PRECISE TEMPERATURE CONTROL FOR REFRIGERATORS

(BUZDOLAPLARI IÇIN HASSAS SICAKLIK KONTROLÜ)

by

MEHMET KERİM PEKER, B.S.

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

in

COMPUTER ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

Jun 2024

ACKNOWLEDGMENTS

Primarily, I represent my sincere thanks to all my instructors who built my technical background that plays a essential role in my success.

Foremost, I would like to express my special gratitude to Assist. Prof. Dr. Ahmet Teoman Naskali for his substantial guidance and supervision, as well as his contributions to my vision and work through the research process.

Also, I would like to thank Sanem Kılıçaslan, Nuri Özbey and Gökhan Göl for their contributions and continuous help.

Last but not least, I owe a debt of gratitude to my wife, Ruba Pekdemir Peker, my daughter, Neva Peker, and my family for their endless belief, support, and love.

Jun 2024
Mehmet Kerim PEKER

TABLE OF CONTENTS

LIST OF SYMBOLS

IoT	Internet of Things
AI	Artificial Intelligence
ML	Machine Learning
RL	Reinforcement Learning
PID	Proportional-Integral-Derivative
TD	Temporal Difference
RBF	Radial Basis Function
FF	Fresh Food Compartment
FRZ	Freezer Compartment
MCU	Microcontroller Unit
IP	Intellectual Property
ReLU	Rectified Linear Unit

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This thesis explores the development of an advanced adaptive PID controller for optimal temperature control in refrigerators while maintaining low power consumption. The proposed controller combines PID and reinforcement learning with a Radial Basis Function network, designed under an Actor-Critic structure. The framework adapts PID parameters based on reinforcement learning techniques, enabling real-time adjustments in response to system conditions. The paper's focus on leveraging reinforcement learning ensures that the controller can learn from experience and fine-tune its behavior, resulting in improved efficiency and reduced power consumption. The Proportional, Integral, and Derivative components are crucial for achieving precise control for PID. The reinforcement learning approach introduces a novel adaptive weight updating rule adjustment over RBF network, allowing for on-line tuning of the PID parameters based on predictive outputs and reinforcement signals. By sharing the hidden layer between the Actor and Critic, the storage space requirement is minimized, and computational costs are reduced. Moreover, customer profiling through door actuation and ambient temperature data from IoT refrigerators allows for user-adaptive cooling algorithms. These algorithms improve cooling performance, reduce energy consumption, and minimize temperature fluctuations. This integration of reinforcement learning into adaptive PID controllers reflects a shift toward more intelligent and flexible control systems in modern refrigeration technology. The data collection phase involves obtaining anonymized information on refrigerator door openings, ambient temperature, and humidity levels from IoT refrigerators. The deployment stage consists of both Cloud Deployment and Control Board Deployment, with clustering algorithms like K-means used to categorize data and generate user profiles. By deploying these profiles to the cloud and control board, the learning rate and PID parameters are optimized, resulting in a more efficient cooling system with reduced energy consumption.

Keywords : Reinforcement Learning, IoT, PID

ÖZET

Bu tez çalışmasında, düşük güç tüketimini korurken buzdolaplarında optimum sıcaklık kontrolü için gelişmiş bir uyarlanabilir PID kontrol cihazının geliştirilmesi amaçlanmıştır. Önerilen kontrolör, PID ve takviyeli öğrenmeyi, Aktör-Eleştirmen yapısı altında tasarlanan Radyal Temel Fonksiyon ağıyla birleştirir. Çerçeve, PID parametrelerini takviyeli öğrenme tekniklerine göre uyarlayarak sistem koşullarına yanıt olarak gerçek zamanlı ayarlamalara olanak tanır. Çalışmanın takviyeli öğrenmeden yararlanmaya odaklanması, denetleyicinin deneyimlerden ders alabilmesini ve davranışında ince ayar yapabilmesini sağlayarak verimliliğin artmasını ve güç tüketiminin azalmasını sağlar. PID kontrolöründeki Oransal, İntegral ve Türev bileşenleri hassas kontrol için çok önemlidir. Takviyeli öğrenme, RBF ağındaki ağırlıkların ayarlanması yoluyla yeni bir uyarlanabilir güncelleme kuralı getirerek, tahmine dayalı çıktılara ve takviye sinyallerine dayalı olarak PID parametrelerinin çevrimiçi olarak ayarlanmasına olanak tanır. Gizli katmanın Aktör ve Eleştirmen arasında paylaşılmasıyla depolama alanı gereksinimi en aza indirilir ve hesaplama maliyetleri azalır. Ayrıca, IoT buzdolaplarından kapı çalıştırma ve ortam sıcaklığı verileri aracılığıyla müşteri profili oluşturma, kullanıcıya uyarlanabilir soğutma algoritmalarına olanak tanır. Bu sayede enerji tüketiminde ve sıcaklık dalgalanmalarında azalma elde edilir. Veri toplama aşaması, IoT buzdolaplarından kapı açıklıkları, ortam sıcaklığı ve nem seviyeleri hakkında anonimleştirilmiş bilgilerin toplanmasını içerir. Çalışma iki aşamadan oluşmaktadır. Bulut tarafında çalışan K-means kümeleme algoritması ve gömülü sistem üzerinde çalışan kontrol algoritması. Verileri kategorilere ayırmak ve kullanıcı profilleri oluşturmak için kullanılan K-means gibi kümeleme algoritmaları bulutta çalışır. Bu profillerin gömülü sisteme gönderilmesiyle öğrenme hızı ve PID parametreleri optimize edilir, bu da daha az enerji tüketimiyle daha verimli bir soğutma sistemi sağlar.

Anahtar Kelimeler : Takviyeli Öğrenme, IoT, PID

1 INTRODUCTION

Storage plays a vital role in food preservation and one of the best way to store food is cooling. It is the process of reducing the temperature of a substance or system to a level lower than that of its surroundings. This process is accomplished by removing heat from one medium and transferring it to another medium. It is used in many areas such as refrigeration, food preservation, medical treatments, industrial processes and comfort. With the advancement of technology, cooling methods have been changed and replaced by modern refrigerators found in every home.

The development of refrigeration technology is very effective in the formation of the history of refrigerators. Although the principle of refrigeration was first discovered in the 18th century, widespread use of refrigerators did not occur until the 20th century. The first domestic refrigerators appeared in the early 1900s. General Electric's Monitor Top refrigerator, introduced in 1927, became one of the first widely accepted home refrigerators, using a more reliable and safer sealed refrigeration system. Nowadays, refrigerators have become an indispensable part of modern kitchens. They focus on energy efficiency and sustainability. Advanced technologies such as variable-speed compressors, digital temperature controls, and energy-efficient lighting are common. Refrigerators, available in different sizes, features and price ranges, make our lives easier by keeping food fresh for a long time.

Refrigerators have undergone significant evolution over the centuries, transitioning from simple ice-based cooling methods to advanced, energy-efficient appliances. In the past three decades, there has been a significant rise in the usage of household electrical appliances due to advancements in technology and reduced costs of these devices. Continuous development of refrigerators aims to improve food preservation. In recent years, global warming and new regulations have become important factors that refrigerator manufacturers consider when designing their products. Energy efficient precise temperature control is the key point for both reducing food waste and energy consumption. Refrigerators are among the most energy-consuming appliances in modern households. According to 2020 data from United States Energy Information Administration, refrigerators account for 7.9% of the average residential site electricity consumption (U.S Energy Information Administration, 2020). This situation poses both a significant economic burden and an environmental problem. Improvements in refrigerator design can help lower electricity bills and limit greenhouse gas emissions by significantly reducing energy consumption.

At this point, it is important to develop and implement a control algorithm for temperature oscillation. This approach prioritizes minimizing temperature swings, adapts to user behavior and environmental factors, and ensures energy efficiency and affordability. With the combination of Proportional-Integral-Derivative (PID) which is traditional but effective control technique and Reinforcement Learning (RL), adaptive cooling algorithm that minimizes energy consumption and temperature oscillations can be made available to customers. RL allows the PID controller to learn and adjust its parameters in real-time, enabling it to overcome these challenges and achieve optimal performance even in dynamic environments.

PID is a fundamental control technique for automation and industrial processes (IEEE, 2006). A control system employs PID control to guide a process towards a desired outcome and maintain that state. An error signal is computed in PID control by the controller, representing the disparity between the desired and actual values in the system. By minimizing this error, PID control ensures that the system remains at the desired state. Widely utilized across various industrial processes and automation tasks, PID control enhances system stability and accuracy. Its applications range from temperature regulation, velocity adjustment, and pressure management to water level supervision and robotic operations. Traditional PID control can struggle with complex systems or those with significant disturbances. These challenges can arise from factors like nonlinear dynamics, time delays, or unpredictable environmental conditions.

Reinforcement Learning, a subset of ML, focuses accomplishing tasks or reaching objectives within a given environment by training an agent (Sutton and Barto, 1998). This approach guides the agent in choosing actions and improving possible outcomes based on the actions. Essentially, Reinforcement Learning involves the agent striving to excel at a given task by comprehending agent's surroundings, then maximization the rewards it receives as well. Through continuous interaction with the environment, the agent learns, earns rewards, and refines its action strategies iteratively.

Understanding user habits and the control mechanisms employed is crucial for optimizing refrigerator operation, considering its direct interaction with users. This knowledge allows for the efficient management of energy consumption and heat distribution across racks, resulting in enhanced performance. Understanding usage patterns can inform the design of smarter appliances. Smart refrigerators that learn from user behavior can automatically adjust cooling settings to match usage patterns, further improving energy efficiency. They can also suggest optimal food storage practices based on user habits, reducing food waste.

2 LITERATURE REVIEW

Thermostatically controlled loads, particularly refrigerators, in global power consumption and their potential for improved control and demand response with the rise of IoT technology are significant. A study was carried out to evaluate the effect of temperature hysteresis bands on optimizing refrigerator energy consumption through the application of Model Predictive Control (Sabegh and Bingham, 2019). The influence of various hysteresis limits for domestic refrigerators' yearly energy consumption and the potential effects on the lifetime of their components is taken into consideration. Investigations are conducted to analyze the power consumption of refrigerators under various hysteresis conditions, considering scenarios with and without internal product. The results suggest that significant energy savings can be achieved by adapting hysteresis bands based on product mass. Moreover, the potential for real-time adaptation schemes and the utilization of IoT in large-scale refrigeration systems for optimal energy efficiency are discussed. Additionally, consideration is hinted at for variable speed drives and compressors for further efficiency improvements, paving the way for future research in this area. Implementing conventional hysteresis band control may result in a lack of flexibility. Ensuring consistent temperature levels is important in different use cases and different use environments, alongside the goal of energy conservation.

Temperature control in refrigerators varies from old-style bimetal thermostats to modern microcontroller-based algorithms. Research has been conducted on various control strategies such as adaptive relay controllers and PID controllers to increase energy efficiency and improve temperature regulation. However, most existing refrigerator control algorithms work in limited situations, which limits their potential to fully exploit control theory. The necessity of adaptive algorithms arises from factors such as variations in production, aging of ingredients, and variable food storage. It is aimed to use a model reference adaptive controller derived from a validated dynamic model to fully utilize the capabilities of refrigerator actuators. It is focused on the development of an adaptive controller to maintain consistent temperatures in refrigerator compartments while minimizing variations in commercially available refrigerators. It also provides insight into the potential for further improvements in refrigerator efficiency by considering variable speed compressors and advanced control strategies (Wait, 2012).

A research was made in (Hamid et al., 2009) on a PID control algorithm designed for regulating speed of fan based on a refrigeration system simulation. This work involves a comparison between PID control algorithm and the conventional on-off relay control method. It explains the components and operation of both controllers and underscores the importance of selecting appropriate parameters for optimal performance. It aims to compare the effectiveness of PID and on-off controllers in maintaining refrigerator temperature. The components and operation of the cooling system, including the compressor, coils, expansion valve, and fan, are detailed. Performance of both controllers under varying ambient temperatures, with the PID controller generally outperforming the on-off controller in maintaining the set temperature. It is confirmed the superior efficiency of the PID controller over the on-off controller and importance of adjusting PID parameters for optimal performance under different environmental conditions. Moreover, while the PID controller operates with predetermined parameters typically adjusted manually, it is also a necessity for enhanced adaptability and effectiveness in fluctuating conditions.

It is discussed the development of a data-driven supervisory Reinforcement Learning (RL) Proportional Integral Derivative (PID) controller using a modified Proximal Policy Optimization (m-PPO) algorithm for managing unstable industrial processes (Shuprajhaa et al., 2022). This controller addresses the complexities of unstable systems by integrating early stopping criteria, modified reward functions, and action repeat mechanisms into the m-PPO algorithm. Specifically designed to fine-tune the parameters of adaptive PID controllers for unstable systems while ensuring stability and performance, the study showcases the efficacy of the m-PPO algorithm through comparative performance assessments with alternative controllers on unstable Continuous Stirred Tank Reactor (CSTR) processes. Furthermore, the research underscores the versatility of the proposed approach across both stable and unstable industrial processes, highlighting its data-driven and universally applicable characteristics.

Conventional PID control techniques may lack of adaptability when dealing with complex and dynamically changing systems. To address this, an adaptive PID controller was formulated utilizing Reinforcement Learning, aiming to create a robust PID controller applicable to nonlinear systems, with its effectiveness validated through numerical simulations (Guan and Yamamoto, 2020). Employing a single RBF network for calculating the value and policy function simultaneously, the study incorporates a Temporal Difference (TD) error to integrate the error criterion, with gradient descent utilized for adaptive updates of network weights and kernel function. It has been demonstrated through numerical simulations, with an emphasis on its model-independent design, making it well-suited for complex real-world systems. The research offers insights into the fusion of RL technology with traditional control methodologies to enhance overall system performance.

A RL supported adaptive PID controller provides improved performance than conventional PID control techniques when considering self-tuning PID parameters. Leveraging Actor-Critic learning, it is proposed a method which dynamically adjusts PID parameters, capitalizing on the on-line and model-free learning attributes of RL (song WANG et al., 2007). To enhance learning efficiency and reduce storage requirements, a single RBF neural network is employed to concurrently approximate the Actor's policy function and the Critic's value function. The RBF network takes inputs such as system error, first-order difference of error, and second-order difference of error. System states are mapped to PID parameters by the Actor, while the Actor's outputs are evaluated and the TD error is generated by the Critic. Updating rules for the RBF kernel function and network weights are determined based on TD error performance index and gradient descent. Simulation results demonstrate the efficiency, adaptability, and robustness of the proposed controller for complex nonlinear systems, outperforming traditional PID controllers.

Another adaptive PID control approach utilizing reinforcement learning tailored for wind energy conversion systems (WECS) has been carried out (Sedighizadeh and Rezazadeh, 2008). Actor-Critic learning optimizes PID parameters adaptively, leveraging reinforcement learning's model-free and online learning capabilities. To enhance storage efficiency and learning speed, a single RBF neural network approximates both actor and critic function simultaneously, with inputs including system error and its first and second-order differences. System states are mapped to PID parameters by the Actor, while the Actor's outputs are evaluated and the TD error is generated by the Critic. Updating rules for network weights and kernel function of RBF are determined using performance index based on TD error and gradient descent. Simulation results highlight the effectiveness, adaptability, and robustness of the proposed controller for

WECS, surpassing conventional PID controllers.

PID controllers are commonly utilized in industrial processes and automation because they are practical and generally provide good performance. However, optimal tuning of PID parameters can be challenging due to varying operational conditions and process uncertainties. Another approach involves first creating an initial approximate step-response model, allowing the reinforcement learning agent to train offline with minimal effort (Dogru et al., 2022). Once a satisfactory training level is achieved on the model, the agent is fine-tuned online with the actual process, enabling adaptation to real-world dynamics. This minimizes the training time on the real system and reduces unnecessary wear, making it suitable for industrial applications. The effectiveness of this sample-efficient method is demonstrated through a pilot-scale multi-modal tank system, with performance validation achieved through setpoint tracking and disturbance regulatory experiments.

After the studies described inside this section are examined, PID controllers, which are the traditional control method, are widely supported by using reinforcement learning. Inspired by the mathematical models and approaches used in these studies, it is aimed to realize a working algorithm that will make the refrigerator control algorithm smarter. In order to increase the effectiveness of the new algorithm to be studied and to eliminate differences that may occur depending on usage, it is also aimed to create an algorithm that will constantly adapt itself according to usage in order to classify users and add this classification to the newly designed algorithm.

3 REFRIGERATORS

A refrigerator operates continuously, 24 hours a day and seven days a week, to maintain foods or items at the desired temperature to preserve freshness and prevent spoilage. Although there are many different models and varieties of refrigerator products available, a refrigerator product generally consists of two compartments, freezer compartment (FRZ) and fresh food compartment (FF). A FRZ set to operate at temperatures between -15°C and -24°C is used for foods intended to be frozen and stored for long periods, while an FF set to operate at temperatures between 1°C and 8°C is used for preserving foods that will be refrigerated and consumed quickly.

Refrigerators can also be classified based on various factors, such as the type of application, type of controller, defrosting method, cooling circuit type, and design configuration. In terms of application, refrigerators can be for domestic or commercial use. When it comes to defrosting, they can be manual or automatic. As for compartment configuration, the refrigerator market is primarily comprised of six types of cold appliances : single-door refrigerators without freezers, single-door refrigerators with freezers, double-door fridge-freezers, side-by-side fridge-freezers, upright freezers, and chest freezers (Barthel and Götz, 2012). On the other hand, the most common ones are illustrated in Figure 3.1 (ASHRAE, 2014)

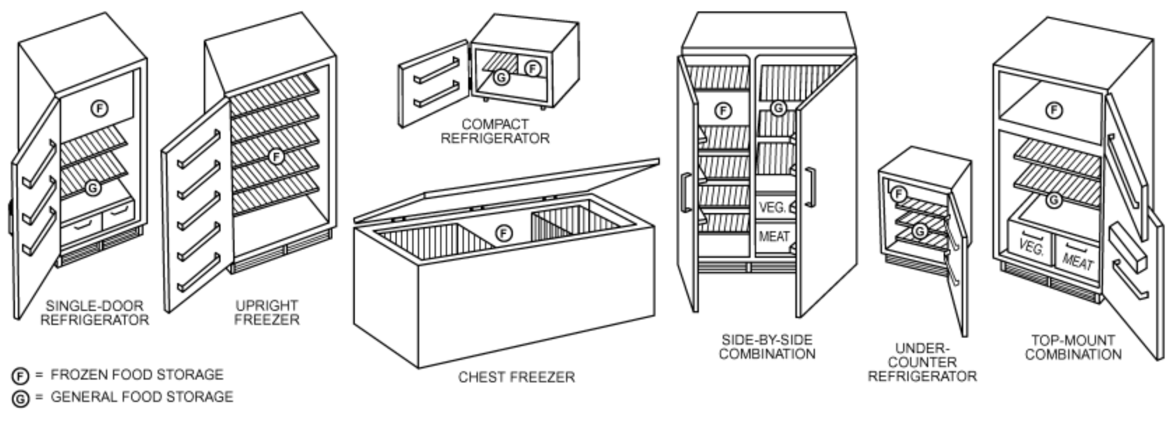

FIGURE 3.1 – Common Refrigerators and Freezers

3.1 Working Principles of Refrigerators

Just as there are many different models of refrigerators, there are also differences in the cooling systems of these models which basically are serial cooling system, semi-parallel cooling system and parallel cooling system for two or more compartments refrigerators. A serial cooling system in a refrigerator generally indicates a configuration where cooling components are linked in a sequential order. In this setup, a single cooling circuit provides cooling to multiple compartments, such as the FF and FRZ, in a linear sequence. Semi-parallel cooling system for refrigerator refers to the combination of series and parallel cooling systems. In this type of system, the cooling circuit serves multiple compartments, but maintains a balance between the independence of the parallel cooling system and the efficiency of the series cooling system. A parallel cooling system for a refrigerator is a type of cooling system that allows controlling the temperatures of different compartments by distributing the refrigerant flow across multiple independent cooling circuits. In such systems, each cooling circuit usually has its own evaporator, and these circuits can be driven by a single compressor or multiple compressors. Since the cooling circuits operate in parallel, each compartment can control its own refrigerant flow.

Although the mechanical design differences of these cooling systems were not focused on, a cooling system that is controlled by an electronic board design choice was made that would serve the purpose of this study. The control algorithms embedded in MCU control the operation of the compressor, fans, and defrost operation, among other functions. Their main objective is to maintain the desired air temperature within the compartments. The refrigerator cooling system that is used for this work has parallel cooling system illustrated in Figure 3.2. This system has multiple components. They are basically electronic control board, compressor, condenser, gas valve, FRZ evaporator, FF evaporator, FRZ fan, FF fan and temperature sensors illustrated in Figure 3.3. One of the key points of parallel cooling system is valve. It is essential to direct the refrigerant gas in the closed system within the cooling system to the relevant compartment. The compressor ensures that the refrigerant gas is conveyed to the evaporators, the heat of the hot refrigerant circulating through the condenser cooling system and returning to the compressor is discharged to the outside environment, the evaporators ensure that the compartment cools by transferring heat within the compartment, and the fans ensure that the cold air mass provided by the refrigerant gas in the evaporator is distributed within the compartment. In the parallel cycle cooling system, the FRZ evaporator and FF evaporator are connected to each other with a gas valve. Thus, the refrigerant gas is directed to whichever evaporator it is intended to be sent to.

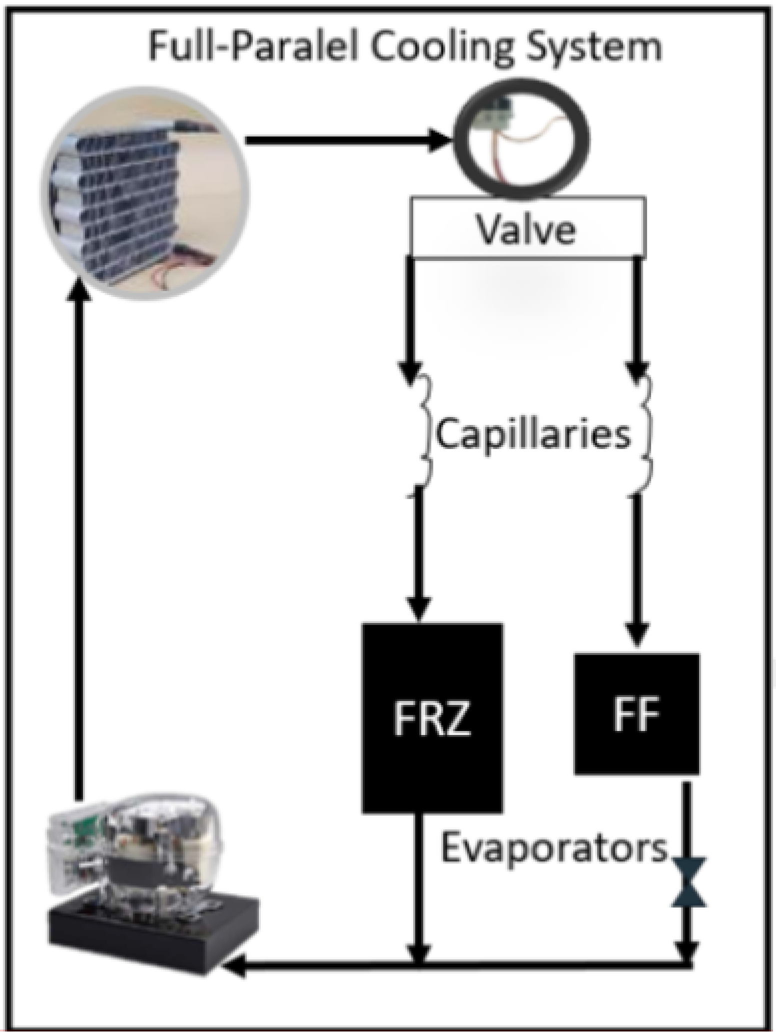

FIGURE 3.2 – Parallel Cooling System

In order to carry out the cooling process in refrigerators with a parallel cycle cooling system, the refrigerant gas coming out of the compressor first enters the condenser, then enters the gas valve which is controlled by electronic board. Control algorithm that is designed for this system decides where gas will be transferred. If the gas valve position is directed to the FF, the refrigerant gas reduces the temperature of the FF evaporator. If the gas valve position is directed to the FRZ, the refrigerant gas reduces the temperature of the FRZ evaporator. When the related fan placed in the compartment is activated according to the gas valve state, the cold air formed on the evaporator is

(a) Compressor (b) FF Evaporator (c) FF Fan (d) FRZ Evaporator

(e) FRZ Fan (f) Gas Valve (g) NTC Temperature Sensor (h) Condenser

FIGURE 3.3 – Some Refrigerator Components

directed into the related compartment and the air temperature of the compartment is reduced. There are two temperature sensors in both compartments in order to carry out the cooling process in a controlled manner. While one of these sensors is placed on the evaporator, the other is positioned in a suitable place to measure the air temperature of the compartment.

3.2 Importance of Energy Efficiency for Refrigerators

With the new regulations made by countries, refrigerator products are required to increase their efficiency and consume less energy. Various factors can impact a refrigerator's energy consumption. One of the most important factors is the usage habits of consumers. A study surveyed 200 domestic refrigerator users to understand the effect of consumer usage behaviours (Belman-Flores et al., 2019). The research revealed variations in general refrigerator characteristics and consumer behavior, including refrigerator age, frequency of door openings, food load, cleaning habits, and refrigerator position in the house. The survey found that FF door opened between 21 and 40 times daily by 40% of respondents. On the other hand, FRZ door opened less than five times by 46% of respondents. Consumer habits could inform the design of energy-efficient refrigerators. It was highlighted the impact of consumer behaviors, such as door opening frequency and thermal load, on domestic refrigerator energy consumption. It also emphasized that usage activities are unique to each user, playing a significant role in shaping air temperature patterns within refrigerators. Consequently, even with the same refrigerator model, actual energy consumption can vary based on consumer usage patterns.

Another study has been conducted to understand the effect of open and closed door situation for refrigerators (Hasanuzzam et al., 2008). The results indicate that various factors significantly impact energy consumption in refrigerators, with the average usage around 3.3 kWh per day. Key influences include the number of door openning times, ambient temperature, and the load within the compartments. Energy usage with the door open is approximately 40% higher than when the door remains closed. Additionally, the average maximum energy consumption is 27.3% greater than the overall average, and 55.6% greater than the average minimum.

4 SMART CONTROLLER DESIGN

4.1 Smart Controller for Embedded System

The objective was to develop an optimal refrigerator with precise temperature control capabilities while maintaining low cost hardware with lower power consumption. The process of model creation and application, as illustrated in Figure 4.1, is elaborated in detail in following sections. Results from real-world implementation of the algorithm on refrigerators are presented in section 5.

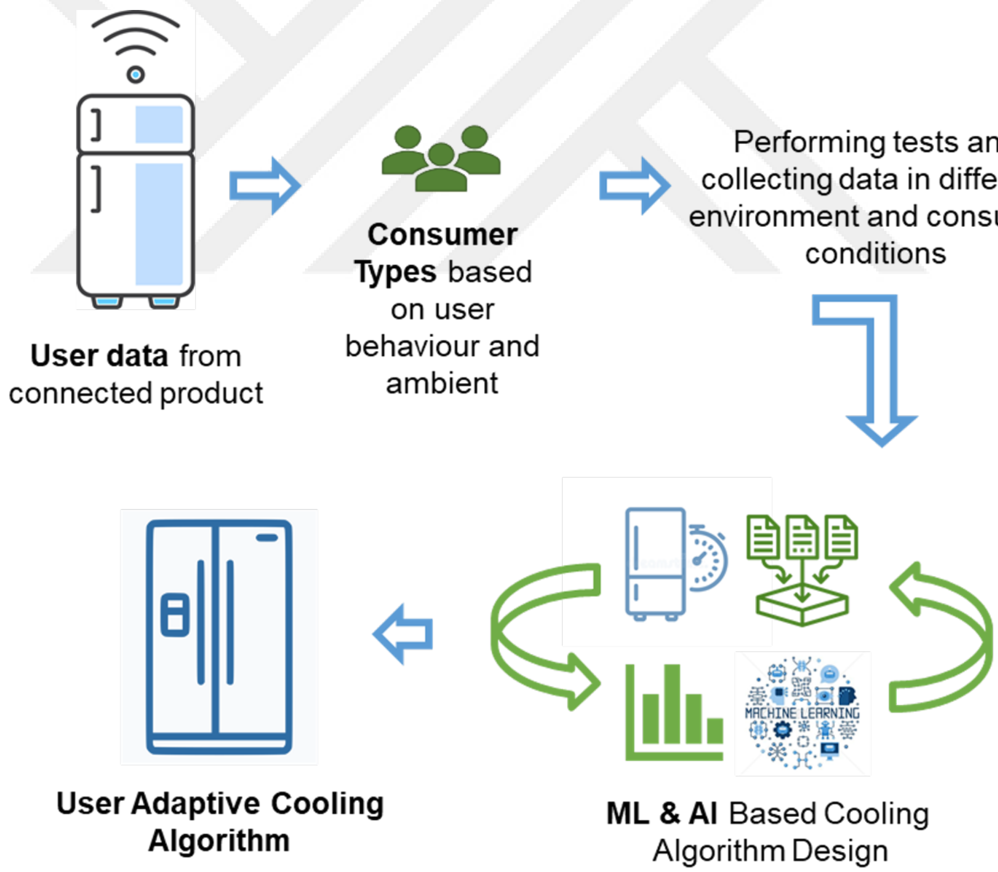

FIGURE 4.1 – User Adaptive Cooling Algorithm

There are two development and deployment processes. One is for embedded system and the other one is for cloud. New controller to be used in the embedded system has been prepared using PID controller based on RL. On the other hand, K-means clustering has been used to cluster the consumers according to their usage habits. The coefficient resulting from the clustering was sent to the control algorithm as a parameter in order to customize and increase efficiency of the control algorithm running in the embedded system according to the user.

4.1.1 PID Controller

The PID controller uses feedback to compare the signal from the output with the reference signal. The difference between these two signals generates an error. This error is then fed to the PID controller, which generates an output to the system input to keep the error at a minimum level. To understand how PID control works, its three key components, Proportional, Integral, and Derivative, should be examined.

The proportional component is the most basic element of the PID controller. It creates a control output that is proportional to the current control error, which is the difference between the desired target point and the actual process variable(Visioli, 2010).

$$P = u(t) = K_p e(t) = K_p(r(t) - y(t)), \tag{1}$$

where K_p is the proportional gain and the error at a given time is $e(t)$. Higher proportional gains result in stronger corrective actions, but they can also lead to oscillations if set too high. The proportional component determines how aggressively the controller responds to errors, thereby influencing the speed of the system's response.

The integral component addresses the issue of steady-state error, which is the residual error that remains even when the proportional term is applied. It does this by integrating the error over time, essentially accumulating the error and using it to adjust the control output(Visioli, 2010).

$$I = u(t) = K_i \int e(t)dt, \tag{2}$$

where K_i is the integral gain. By integrating the error, the integral term helps drive the system toward zero steady-state error. However, if the integral gain is too high, it can cause the system to overshoot and oscillate.

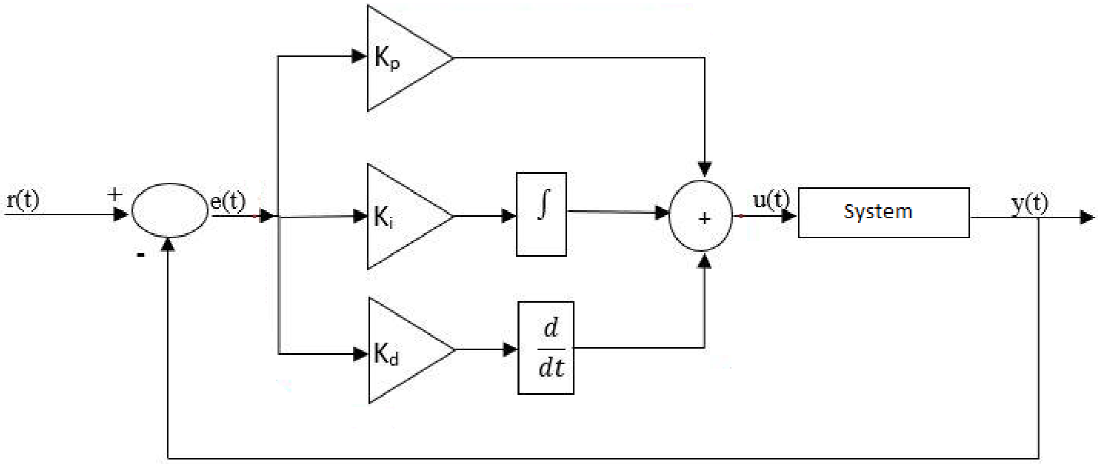

FIGURE 4.2 – Block Structure of Closed Loop Control System Including PID Controller

The derivative component anticipates future error trends by computing the rate of change of the error. It serves as a dampening factor, helping to smooth out rapid changes in the control output and reducing overshoot(Visioli, 2010).

$$D = u(t) = K_d \frac{d}{dt} e(t), \tag{3}$$

where K_d is the derivative gain. By incorporating the derivative term, the PID controller can react more quickly to changing conditions, improving system stability and reducing oscillations.

To achieve optimal control, the proportional, integral, and derivative gains must be carefully tuned. The objective is to find a balance that provides a fast and stable response while minimizing overshoot and oscillation. The block structure of a closed-loop control system containing a PID controller is shown in Figure 4.2. The common equation for PID is as follows :

$$u(t) = K_p e(t) + K_i \int e(t)dt + K_d \frac{d}{dt} e(t), \tag{4}$$

4.1.2 Reinforcement Learning

Reinforcement learning encompasses a range of concepts, including problem solving, solution methods, and the broader field of study that investigates these problems and solutions. At its core, RL is about learning what actions to take to maximize a numerical reward signal, with an emphasis on exploring and exploiting the environment's feedback. Unlike other learning paradigms, RL doesn't prescribe which actions to take, and instead requires agents to discover the best course of action through trial and error, leading to an inherently closed-loop process where the agent's actions affect future states and rewards.(Sutton and Barto, 2018).

In supervised learning, the system is trained on labeled examples, while unsupervised learning focuses on finding hidden structure in data. On the other hand, three key features distinguish RL from other types of learning. First, it involves continuous feedback from the environment, requiring the agent to adapt its actions based on outcomes. Second, RL doesn't provide explicit instructions on what actions to take, allowing the agent to explore different strategies. Third, the consequences of actions can play out over extended periods, requiring agents to consider not just immediate rewards but also their impact on future states.

A significant challenge in RL is balancing exploration and exploitation. To achieve optimal performance, an RL agent must experiment with new actions (exploration) while also leveraging known effective strategies (exploitation). This trade-off is unique to RL and does not arise in supervised or unsupervised learning. In addition, RL explicitly considers the full problem of a goal-oriented agent operating in an uncertain environment, unlike many approaches that focus on isolated subproblems without addressing their larger context.

Overall, reinforcement learning is a robust and dynamic field that addresses complex problems through adaptive learning, balancing exploration and exploitation, and integrating knowledge from various domains. It represents a paradigm shift in machine learning, emphasizing interactive learning from experience to achieve long-term goals. As a matter of fact, RL is defined by a specific type of problem, and all its solutions are classed as RL algorithms. Some of them :
Q-Learning : A model-free approach that uses a Q-table to estimate the expected re-

wards for taking certain actions in specific states. It updates the Q-values based on a learning rate and a discount factor, with the goal of finding the optimal policy(Watkins and Dayan, 1992).
SARSA (State-Action-Reward-State-Action) : Similar to Q-Learning but considers the next action in the update rule. This approach incorporates the on-policy learning mechanism, focusing on the current policy's performance rather than exploring alternative actions(Rummery and Niranjan, 1994).
Deep Q-Learning : An extension of Q-Learning that uses neural networks to approximate the Q-values. This approach is useful for handling large state spaces and complex problems, leveraging the power of deep learning(Mnih et al., 2013).
Policy Gradient Methods : These methods directly optimize the policy (a mapping from states to actions) by computing the gradient of the expected reward(Sutton et al., 2000). Common examples include REINFORCE, which uses a Monte Carlo approach to update the policy, and Advantage Actor-Critic (A2C), which combines value-based and policy-based methods.
Actor-Critic Methods : These methods maintain separate networks for the policy (the actor) and the value function (the critic)(Konda and Tsitsiklis, 2001). The critic evaluates the actions chosen by the actor and provides feedback to improve the policy. Proximal Policy Optimization (PPO) and Asynchronous Advantage Actor-Critic (A3C) are popular variants.
Temporal Difference (TD) Learning : A blend of Monte Carlo methods and dynamic programming, TD learning updates the value function based on the difference between estimated values from consecutive states(Sutton, 1988). TD extends this approach with eligibility traces, allowing for a balance between one-step and multi-step updates.
Monte Carlo Methods : These methods rely on complete episodes to calculate returns and update policies or value functions(Sutton and Barto, 2018). They are particularly useful for problems where the full trajectory is needed to understand the structure.
Hierarchical Reinforcement Learning : This approach involves breaking down complex tasks into simpler sub-tasks, each with its own policy and value function(Dietterich, 2000). Methods like the Options Framework and Feudal Reinforcement Learning fall into this category.

Actor-Critic learning allows for the consolidation of input and hidden layers within the RBF network, resulting in reduced storage space demands and eliminating repetitive computations for hidden unit outputs, consequently improving learning efficacy. Actor-Critic Method has been chosen with RBF network for this work due to cost constraints in the embedded system to create adaptive PID controller that can dynamically adjust to changing system conditions and improve control performance. The flexibility and adaptability provided by reinforcement learning make it an ideal approach for designing controllers that can learn from experience and optimize their behavior over time.

4.1.3 Smart Controller : Adaptive PID Controller Based on RL

By incorporating with RL, it is devised an approach to enhance the conventional PID controller, resulting with user-adaptive cooling algorithm. This algorithm was specifically tailored for preserving food and any other things that need to be stored. Leveraging user habits, the algorithm not only improves cooling performance but also reduces energy consumption. Additionally, energy-efficient cooling algorithm successfully minimizes temperature fluctuations, ensuring optimal storage conditions.

The research integrates creating consumer profiles based on ambient temperature data, door openning and door closing obtained from IoT refrigerators deployed in the field. User models dynamically adjust the refrigerator behavior by modifying parameters within the RL algorithm. RL algorithm learning rate changes according to user behaviour parameter from the cloud. When refrigerator works in extreme conditions like high ambient temperatures and frequent door openings, learning rate increases for quick adaptation for new conditions. This provides better user experience in rapid temperature and load changes. When refrigerator works in normal conditions like low ambient temperatures and less door openings, learning rate remains or decrease for preventing unnecessary changes in compressor and fan speeds. This provides more stable compressor and fan speeds. Thus energy consumption is optimized in stable temperature and load conditions. To implement adaptive cooling algorithm, it is employed an adaptive PID controller structure alongside a learning algorithm based on user habits, utilizing the Actor-Critic learning method (song WANG et al., 2007). This method employs a RBF neural network to approximate both the Actor's policy function and the Critic's value function. The RBF network takes inputs such as system error and its derivatives, enabling the Actor to map system states to PID parameters, while the Critic evaluates Actor outputs and generates the Temporal Difference (TD) error used for network updates.

$$u(t) = u(t-1) + \Delta u(t) = u(t-1) + K(t)x(t) =$$
$$u(t-1) + k_I(t)x_1(t) + k_P(t)x_2(t) + k_D(t)x_3(t) = \tag{5}$$
$$u(t-1) + k_I(t)e(t) + k_P(t)\Delta e(t) + k_D(t)\Delta^2 e(t),$$

The architecture of an adaptive PID controller utilizing Actor-Critic learning is inspired by the concept of the incremental PID controller outlined in Equation (5).

$x(t) = [x_1(t), x_2(t), x_3(t)] = [e(t), \Delta e(t), \Delta^2 e(t)]^T$; $e(t) = y_d(t) - y(t)$, $\Delta e(t) = e(t) - e(t-1)$ and $\Delta^2 e(t) = e(t) - 2e(t-1) + e(t-2)$ represents the system output error, the first-order difference of error and the second-order difference of error respectively ; $K(t) = [k_I(t), k_P(t), k_D(t)]$ is a vector of PID parameters (song WANG et al., 2007).

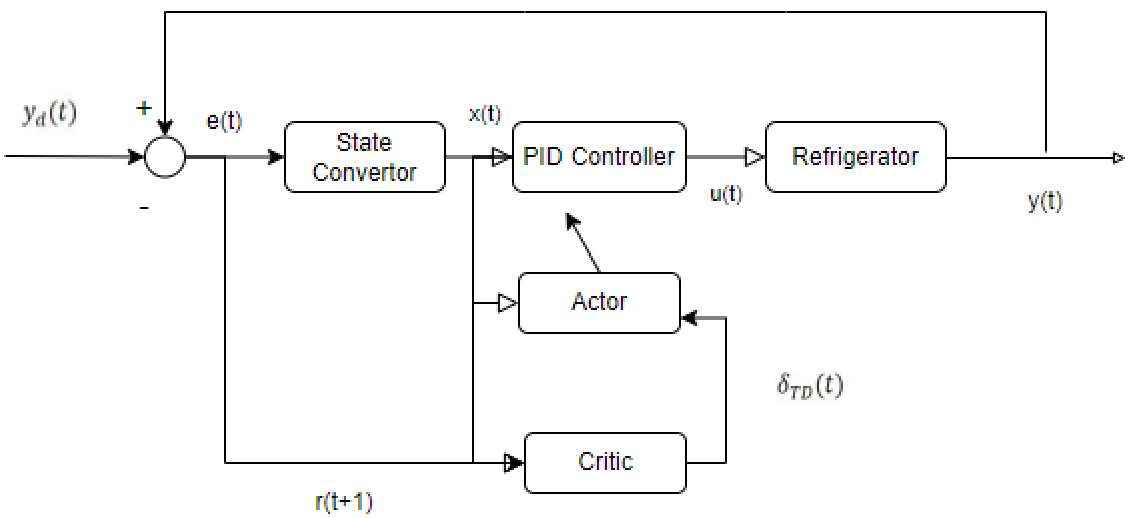

FIGURE 4.3 – Self-learning PID Control Diagram Based on RL

The symbols $y(t)$ and $y_d(t)$ denote the target and current outputs for the system, respectively. The error $e(t)$ is converted into a system state vector $x(t)$ through a state converter in order to enable Actor-Critic learning as seen in Figure 4.3 (Guan and Yamamoto, 2020). An Actor-Critic learning architecture comprises three key elements : an Actor, a Critic, and a stochastic action modifier (SAM).The Actor estimates a policy function, mapping the current system state vector to suggested PID parameters $K'(t) = [k'_I t(t), k'_P(t), k'_D(t)]$, which are not directly participated in PID controller design (Sedighizadeh and Rezazadeh, 2008). The SAM is used to generate stochastically the actual PID parameters $K(t)$ according to the recommended PID parameters $K'(t)$ suggested by the Actor and the estimated signal $V(t)$ from the Critic. The Critic receives a system state vector and an external reinforcement signal (immediate reward) $r(t)$ from the environment and produces TD error (i.e., internal reinforcement signal) $\delta_{TD}(t)$ and an estimated value function $V(t)$. $\delta_{TD}(t)$ is provided for the Actor and the Critic directly and is viewed as an important basis for updating parameters of the Actor and the Critic. $V(t)$ is send to the SAM and is used to modify the output of the Actor.

Simultaneously considering the impact of system error and error change rate on control performance is imperative when designing the external reinforcement signal $r(t)$. Therefore, $r(t)$ is defined as

$$r(t) = \alpha r_e(t) + \beta r_{ec}(t) \tag{6}$$

where α and β are weighted coefficients,

$$r_e(t) = \begin{cases} 0 & |e(t)| \leq \epsilon \\ -0.5 & \text{otherwise} \end{cases}, \quad r_{ec}(t) = \begin{cases} 0 & |e(t)| \leq |e(t-1)| \\ -0.5 & \text{otherwise} \end{cases}$$

and ϵ is a tolerant error band.

RBF network is utilized for parameter identification via function mappings. Its simplicity in structure, convergence of parameters, and efficient learning process are recognized as advantages and are further detailed in (V.T. and Shin, 1994). Due to hardware memory limitations of the low-cost embedded system operating the refrigerator in this project, only three kernel functions are defined, as seen in Figure 4.4.

Alternatively, both the Actor and Critic receive the same state vector as input, derived from the environment, with a slight distinction in their outputs. Consequently, there's only one RBF network, depicted, which concurrently implements both the Actor's policy function learning and the Critic's value function learning. This setup allows the Actor and Critic to share the input and hidden layers of the RBF network, reducing the

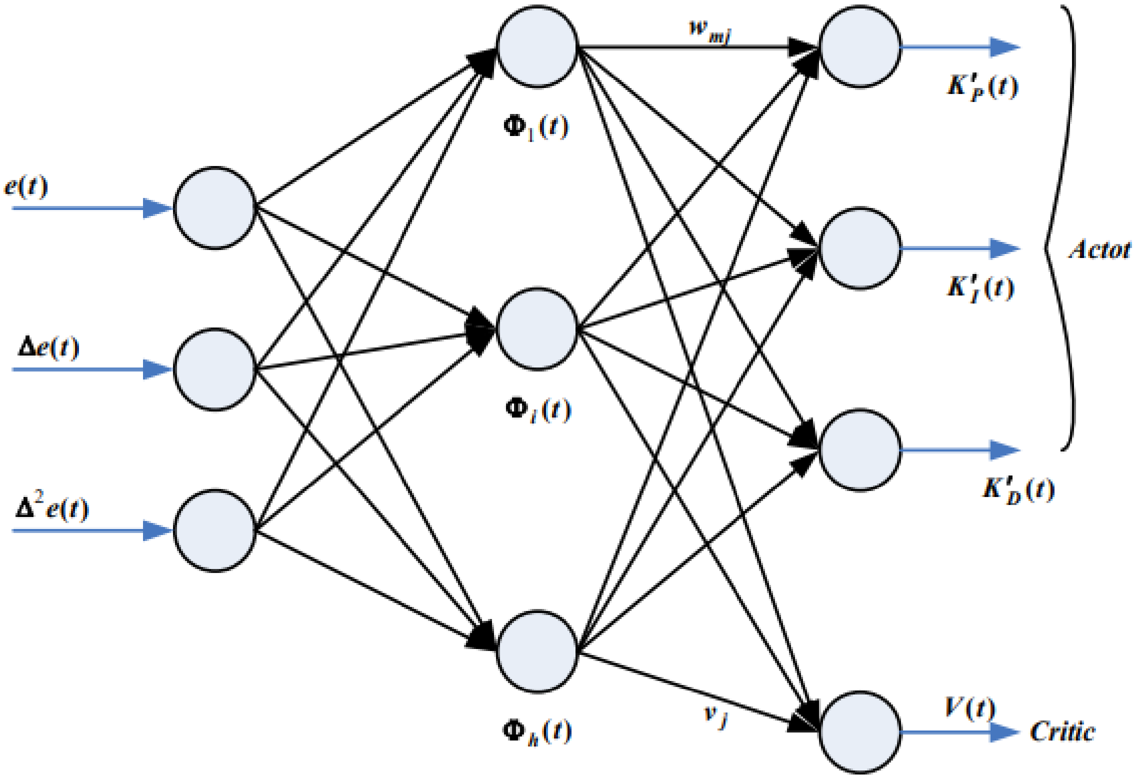

FIGURE 4.4 – Actor-Critic learning built from RBF network (Sedighizadeh and Rezazadeh, 2008)

storage space requirements and avoiding redundant computations for the outputs of the hidden units, thereby enhancing learning efficiency. The specific role of each layer is explained as follows :

Layer 1 : input layer. Each unit in this layer represents a system state variable x_i with i indicating the index of the input variable. The input vector $x(t) \in R^3$ is transmitted directly to the following layer.

Layer 2 : hidden layer. The kernel function for the hidden units in the RBF network is a Gaussian function. The output of the j th hidden unit is

$$\Phi_j(t) = exp(-\frac{(\| x(t) - \mu_j(t) \|)^2}{2\sigma_j^2(t)}), j = 1, 2, ..., h \tag{7}$$

with $\mu_j = [\mu_{1j}, \mu_{2j}, \mu_{3j}]^T$ and σ_j are the center vector and the width scalar for the j th hidden unit. h represents the total number of hidden units.

Layer 3 : output layer. This layer consists of an Actor part and a Critic part. The m th output of the Actor part, $K'_m(t)$ and the value function of the Critic part, $V(t)$ are calculated as

$$K'_m(t) = \sum_{j=1}^{h} w_{mj}(t)\Phi_j(t), m = 1, 2, 3 \tag{8}$$

$$V(t) = \sum_{j=1}^{h} v_j(t)\Phi_j(t) \tag{9}$$

where w_{mj} represents the weight between the j th hidden unit and the m th Actor unit, and v_j represents the weight between the j th hidden unit and the single Critic unit. To address the challenge of balancing 'exploration' and 'exploitation', the Actor's output doesn't directly influence the PID controller. Instead, a Gaussian noise term η_k is introduced to adjust the suggested PID parameters $K'(t)$ from the Actor, thereby modifying the actual PID parameters $K(t)$ as per Equation 10. The amplitude of the Gaussian noise is contingent upon $V(t)$; when $V(t)$ is significant, η_k is minimal, and vice versa.

$$K(t) = K'(T) + \eta_k(0, \sigma_V(t)) \text{ where } \sigma_V(t) = \frac{1}{1 + exp(2V(t))} \tag{10}$$

Actor-Critic learning is distinguished by the Actor learning the policy function and the Critic learning the value function concurrently through the TD method. The TD error $\delta_{TD}(t)$ is computed by the temporal difference in the value function between consecutive states during state transition.

$$\delta_{TD}(t) = r(t) + \gamma V(t+1) - V(t) \tag{11}$$

The external reinforcement reward signal is represented by $r(t)$, where $0 < \gamma < 1$ represents the discount factor determining the weight of future rewards. The TD error reflects the quality of the current action, thus defining the performance index function for system learning.

$$E(t) = \frac{1}{2}\delta_{TD}^2(t) \tag{12}$$

The Actor and Critic weights are adjusted based on the TD error performance index using the gradient descent method and chain rule, as outlined in the following equations.

$$w_{mj}(t+1) = w_{mj}(t) + \alpha_A \delta_{TD}(t) \frac{K_m(t) - K'_m(t)}{\sigma_V(t)} \Phi_j(t) \tag{13}$$

$$v_j(t+1) = v_j(t) + \alpha_C \delta_{TD}(t) \Phi_j(t) \tag{14}$$

where α_A represent the learning rate for Actor and α_C represents the learning rate for Critic.

Since the Actor and the Critic utilize the same input and hidden layers in the RBF network, updates to the centers and widths of hidden units occur just once.

$$\mu_{ij}(t+1) = \mu_{ij}(t) \eta_\mu \delta_{TD}(t) v_j(t) \Phi_j(t) \frac{x_i(t) - \mu_{ij}(t)}{\sigma_j^2(t)} \tag{15}$$

$$\sigma_j(t+1) = \sigma_j(t) \eta_\sigma \delta_{TD}(t) v_j(t) \Phi_j(t) \frac{\| x(t) - \mu_j(t) \|}{\sigma_j^3(t)} \tag{16}$$

where μ_μ and μ_σ are learning rates of center and width respectively.

The described PID controller with RL can be expressed by the following steps.

— Step 1. Initializing parameters of Actor-Critic learning controller, including $w_{mj}(0)$, $v_j(0)$, $\mu_{ij}(0)$, $\sigma_j(0)$, η_μ, η_σ, α_C, α_A, γ, ϵ, α, β.

— Step 2. Detecting the actual system output $y(t)$, calculating the system error $e(t)$, constituting system state variables $e(t)$, $\Delta e(t)$, $\Delta^2 e(t)$.

— Step 3. Receiving an immediate reward $r(t)$ from Equation (6).

— Step 4. Calculating the Actor output $K'(t)$ and the Critic value function $V(t)$ from Equation (8) and Equation (9) at time t respectively.

— Step 5. Calculating the actual PID parameters $K(t)$ from Equation (10), and consequently calculating the control output of PID controller $u(t)$ from Equation (5).

— Step 6. Applying $u(t)$ to the controlled plant and observing the system output $y(t+1)$ and the immediate reward $r(t+1)$ at the next sampling time.

— Step 7. Calculating the Actor output $K'(t+1)$ and the Critic value function $V(t+1)$ from Equation (8) and Equation (9) at time $(t+1)$ respectively.

— Step 8. Calculating the TD error $\delta_{TD}(t)$ from Equation (11).

— Step 9. Updating the weights of the Actor and the Critic from Equation (13) and Equation (14) respectively.

— Step 10. Updating the centers and the widths of RBF kernel functions according to Equation (15) and Equation (16) respectively.

— Step 11. Judging whether the control process is finished or not. If not, then $t \leftarrow t+1$ and turn to Step 2.

4.2 Machine Learning Approach for Cloud : K-means Clustering

Clustering, essentially, is a sub-branch of unsupervised machine learning technique. Unsupervised ML involves analyzing datasets composed of input data without predefined labels. This method is typically used to discover significant structures, underlying explanatory mechanisms, generative properties, and clusters within a collection of examples (Surya Priy, 2024). Clustering involves segmenting a population or dataset into distinct groups, where data points within the same group are more similar to each other than to those in other groups. Clustering is crucial because it determines how unlabeled data is organized into meaningful groups. There are no universally accepted criteria for what constitutes good clustering—it depends on the user's goals and the criteria they choose to evaluate the groupings. Factors influencing clustering choice can include the nature of the data, the desired outcomes, the complexity of the clustering algorithm, and the user's interpretation of the data's structure.

The K-means clustering algorithm was used as an unsupervised machine learning algorithm to separate the elements into groups. The 'K' in the algorithm name represents the number of groups/clusters into which is required to classify requested items. The algorithm will categorize items into k groups or similarity clusters. Euclidean distance is used as the metric to calculate this similarity. The basic working logic of the algorithm is as follows. First, k points called means or cluster centers are randomly initialized. Each item is categorized based on the nearest mean and is thus far in that cluster. The coordinates of the mean, which is the average of the categorized items, are updated. The process is repeated for a certain number of iterations and eventually clusters are obtained. Customer segmentation was performed using data obtained from connected refrigerators, leading to the formation of user profiles. The collection and refinement of data are the most important part for creating consumer clusters. Detailed explanation of the data collection, model training and fine tuning processes are provided in following sections.

4.2.1 Data Collection

In order to gather data, customers need to ensure that their devices are actively connected to the internet. Refrigerators that are used for this project are equipped with an additional controller for data collection purposes. The Homewhiz mobile application is essential for establishing connectivity between the controller and the cloud (Arçelik A.Ş., 2020). This application facilitates Wi-Fi connection, enabling the device to link up with cloud systems. Through the Homewhiz app, users can access various details regarding their refrigerators, including cooling status, internal temperature, and other metrics. The Homewhiz platform stores this data on its servers, offering options for analysis, long-term storage, or reviewing device performance for the past upon user request. Users can also manage their devices remotely using these applications, such as adjusting refrigerator temperature or controlling content viewing. Data collection via Homewhiz adheres to user consent guidelines and implements appropriate security measures to safeguard personal privacy. Users are provided with transparency regarding the collection and utilization of data about to their household appliances, empowering them to make informed decisions regarding data sharing.

Data on door opening activities, door opening durations, ambient temperature, and humidity levels for refrigerator are sanitized to remove personal information before being stored and processed in this thesis. The data was gathered over a span of 2 months from 1680 connected refrigerators, with a data transmission rate of 1 message per hour. Each message comprises 460 bytes of data, including minute temperature sensor readings for the corresponding hour. The temperature sensors boast a precision and accuracy rate of 97%. All data is synchronized via WiFi with the IoT module installed in the refrigerator (Peker et al., 2023).

4.2.2 Model Training and Fine Tuning

It is utilized the K-means clustering algorithm to partition the gathered data in the modeling phase. K-Means clustering is an unsupervised machine learning approach that segments a dataset into a predetermined number of clusters (Sinaga and Yang, 2020). Typically, unsupervised algorithms analyze datasets only based on inputs without any prior knowledge or labeled outcomes. The initial step of the K-Means Algorithm involves determining the number of clusters to create (K value), which depends on the application and data. Next, initial cluster centers are randomly chosen, and data points are assigned to the nearest cluster center to initiate clustering. Subsequently, the cluster centers are updated by calculating the average position of data points within each cluster. By iteratively repeating these steps, the assignment of data points to clusters and the updating of cluster centers are continually refined. The algorithm terminates either when the specified number of iterations is reached or when the change in cluster centers falls below a predetermined threshold value (Peker et al., 2023).

K-Means clustering is sensitive to the initial selection of centroids and may converge to local optima depending on the initial configuration. To mitigate this issue, the algorithm is often run multiple times with different initializations, and the solution with the lowest total within-cluster variation (also known as inertia) is selected. K-Means clustering is a versatile and widely used algorithm for partitioning datasets into clusters, making it valuable in various applications such as image segmentation, customer segmentation, and anomaly detection.

From the collected data, it is observed that the number of door openings and door opening times fall into three distinct clusters. Upon examining these clusters, it is identified the following characteristics :

— Cluster 1 : Short duration with infrequent door openings

— Cluster 2 : Short duration with frequent door openings

— Cluster 3 : Long duration with infrequent door openings

Visualized version of the data for the clustering is illustrated in Figure 4.5a.

Additionally, a separate clustering analysis was conducted based on ambient humidity and temperatures. Clustering the sensor values for temperature and humidity yielded four distinct clusters :

— Cluster 1 : low temperature and medium humidity

— Cluster 2 : low temperature and low humidity

— Cluster 3 : high temperature and low humidity

— Cluster 4 : high temperature and medium humidity

Visualized version of the data for the clustering is illustrated in Figure 4.5b. Combining the clusters derived from ambient conditions and door open metrics resulted in 12 unique clusters. Table 4.1 and Table 4.2 show the centroids of clusters separately for both door actuation and ambient values respectively.

TABLE 4.1 – Door Actuation Based Clustering

Cluster	**Daily Door Opening Count**	**Door Open Time per Opening**
1 (Yellow)	31	13
2 (Purple)	26	20
3 (Green)	65	14

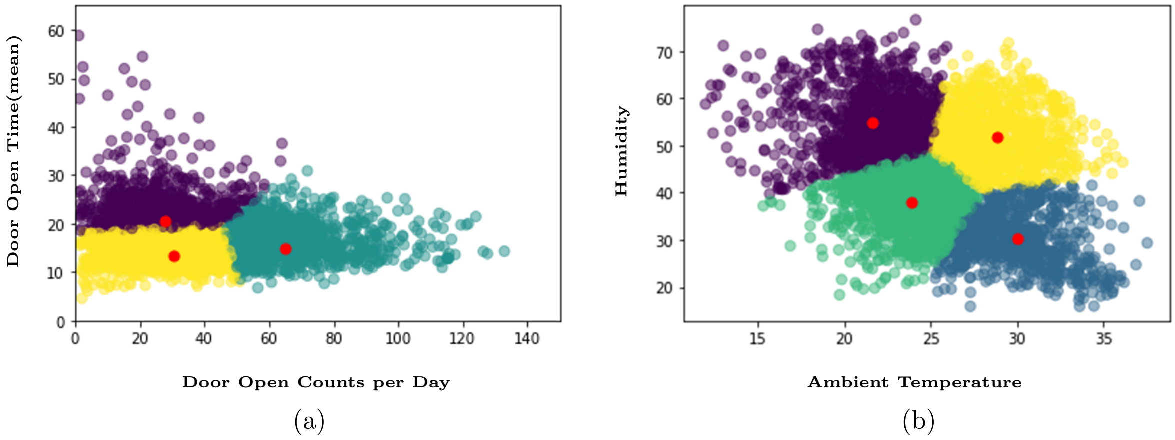

FIGURE 4.5 – (a) door open counts and time values (b) ambient and humidity values

TABLE 4.2 – Humidity and Ambient Temperature Based Clustering

Cluster	Ambient Temperature	Humidity
1 (Blue)	30	31
2 (Purple)	21	54
3 (Yellow)	29	51
4 (Green)	23	37

4.3 Deployment

The deployment process comprises two main stages : Cloud Deployment and Control Unit Deployment. Deployment for cloud focuses creating specific data clusters. Deployment for embedded system focuses application and adaptation of the new algorithm to the cooling system.

4.3.1 Cloud Deployment

After analyzing the data and generating specific clusters, deploying these clusters is a crucial step. This involves ensuring that the application operates on cloud systems. The consumer type, determined based on clusters, is then forwarded to the embedded system as a parameter. This parameter is utilized to optimize the learning rate. It is used as factor multiplier for new algorithm. The learning rate was set as a parameter to be between 1 and 2 as the number of clusters which is 12, and it is sent to the control unit via the cloud using the IoT infrastructure provided by the Homewhiz application. Consumer profiles are computed by iteratively daily processing of weekly data. Each day, the data from the earliest day of the week is removed, and consumer types for the related week are determined. These iteration operations are specified in Figure 4.6.

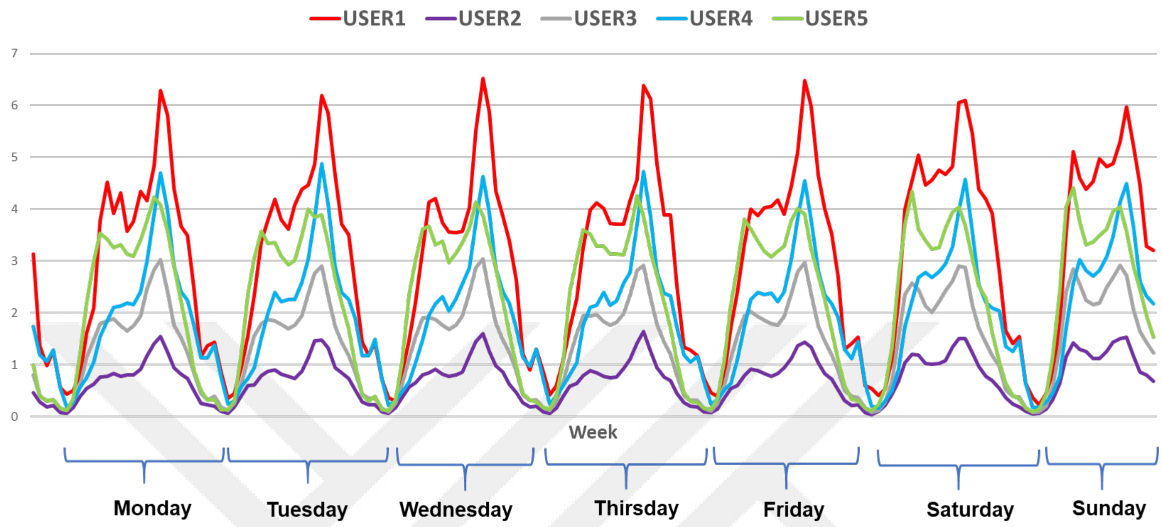

FIGURE 4.6 – Weekly Consumer Usage Activities for Refrigerator

4.3.2 Electronic Control Unit Deployment

In the project, an electronic control unit with NXP MKE04Z128xxx4 MCU was used and the embedded software was prepared accordingly using the C programming language in the IAR development environment.

On the control unit, the RL algorithm dynamically adjusts the PID coefficients. At the same time, the learning rate parameter which is obtained from cloud system through customer cluster analysis is sent to embedded system for the control algorithm as a factor. Shortly and simply, it is better to say that PID is used for control. AI is used to calculate the Proportional (P), Integral (I), and Derivative (D) coefficients. The adaptive PID control algorithm integrates information from both the FF and FRZ evaporator temperatures to determine the compressor's operational status and speed. Meanwhile, the FF and FRZ air temperatures are used for regulating the respective cabinet's fan as seen in Figure 4.7. When determining fan speeds, the cooling demand is calculated considering the desired set-point value and the measured ambient air temperature. If this calculated value exceeds a threshold, the PID coefficients are updated;

otherwise, the PID cycle continues with the current coefficients. The fan speed derived from the PID output is then applied to the refrigerator (Peker et al., 2023).

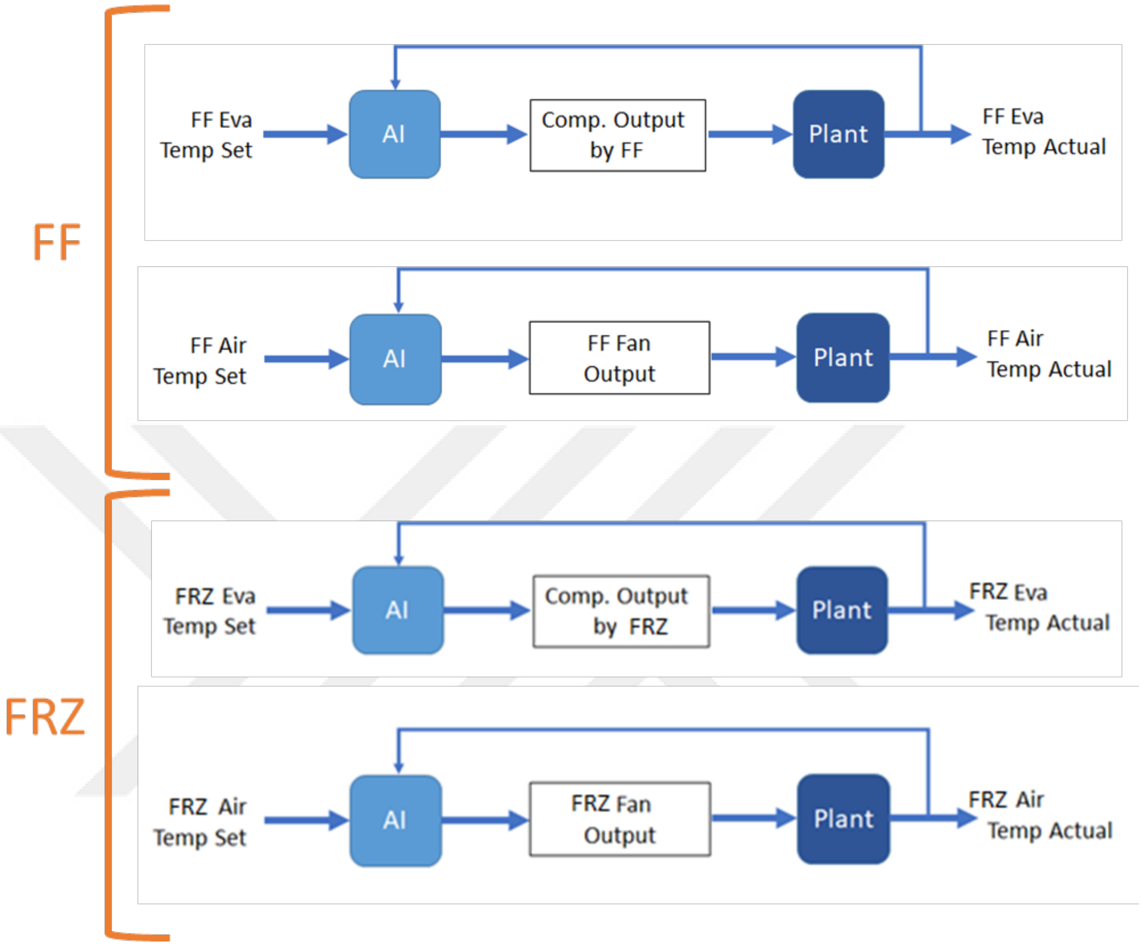

FIGURE 4.7 – FF and FRZ Cabinets Set Value Optimization

The evaporation set temperature required for compressor frequency calculation is determined based on the fan speed. Fan speeds are reduced when they exceed 99% and increased when they fall below 90%. This maintains the fan speed within the range. With high value fan working range, desired temperatures can be reached by low compressor speed. As a result, total energy consumption decreases. Speed of the fans stays within the range of 90%-99%. The error value is calculated by comparing the obtained evaporation set temperature with the measured evaporation set temperature. If this value exceeds a threshold, the PID coefficients are recalculated ; otherwise, the compressor speed is calculated using the existing PID coefficients (Peker et al., 2023). Details on how compressor and fan speeds are calculated can be found in Figure 4.9.

The valve algorithm is also a dynamically changed decision algorithm as seen in Figure

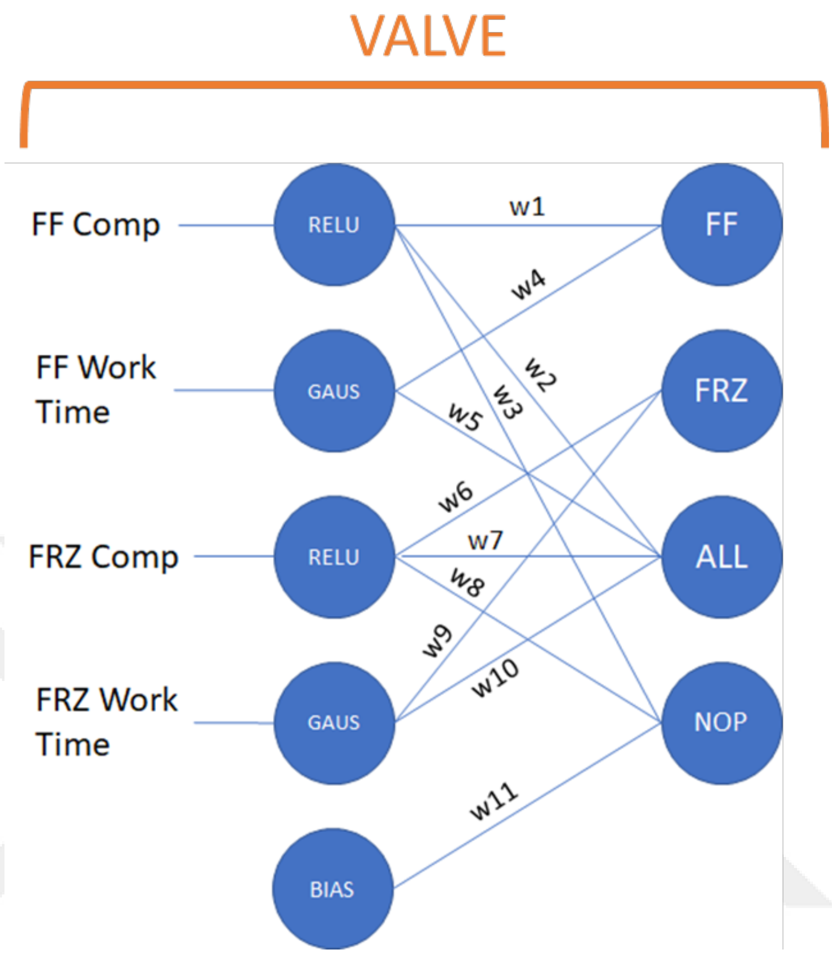

FIGURE 4.8 – Valve Operation Work Flow

4.8. The cabin, which has a compressor, loses its power over time by using the gaussian distribution feature. So the other cabin can have a chance for compressor work. The valve decision algorithm based on compressor speeds and working times. Compressor speed effects with Rectified Linear Unit (ReLU) function and compressor working time effects with a gaussian function to valve position decision score. If a cabin works in low speeds in a long time, score decreases. Meantime other cabin's score increases and takes the priority of compressor work. The priority can be given a specific cabin by adjusting weights in the algorithm. With higher weights, in selected cabin (FF or FRZ) more precise temperatures can be provided. Because the compressor has much more chance to optimize speed in a prior cabin. This decision algorithm also efficient to prevent high temperature oscillations in one cabin while valve position changes.

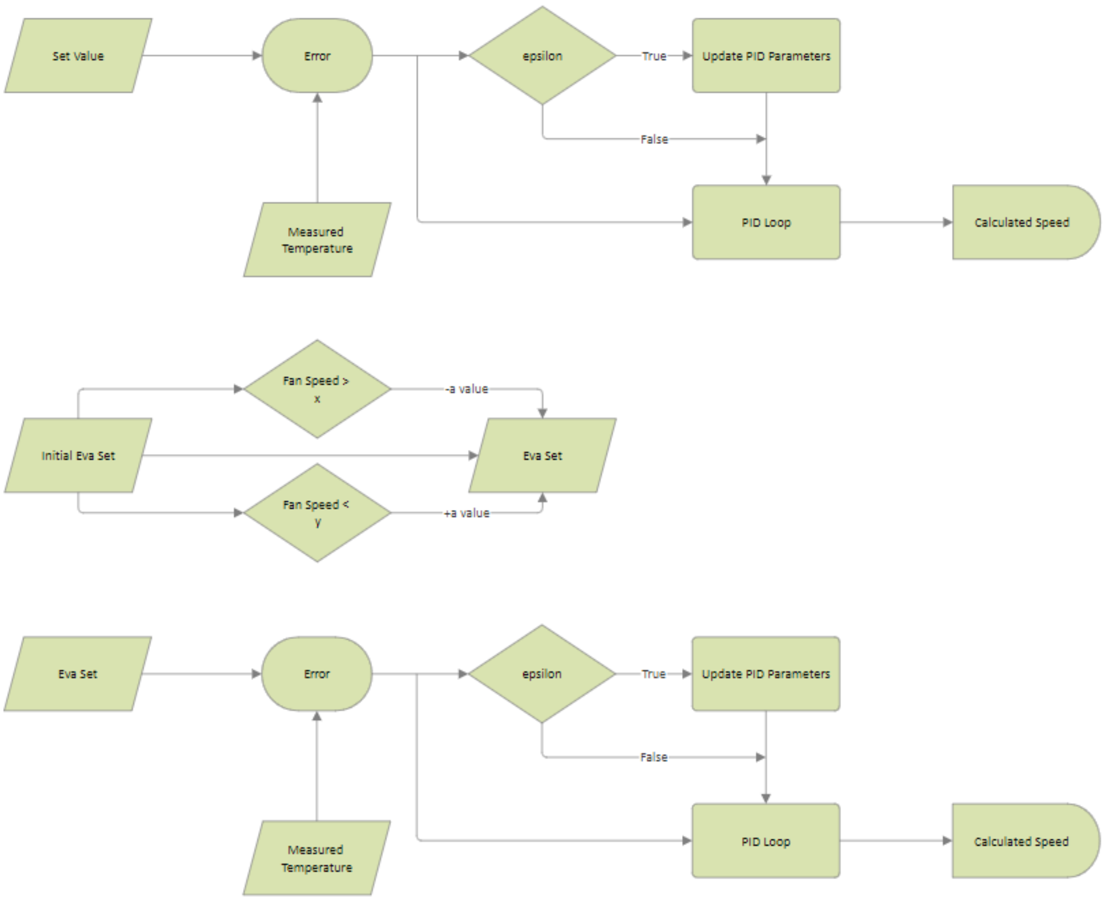

FIGURE 4.9 – Flow Chart of Adaptive PID Control Algorithm

5 TEST RESULTS

Once the machine learning and embedded software were developed, tests were conducted on both the traditional cooling algorithm and the new cooling algorithm which utilizes PID control with reinforcement learning and supported by a cloud-based clustering algorithm. The tests were performed using two different methodologies and the same test procedures and conditions were applied to both algorithms. The first one was carried out to measure energy consumption according to regulations. The second one was carried out according to the consumer usage conditions to simulate their usage habits. Energy consumption tests which were applied to the algorithms are referenced according to IEC standards (International Electrotechnical Commission, 2015a), (International Electrotechnical Commission, 2015b) and (International Electrotechnical Commission, 2015c). These standards comprises three international standards that outline the testing and evaluation procedures for refrigerators and freezers, providing consistency across the industry. (International Electrotechnical Commission, 2015a) covers the general definitions, terms, and requirements for refrigerators and freezers. It establishes the basic concepts and classification criteria for these appliances, including the terminology and technical specifications needed to maintain a common understanding across the industry. (International Electrotechnical Commission, 2015b) focuses on performance testing and measurement methods. This standard describes the procedures for evaluating energy consumption, storage temperatures, cooling capacities, and other key performance metrics for refrigerators and freezers. It aims to ensure consistency in testing methods and to provide a basis for energy efficiency and overall performance assessment. (International Electrotechnical Commission, 2015c) deals with the more specific aspects of performance evaluation, such as air temperatures, storage temperatures, and freezing capabilities. It describes the test methods used to assess the stability of internal temperatures and the performance of refrigerators and freezers under specified conditions. The tests which were applied to simulate usage conditions for consumers are referenced to Arçelik quality standards which are a proprietary standard protected under IP. It can simply be called a test based on opening and closing doors at certain periods. These tests are named "Energy Test" and "User

Test" respectively in the following figures and tables.

All tests were repeated on the same product to obtain stable and product-independent results for this study. Temperature measurements were taken from different places from FF and FRZ. In order to make the temperature measurements in the following figures more understandable, Figure 5.1 and Figure 5.2 represent where the temperature measurements were taken for FF and FRZ in the refrigerator respectively. All temperature measurements were taken with the help of thermocouples.

The results obtained from the performed tests can be seen in the following figures. Different colors in the figures represent measurements taken from different points of the compartments. These points are the same points for both traditional and new control algorithms, and remain the same during the tests. Blue colour represents crisper, orange colour represents bottom, green colour represents top and red colour represents crisper cover of FF in the refrigerator.

Figure 5.4 illustrates the outcomes obtained from the traditional cooling algorithm applied to FF at 4°C for energy test. 4°C is the set temperature for FF which is arranged from the display of the refrigerator. In Figure 5.6, the outcomes of the traditional cooling algorithm applied to the FRZ at -18°C for energy test are depicted. -18°C is the set temperature for FRZ which is arranged from the display of the refrigerator. It is evident according to the results that the FF exhibits oscillation of approximately 1,67°C, while the FRZ shows oscillation of approximately 3,35°C. The tests applied to the product were conducted as mentioned according to the IEC energy consumption measurement standards.

On the contrary, Figure 5.3 illustrates the results obtained from the new cooling algorithm for the FF, while Figure 5.5 presents the results for the FRZ. The new cooling algorithm successfully achieves the reduction in temperature oscillation, with the FF exhibiting an average oscillation of approximately 0,48°C and the FRZ showing oscillation of approximately 0,81°C.

Energy tests were repeated separately at 3°C for FF and -22°C for FRZ to verify the results gathered from test which has been performed at 4°C/-18°C. Figure 5.7 and 5.8 illustrate the outcomes for the FF while Figure 5.9 and 5.10 illustrate the outcomes for the FRZ. The new cooling algorithm successfully achieves the reduction

Figure 5.1 – Temperature Measurements Points Taken from FF

Figure 5.2 – Temperature Measurements Points Taken from FRZ

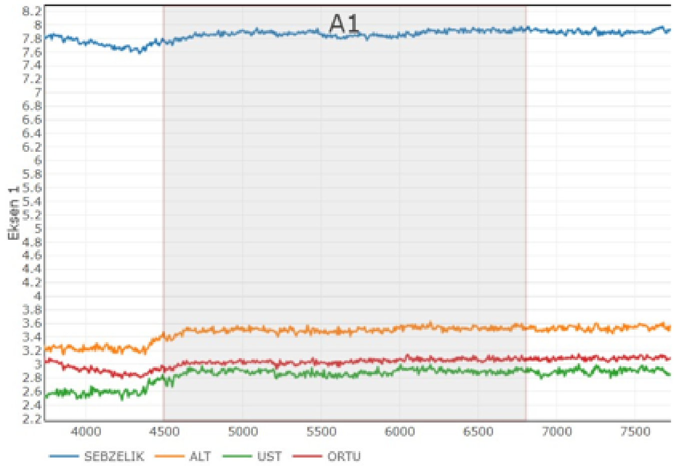

FIGURE 5.3 – New Cooling Software Results for FF at 4°C for Energy Test

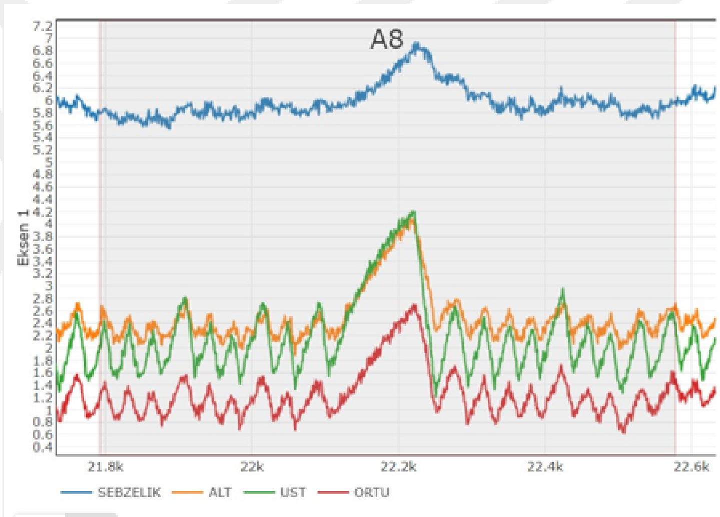

FIGURE 5.4 – Traditional Cooling Software Results for FF at 4°C for Energy Test

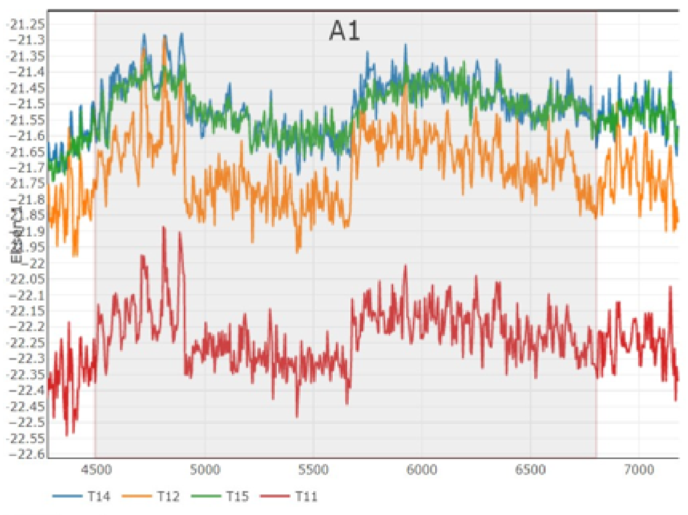

FIGURE 5.5 – New Cooling Software Results for FRZ at -18°C for Energy Test

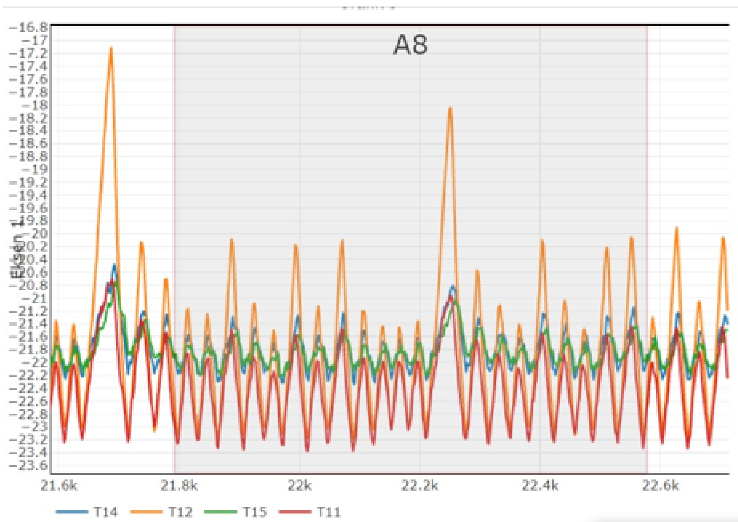

FIGURE 5.6 – Traditional Cooling Software Results for FRZ at -18°C for Energy Test

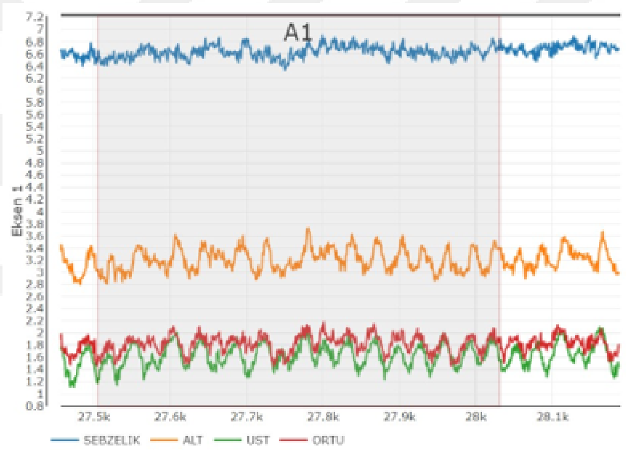

FIGURE 5.7 – New Cooling Software Results for FF at 3°C for Energy Test

in temperature oscillation, with the FF compartment exhibiting an average oscillation of approximately 0,85°C and the FRZ compartment showing an average oscillation of approximately 1,93°C. As it can be seen from the figures, the new cooling algorithm optimized temperature oscillations better than the traditional cooling algorithm at 3°C/-22°C set points as well as at 4°C/-18°C.

Consumer usage tests has been performed at 4°C/-18°C. Figure 5.11 and 5.12 show the result for the FF while Figure 5.13 and 5.14 show the result for the FRZ. The new cooling algorithm successfully achieves the reduction in temperature oscillation for the FF compartment exhibiting an average oscillation of approximately 4,18°C while the FRZ compartment showing an average oscillation of approximately 13,64°C which is about the same with current software.

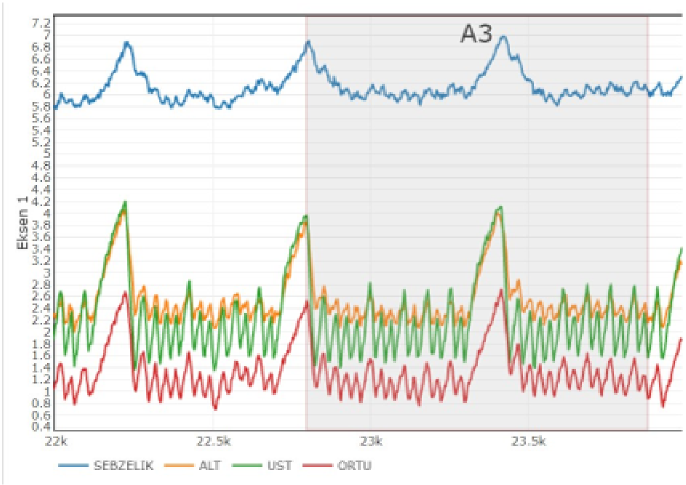

FIGURE 5.8 – Traditional Cooling Software Results for FF at 3°C for Energy Test

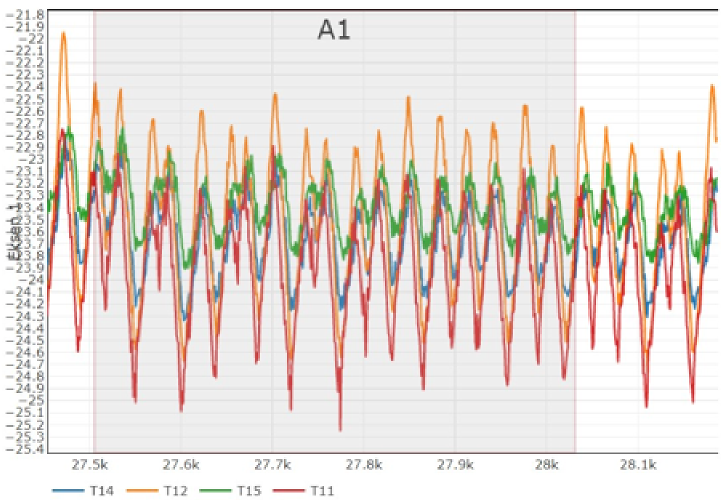

FIGURE 5.9 – New Cooling Software Results for FRZ at -22°C for Energy Test

The same test procedure has been performed at 3°C/-22°C to verify the results gathered from test which has been performed at 4°C/-18°C. Figure 5.15 and 5.16 shows the result for the FF while FRZ test results can be examined in Figure 5.17 and 5.18 which is new cooling software and traditional cooling software respectively. The new cooling algorithm successfully achieves the reduction in temperature oscillation, with the FF compartment exhibiting an average oscillation of approximately 3,18°C and the FRZ compartment showing an average oscillation of approximately 9,39°C.

Since the test room temperature is also important for the tests, all tests were carried out at 25°C ambient temperature as specified in IEC test standards which is referenced at the beginning of the section. The results in the figures are summarized in the tables

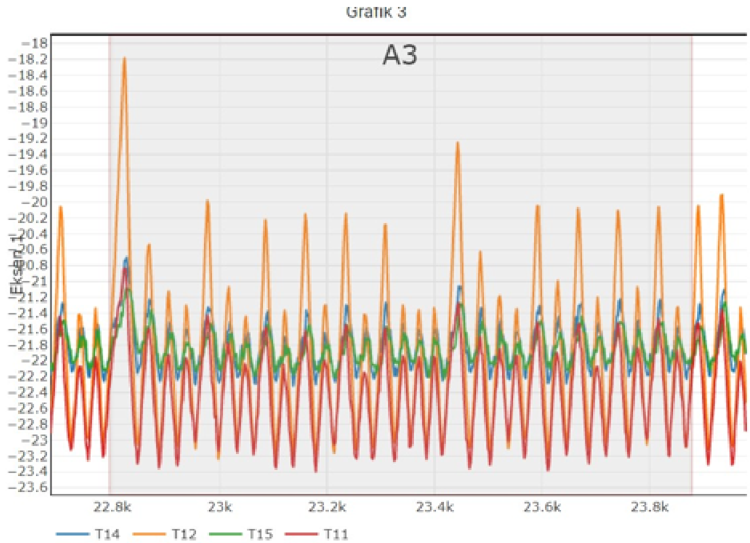

FIGURE 5.10 – Traditional Cooling Software Results for FRZ at -22°C for Energy Test

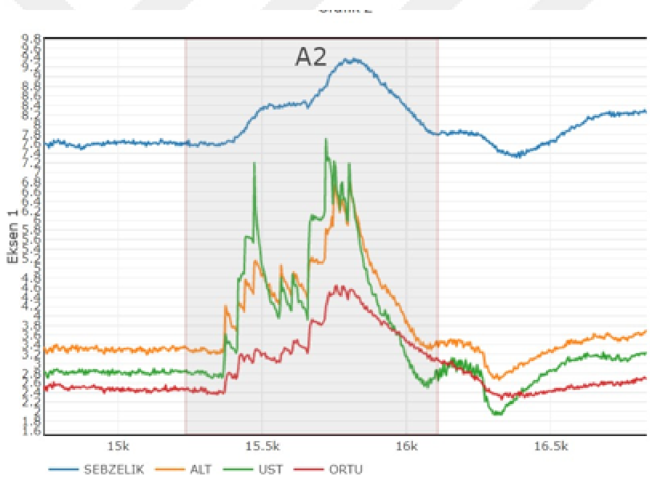

FIGURE 5.11 – New Cooling Software Results for FF at 4°C for Consumer Usage Test

below. The outcomes presented in Table 5.1 and Table 5.2 demonstrate that the new cooling algorithm can eliminate up to 75% of oscillations, with a a slight rise for power consumption of nearly 1.5%. Moreover, when the tests are conducted again with varying temperature value in the same environment, the results displayed in Table 5.3 and Table 5.4 indicate a decrease in both temperature fluctuations which is about 70% for FF while 50% for FRZ and energy consumption which is about 1.6%.

Furthermore, the test results in the figures from consumer usage tests conducted under similar environmental conditions and usage patterns are summarized in tables below. According to these findings in Table 5.5 and Table 5.6, although there was no improvement in temperature oscillation for FRZ, an improvement of approximately 32% was

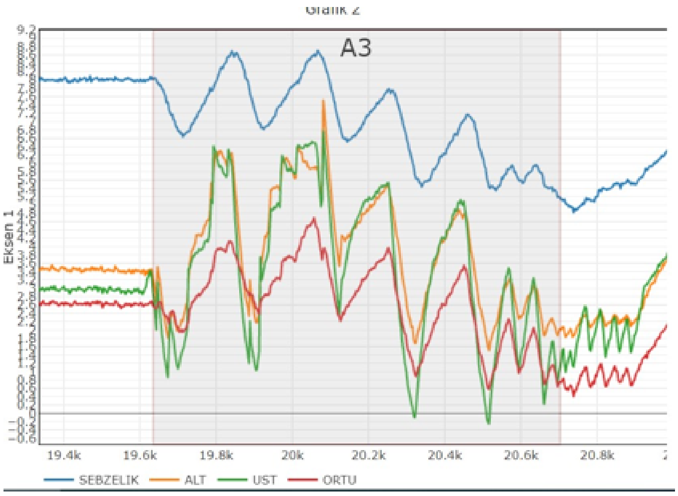

FIGURE 5.12 – Traditional Cooling Software Results for FF at 4°C for Consumer Usage Test

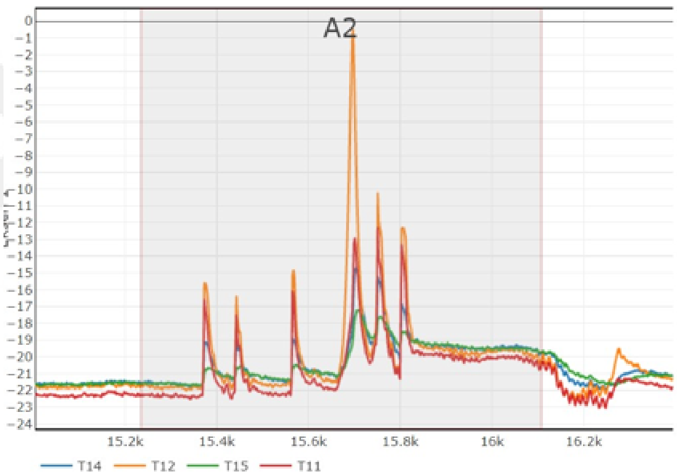

FIGURE 5.13 – New Cooling Software Results for FRZ at -18°C for Consumer Usage Test

achieved for FF at 4°C. Moreover, an improvement in energy consumption of approximately 4% was achieved. Table 5.7 and Table 5.8 summarize that an achievement has been processed for temperature oscillation which is up to 25% and energy consumption improvement which is up to 12% at 3°C/-22°C.

It is important to note that the data utilized for this thesis were anonymized and the measurements were conducted in a controlled laboratory setting.

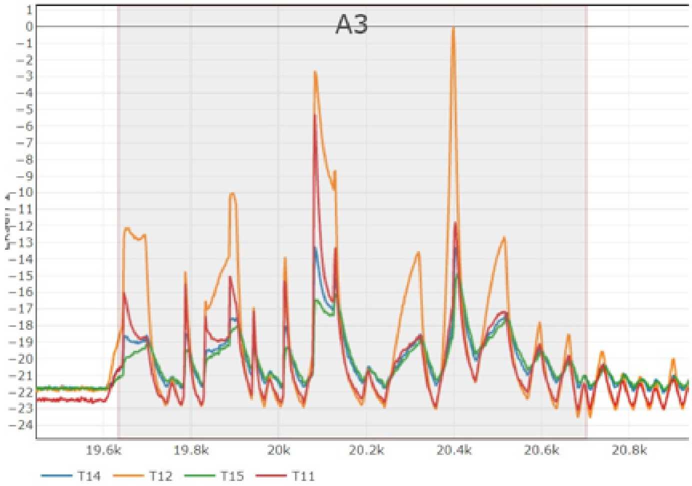

FIGURE 5.14 – Traditional Cooling Software Results for FRZ at -18°C for Consumer Usage Test

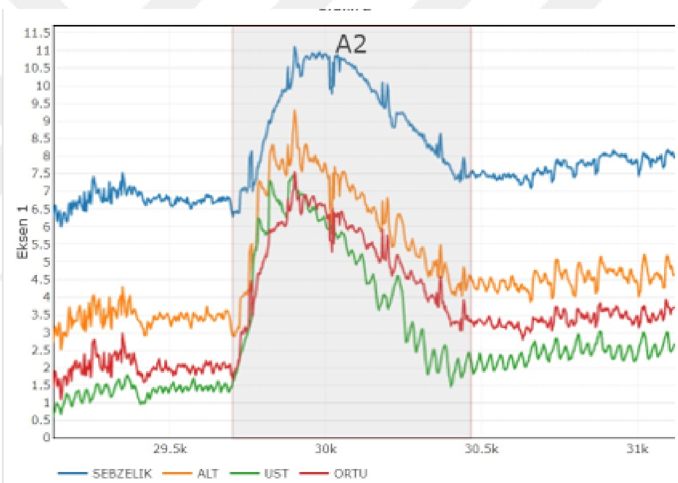

FIGURE 5.15 – New Cooling Software Results for FF at 3°C for Consumer Usage Test

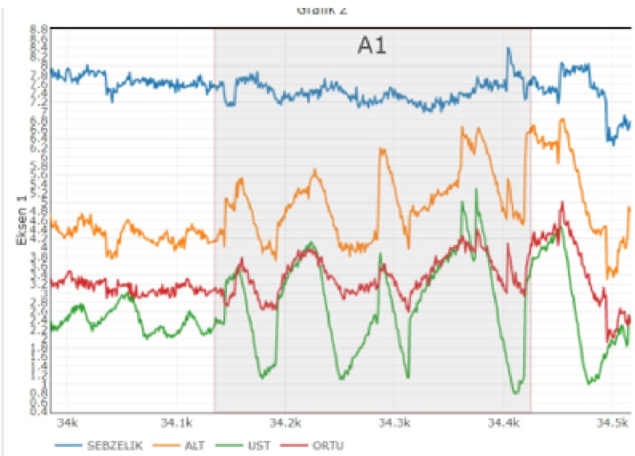

FIGURE 5.16 – Traditional Cooling Software Results for FF at 3°C for Consumer Usage Test

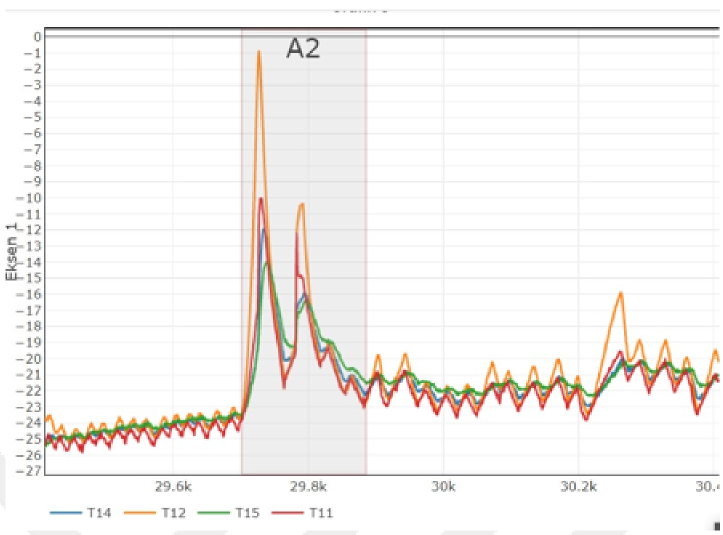

FIGURE 5.17 – New Cooling Software Results for FRZ at -22°C for Consumer Usage Test

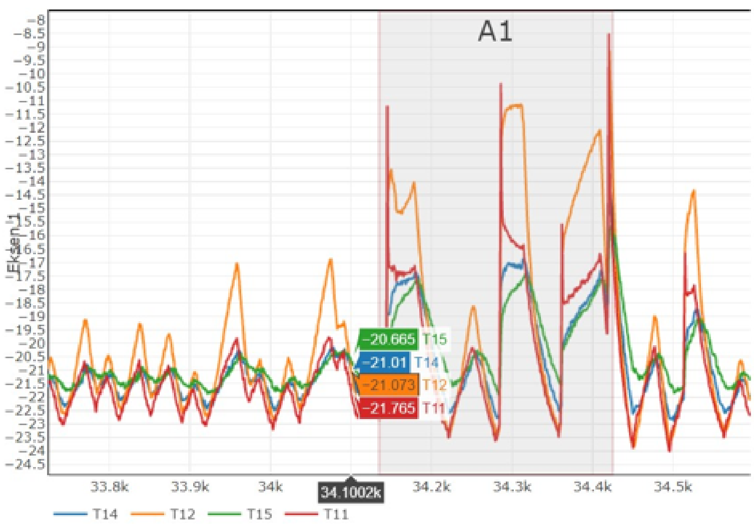

FIGURE 5.18 – Traditional Cooling Software Results for FRZ at -22°C for Consumer Usage Test

Table 5.1 – Traditional Software Energy Test Results in 25°C Environment on -18°/4°C set

Ambient 25° C			
Refrigerator Set	4°C/-18°C		
Energy Consumption	354,95 kWh/y		
Compressor Work Time	100%		
	Min. (° C)	***Max. (° C)***	***Oscillation (° C)***
Avarage Temp. (FF)	1,99°C	3,66°C	1,67°C
Avarage Temp. (FRZ)	-23,17°C	-19,82C	3,35°C

Table 5.2 – New Software Energy Test Results in 25°C Environment on -18°/4°C set

Ambient 25° C			
Refrigerator Set	4°C/-18°C		
Energy Consumption	360,67 kWh/y		
Compressor Work Time	100%		
	Min. (° C)	***Max. (° C)***	***Oscillation (° C)***
Avarage Temp. (FF)	2,89°C	3,37°C	0,48°C
Avarage Temp. (FRZ)	-22,27°C	-21,46	0,81°C

Table 5.3 – Traditional Software Energy Test Results in 25°C Environment on -22°/3°C set

Ambient 25° C			
Refrigerator Set	3°C/-22°C		
Energy Consumption	393,08 kWh/y		
Compressor Work Time	63 minute ON / 10 minute OFF		
	Min. (° C)	***Max. (° C)***	***Oscillation (° C)***
Avarage Temp. (FF)	2,68°C	5,53°C	2,85°C
Avarage Temp. (FRZ)	-23,32°C	-19,27C	4,05°C

Table 5.4 – New Software Energy Test Results in 25°C Environment on -22°/3°C set

Ambient 25°C			
Refrigerator Set	3°C/-22°C		
Energy Consumption	386,67 kWh/y		
Compressor Work Time	100%		
	Min. (°C)	***Max. (°C)***	***Oscillation (°C)***
Avarage Temp. (FF)	1,81°C	2,66°C	0,85°C
Avarage Temp. (FRZ)	-24,65°C	-22,72	1,93°C

Table 5.5 – Traditional Software User Test Results in 25°C Environment on -18°/4°C set

Ambient 25°C			
Refrigerator Set	4°C/-18°C		
Energy Consumption	482,58 kWh/y		
Compressor Work Time	72 minute ON / 8 minute OFF		
	Min. (°C)	***Max. (°C)***	***Oscillation (°C)***
Avarage Temp. (FF)	0,51°C	6,72°C	6,21°C
Avarage Temp. (FRZ)	-19,38°C	-6,22°C	13,16°C

Table 5.6 – New Software User Test Results in 25°C Environment on -18°/4°C set

Ambient 25°C			
Refrigerator Set	4°C/-18°C		
Energy Consumption	464,53 kWh/y		
Compressor Work Time	100%		
	Min. (°C)	***Max. (°C)***	***Oscillation (°C)***
Avarage Temp. (FF)	2,61°C	6,79°C	4,18°C
Avarage Temp. (FRZ)	-22,27°C	-8,63°C	13,64°C

Table 5.7 – Traditional Software User Test Results in 25°C Environment on -22°/3°C set

Ambient 25°C			
Refrigerator Set	3°C/-22°C		
Energy Consumption	435,01 kWh/y		
Compressor Work Time	100%		
	Min. (°C)	***Max. (°C)***	***Oscillation (°C)***
Avarage Temp. (FF)	3,45°C	7,67°C	4,22°C
Avarage Temp. (FRZ)	-23,09°C	-11,77°C	11,32°C

Table 5.8 – New Software Energy Test Results in 25°C Environment on -22°/3°C set

Ambient 25°C			
Refrigerator Set	3°C/-22°C		
Energy Consumption	386,67 kWh/y		
Compressor Work Time	100%		
	Min. (°C)	***Max. (°C)***	***Oscillation (°C)***
Avarage Temp. (FF)	2,36°C	5,54°C	3,18°C
Avarage Temp. (FRZ)	-22,65°C	-13,26	9,39°C

6 CONCLUSION

Through the exploration and discussion of adaptive PID controllers integrating with reinforcement learning, it is clear that this junction represents a significant stride towards more efficient, responsive, and intelligent control systems across various sectors. The adaptation and enhancement of PID control through machine learning, particularly reinforcement learning, herald a new era where control mechanisms can dynamically adjust and optimize in real-time, demonstrating superior performance over traditional methods. These advancements promise to transform industries by making systems not only more adaptable to changing environments but also more capable of preemptive and precise decision-making, thereby enhancing overall efficiency and effectiveness.

The implications of adapting reinforcement learning in PID control are vast, indicating a potential shift in how we approach automation and control in various domains, from industrial processes to autonomous vehicles. The flexibility and continual learning offered by these systems suggest a future where predictive adjustments and improved control strategies become standard.

This thesis explores an adaptive PID controller designed within an Actor-Critic framework using a RBF network on refrigeration system. It presents an adaptive updating rule based on adjusting the network's weights. The traditional PID controller is integrated with RL on an RBF network basis, allowing online tuning. The reinforcement signal is based on predictive output, facilitating accurate updates. In this approach, the hidden layer of the RBF network is shared by both the Actor and Critic, reducing storage needs and computational costs for hidden unit outputs. The PID parameters are defined as zero at initialization, indicating no need for prior system knowledge.

In conclusion, it is confidently stated that the new cooling algorithm which is developed and adapted to refrigerators has a positive impact on the temperature stability which is beneficial for storage. According to energy test results, it is observed an improvement in compartment temperature oscillation up to 75%. In user tests, refrigerators' energy consumption is decreased by up to 12%, and up to 24% decrease in compartment temperature oscillation is obtained. The implemented algorithm is lightweight, allowing it to operate on current cost-optimized hardware. Moreover, it enhances the energy efficiency of refrigerators. Given the widespread use of refrigeration units in household, medical, and industrial settings, improvement in energy consumption can positively impact global energy usage.

For future research, an alternative optimization could involve integrating real-time power consumption data from the refrigerator into the the control system to further reduce energy usage. Moreoever, by integrating with electricity distribution systems such as solar power plants, wind power plants and any other power plants, with the help of IoT, the systems can communicate within themselves and thanks to this communication, algorithms such as those implemented in the study can be activated and deactivated for instant or long-term optimization of power consumption.

REFERENCES

Arçelik A.Ş. (2020). Homewhiz Mobile Application.
URL: *https ://www.homewhiz.com/*

ASHRAE (2014). *Ashrae Handbook-Refrigeration*, si edn, www.ashrae.org. ISBN : 978-1-936504-72-5.

Barthel, C. and Götz, T. (2012). *What Users Can Save with Energy-efficient Refrigerators and Freezers*, bigEE.

Belman-Flores, J., Pardo-Cely, D., Elizalde-Blancas, F., Gallegos-Muñoz, A., Pérez-García, V. and Gómez-Martínez, M. (2019). Perspectives on consumer habits with domestic refrigerators and its consequences for energy consumption : Case of study in guanajuato, mexico, *Energies* **12**(5) : 860.
URL: *http ://dx.doi.org/10.3390/en12050860*

Dietterich, T. (2000). Hierarchical reinforcement learning with the maxq value function decomposition, *The Journal of Artificial Intelligence Research (JAIR)* **13**.

Dogru, O., Velswamy, K., Ibrahim, F., Wu, Y., Sundaramoorthy, A. S., Huang, B., Xu, S., Nixon, M. and Bell, N. (2022). Reinforcement learning approach to autonomous pid tuning, *Computers & Chemical Engineering* **161** : 107760.
URL: *https ://www.sciencedirect.com/science/article/pii/S0098135422001016*

Guan, Z. and Yamamoto, T. (2020). Design of a reinforcement learning PID controller, *2020 International Joint Conference on Neural Networks (IJCNN)*, IEEE.
URL: *https ://doi.org/10.1109/ijcnn48605.2020.9207641*

Hamid, N. H. A., Kamal, M. M. and Yahaya, F. H. (2009). Application of PID controller in controlling refrigerator temperature, *2009 5th International Colloquium on Signal Processing & Its Applications*, IEEE.

Hasanuzzam, M., Saidur, R. and Masjuki, H. (2008). Investigation of energy consumption and energy savings of refrigerator-freezer during open and closed door condition,

Journal of Applied Sciences **8**(10) : 1822–1831.
URL: *http ://dx.doi.org/10.3923/jas.2008.1822.1831*

IEEE (2006). Advanced PID control - [book review], *IEEE Control Systems* **26**(1) : 98–101.
URL: *https ://doi.org/10.1109/mcs.2006.1580160*

International Electrotechnical Commission (2015a). IEC 62552-1 :2015 Household refrigerating appliances - Characteristics and test methods - Part 1 : General requirements.
URL: *https ://webstore.iec.ch/publication/21805*

International Electrotechnical Commission (2015b). IEC 62552-2 :2015 Household refrigerating appliances - Characteristics and test methods - Part 2 : Performance requirements.
URL: *https ://webstore.iec.ch/publication/21804*

International Electrotechnical Commission (2015c). IEC 62552-3 :2015 Household refrigerating appliances - Characteristics and test methods - Part 3 : Energy consumption and volume.
URL: *https ://webstore.iec.ch/publication/21803*

Konda, V. and Tsitsiklis, J. (2001). Actor-critic algorithms, *Society for Industrial and Applied Mathematics* **42**.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. (2013). Playing atari with deep reinforcement learning.

Peker, M. K., Naskali, A. T., Kılıçaslan, S. and Özbey, N. (2023). Precise temperature control for sensitive medicines, *2023 IEEE International Conference on Big Data (BigData)*, IEEE.
URL: *http ://dx.doi.org/10.1109/BigData59044.2023.10386421*

Rummery, G. and Niranjan, M. (1994). On-line q-learning using connectionist systems, *Technical Report CUED/F-INFENG/TR 166* .

Sabegh, M. R. Z. and Bingham, C. M. (2019). Impact of hysteresis control and internal thermal mass on the energy efficiency of IoT-controlled domestic refrigerators, *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*,

IEEE.
URL: *https ://doi.org/10.1109/sege.2019.8859886*

Sedighizadeh, M. and Rezazadeh, A. (2008). Adaptive pid controller based on reinforcement learning for wind turbine control, *Proceedings of World Academy of Science, Engineering and Technology (CESSE2008)* **27**.

Shuprajhaa, T., Sujit, S. K. and Srinivasan, K. (2022). Reinforcement learning based adaptive PID controller design for control of linear/nonlinear unstable processes, *Applied Soft Computing* **128** : 109450.
URL: *https ://doi.org/10.1016/j.asoc.2022.109450*

Sinaga, K. P. and Yang, M.-S. (2020). Unsupervised k-means clustering algorithm, *IEEE Access* **8** : 80716–80727.

song WANG, X., hu CHENG, Y. and SUN, W. (2007). A proposal of adaptive PID controller based on reinforcement learning, *Journal of China University of Mining and Technology* **17**(1) : 40–44.
URL: *https ://doi.org/10.1016/s1006-1266(07)60009-1*

Surya Priy (2024). Clustering in Machine Learning.
URL: *https ://www.geeksforgeeks.org/clustering-in-machine-learning/*

Sutton, R. (1988). Learning to predict by the method of temporal differences, *Machine Learning* **3** : 9–44.

Sutton, R. and Barto, A. (1998). Reinforcement learning : An introduction, *IEEE Transactions on Neural Networks* **9**(5) : 1054–1054.
URL: *https ://doi.org/10.1109/tnn.1998.712192*

Sutton, R., Mcallester, D., Singh, S. and Mansour, Y. (2000). Policy gradient methods for reinforcement learning with function approximation, *Adv. Neural Inf. Process. Syst* **12**.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning*, Adaptive Computation and Machine Learning series, 2 edn, Bradford Books, Cambridge, MA.

U.S Energy Information Administration (2020). Residential Energy Consumption Survey.

URL: *https ://www.eia.gov/energyexplained/use-of-energy/electricity-use-in-homes.php*

Visioli, A. (2010). *Practical PID Control*, Advances in Industrial Control, Springer, London, England.

V.T., S. E. and Shin, Y. (1994). Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems, *IEEE Transactions on Neural Networks* **5**(4) : 594–603.
URL: *https ://doi.org/10.1109/72.298229*

Wait, K. W. (2012). Adaptive temperature control of a class of home refrigerators, *2012 American Control Conference (ACC)*, IEEE.
URL: *https ://doi.org/10.1109/acc.2012.6314736*

Watkins, C. J. C. H. and Dayan, P. (1992). Q-learning, *Machine Learning* **8**(3–4) : 279–292.
URL: *http ://dx.doi.org/10.1007/BF00992698*