GSJ: Volume 10, Issue 2, February 2022, Online: ISSN 2320-9186

www.globalscientificjournal.com

DESIGN OF A DISTRIBUTED HEATING SYSTEM USING Q-LEARNING

[1]FAGBOHUNMI, Griffin Siji [2]UCHEGBU Chinenye E.

[1]Computer Engineering Department
Abia State University, Uturu, Abia State, Nigeria.
fagbhume,griffn@abiastateuniversity.edu.ng

[2]Department of Electrical and Electronics Engineering
Abia State University, UturuAbia State Nigeria
ceuche@gmail.com

KeyWords

District heating system, demand-side management, reactive, Just in time, Q-learning, distributed heating system, software agent,

ABSTRACT

The aim of this paper is to design an energy efficient monitoring and control of district heating systems through the use of Q-learning agent technology. This research is of utmost importance because as the world grapples under the impact of COVID-19 pandemic and its attendant effects, it becomes pertinent for the stakeholders in the academia to come up with solution to various socio-economic events bedeviling human existence. A district heating system comprises of production units, a distribution network, and a host of consumer substations. The operations of district heating system usually involve conflicting goals, e.g., to satisfy customers, minimize production costs and maximize profit. Hence an intelligent 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 paper is to equip substations with software agents to form a multi-agent system using Q-learning. The purpose is to dynamically control the district heating system using demand-side-management strategies. Results from simulation studies indicate that the approach makes it possible to reduce production cost by 18% when compared to predictive heating system, while at the same time maintaining the quality of service. 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. Finallyan experiment was conducted in a controlled physical environment, where the Q-learning approach used in this paper was compared to two agent-based approaches. The experiment shows an improvement of 12% in automatically load balance of a small district heating network over predictive systems.

I INTRODUCTION

As the world grapples under the impact of COVID -19 pandemic and its attendant effects, it becomes pertinent that stakeholders in the academia comes up with solution to various socio-economic events bedeviling human existence. This is particularly important where it is necessary to regularly wash hands with warm water .One of this is in the area of having a more economical way of having access to hot water in various households. In Nigeria a lot of money is spent by various household for buying either kerosene or gas for boiling water on a daily basis.

This brings to the fore the need for an energy efficient district heating system which not only enable each substation make an informed decision on efficient control mechanism but also translates to a lot of savings by each household where instead of the huge daily expenditure on kerosene or gas, will only need to pay monthly stipend for a distributed heating system provided for by water corporation.

This paper mainly investigates the applicability of Multi-Agent Systems (MAS) as a distributed control approach for District Heating Systems (DHS). The consumers, i.e., the heat exchange systems, in current DHS employ purely reactive devices and have typically no communication capabilities. They are only able to make local decisions without taking into account the global situation in the system. 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 is the minimization of production cost to any variation in the consumption level of the substation.

The basic idea behind district heating is to use cheap local heat production plants to produce hot water (instead of using kerosene or gas to boil water). The water is then distributed by using pumps at approximately 1-3 m/s through pipes to the customers where it may be used for heating both tap water and the radiator water. The cooled water then returns to the production plant forming a closed system (see Fig 1

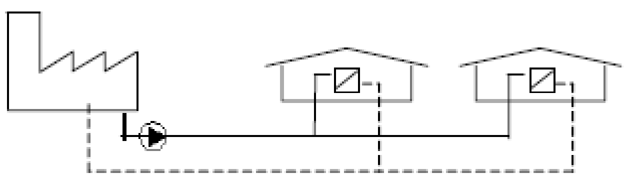

Fig 1. A simple district heating network containing one heat producer andtwo consumers

At the customer side, there is a substation (see Fig 2). It is normally composed of two or three heat exchangers and a control unit, which receives hot water from the district heating network. The substation heats both cold tap water and the water in the radiator circuit by exchanging the required heat indirectly from the primary flow of the distribution network. The hot water is returned to the network at a somewhat lower temperature. Both the temperature of the returning water and the flow rate in the network are dependent on the consumption of substations. When the water, returned by substations, arrives at the heat production plant it is heated and again pumped into the distribution network. Several different energy sources may be used for heating, e.g., waste energy, byproduct from industrial processes, geothermal reservoirs, otherwise combustion of fuels as oil, natural gas etc. is used. If the demand from the customers is high several heat producing units must be used. A district heating system in a large city can be very complex, containing thousands of substations and hundreds of kilometers of distribution pipes. In addition, they are dynamic as new substations may be added or old substations may be replaced by new ones with different characteristics.

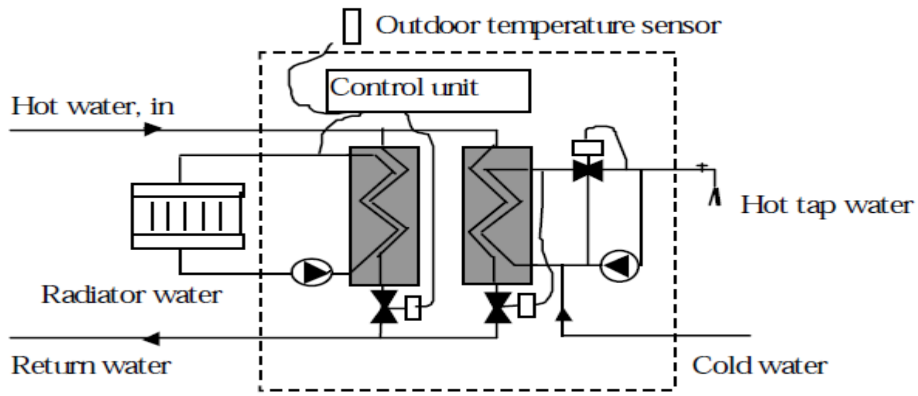

Fig 2. A substation consisting of heat exchangers control valves, pumps and a control unit.

A distributed control system usually refers to systems composed of interconnected components like sensors, actuators and controllers. The availability of small computational units has led to an increasing decentralization within automation systems and a distribution of functionality into geographically dispersed devices. However, resulting from this distribution of devices is an increasing amount of communication and increasing effort for the configuration of individual devices as well as of the complete system [1] The development of distributed and heterogeneous systems, such as software for automation and control, poses significant challenges for system developers. In general, the functions that can be automated in distributed systems are classified into two categories, monitoring functions and control functions. The properties to consider are comparable to the general properties of complex decentralized systems as described by [2]. Also, there is a strong requirement on flexibility and adaptability of the software as automation systems typically are subject to an ongoing partial modification, e.g., by introduction of new hardware. They are modeled as reactive systems, hence they do not have the ability to learn the long term characteristics of the environment. The most common model for distributed automation and control systems is the Supervisory Control AnD Acquisition (SCADA) model. The SCADA model is centralized by nature. From a central reading location, a master station monitors a number of remote sites (substations) equipped with remote telemetry units. The remote units measure various conditions and report the data back to the master station, which is carrying out the necessary analysis and control functions. While the SCADA model provides acceptable performance and reliability, experience has shown that this approach can lead to a lack of system fault tolerance, reconfigurability, extensibility, and adaptability [3]. Also, a general argument against centralized approaches to complex distributed problems is that when the problems are too extensive to be analyzed as a whole, solutions based on local approaches often allow them to be solved more quickly during run-time [4]. An alternative approach to deal with high complexity and the inadequacies of centralized approaches is to specialize and decentralize. The general idea is to divide and conquer, i.e., to partition the complex problem into a number of simpler sub-problems that can be solved in a distributed manner, and whose individual solution contribute to the solution of the original complex problem. In this way, a number of interacting decision-makers takes the place of a single centralized decision-maker. Recently many manufacturers have introduced advanced Intelligent Electronic Devices (IED) into their products that can perform functions such as parameter configuration and monitoring. The possibility of connecting these distributed IED in a Local Area Network (LAN) promise highly dynamic systems. However, the problem of providing a suitable framework for managing the connected devices remains. Also the IED are reactive and is not enabled to learn the long term characteristics of the system. There is a continuous search for new concepts and abstractions to facilitate the design and implementation of systems of this kind. One such concept is software agents [5].

A software agent can be defined as a natural extension of the concept of software objects, i.e. object oriented programming with added abstraction entities, i.e., objects that add persistent local states to the structured programming paradigm with the added functionality of interacting with the environment and responding to changes thereof. 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 an object, an

agent is able to act in such a way to maximize a given objective., e.g., by interacting with other agents, reading sensors, or sending commands to effectors, 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 [6]: “An agent can be defined as a computer system that is situated in some environment, and that is capable of autonomous action in its environment in order to meet its design objectives”. The complexity of agents depends on their tasks. Purely reactive agents only perform a mapping from sensor data (input) to effector signals (output). (Sensing refers to taking input from the environment and responding to changes in it, while effecting refers to changing the status of an actuator in order to give an output response to the input from the environment. In the most basic case, the behaviour of a reactive agent can be specified by a collection of independent situation-action rules. A more sophisticated approach is the subsumption architecture [7] 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 an explicitly represented model of the world that is used for decision making. The working of a deliberative agent can be described as a sense-model-deliberate-act cycle [8] . The sensors sense the environment and receive messages, which are used to update its model of the world. The model of the world is used by the deliberation module to decide which actions to take, this action serve as input to the effectors that carry out the actions. 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 [9], employ a single representation and control scheme for both reaction and deliberation, whereas layered agent architectures, such as InteRRaP [10], use different representations and algorithms (implemented in separate layers) to perform these functions.
A single agent may sometimes be operating usefully by itself. However, in most situations there can be an increase in functionality and productivity by letting several agents interact with each other forming a MAS. Such systems are often heterogeneous, i.e., they are composed of agents of different kinds that have different roles in the system. In any environment where software agents participate, the agents need to engage in cooperative and/or competitive tasks to effectively achieve their design objectives. The remainder of this work is organized asfollows, section 2 discusses the related works, section 3 describes the Q-learning protocol for the distributed multi-agent system. Section 4 shows the results and analysis of experiments conducted in a simulated environments while section 5 concludes the paper. Also areas for further research is included.

2 RELATED WORKS

District heating systems are inherently distributed both spatially and with respect to control. A customer, or more commonly, a set of customers, is represented by a substation embedded within the district heating network. In [11] the authors designed a system in which, the substation instantaneously attempts to satisfy the demands of its customers without considering the amount of available resources or the demands of other substations. Each substation was viewed as a "black-box" where local decisions are made in each substation without communication with the central heating station, thus the global situation is not taken into consideration. Thus, in such systems a district heating network is basically a collection of autonomous entities, which may result in a system that is only locally optimal. For instance, during a shortage in the network, resource allocation is unfair since consumers close to the production source will have sufficient amount of heat, while those distantly located will suffer. Secondly since the global information regarding the entire system is unknown, there will be unequal distribution of heat energy to the various substations. Another consequence of the way that these substations work is that the producers only have very limited information concerning the current state of the district heating system. In addition to this, the considerable long distribution time, results in the increase in the amount of heat to produce and deliver to the substations which is typically based on uninformed estimates of the future heat demand, made by the control engineer at the production plant. In order to ensure sufficient heat supply, According to researchers in {12], [13], their design of district heating system is based on a centralized control. In the design, a central heating system controls the overall consumption in the network. In order to deal with the problem of variation in heat supply needed for the different substations, the central heating system was made to produce more heat than was necessary. This approach has three setbacks, (i) The heat supply was not optimum leading to wastage of energy (ii) the required heat energy needed for the different substations was not ideally catered for and (iii) the discharge of the surplus heat leads to an increase in the greenhouse effect, a feature that should be discouraged.

In the work of [14], the authors designed ABSINTHE (Agent-based monitoring and control of district heating systems) that deal with the problem of unequal distribution of heat supply to all substations by defining an open substation architecture comprising of different modes. The main goal of ABSINTHE was to develop a decision support system for the district heating system operators that makes it possible to reduce the surplus production by increasing the knowledge about the current and future state of the system, this was achieved by equipping each substation with an agent that continually makes predictions of future consumption and monitors current consumption. The different modes are to implement restrictions on decisions to be taken under different scenarios that may lead to unequal heat distribution to all substations. However this method may not offer solutions to all environment induced distribution scenario. The software is also cumbersome, and may result in delayed decision making which may impact the performance of the system. There is a need for an intelligent agent system which is capable of interacting with the environment and gives decisions based on the outcome of these interactions. This will present a more robust solution to the problem than proffering solution to only specific sets of environment constraints.

The authors in [15] designed agents that makes a crude estimate of future consumption based mainly on experience and rules-of-thumb due to the difficulty in predicting future estimates of the heating requirements since many factors are unknown or uncertain. As a consequence, and in order to be sure to satisfy the consumers, district heating systems are typically run with large margins producing more heat than necessary, however this leads to energy wastage or shortage. Furthermore, operators are usually busy with keeping the heating plants running and the time available for making production decisions is therefore limited leading to erroneous prediction which may be counter=productive to the objective of the system

According to [16], the authors designed a multi-agent system architecture for Just-in-time production and distribution. The aim of the design is to produce the right amount of heat energy required for all substations and secondly that each substation should get the right amount of heat energy to cater for its various consumers. This is achieved by equipping each substations with an intelligent agent that makes predictions of future needs, which it sends to the production agent. Also, the substations are formed into clusters based on the consumption needs of its consumers. Substations having approximately the same heat energy consumption rate are made to belong to the same cluster. Each substations are then supplied energy via the means of a redistribution agent where the appropriate heat energy is supplied so as to cater for discrepancies in energy requirement for each substation. Their contribution was a design that would offer higher quality of service, i.e. better ways of dealing with shortages of heat energy from the various substations which is facilitated by the use of redistribution agent. The shortcoming here is the complex nature of the software which may lead to a longer time for decision making that may impact the performance of the system.
The aim of this paper is to proffer a design that provides up to date information regarding the consumptions of the various substations using light weight protocol that can mitigate the shortcomings earlier stated.

3. THE PROPOSED DISTRICT HEATING SYSTEM

The Q-learning technique was used in this paper to describe the core subsystems suitable to be modelled as software agents for the district heating system. In this technique an agent moves 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.

Unlike in other Q-learning environment where the model of the environment is not explicitly designed, in this research, the abstract concepts, e.g., roles and permissions, to be used during analysis to conceptualize the system, will be modelled using MATLAB toolbox.

Q-learning is a reinforcement learning method in which an agent interacts with the environment in order to maximize a cumulative reward. The agent learns through continuous interaction with the environment by going through the state spaces. In a dynamic sequential decision-making process, the *state* $S_t \in \mathcal{S}$ refers to a specific condition of the environment at discrete time steps t=0,1,…. The states in this research are the different substations. By interacting and responding to the environment, the agent chooses a deterministic or stochastic *action* $A_t \in$ S (selecting from a pool of consumption levels) that tries to maximize future returns (maximizing heat supplies to substations at all times) and receives an instant *reward* $R_{t+1} \in \mathcal{R}$ as the agent transits to the newstate S_{t+1}. The reward is usually represented by a quantitative measurement. A sequence of state, action and reward is generated to form an MDP (Fig. 3 [17].

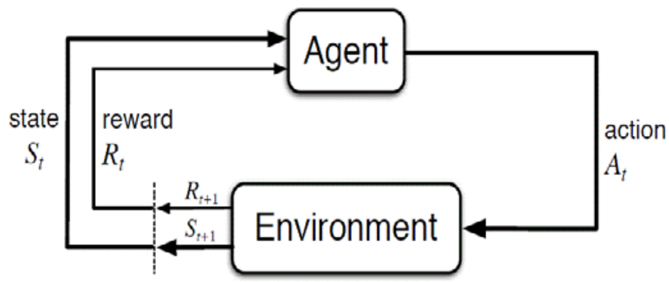

Fig. 3.The interaction between agent and environment in an MDP

The generalization of the Markov Decision Process to the multi-agent distributive heating system is the stochastic game. A stochastic game is a tuple (X,U_1, ……,U_n,) $Ƥ_1$, …..$Ƥ_n$) where *n* is the number of agents, (substations) *X* is the finite set of environment states, (i.e. different heat consumption) U_i, *i*= 1, . . . ,*n* are the finite sets of actions available to the agents, (updating the consumption level at each substation to reflect the current state of the substations, this yields the joint action set $U = U_1 \times \cdots \times U_n$, $f : X \times U \times X \rightarrow [0,1]$ which is the state transition probability function, Here the probability of changing the current status in each substation is between 0 and 1.and U*i*: *X* ×*U* ×*X* →R, *i*= 1, . . . ,*n* are the reward functions of the agents.

It is assumed here that the reward functions are bounded. In the multi-agent case, the state transitions are the result of the joint action of all the agents, $u_k = u_{1,k}^T, \ldots \ldots u, u_i, k \in U_i$

(where T denotes vector transpose). The policies *hi* :*XUi*→[0,1] form together the joint policy *h*. Because the rewards $r_{i,k+1}$ of the agents depend on the joint action, their returns depend on the joint policy as shown in equation 1:

$$R_i^h(x) = E\left(\sum_{k=0}^{\infty} \gamma^k r_{i,k} + 1 \;\middle|\; x_0 = x, h\right) \qquad (1)$$

The

The Q-function of each agent depends on the joint action and on the joint policy. This is depicted in equation 2

$$Q_i^h : X \times U \longrightarrow R, \text{ with } Q_i^h(x,u) = E\{\sum_{k=0}^{n} \gamma^k r_{i,k} + 1 | x_0 = x, u_0 = u, h\} \qquad (2)$$

In fully cooperative stochastic games, the reward functions are the same for all the agents: Ƿ1 = Ƿ2=---------Ƿn.

It follows that the returns are also the same, R_1^h = -------- = R_n^h and all the agents have the same goal: to maximize the common return. If *n* = 2 andǷ1 = - Ƿ2. This signifies that the two agents have opposing goals, and the stochastic game is fully competitive. Mixed games are stochastic games that are neither fully cooperative nor fully competitive. This paper deals with competitive stochastic game because the goals of the distributed system and the substations are opposing. The distributed system wants to minimize energy use in supplying hot water to the substations while the substations want to maximize the consumption needs of the customers. The flowchart for the Q-learning model for the distributed heating system is shown in fir 4.

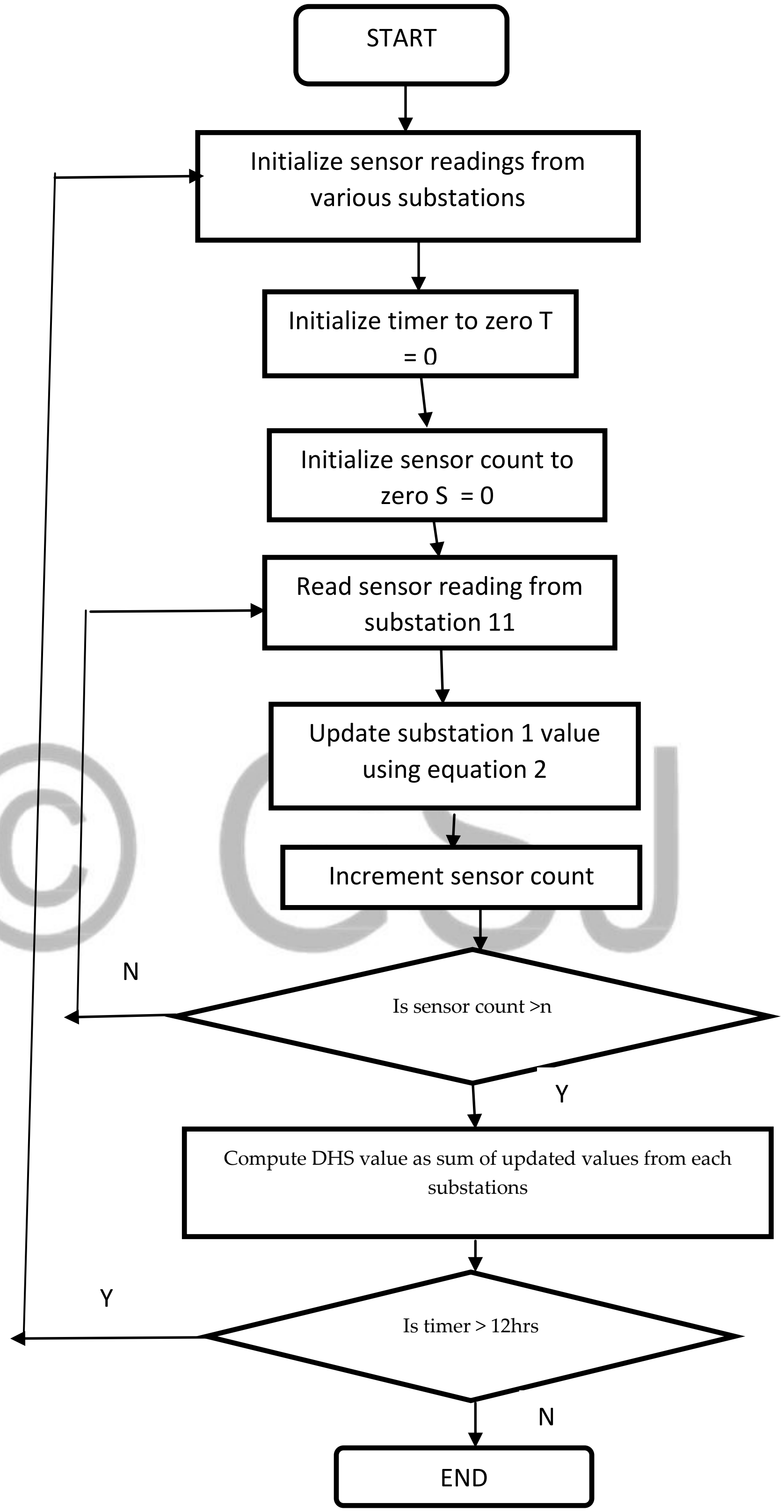

Fig4. Q-learning flowchart of the Distribbuted heating system

3.1 Multi-agent reinforcement learning using Q-learning

Generally, the transition probability for MARL extends equation. (3) below for a single agent Markov decision process to a multi-action case:

$$P(s', r \mid s, a) = P[S_t = s', R_t = r \mid S_{t-1} = s, A_{t-1} = a] \quad (3)$$

The multi-agent system Q-learning model is shown in equation 4

$$P(s', r \mid s, a_i) = P[S_t = s', R_{t,i} = \{r_1, \ldots\ldots r_n\} \mid S_{t-1} = s, A_{t-1,I} = a_1, \ldots\ldots, a_n\}] \quad (4)$$

where$n$ is the total number of agents, r_i is the reward for agent i and $\boldsymbol{s}=\{s_1,\ldots s_n\}$ is the set of individual states. In equation 4, 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 5

$$\sum_{a_i \ldots\ldots a_n} q, (s, a_i) \pi_i \left(a_i \mid s\right) \ldots\ldots \pi_n \left(a_n \mid s\right) \geq \sum_{a_i \ldots\ldots a_n} q, (s, a_i) \pi_i \left(a_i \mid s\right) \ldots\ldots \pi_n \left(a_n \mid s\right) \quad (5)$$

where$q(\boldsymbol{s},a_i)$ is the optimal action-value function for agent i and $\pi i(a_i \mid \boldsymbol{s})$ is the individual probability of taking action a_i given the Nash equilibrium policy.The advantage of the Q-learning approach is that there will be no need for redistribution agents as the Q-learning agent gets updated information from the various substations. The Q-learning protocol is shown in Fig 5

3.2 Q-learning Approach to Monitoring and Control of District Heating Systems

The purpose of this subsection is to improve the monitoring and control of district heating systems through the use of Q-learning agent technology. In order to increase the knowledge about the different states of the district heating system, the different consumer levels for each the various substations is stored by the Q-learning agent. This is opposed to the approximation of only the current and future state in a district heating system as devised by the previous researcher. The advantage here is that Q-learning can be used to store the various states (substation) of the district heating systems. This has the advantage of knowing different states, i.e. consumption levels of the various substations, thereby assigning Q-values to them. This will help the intelligent agent to better ascertain the optimum consumption level of the substation at all times. Additionally since the Q-learning continuously update the data from current interaction with the system, there will be no need for the redistribution agents thereby making the algorithm much simplified. With this technique the agent is able to distribute the excess or shortage in energy supply based on the actual consumption level from the various substations and not on an approximated values as proposed by the previous researcher. From Fig 5,Es = Ei + Δ. Where Es = energy supplied to substation, Ei = Q-value(energy) of ith substation and Δ = increment or decrement in energy based on proportional sharing of the change in energy from the Q-value of computed sum from the various substations.

In the Q-learning system at the producer side, each substation is equipped with an agent that stores all consumption levels of all substations. This makes the prediction of the consumption at any instant in time easy because of the higher number of state spaces present in the model as opposed to limited state space present in the previous researcher's model. The contributions to the consumers, will be higher quality of service, e.g., better ways to deal with major shortages or excess of heat water, which is facilitated by frequent interaction of the agent with the system (the consumption from the substations are updated every 12 hours) as opposed to the use of redistribution agents, used by the previous researcher. This leads to lower costs since less energy is needed for the heat production. The Q-learning protocol or monitoring and control is shown in Fig 5.

GSJ: Volume 10, Issue 2, February 2022
ISSN 2320-9186

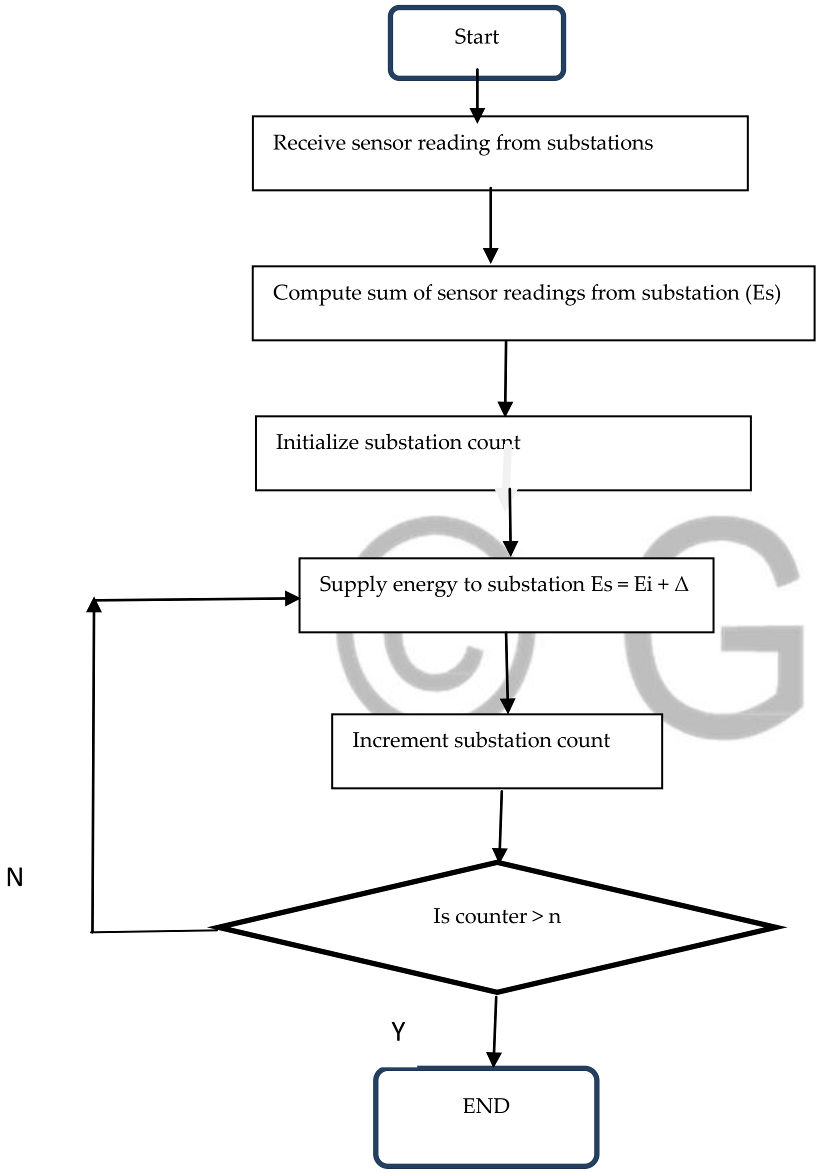

Fig 5. Flowchart for control and distribution of DHS

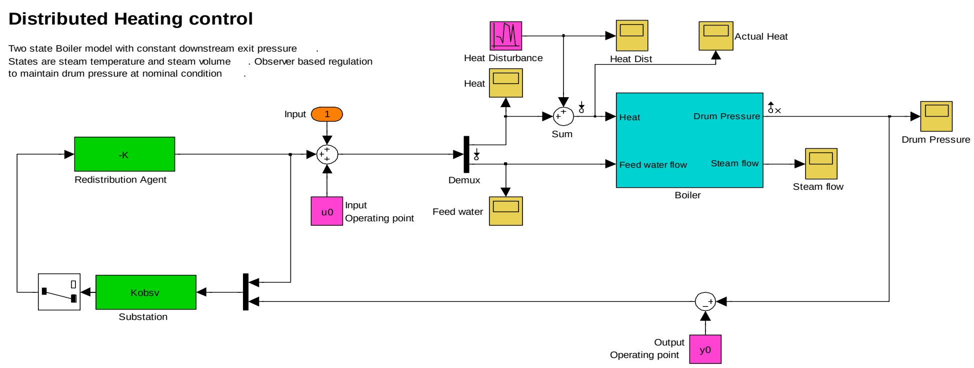

Fig6 Simulink model for multi agent distributed heating system

The Simulink model for the multi agent system is shown in Fig 6.

4 RESULTS AND ANALYSIS

In this section, the simulation results for the DHS will be given, the experiment design, and the results of the experiments will also be discussed. We have assumed that the distribution time from the producer to the consumers is 1 hour and that there is a single production source.

The MAS as well as the simulation environment was implemented in JADE (Java Agent DEvelopment framework) [18]. Thus, we used an agent-based approach that was time-driven, i.e., where the simulated time is advanced in constant time steps, to simulate the environment. Each simulated entity was implemented as a separate agent. Fig 7 shows the different parts of the simulation software.

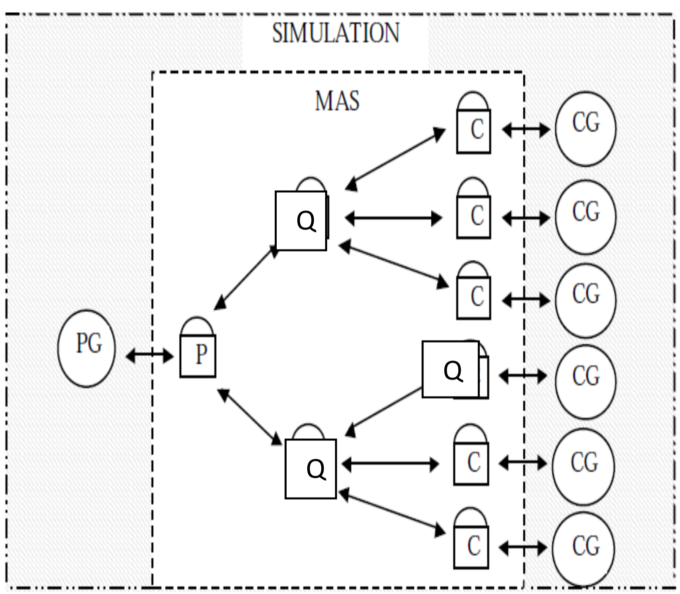

Fig 7. The MAS and Simulation software parts; Q = Q-learning agent, C= Consumer agent, P = Producer agent, CG = consumption generator and PG= Production generator

A simulator for hot water usage, is based on detailed field measurements performed in Abia State University (Mechanical Engineering Department lab), this 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 24 hours for a given substation for a period of five days. This results in a discrepancy between "predicted" and "actual" consumption, which the Q-learning agent needs to handle, as shown in Fig 8

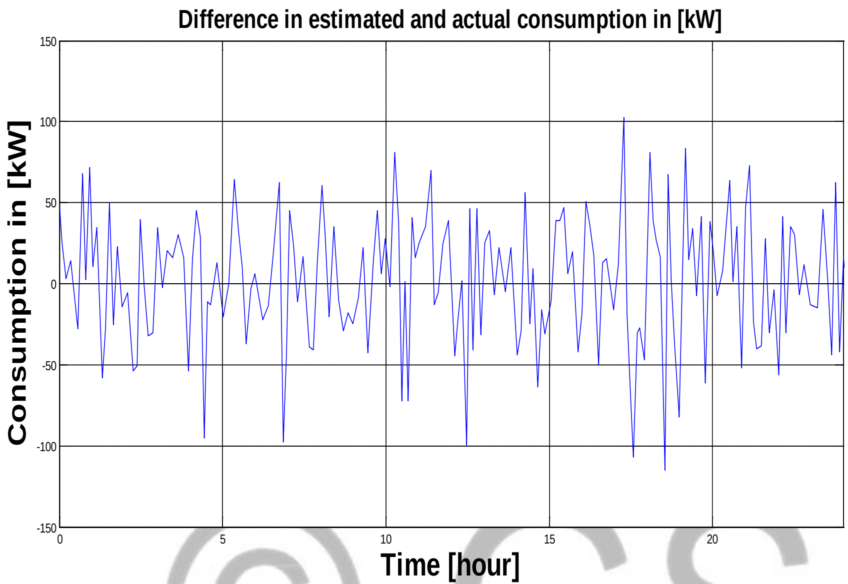

Fig 8: Example of the difference between two consumption sequences over 4 hours

In the first series of experiments, the Q-learning agent was managing five substations consisting of 10 consumer agents, where each of the five substations were serving a building with 5 offices and ten consumers. Experiments was run on different degrees of surplus production (from 0% surplus production, in steps of 1%, to 5%), where surplus production is defined with respect to the predicted consumption. For example, if the predicted total consumption is 200 kW and the surplus production is 2%, 204 kW is produced. The quality of service was measured in terms of the change from actual consumption to the consumption provided by the Q-learning agent.

The suggested approach was compared to the distributed multi agent system approach used the previous researcher and the production reference control scheme, which, is believed to be a very optimistic approximation of the current production strategy for district heating. Also here, different degrees of surplus production were tested. The result from this simulation is shown in Fig 9.
From the results it can be seen that despite variation in the consumption from the various substations, the multi-agent system using the Q-learning model was able to regulate the amount of heat production even with fluctuation in the reference production level.
This is due to the updated information from the redistribution agent embedded with Q-learning. On the other hand the distributed model was not able to adjust to changing reference

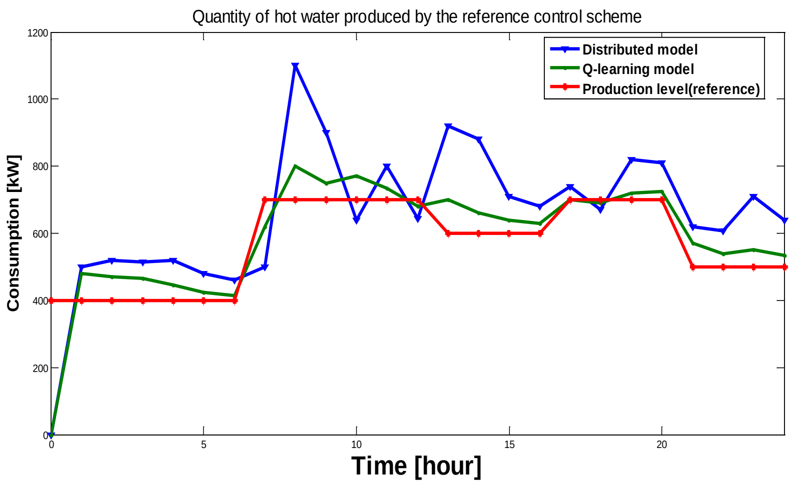

Fig 9: The amount of hot water produced by the reference control scheme indicated by the red line, indicating 0% surplus production, compared to the 5 day average consumption

Fig 9 shows the quantity of hot water produced and the consumption level as a result of the time lag necessary to compute new reference consumption level when it is varied. The idea here is to verify how the two compared techniques can adapt to changing consumption levels. The Q-learning was able to improve its adaptability to changing consumption level by 18% compared to the distributed model.

In the second series of experiments, the sizes of the clusters were varied in order to study how the cluster size affects the quality of service. The number of consumer agents in the cluster (substation) that was studied were 5, 10, and 15, where each consumer agent served a building with 5 offices. In this series of experiments there were no surplus production. The result is shown in Fig 10.

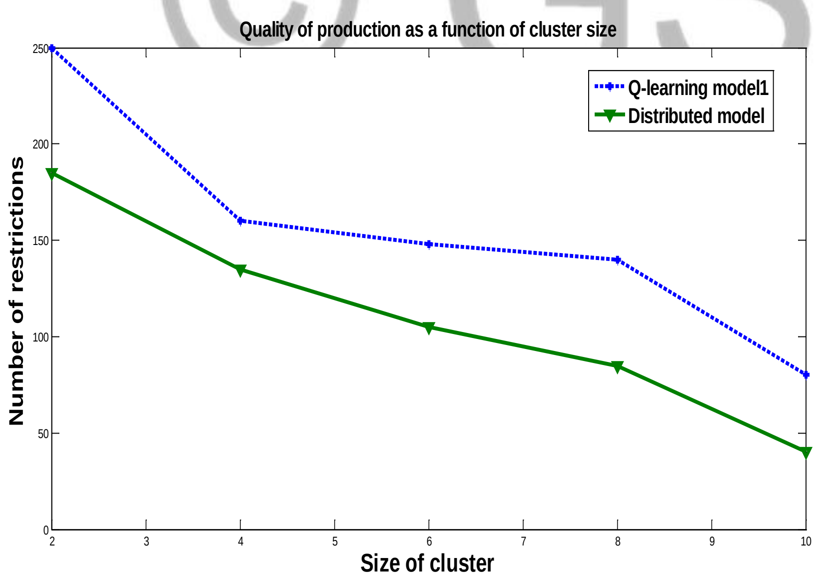

Fig 10. The quality of service as a function of cluster size

From the result in Fig 7, it was found out that the Q-learning multi-agent system performs well, coping with variation in the number of consumers, it can also be seen that the Q-learning model was able to adjust its supply as the number of consumer increases through appropriate update of the Q values. This is again due to the fact that the Q-learning agent was able to receive up to date consumer needs from the consumer agents and therefore able to adjusts the supply to minimize shortage to each consumer agent. The Q-learning model outperforms the distributed model by 19% over the variation in the sizes of the clusters.

Fig 11 shows the distribution of tap water and radiator (heating water) to the substations 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 (optimal supply) to the amount of surplus production .From Fig 11, it can be seen that there are almost no shortage to the various substations when 4% more hot water than the predicted consumption is produced for the Q-learning model. This is an improvement over the previous researcher distributed model in which a surplus production of same amount 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 49% compared to the distributed model.

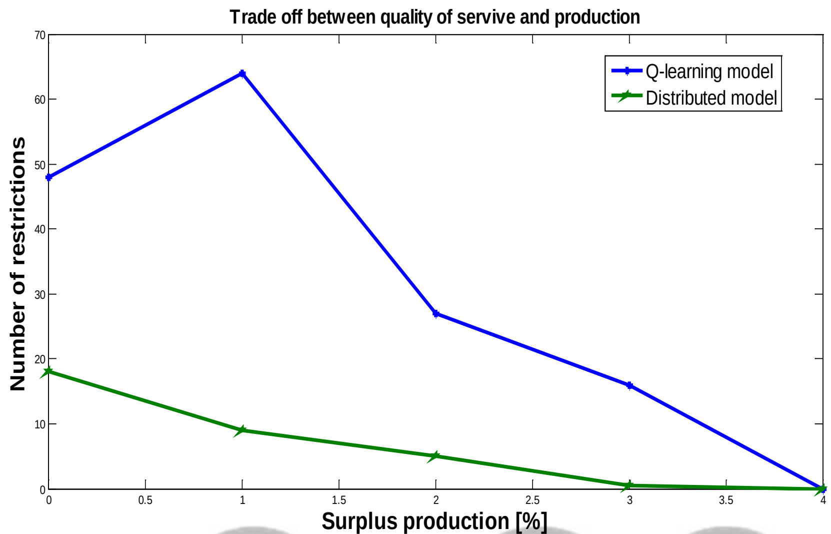

Fig 11. Surplus production [%]

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.

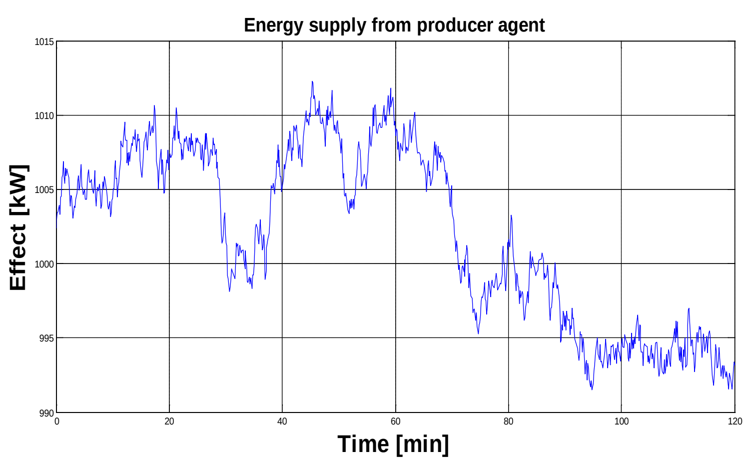

Fig12: Household heating and domestic hot water tapping from Q-learning model

4.2 Result and Analysis for Distributed Control of District Heating System

Consumption Scenario

In each experiment the scenario illustrated in Fig 12 was utilized. The substation is set to have a constant radiator demand. The Substation has a set value for radiator temperature of 48°C (approximately 25kW) the system is first allowed to reach a steady state during five minutes. After five minutes the substation initiates a domestic hot water tapping of 0,2 kg/s for a duration of two hours. The system is then *give*n ten minutes to stabilize.

From Fig 12 it can be seen that the Q-learning agent is able to stabilize the injection of additional heat by adjusting its put production accordingly within the two hours INTERVAL

4.3 Result and Analysis for different Control Strategies Employed

The control strategies used in the experiment are described in table 1.

Table 1 Control strategies

No restrictions	Substations are free to consume the amount requested
Q-learning restriction	Consumer agents individually enforce restriction based on the up to date consumption level of the end users
Local restrictions	The consumer agents enforce reductions individually, when the consumptions attains a particular limit.
Distributed / Hierarchy restrictions	The consumer agents enforce restrictions individually, when the consumptions attains a particular limit. This may require assistance from other consumer agent

Fig 13 shows the total energy consumption for the four different control strategies.

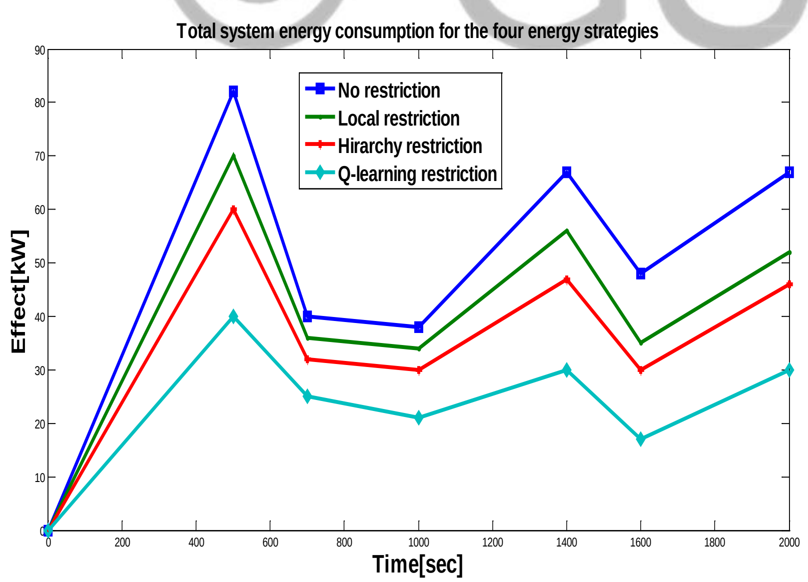

Fig 13 : Total system energy consumption for the four control strategies. Thedesired global consumption is 50 kW

The results from simulation experiments where different modes of restriction was applied to the distributed heating system is shown in Fig 13, it can be seen that the strategy that used local restrictions clearly reduces the consumption peaks by 6% and that the strategies to use hierarchical as well as distributed multi-agent based approaches reduces the peaks by 13%. However, the Q-learning based approaches requires slightly more time (6 seconds) to assume the stable level after reductions. It should be noted also that in tapping continuously for five minutes, the strategy with local restrictions are unable to reach the limit level of 50 kW. In Fig 14 it was shown that the primary flow of the energy consumption indicated was similar to that in Fig 13. The reason to study the flow is that it is important to keep the flow down for reasons of both production and potential flow limitations in the network.

Fig 14 shows the amount of time of the total experiment that the consumption has reached and been above effects from 45 kW and 80 kW, e.g., both agent systems had a consumption of 60kW and above during 15% of the duration of the experiment.

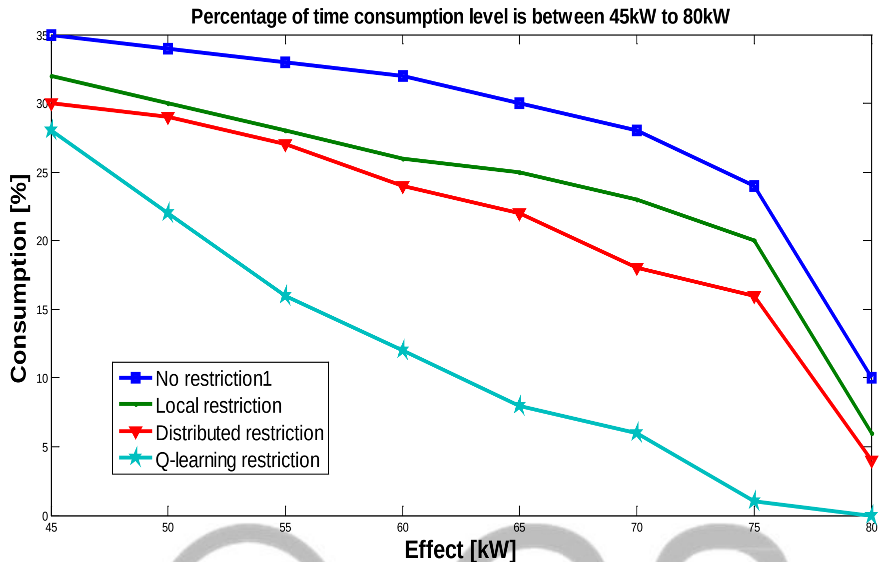

Fig 14 Amount of time of the experiment that the consumption has reached and been above effects from 45 kW and 80 kW

The Q-learning based approaches only consumes 25 kW or more during approximately 10% of the duration of the experiment. This is contrasted to 20% for the strategy with local restrictions and to 30% for no load balancing at all.

5 CONCLUSION

This paper demonstrates the application of reinforcement learning and in particular, Q-learning to distributed heating system. The approach leads to an energy efficient design as it doesn't require the need for redistribution agent. The protocol is also able to adjust the equilibrium of the system when there are variation in the consumption of the substations. The use of this technique enables the final users, i.e., the operators and the consumers, to have higher quality of service, e.g., faster feedback form the Q-learning agent to deal with major shortages of heat water, and lower costs, i.e., less energy is needed to produce the heat water. Since the heating of water often is associated with burning fuel (kerosene and gas), that pollutes the air in one way or another, the project obviously contributes to increase the quality of life for Nigerians. We also believe that with the introduction of advanced information and communication technology, this will help enhance the work situation for the network operator staff, e.g., through the new possibility for remote diagnosis of heat exchanger systems. Future work includes:

- Trying out other restrict ion policies than fairness, for example, based on priority.
- Improve the simulation environment by simulating the water flow.
- Testing the approach in actual field tests.
- Performing experiments in full-scale district heating systems.

References

[1] Brennan, R.W., and Norrie, D.H.(2018), "Evaluating the Relative Performance of Alternative Control Architectures for Manufacturing", in *Proceedings of the IEEE ISIC/CIRA/ISAS Joint Conference*, Gaithersburg.

[2] Rinaldo, J., and Ungar, L.(2017) "Auction-Driven Coordination for Plantwide Optimization", in *Proceedings of Foundations of Computer-Aided Process Operation, FOCAPO*.

[3] Ferber, J.(2018) "*Multi-Agent Systems*", Addison-Wesley, ISBN 0-201-36048-9.

[4] Weiss, G.,(2018) "*Multi-Agent Systems*", MIT Press, Cambridge, MA, ISBN 0-262-23203

Anwar, M.F., and Nagi, R., (2012)"Integration of Just-In-Time Production and Material Handling for an Assembly Environment", in *Proceedings of the 5th Industrial Engineering Research Conference*, Minneapolis.

[5] Wooldridge, M., Jennings, N.J., and Kinny, D., (2016) "The Gaia Methodology for Agent-Oriented Analysis and Design", *Journal of Autonomous Agents and Multi-Agent Systems*, Vol. 3(3): pp 285-312.

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

[7] Bøhm, B., Lucht, M., Park, Y., Sipilä, K., Ha, S., Won-tae, K., Bongkyun, K., Koljonen, T., Larsen, H., Wigbels, M., and Wistbacka, M.(2018), "Simple Models for Operational Optimization", *Report S1, NOVEM*, ISBN 9057480212.

[8] Davidsson, P., Johansson, S.J., Persson, J.A., and Wernstedt, F.,(2019) "Agent-based Approaches and Classical Optimization Techniques for Dynamic Distributed Resource Allocation" in *Proceedings of the Workshop on Representations and Approaches for Time-Critical Decentralized Resource/Role/Task Allocation at the Second International Joint Conference on Autonomous Agents & Multi-Agent Systems*, Melbourne, Australia.

[9] Tamminen, E., and Wistbacka, M., (2017)"Capacity and Cost Models for Thermal Power Systems with Random Outages of Plants", *VTT Energy, Espoo, Research Report ENE6/44/01*.

[10] Arvastsson, L.,(2017) "*Stochastic Modeling and Operational Optimization in District Heating Systems*", Lund Institute of Technology, Sweden, ISBN 91-628-4855-0.

[11] Bellifemine, F., Poggi, A., 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 (2017) pp103-128.

[12] Mirsky, M.J., (2012) "Direction of Optimization Technologies", *Presentation at the APICS International Conference*, New Orleans, [http://www.supply-chain-systems.com]

[13] Aringhieri, R., and Malucelli, F.,2019) "Optimal Operations Management and Network Planning of a District Heating System with a Combined Heat and Power Plant", *Annals of Operations Research*, Vol.120, pp.173-199.

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

[15] Lehtoranta, O., Seppälä, J., Koivisto, H., and Koivo, H., (2016) "Adaptive District Heat Load Forecasting using Neural Networks", in *Proceedings of Third International Symposium on Soft Computing for Industry*, Maui, USA.

[16] Fredrik Wernstedt and Paul Davidsson (2018) A Multi-Agent System Architecture for Coordination of Just-in-time Production and Distribution in proceedings of The Knowledge Engineering Review, Cambridge University Press, Volume 17, Issue 4, pp. 317-329,

[17] Zhou, Y and Gans, N.(2013). A Single-Server Queue with Markov Modulated Service Times". Financial Institutions Center, Wharton, UPenn. Retrieved from
http://fic.wharton.upenn.edu/fic/papers/99/p9940.html. Retrieved 2022-01-11

[18] Bellifemine, F., Poggi, A., and Rimassa, G., (2016) "Developing multi-agent systems with a FIPA-compliant agent framework", Software: Practice and Experience , Vol. 31(2), John Wiley & Sons, Ltd, New York (2016) pp 103-12