International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET1229152

Optimization of Multi Agent System for Distribution Control of Distinct Heating System Using Improved Q-Learning Controller

Ehibe Prince[1], Uchegbu Chinenye E[2]

[1]Electrical and Electronic Engineering Department Abia State Polytechnic Aba Nigeria.

[2]Electrical and Electronic Engineering Department, Abia State University Uturu Nigeria.

Article Info

Volume 9, Issue 1

Page Number : 235-249

Publication Issue :

January-February-2022

Article History

Accepted : 17 Feb 2022

Published: 26 Feb 2022

ABSTRACT

The aim of this research is the use of multi-agent systems for the optimization of a distributed control for district heating systems. A district heating system comprises of production units, a distribution network, and a host of consumer substations. The operations of district heating system usually involves conflicting goals, e.g., to satisfy customers and to minimize production costs. Hence the agent must be capable of optimizing between maximizing supply to substations and minimizing production cost. Current substations employ purely reactive devices, making local decisions without taking into account the global state. Moreover the substations determine the flow in all parts of the district heating system. The optimal operation of the district heating system is therefore limited to providing sufficiently high temperature and pressure to all customers by taking local measurement to achieve this goal without considering other factors such as cost of production and time. The approach studied in this research is to equip substations with software agents to form a multi-agent system using Q-learning. The study also shows that it is possible to control the trade-off between quality of service and degree of surplus production as well as the possibility of extending the system with new consumers without increasing production capacity. In another study, an experiment was conducted in a controlled physical environment, where two agent-based approaches were evaluated and compared to existing technologies. The experiment shows that it is possible to automatically load balance a small district heating network using agent technology.. Finally, a generalized formal characterization of the problem space under investigation is provided, i.e., production and logistics network management, together with a preliminary evaluation of the applicability of the suggested multi-agent system approach for this general problem area.

Keywords : Q- learning, multi agent, optimization ,controller, simulation, District heating system

I. INTRODUCTION

This paper is basically on investigation of the applicability of Q-learning controller to model a system of Multi-Agent Systems MAS as a distribution control approach for District Heating Systems DHS. A multi-agent system MAS or "self organized system" is a characterized system composed of multiple interacting combine to achieve a certain goal, and the agent has ability to solve problem which individual agent can't handle and has the capacity to monitor and give accuracy that is impossible for an individual agent or a monolithic can solve. A multi-agent system (MAS) is organized to meet objectives based on rules and regulations of agent. An MAS is a system that combine integrates a set of agents that interact, communicate, coordinate themselves to achieve the established oriented objectives goal (OOG).

And also agents is an intelligent entities with social skills allocation which (interactive, interaction, collaboration, communication, coordination, competence, negotiation, intelligence integrative) that encapsulate a functionality to solve a problem within the domain, (Wooldridge and Jennings 2018), and (Jennings *et a*l*.,* 2014). DHS involves the burning or combustion of renewable biomass in a facility to produce steam based on demand, which is subsequently passed through a network of pipe and used to supply to buildings (Bauen *et a*l*.,* 2009). The heat exchange systems, in current DHS employ purely reactive devices and have typically no communication capabilities. In this note they are able to make local decisions without taking into account any situation in the domain. In this work, an intelligent multi-agent system is designed using Q-learning this has the advantage of interacting with the environment (heating system) in order to maximize for a reward which the minimization of production cost to any variation in the consumption level of the substation. Q-Learning comes under Value-based learning controller.

The objective function is to optimize a value function suited to a given problem environment for state reward of action. The 'Q' stands for quality; it helps in finding the next action resulting in a state of the highest quality reward in a step. This approach is rather simple and intuitive. It a very good place to start the QL journey. The values are stored in a table, called a Q Table implementation.

A software agent is a natural extension of the concept of software Object-oriented programming supplementary, i.e., objects that have persistent local states to the structured programming paradigm. Similarly, agent based programming adds abstraction entities, i.e., agents that have an independent execution thread and pro-activity to the object-oriented paradigm. Thus, compared to other agent which is able to act upon in a goal directed by an action of reward, e.g., by interacting with other agents, on actuators reading sensing actions, or sending commands to effectors for policy, rather than only passively react to procedure call. There is no strict consensus within the agent community on the definition of an agent but a commonly used is the definition by Wooldridge and Jennings (Fredriksen and Werner, 2009): "An agent is a programme software computer system that is installed in an environment, and that is capable of autonomous action in its domain in order to meet its design objectives goal". Based on their tasks, the complexity of the agents varies depend on rules. Purely reactive agents only perform a mapping from sensor data to effect or signals. (Sensing and command is given a very general interpretation and coordination, including receiving and sending messages based on code mapping.) In the most cases, the various behaviour of a reactive agent can be specified by a collection of independent situation action base rules. A more sophisticated approach is the system architecture (Bøhm *et al.,* 2018) which consists of a hierarchy of behaviours where each behaviour is a rule-like structure that" competes" with other behaviours to exercise control over the agent. Reactive agents have been proved to be good at doing a number of simple tasks in real world domains. In contrast to reactive agents, deliberative agents have modularized cognitive abilities (perception, world modeling, planning etc.). Purely deliberative agents contain add code explicitly represented model that is used for decision making. The working principle of a deliberative agent can be described as a sensing model deliberate act cycle. The sensors sensed the environment through and agent and receive messages, which are used to predict and update model. The model is used in this research is deliberation module to decide which actions to be taken, which serve as input to the effectors that carry out the actions execution. Although purely deliberative agents may be suitable for more complex tasks, they have problems with "simpler" tasks such as routine reaction that require fast action but no extensive deliberation since planning is typically very time consuming, requiring exponential search through potentially enormous problem spaces. Consequently, deliberative agents tend not to work well in highly dynamic environments that require fast reaction. Hybrid agents try to integrate the abilities of reactive agents for routine tasks with the power of deliberation necessary for more advanced or long term tasks. Two categories of hybrid agents can be distinguished. Uniform agent architectures, such as the Procedural Reasoning System (Tamminen and Wistbacka, 2017), employ a single representation and control scheme for both reaction and deliberation, whereas layered agent architectures, (Arvastsson, 2017), use different representations and algorithms (implemented in separate layers) to perform these functions.

II. OVERVIEW

2.1 Conceptual Frame Work

In this part of the review, the research works related to the performance enhancement for different researchers on multi-agent system for control of district heating system application using improved Q-learning intelligence, various line and support supply of hot water in different building is introduced. The concept of comprehensive phenomenon literatures used in this part is summarized in sections of linked according to the control operation of author's strategies. To implements the concepts of the application in automation controller. Researchers in several fields as well as control automation, computer science, electronic engineering have studied in conceptual framework approach in the area of coordination therein.

2.2 Theoretical Frame Work

The design of agent systems deals with building interaction and collaboration in distribution and control system domains, theoretical frame work is been made for overall where each autonomous agent works cooperatively to improved district heating system. However, an agents grouped is more flexible and efficient than a single agents only when a flexible and efficient ways of coordinating the agents. In many ways, the agent system design problem is similar to that of parallel computing, increasing the number of controller used in a computation usually wills not double the speed with which the solution is found. The extra processing power become an advantage until a sophisticated means of cooperative processing is found with reward of policy of state environment for effective action. This challenges causes many design patterns and rule base definitions. "Using updated mathematical model to investigate ways in which multi agents with multiple goals interact to achieve a certain goals. The theory includes all forms of communication, recognition, cooperation, and coordination for engineering and social agent. Using a model of mathematical optimization technique rules for tactical planning and agents operational re-planning and performing theoretical adjustments plan in real-time to handle the actual conditions planed expected. Also a refinement of the above approach would just construct partial plans when the discrepancies are large only in a part of the system. In this case the corresponding part of the multi-agent system passes the control of that area to the optimization algorithm for various theoretical approaches for result of agent control of system.

2.3 Related works

In this part of the review, the works related to the performance enhancement for different multi agent system control for dispense of hot water are introduced. The related works used in this part are summarized in sections according to its operation for details of engagements. To implement this concept of the application in operation of district heating system for buildings, hospital and industrials services has been done by various Researchers in several fields as well as instrumentation and control automation system, computer Engineering, Electronic engineering Automation controller have studied in related works approach in the area of coordination therein etc, for reference purpose in related work.

2.3.1 Combined District Heating and Electrical Network Simulation

''Unified Energy Agents for Combined District Heating and Electrical Network Simulation" (Nils Loose *et a*l., 2020) for energy valuation of heating. Based on this scenario a variety of mechanisms have been developed to manage and coordinate problems in heating application. On one side they use an organizational structures and social laws (Davidsson and Wernstedt, 2018), long term rules that governed the behaviour of the society of agents. At the other end are the black board model (Ferber, 2015) and the one shot protocols, e.g., contract net (Huhns and Stephens, 2015). In between are techniques such as partial global planning (Malmström *et a*l., 2011) and various negotiation techniques, e.g., market based (Nwana *et a*l., 2011) and game theoretic (Paranak 2015) negotiation. Several researchers have shown that there is no single best organization or coordination mechanism for all environments (Parunak, 2016). Based on concentrated on MAS application in which the frame of the system can be divulge at design time, and where agents basically cooperate in order to fulfill a goal on the system level dated.

2.3.2 Responsive Control Strategy for District heating systems.

A novel request for control of responsive strategy for district heating systems, featuring return temperature reduction (Hakim Ibrahim et al 2020) the authors present a novel demand responsive control strategy to equip centrally the district heating systems level. (Geo Watt 2016) They demanded responsive features was maintained based on direct and the indirect substation configurations for network supply to various customer on cluster (based on their rating measures) in order to achieve energy return temperature degrees from the end-user substations. Different than the normal weather compensation based supply temperature resetting was used; the new control strategy was formulated to adjust the supply temperature at the district level as to the heating performance at the end-user substations. Different simulations were conducted in order to quantify the benefits of the novel control strategy as compared to the original weather compensation, equipped both at the substation level and the district level. The results obtained showed that the new control strategy, when considering the electricity loss at the heat production plant and gain, shows superiority when compared to other control strategies.

2.4 Theory of Technique

An agent based approach for control of district heating system approach, from various studied design perspective the concept did not cover abstraction that provides an easier and more natural conceptualization of the problem domain. Other advantages are increased, the production, distribution of control to a number of agents various substation to cluster and customer efficiency, less complex computations and communication are necessary if control is distributed, flexibility, the use of agent communications languages that support complex interaction between entities provides a after various work done a communication protocols, openness, by having a common communication language, agents implemented by different developers are still able to interact with each other, scalability, a multi-agent system, and Q-learning controller of reinforcement learning approach for district heating system is adopted as technique of this work.

III. METHODOLOGY

3.1 Design Method

The Q-learning technique is used in this thesis to describe the core subsystems suitable which is modeled as software agents for the district heating system. Q-learning is an Reinforcement learning controller to learn the value of an action in a particular state of function. It studied an environment through an agent which perceive and environment through a sensor and acted upon an environment through an actuator, and it can handle problems with stochastic transitions and rewards without requiring adaptations for policy, reward, state and value function of optimal model of and environment. For a scenario code of finite Markov decision process (FMDP) adaptation, Q-learning finds an optimal policy which performs specific action in the sense of maximizing expected value of the total reward over any successive sequential steps, starting from the current state. Q-learning can identify an optimal action selection policy for any given FMDP, given infinite exploration time and a partly random policy. "Q" means quality function that the algorithm controller computes in the expected rewards which performs actions in an environment that follow policy in a state.

3.2 Materials Used

In this technique a system is designed to move from no knowledge of the environment conditions to increasingly concrete concepts where the interaction with the environment leads to an optimal solution of the environment variables. Rewards are assigned to goal value in each state (substation consumption level), and after an exhaustive interaction with the environment, the state with the highest reward is chosen for each substation. Q-learning environment where the model of the environment is explicitly designed, in this research, the abstract concepts, e.g., roles and permissions, to be used during analysis to conceptualize the system, are modeled using MATLAB. The various objectives of this research will be implemented sequentially using by characterized the first the analysis, both theoretical approach and experimentation. This implies that some type of experimentation will be done. To experiment with actual DHS is both difficult and expensive. This is as a result of the District Heating System is distributed over a large area of confine space and which, would as a sequential order modification of numerous heat exchange systems of HVAC which would be unreasonably expensive for this research. A more common method for analyzing systems as complex as DHS is to perform simulation experiments. Consequently, the work presented in this thesis is mostly based on empirical work with a simulation model. An initial simulation model will be constructed and implemented using an initial simulation model to perform simulation experiments as a case study for the more general problem area of just-in-time production and distribution. To obtain an answer to the research objective questions on what is gained by introduction of cooperation and communication between the Systems and the distributed units, an experiment for analysis was done in a laboratory at Glass Force

Industry Company Aba, where a small DHS was used to perform a controlled small scale experiments, in which the existing technologies was compared with solutions based on different MAS architectures. In order to further evaluate the applicability of MAS for control of DHS a high version simulation tool was needed where every part of the DHS was modeled at a detailed level, e.g., the water flow and temperature propagation in every part of the distribution network is modeled. The Simulink tool in MATLAB has the features for simulating all this control structures. This thesis also includes a theoretical analysis based on a literature study regarding the properties of agent based and classical optimization techniques to evaluate the advantages and disadvantages of distributed agent-based approaches compared to classical mathematical optimization techniques. Finally, to generalize the domain of DHS develop a formal characterization of logistics network and describe DHS in term of this general framework which the steps are as follows.

Implementations and method of the study:

3.3 To Characterize a Multi agent Environment of DHS based Q-Learning through Distribution control system.

Optimizing District heating system (DHS) is considered a promising to improved the efficiency of supply of hot water in buildings through multi agent system (i.e. Hospitals Residential, commercials, hospitals and industries) etc (Olsthoorn *et al.,* 2016). Thus, a greater interest in installing DHS has arisen because the demand for hot water in Nigeria is commendable. DHS have much value of advantages, for an optimum use of biomass (Benonysson *et al.,* 1995) for it significant.

Generally, a DHS is represented by transmission networks employed to supply heat from supply side source (i.e. Generation site) directly to demand side (i.e. end users) to meet the demand of users been consumer. Often, DHS can be engaged either with centralized heating substations and or distributed heating system units facility. Thus, numerous kinds of heat generation technologies (renewable energy, boilers, heat pumps, cogeneration plants, etc which communicated through actuators, sensors and controllers.) can be adopted (Joelsson *et al.,* 2008), combined heat and Q-learning controller based DH systems are often seen as a key solution to meet the local heat demand in buildings by Appling the multi-agent approach to quashing effect and, therefore, these agent units are frequently heat-driven (Elci *et al.,* 2015). This mode is often seen as the most economical option and therefore preferred by most operators of distributed energy systems due to the utilization of produced heat to meet the local heat demand and, therefore, no heat is wasted (Shipley *et al.,* 2008) This is integrated through agent systems by distribution control system controller by (Bracco *et al.,* 2013), process controllers configured for supervisory control and data acquisition set point typically operating mode of automatic system with remote SP: Controller automatically switch its output to keep PV to steady state where Set point value are mimic to supervising computer multi sensory application.

The supervisory control and data acquisition system can include many such controllers and can be connected and automatically adjust to master computer for action shown in figure 3.1.

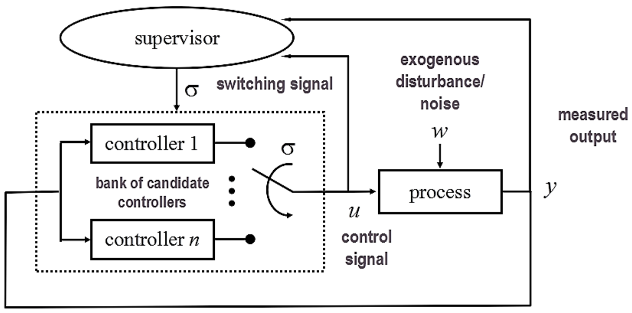

Figure 3.1 Supervisory Controls of MAS

The system controller makes the set point SP and gives to the controllers over Q-learning reward. Controllers adjust the PV with respect to the set point using DHS controller by combing the controller of multi agent system and varying the Q-learning by checking the reward of the environment to issue command of set point of action. This forms an N-layer process control system: the "base" or "regulatory" layer of control (DHS loop controllers) of various substations and the "high" or "supervisory" level of control (the powerful computer with the mathematical process models).

So a series of digital data lines are implemented to transfer set-point from supervisory to controller. But it may also carry process variable information from those controllers back to the supervisory data for its optimization control to operate as a reward of duty of state.

3.4 Multi-Agent Reinforcement Learning using Q-learning Control.

Generally, the transition probability for Multi-agent reinforcement learning (MARL) extends equation. (3.1) below for a single agent Markov decision process to a multi-action case:

$$p(s', r|s, a) = \mathbb{P}[S_t = s', R_t = r|\, S_{t-1} = s, A_{t-1} = a], \qquad 3.1$$

The multi-agent system Q-learning model is shown in equation 3.2

$$p(s', r_i|s, a_i) = \mathbb{P}[S_t = s', R_{t,i} = \{r_1, \dots, r_n\}|\, S_{t-1} = s, A_{t-1,i} = \{a_1, \dots, a_n\}], \qquad 3.2$$

Where n is the total number of agents, r_i is the reward for agent i and $\boldsymbol{s} = \{s_1 \dots s_n\}$ is the set of individual states. In equation 3.2, the stochastic transition is a probability distribution over the next vector of states, $\boldsymbol{s}'$, given the current vector of states, $\boldsymbol{s}$, and joint action, a_i. For the policy $\pi i \in \Pi i$, the optimal policy $\pi i*$ for agent i fulfills the Nash equilibrium shown in equation 3.3

$$\begin{aligned} &\sum_{a_1,\dots,a_n} q_*(\boldsymbol{s}, a_i)\, \pi_1^*(a_1|\boldsymbol{s}) \cdot \dots \cdot \pi_i^*(a_i|\boldsymbol{s}) \cdot \dots \cdot \pi_n^*(a_n|\boldsymbol{s}) \\ \geq &\sum_{a_1,\dots,a_n} q_*(\boldsymbol{s}, a_i)\, \pi_1^*(a_1|\boldsymbol{s}) \cdot \dots \cdot \pi_i(a_i|\boldsymbol{s}) \cdot \dots \cdot \pi_n^*(a_n|\boldsymbol{s}), \end{aligned} \qquad 3.3$$

Where $q*$ $(\boldsymbol{s},)$ is the optimal action value function for agent i and $\pi i*(a_i|\boldsymbol{s})$ is the individual probability of taking action a_i given the Nash equilibrium policy from figure 3.2 above

3.5 The MAS Distribution Heating System Simulator (Operational) Model

In the following sections the simulation model for the real time control of the district heating system for optimization of Q-learning and its different modules are described. First the description of the requirements will be outlined; this is followed by an analysis of these requirements. Then the design and implementation of the simulator are presented. Finally, the validation of the decision making control hierarchy for DHC systems operation is shown in figure 3.2. A typical Mile prediction horizon is 6–12h controller. The lower control layer corresponds to a fast time scale model that handles the basic regulation action of process variables of model domain, such as configuration of action, temperature supply, mass flow rate and pressure, at the substation building level is based on decision making which show control operation supply for customers.

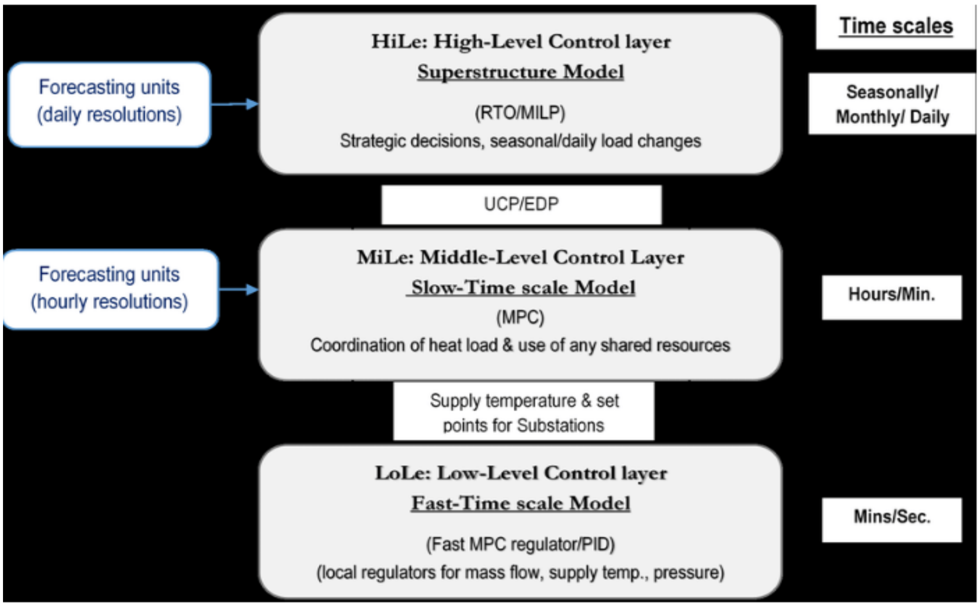

Figure 3.2 Decision making control hierarchy for DHC systems operation.

Supporting Q-Learning-Based Tools efficient forecast are necessary components for DHC planning in networks supply for customer playing a key role in operational decision making. Hot water forecasting for customers could be a crucial contributing factor for estimating future district heating demand in DHS. Several weather parameters, such as temperature, solar radiation, atmospheric pressure, humidity and precipitation, are widely available are been required. The availability of this information base sets environment goal reward the stage for a great adoption of more advanced data-driven computer modeling and automated processing of network data. However, data driven demand prediction remains a challenge due to the inherent uncertainties of the existing weather forecast methodologies. Big data and advances in sensors and smart meters have resulted in massive growth of energy datasets, thus creating new opportunities for energy prediction in DHS. Taking advantage of the vast availability of data, a machine learning (ML) of Q-learning approach is proposed that relies on readings from smart meters and contextual information such as weather data or forecasting weather parameters that recorded in SCADA which will be display on HMI, DCS process for SCADA Communication which show in Figure 3.3 for control of configuration of operations.

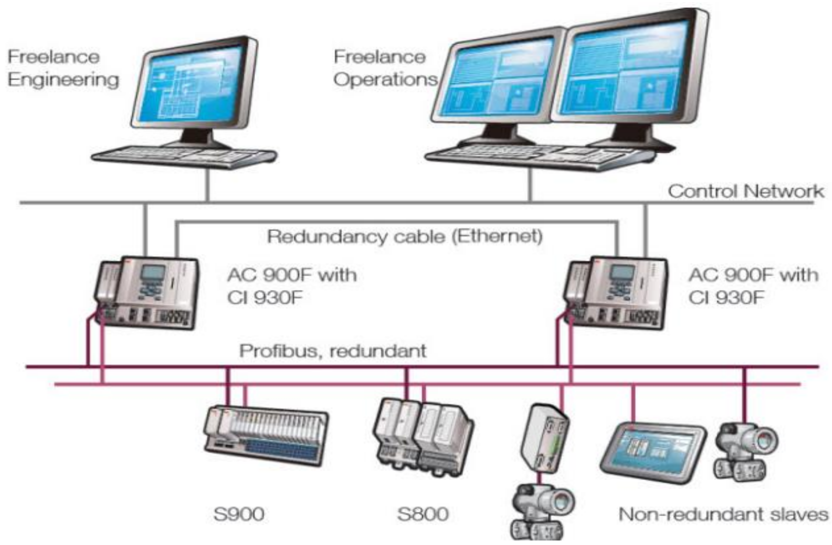

Figure 3.3 DCS process for SCADA Communication

This has a line process that categorized into two stages prediction tool comprises of:

(a) A data driven demand for hot water forecasting unit that is trained on local weather data for individual that need hot water and predicts future weather parameters within a horizon of 24 h at maximum; and

(b) A ML-based short-term, real-time energy demand prediction unit that uses both weather predictions and energy consumption historical data to predict accurately the forthcoming consumer energy needs within the same time horizon of 24 h.

3.6 Monitoring and Control of Real Time District Heating Systems

The technique of this objective is to improve the monitoring and control of district heating system DHS networks through the use of agent controller using Q-learning. One of the means for achieving this is to equip the intelligent agent to store all the states (consumption levels) of consumers in the network at the producer side. In current district heating networks, the operators usually have very little knowledge about the actual state of the network and even less knowledge about predicting the future states. Instead, the operators control the network by using rules of programming based on the HVAC temperature and what time of the day it is. In contrast, approach makes use of current consumption (and therefore more informed) predictions of future needs by Q commands form figure 3.4

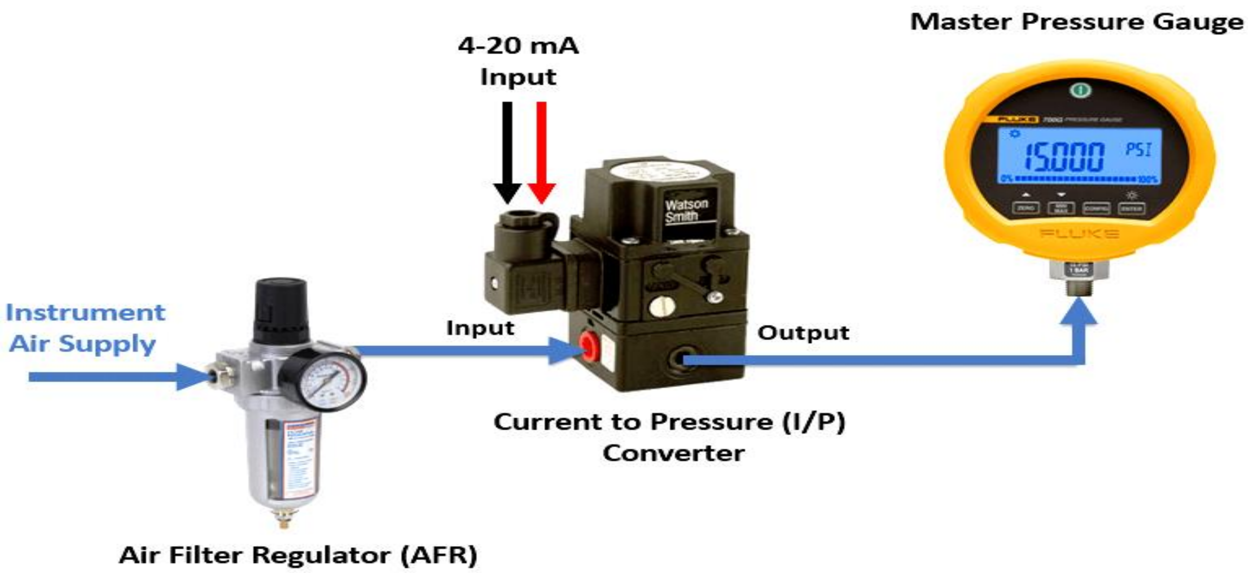

Figure 3.4 Pressure Monitoring and Control

Flow of Current to pressure converter of input/output (I/P) converts an analog signal of (4 to 20 mA) to a proportional pneumatic output (3 to 15 psi) to keep process steady in operations which is for monitoring and control of process. It is one of the main pneumatic instruments used in this work for the district heating system supply for various substations to execute results and update.

IV. DATA PRESENTATIONS

4.1 Results and Discussions of Characterization of Multi agent DHS Q learning

In this module, the simulation results will be analyzed, the characterization design, and the results of it will also be discussed. Assumed that the distribution time from the producer to the consumers is on demand and there is a single production source for distribution to various substations for effective service to customer. In the model the MAS field excitation is provided with constant source and the network is provided by selecting the controlled substations selected for the same with a step signal as input. The mechanical system is modeled to provide load customers demand. The dynamic response from Q-learning model is captured using scope with step change to distribute control and monitoring. Also the simulation results are generated with graphs tabulated and conclusions are drawn from the results based on updating of data.

MATLAB / SIMULINK were used to obtain the optimal specifications, which are for fast decision with improved Q-learning controller depending on the demand from various substations to customer. A comprehensive MATLAB code for the multi-agent system in district heating is shown in appendix A

4.2 Data Presentation Analysis

The MAS as well as the simulation environment was implemented in simulink (Agent Development framework) (Bellifemine et al 2017). Thus, use an agent-based approach that was time driven, i.e., where the simulated time is advanced in constant with time step by step which generate the data in the environment based on policy. Each data entity is implemented as a separate agent. Figure 4.1 shows the different parts of the generation reward policy.

The parametric tests are on a quantitative (data) scale, with a normal distribution of the system. The samples have the same variance (homogeneity of variances). The samples are randomly drawn from model of simulation, and the observations within a group are independent of each other when run and graph was generated.

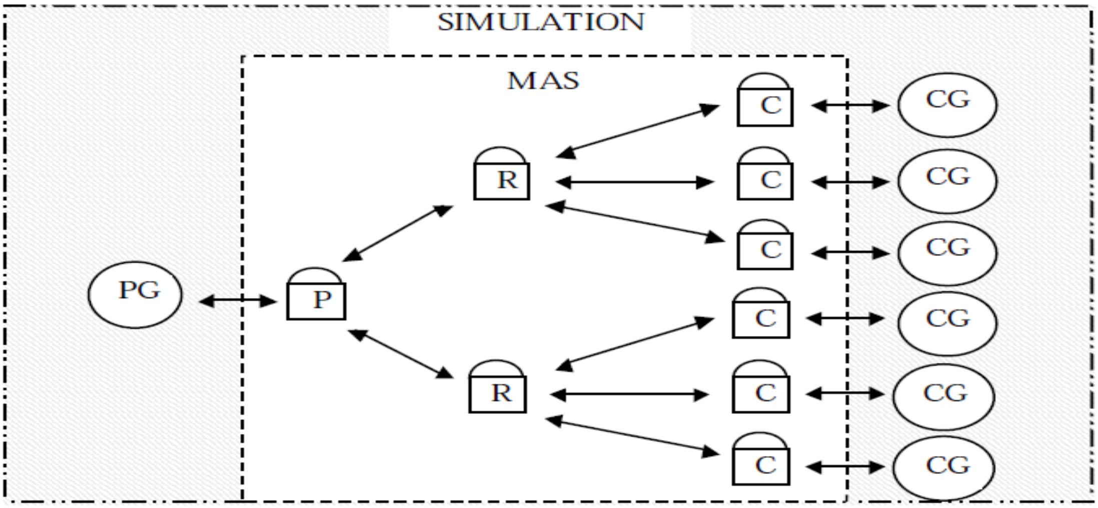

Figure 4.1: The MAS and flow parts; RA = Redistribution Agent, CA = Consumer Agent, PA= Producer Agent, CG = Consumption Generator and PG = Production Generator

Analysis for hot water usage, based on detailed field measurements performed in glass force industry Aba (Control Engineering Department Laboratory), was used to generate consumption values. The predictions of consumption for each substation were calculated as the average over five generated consumption sequences (these are the consumption value obtained through simulation). This may be thought of as corresponding to averaging the consumption for a time interval of ten minutes for a given substation for a period of five days. The optimal operation of the district heating system is therefore is to providing sufficiently high temperature and pressure to all customers by taking local measurement to achieve the results of system.

Table 4.1 : District heating flow Regulator

Time (s)	(KW)
0	67.0453
0.1000	88.4961
0.2000	100.7779
0.3000	136.5707
0.4000	209.8209
0.5000	220.0000
0.6000	220.0000
0.7000	220.0000
0.8000	220.0000
0.9000	220.0000
1.0000	220.0000
1.1000	220.0000
1.2000	220.0000
1.3000	220.0000
1.4000	220.0000
1.5000	220.0000
1.6000	220.0000
1.7000	220.0000
1.8000	220.0000
1.9000	220.0000
2.0000	220.0000
2.1000	220.0000
2.2000	220.0000
2.3000	220.0000
2.4000	220.0000
2.5000	220.0000
2.6000	211.3222
2.7000	201.3222
2.8000	191.3222
2.9000	181.3222
3.0000	171.3222
3.1000	161.3222
3.2000	151.3222
3.3000	141.3222
3.4000	131.3222
3.5000	121.3222
3.6000	111.3222

3.7000	101.3222
3.8000	91.3222
3.9000	81.3222
4.0000	71.3222
4.1000	61.3222
4.2000	51.3222
4.3000	41.3222
4.4000	31.3222
4.5000	21.3222
4.6000	11.3222
4.7000	1.3222
4.8000	-8.6778
4.9000	-18.6778
5.0000	-28.6778
5.1000	-38.6778
5.2000	-48.6778
5.3000	-58.6778
5.4000	-68.6778
5.5000	-78.6778
5.6000	-88.6778
5.7000	-98.6778

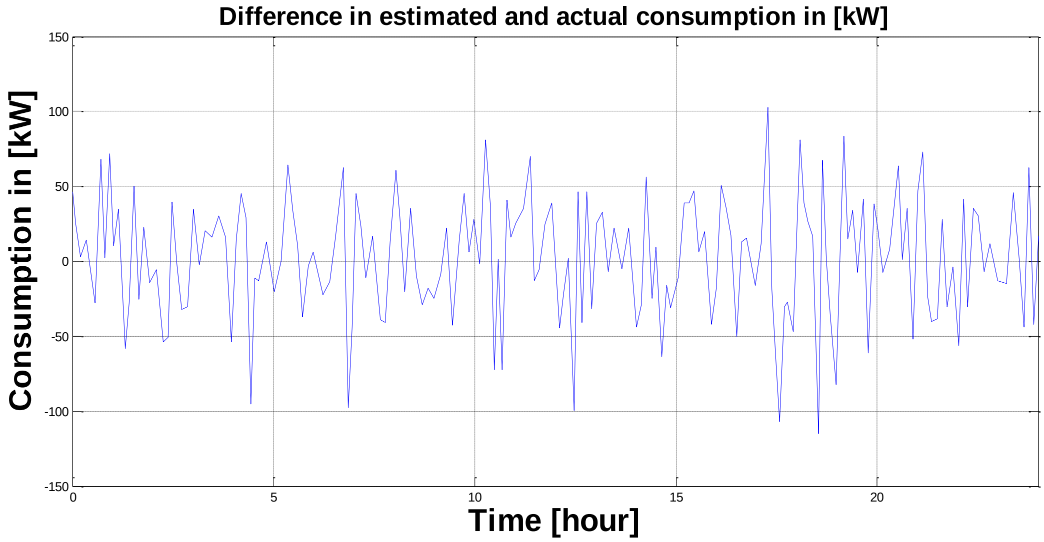

Figure 4.2: Example of the difference between two consumption sequences over 24 hours

This results in a discrepancy between "predicted" and "actual" consumption, which the redistribution agent needs to handle, as shown in figure 4.2. It can be seen that the difference between the predicted and actual consumption is high only at few points within the twenty four hour interval used for the experiment. For this experiment the simulated value was compared with the actual consumption. The difference here shows that the Q-learning approach used for the modeling gave a very good estimate of the expected consumption. The difference at scattered time interval is due to bursty consumption of end-users, but even at this the Q-level model is able to stabilize the difference because of its up to date interaction with the consumer agent. This leads to optimum update of the Q-values and good estimate of the consumption level. This result gives 50% improvement over the previous model that employs distributed modeling technique. In the previous researcher's work an estimate is given based on statistical estimation of the end-user consumption pattern over a certain period. The shortcoming is that the end-user consumption pattern cannot be efficiently modeled in this pattern due to unpredictable change in user's need as a result of climate change which may change quite often. The Q-learning is able to address this shortcoming as the Q-value of the end-user is updated at regular interval thereby giving an up to date consumption estimate.

4.2 Results and Discussions of Distribution between Quality of Service and Surplus Production.

Table 4.2 Available Utilities Quality of Service and Surplus Production.

TEMPERATURE			
UTILITY NAME	SUPPLY T_S[°C]	RETURN T_R[S]	ANNUAL COST [N/KW-%]
HEAT	70.0	70.0	3.5
HP-EX	50.0	50.0	2.5
HP-CYCLE	40.0	30.0	2.0
RECYCLE WATER	30.0	20.0	1.0
TIME	10.0	10.0	0.5

An approach towards coordination of just-in-time production and distribution between quality of service and surplus makes it possible to control the trade-off value between (measured in terms of the number of restrictions) and degree of surplus production of resources of quality-of-service. Compared to a reference control scheme, which is categorical in Figure 4.5a and 4.5b shows the total number of restrictions to tap water for household heating (radiator) during one day for different degrees of surplus production. It can be seen here that there is a clear trade-off between the quality of service (number of restrictions) and the total surplus in production are almost with no restrictions of any kind when 5% more hot water than the predicted consumption is produced for costumer. This is an improvement over the previous researcher distributed model in which a surplus production altered the equilibrium in the quality of service where some apartments had excess radiator water making the apartment warmer than expected. The Q-learning model was able to improve its adjustments to changing surplus production by 19% compared to the distributed model which is shown in figure 4.5a below

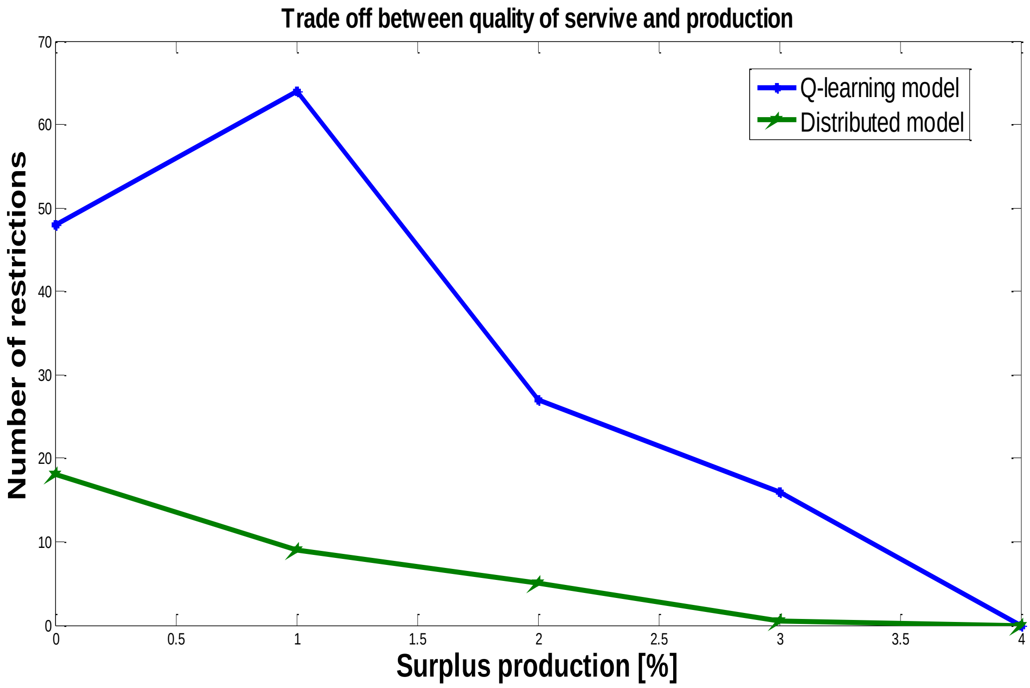

Figure 4.3 Surplus production [%]

In Figure 4.3 it was noticed that there is a reduction in the number of restrictions when the size of the cluster increases. Based on this experimental work there was no surplus in production. However, it should be noticed that the upper limit of the cluster size is determined by the structure of the district heating network. For instance if the distance between the various substations within a cluster is large in interval, the assumption of free and fast redistribution does not hold based on loss of product in the line. Trade-off between quality of service and surplus production. The y-axis corresponds the number of restrictions for the radiator water. The green line shows the variation to the actual consumption of tap water when using the distributed model while the blue line corresponds to the number of restrictions when using the Q-learning model.

Figure 4.4 shows the stand-alone variation in the quality of tap water supply to variation in the consumer requirements. The reference production level for this simulation was 240kW. it can be seen that the Q-learning model was able to adjust the production through the action of the redistribution agent without adversely affecting the supply to the consumer agents.

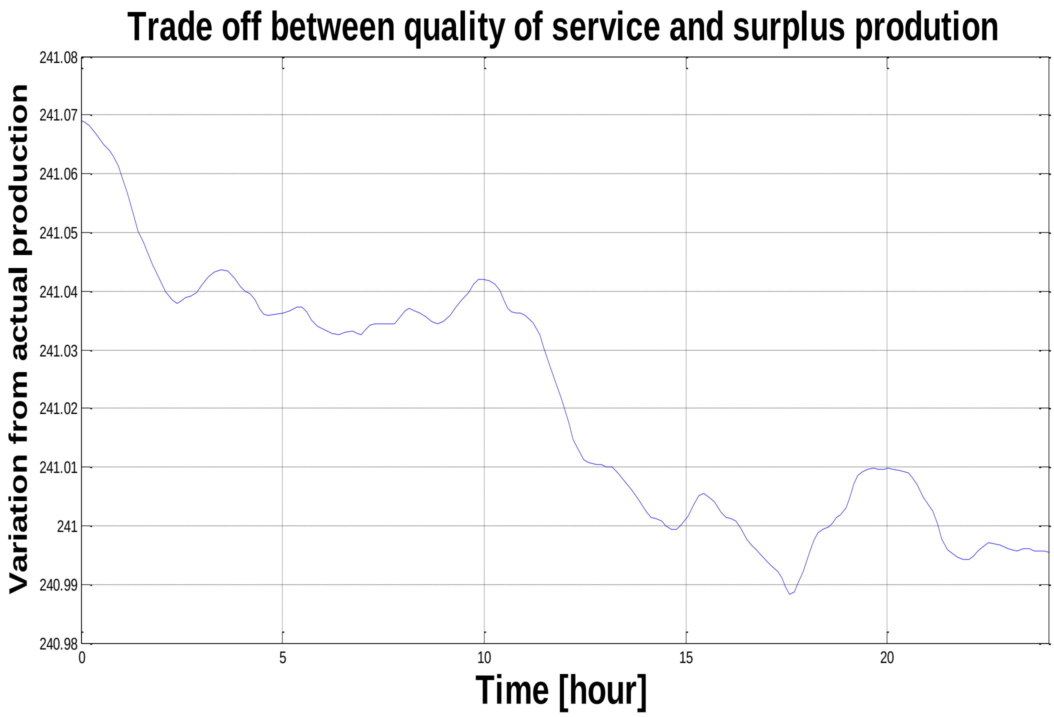

Figure 4.4 Stand alone variation in actual production from surplus production

Using the reference control scheme, 7% surplus production was needed to achieve this level of quality of service as opposed to 17% surplus production from stand alone work on distributed modeling. It should be noted that it is in current practice very difficult to find this minimum degree of surplus prediction. This is because, the operator of district heating system networks typically does not have exact information on neither the quality of service delivered to the customer, nor the amount of the actual degree of surplus production.

V. CONCLUSION

From applicability of Q-learning controller to model a system of Multi-Agent Systems (MAS) as a distribution control approach for District Heating Systems (DHS). The results have been compared to the conventional controller. The design of the Q-Learning controller has been explained and the performance was evaluated by simulation. The simulation results indicate that multi-agent system provides the best performance in comparison with DHS controller that encapsulate a functionality to solve a problem within the domain, in the short term the suggested project will adapt and introduce multi-agent technology in an industrial application were the use of leading edge information technology currently is very low. To the best of my knowledge, an application of Q-learning in agent technology has never been applied to monitoring and control of district heating systems. It will provide a novel significant combination and integration of existing technologies, which will open up new possibilities

VI. REFERENCES

[1]. Anaheim Parunak, et al (2014) "Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide", Multi-Agent Systems and Agent-Based Simulation, LNAI, Springer Verlag Vol. 1534,pp 10-25.

[2]. Arntzen, B.C. and Trafton, L.L., (2011) "Global Supply Chain Management at Digital Equipment Corporation", Interfaces, Vol. 25(1): pp 69-93.

[3]. Arvastsson, L. et al (2017) "Stochastic Modeling and Operational Optimization in District

Heating Systems", Lund Institute of Technology, Sweden, ISBN 91-628-4855-0.

[4]. Bellifemine, F. and Rimassa, G. (2017) "Developing multi-agent systems with a FIPA-compliant agent framework", Software: Practice and Experience, Vol. 31(2), John Wiley & Sons, Ltd, New York (2001) 103-128.

[5]. Bøhm, B., et al (2018), "Simple Models for Operational Optimization", Report S1, NOVEM, ISBN 9057480212.

[6]. Busoniu, R. Babuska, and B. de Schutter, (2008). "A comprehensive survey of multi agent reinforcement learning," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 38, no. 2, pp. 156–172, View at: Publisher Site Google Scholar

[7]. Clearwater, S.H., (2017.), "A paradigm for distributed resource allocation", Xerox PARC, ISBN 981-02-2254-8, 1995.

[8]. Canu, S., and Ding, X. (2010) "District Heating Forecast using Artificial Neural Networks", International Journal of Engineering, Vol. 2(4).

[9]. Cohen, M.A., and Lee, H.L. (2017) "Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods", Journal of Operation Research, Vol. 36(2): pp 216-228.

[10]. Durfee, E. and Lesser, V. (2015), "Partial Global Planning: A coordination framework for distributed hypothesis formation", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 21(5) pp.1167-1183, 1991.

[11]. Derksen, C and Unland, R. (2015) Fay, A. Structure and classification of unified energy agents as a base the systematic development of future energy grids. Eng. Appl. Artif. Intell., 41, 310–324.

[12]. Erman, D.L., et al (2015) "The HER- SAY-II speech understanding system: Integrating knowledge to resolve uncertainty", ACM Computing Survey, Vol. 12 pp.213-253, 1980.

[13]. Fan, M., and Whinston, A.B., (2018) "Decentralised Mechanism Design for Supply Chain Organisations Using an Auction Market", Addison-new deihi, ISBN 2-101-36749-9.

[14]. Ferber, J. (2015) "Multi-Agent Systems", Addison-Wesley, ISBN 0-201-36048-9.

[15]. Fox, M.S. and Teigen, R. (2016) "Agent-Oriented Supply- Chain Managemen1t", The International Journal of Flexible Manufacturing Systems, Vol. 12: 165-188, 2000.

[16]. Johansson, C. and Towards, Y (2010). Intelligent District Heating. Ph.D. Thesis, Blekinge Institute of Technology, Karlskrona, Sweden,

[17]. Jamal, B (2005). "A Pragmatic and Semantic Unified Framework for Agent Communication", a PhD thesis, Université Laval, Québec,

[18]. Wooldridge, M., (2018) "An Introduction to Multi-Agent Systems", John Wiley & Sons, ISBN 0 471-49691.

[19]. Wang, L. Liu, H. Zhang, and G. Xiao (2016)., "Fault-tolerant controller design for a class of nonlinear MIMO discrete-time systems via online reinforcement learning algorithm," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 5, pp. 611–622, View at: Publisher Site Google Scholar

[20]. Zhao, Z. Xia, and D. Wang, (2015). "Model-free optimal control for affine nonlinear systems with convergence analysis," IEEE Transactions on Automation Science and Engineering, vol. 12, no. 4, pp. 1461–1468, View at: Publisher Site | Google Scholar

Cite this article as :

Ehibe Prince, Uchegbu Chinenye E, "Optimization of Multi Agent System for Distribution Control of Distinct Heating System Using Improved Q-Learning Controller", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 9 Issue 1, pp. 235-249, January-February 2022. Available at

doi : https://doi.org/10.32628/IJSRSET1229152

Journal URL : https://ijsrset.com/IJSRSET1229152