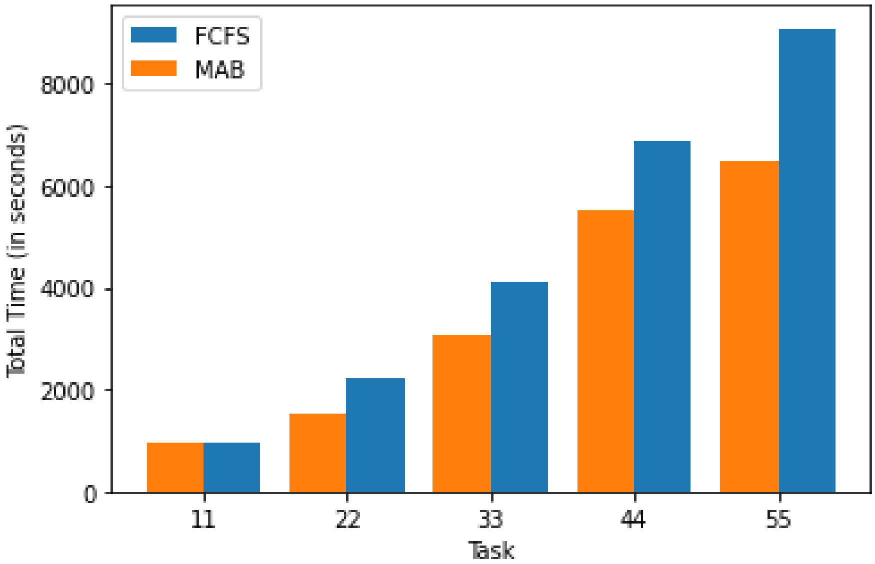

Figure 4.7: Cumulative task duration in Simulation 1 for Robot 1

However, for higher sets of consecutive 11 tasks, the stochastic multi-armed bandit-based approach proves to be faster in Simulation 1, whereas in Simulation 2, the FCFS approach is more efficient (as shown in Figure 4.7 and 4.8).

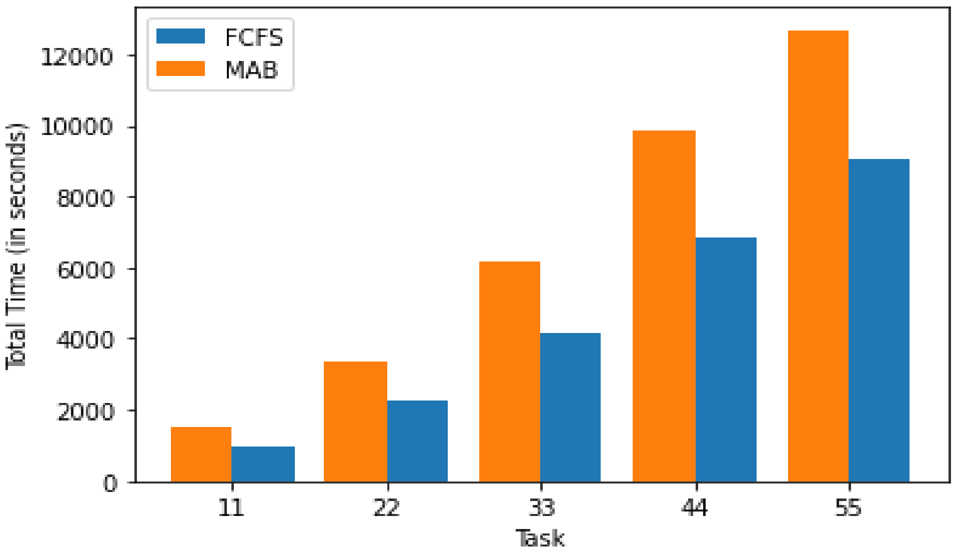

Figure 4.8: Cumulative task duration in Simulation 2 for Robot 1

Table 4.3: Time (in hours) required for Robot 2 to perform 6 tasks in succession in Simulations 1 and 2

Tasks	Simulation 1		Simulation 2	
	FCFS (hours)	MAB (hours)	FCFS (hours)	MAB (hours)
6	0.09	0.07	0.09	0.14
12	0.24	0.08	0.24	0.38
18	0.14	0.1	0.14	0.25
24	0.2	0.07	0.2	0.15
30	0.45	0.09	0.45	0.22

The total time taken by Robot 2 to complete successive sets of six tasks, using both the deterministic and stochastic methodologies, is shown in Table 4.3. In a similar vein, Table 4.4 displays the overall time required by Robot 2 to finish successive sets of eleven tasks using both methods.

Table 4.4: Time (in hours) required for Robot 1 to accomplish 11 tasks in succession in Simulations 1 and 2

Tasks	Simulation 1		Simulation 2	
	FCFS (hours)	MAB (hours)	FCFS (hours)	MAB (hours)
11	0.26	0.27	0.26	0.43
22	0.62	0.42	0.62	0.94
33	1.15	0.85	1.15	1.71
44	1.91	1.53	1.91	2.73
55	2.52	1.8	2.52	3.52

Overall, it can be observed that the total task completion time for all robots was higher when employing the FCFS approach compared to the multi-armed bandit (MAB) approach in both simulations. The MAB approach showcased its efficiency by significantly reducing task completion time, especially as the number of consecutive tasks increased.

Human in the loop

Figure 4.9 illustrates the integration of mixed reality technology into a human in the loop system for a pickup dispatch task. In this setup, a predefined set of locations on the mobile robot's indoor map in the real environment are presented as options for the human user to interact with the robot in the mixed reality world.

Figure 4.9: Mixed Reality implementation of warehouse scenario for a pickup dispatch task with human in the loop

To ensure efficient navigation, the mobile robot employs the A* algorithm for global path planning and utilises the dynamic window approach for dynamic obstacle avoidance and local path planning.

When the mobile robot reaches the robotics arm, the load is picked up and placed as required. To accurately localise the load within the coordinate system of the robotics arm, a camera attached to the ceiling is utilised. By applying classical computer vision techniques, such as colour recognition, the camera identifies the colour of the load. Additionally, a linear regression model is employed to establish a calibration between the image captured by the camera and the coordinate frames of the robot.

For a more comprehensive understanding of this scenario, the supplementary video provides a demonstration, showcasing the integration of mixed reality, the navigation capabilities of the mobile robot, load pickup and placement, and the use of computer vision techniques for load localization.

The participants in the study were assigned the task of maximising the reward in a bandit environment, specifically the Bernoulli and Gaussian bandit environments described in Lattimore et al. [25]. Figure 4.10 provides an explanation of the reward distributions for two users in the Gaussian Bandit environment.

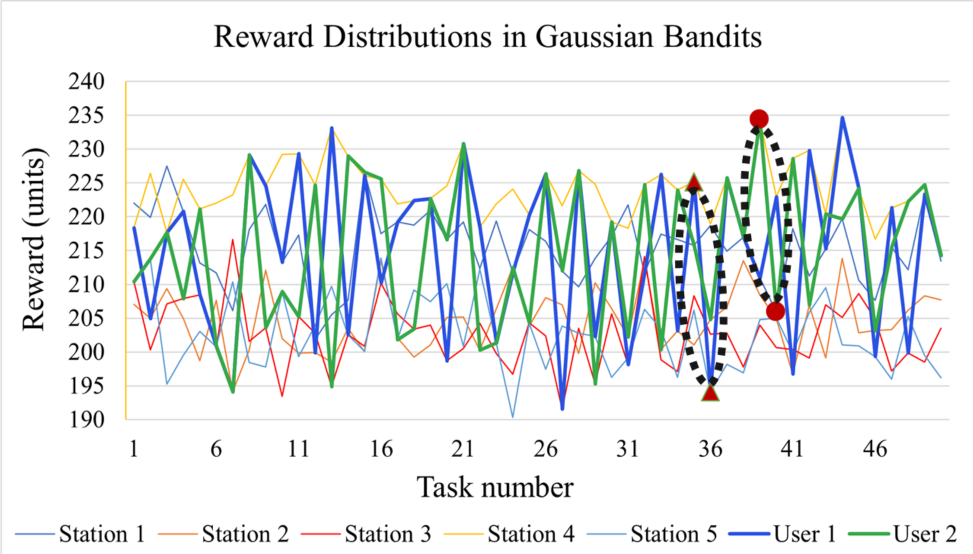

Figure 4.10: Comparison of rewards among users in Gaussian bandit environment

Interestingly, despite having the opportunity to explore different reward stations, the participants consistently opted for sub-optimal reward stations throughout the experiment. Even though the fourth reward station consistently offered the highest reward, the participants did not consistently select it. Instead, they frequently chose other stations, as indicated by the ellipses in the graph.

Upon analysis, it was observed that the participants' decision-making process was influenced by factors other than solely maximising the reward. They considered minimising the distance travelled by the robot as a significant factor in their decision-making process. This led them to select sub-optimal reward stations that were closer in proximity, rather than consistently choosing the station with the highest potential reward.

To enhance the validity of the study, the simulation system incorporated actual robot movement and the associated waiting time. This allowed for a more realistic representation of the task allocation process.

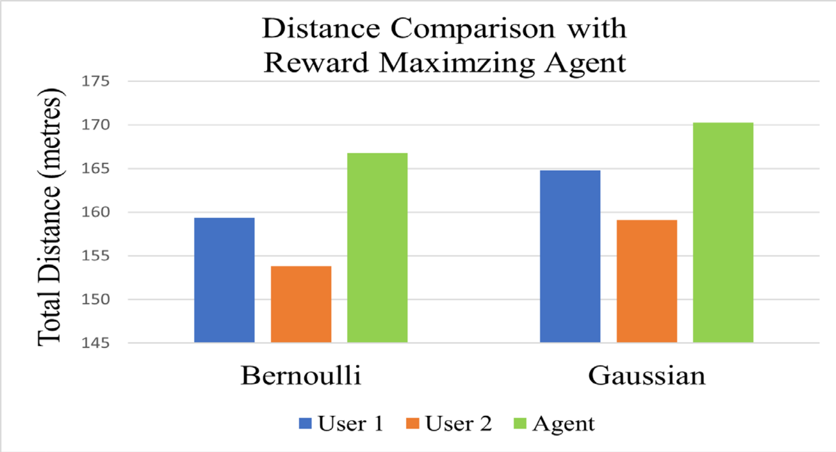

Figure 4.11: Comparison of total distance travelled by mobile robot among users and an imaginary agent which maximises the reward

Furthermore, the study compared the total distance travelled by an ideal agent that maximises the reward with the distance travelled by human participants, as shown in Figure 4.11.

Chapter 05: CONCLUSIONS

5.1 CONCLUSION

This project has addressed the challenges associated with task scheduling in a scenario involving multiple heterogeneous robots. A solution based on the multi-armed bandit technique has been proposed and extensively discussed within the context of this specific application domain.

The user study's findings made an important discovery about how humans allocate tasks. It has been observed that human task schedulers frequently prioritise minimising particular metrics, such as the total distance travelled, rather than rigidly adhering to the best course of action, which is to maximise rewards in a multi-armed bandit situation. The task scheduling procedure is made more complex by this behaviour.

5.5 FUTURE WORK

The following areas can be further researched in order to increase the job scheduling process' adaptability and effectiveness:

- More research may be done on how the multi-agent multi-armed bandit algorithm interacts with human task assigners. This requires examining the most effective ways to incorporate human input into the algorithm to control labour allocation decisions, such as criticism, preferences, and limits. A deeper comprehension of the mechanics of this interaction will make it easier to design flexible and user-friendly task scheduling solutions.
- The analysis and modelling of the decision-making processes and strategies used by human task allocators can improve the performance of the multi-agent multi-armed bandit algorithm. One machine learning method that can be used to learn from human behaviour and adjust the scheduling algorithm is reinforcement learning.
- The system performance can be enhanced by modifying the objective function of the scheduling algorithm to match the priorities of each individual human job allocator. Future research can examine methods

for collecting and simulating these preferences while taking into account factors like efficiency, equity, energy use, or user-specific goals. This customization can be carried out with the aid of machine learning, optimisation techniques, and user input.

- Depending on real-time feedback and system performance, the scheduling algorithm may adaptively modify the job allocation approach. By regularly evaluating the system's efficacy and taking user preferences into consideration, the algorithm may dynamically alter its decision-making process, increasing task allocation efficiency and user happiness.

These study directions will progress the field of work scheduling in heterogeneous MRS by including human factors and preferences into the algorithm design, resulting in more efficient and customised task allocation strategies.

REFERENCES

[1] G. Fragapane, R. de Koster, F. Sgarbossa, and J. O. Strandhagen, "Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda," *Eur. J. Oper. Res.*, vol. 294, no. 2, pp. 405–426, 2021, doi: https://doi.org/10.1016/j.ejor.2021.01.019.

[2] A. Slivkins, "Introduction to Multi-Armed Bandits," *Found. Trends® Mach. Learn.*, vol. 12, no. 1–2, pp. 1–286, 2019, doi: 10.1561/2200000068.

[3] R. Zlot and A. Stentz, "Market-based Multirobot Coordination for Complex Tasks," *Int. J. Rob. Res.*, vol. 25, no. 1, pp. 73–101, 2006, doi: 10.1177/0278364906061160.

[4] A. Viguria, I. Maza, and A. Ollero, "S+T: An algorithm for distributed multirobot task allocation based on services for improving robot cooperation," in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 3163–3168, doi: 10.1109/ROBOT.2008.4543692.

[5] F. Tang and L. E. Parker, "Distributed multi-robot coalitions through ASyMTRe-D," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 2606–2613, doi: 10.1109/IROS.2005.1545216.

[6] S. Stavridis and Z. Doulgeri, "Bimanual Assembly of Two Parts with Relative Motion Generation and Task Related Optimization," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 7131–7136, doi: 10.1109/IROS.2018.8593928.

[7] J. Borrell Méndez, C. Perez-Vidal, J. V. Segura Heras, and J. J. Pérez-Hernández, "Robotic Pick-and-Place Time Optimization: Application to Footwear Production," *IEEE Access*, vol. 8, pp. 209428–209440, 2020, doi: 10.1109/ACCESS.2020.3037145.

[8] R. Szczepanski, K. Erwinski, M. Tejer, A. Bereit, and T. Tarczewski, "Optimal scheduling for palletizing task using robotic arm and artificial

bee colony algorithm," *Eng. Appl. Artif. Intell.*, vol. 113, p. 104976, 2022, doi: https://doi.org/10.1016/j.engappai.2022.104976.

[9] Y.-C. Ho and H.-C. Liu, "A simulation study on the performance of pickup-dispatching rules for multiple-load AGVs," *Comput. Ind. Eng.*, vol. 51, no. 3, pp. 445–463, 2006, doi: https://doi.org/10.1016/j.cie.2006.08.007.

[10] J. Ichnowski *et al.*, "Fog Robotics Algorithms for Distributed Motion Planning Using Lambda Serverless Computing," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 4232–4238, doi: 10.1109/ICRA40945.2020.9196651.

[11] V. C. Kalempa, L. Piardi, M. Limeira, and A. S. de Oliveira, "Multi-Robot Preemptive Task Scheduling with Fault Recovery: A Novel Approach to Automatic Logistics of Smart Factories," *Sensors*, vol. 21, no. 19, 2021, doi: 10.3390/s21196536.

[12] N. Kousi, S. Koukas, G. Michalos, and S. Makris, "Scheduling of smart intra – factory material supply operations using mobile robots," *Int. J. Prod. Res.*, vol. 57, no. 3, pp. 801–814, 2019, doi: 10.1080/00207543.2018.1483587.

[13] M. Korein and M. Veloso, "Multi-armed bandit algorithms for spare time planning of a mobile service robot," *Proc. Int. Jt. Conf. Auton. Agents Multiagent Syst. AAMAS*, vol. 3, pp. 2195–2197, 2018.

[14] H. Claure, Y. Chen, J. Modi, M. F. Jung, and S. Nikolaidis, "Reinforcement Learning with Fairness Constraints for Resource Distribution in Human-Robot Teams," *CoRR*, vol. abs/1907.00313, 2019, [Online]. Available: http://arxiv.org/abs/1907.00313.

[15] A. Dahiya, N. Akbarzadeh, A. Mahajan, and S. L. Smith, "Scalable Operator Allocation for Multirobot Assistance: A Restless Bandit Approach," *IEEE Trans. Control Netw. Syst.*, vol. 9, no. 3, pp. 1397–1408, 2022, doi: 10.1109/TCNS.2022.3153872.

[16] G. Pini, A. Brutschy, G. Francesca, M. Dorigo, and M. Birattari, “Multi-armed bandit formulation of the task partitioning problem in swarm robotics,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7461 LNCS, no. May, pp. 109–120, 2012, doi: 10.1007/978-3-642-32650-9_10.

[17] E. B. Ozgul, S. Liemhetcharat, and K. H. Low, “Multi-agent ad hoc team partitioning by observing and modeling single-agent performance,” in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, 2014, pp. 1–7, doi: 10.1109/APSIPA.2014.7041644.

[18] M. C. Koval, J. E. King, N. S. Pollard, and S. S. Srinivasa, “Robust trajectory selection for rearrangement planning as a multi-armed bandit problem,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 2678–2685, doi: 10.1109/IROS.2015.7353743.

[19] C. Eppner and O. Brock, “Visual detection of opportunities to exploit contact in grasping using contextual multi-armed bandits,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 273–278, doi: 10.1109/IROS.2017.8202168.

[20] S. Krishnasamy, R. Sen, R. Johari, and S. Shakkottai, “Learning Unknown Service Rates in Queues: A Multiarmed Bandit Approach,” *Oper. Res.*, vol. 69, no. 1, pp. 315–330, 2021, doi: 10.1287/opre.2020.1995.

[21] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, “Progress and prospects of the human–robot collaboration,” *Auton. Robots*, vol. 42, no. 5, pp. 957–975, 2018, doi: 10.1007/s10514-017-9677-2.

[22] O. Liu, D. Rakita, B. Mutlu, and M. Gleicher, “Understanding human-robot interaction in virtual reality,” in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-*

MAN), 2017, pp. 751–757, doi: 10.1109/ROMAN.2017.8172387.

[23] J. Patel and C. Pinciroli, “Improving Human Performance Using Mixed Granularity of Control in Multi-Human Multi-Robot Interaction,” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2020, pp. 1135–1142, doi: 10.1109/RO-MAN47096.2020.9223553.

[24] D. Fox, W. Burgard, and S. Thrun, “The dynamic window approach to collision avoidance,” *IEEE Robot. Autom. Mag.*, vol. 4, no. 1, pp. 23–33, 1997, doi: 10.1109/100.580977.

[25] T. Lattimore and C. Szepesv, “Bandit Algorithms.”