

A Cognitive Social IoT Approach for Smart Energy Management in a Real Environment

Claudio Marche, *Student Member, IEEE*, Gian Giuseppe Soma, *Member, IEEE*, and Michele Nitti, *Senior Member, IEEE*

Abstract—Energy usage inside buildings is a critical problem, especially considering high loads such as Heating, Ventilation and Air Conditioning (HVAC) systems: around 50% of the buildings’ energy demand resides in HVAC usage which causes a significant waste of energy resources due to improper uses. Usage awareness and efficient management have the potential to reduce related costs. However, strict saving policies may contrast with users’ comfort. In this sense, this paper proposes a multi-user multi-room smart energy management approach where a trade-off between the energy cost and the users’ thermal comfort is achieved. The proposed user-centric approach takes advantage of the novel paradigm of the Social Internet of Things to leverage a *social consciousness* and allow automated interactions between objects. Accordingly, the system automatically obtains the thermal profiles of both rooms and users. All these profiles are continuously updated based on the system experience and are then analysed through an optimization model to drive the selection of the most appropriate working times for HVACs. Experimental results in a real environment demonstrated the cognitive behaviour of the system which can adapt to users’ needs and ensure an acceptable comfort level while at the same time reducing energy costs compared to traditional usage.

Index Terms—Social IoT, Energy Management, User Comfort, Optimization, Genetic Algorithm, Real Environment.

I. INTRODUCTION

Sustainable development is a core principle of the Treaty on European Union and a priority objective for the Unions internal and external policies. Using energy efficiently works in this direction and also enables to save money and reduce carbon footprint. Consequently, the need for building management has received increasing attention for energy saving since it is responsible for a large portion of energy demand, e.g. 38.5% in the United States and 40% in the European Union [1]. According to the United Nations Environment Programme (UNEP), around 50% of this amount is due to HVAC (Heating, Ventilation and Air Conditioning) systems [2], due to the fact that improper use of appliances and devices causes a significant waste of energy resources. The optimal control of building environmental variables has significant impacts on environmental quality and building energy efficiency. In this scenario, Renewable Energy Systems (RESs) with optimal design and smart operation strategy are regarded as one of the most effective solutions [3]. The sheer quantity of data

produced by everyday objects brings forward several opportunities for the Internet of Things (IoT) paradigm to develop cost-effective and cognitive solutions in the area of building management by providing decision-support tools in order to aid users when utilizing electrical energy. Recently, the community has been extending its interest in these issues, and several works have been proposed with the aim of optimizing their use and reducing energy waste [4]. In general, most of the solutions offered by the literature try to find the best working times for appliances to minimize costs related to energy consumption. However, no necessary attention is given to the final users and their comfort, and so this gradually leads to disaffection and results in switching off the energy management systems.

Therefore, to have a real beneficial impact on the quality of peoples life and a shift from a “device-centric” to a “people-centric” model, the concept of Social IoT (SIoT) appeared, encompassing the idea of things involved in the network together with people [5]. This concept considers a network in which objects can establish social relations autonomously following the rules set by their owners.

In this context, this paper proposes a novel system for smart energy management based on a user-centric approach that takes advantage of the SIoT paradigm to augment the systems scalability and to leverage a social consciousness to allow automated interactions between objects. The paper combines the management of networks (Social IoT), systems (a real building), services (user thermal comfort) and applications (energy management) to offer a cognitive Smart HVAC solution. The system is capable of learning from experience by interacting with the world and to find a trade-off between the costs related to energy consumption, exploiting the energy produced by a RES, and the benefits in terms of users comfort, independently of where the physical system is deployed, since there is no need of any input parameters. The paper provides the following contributions:

- First, we develop a SIoT architecture for a smart building scenario which takes into account users comfort in a multi-room, multi-user environment, considering the presence of a RES to reduce energy costs for the users;
- Second, the system takes advantage of the SIoT paradigm in order to create users’ profiles and thermal characterizations of the considered building and so adapt to them the HVACs functioning;
- Third, we implement the whole system in a real scenario, i.e. three offices at the University of Cagliari, which usually follows a working time from 8.00 a.m. to 8:00

C. Marche, G.G. Soma and M. Nitti are with the Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Italy.

C. Marche and M. Nitti are with National Telecommunication Inter University Consortium, Research Unit of Cagliari, Italy.

Correspondence: C.M. claudio.marche@unica.it, G.G.S. gi-giuseppe.soma@unica.it, and M.N. michele.nitti@unica.it.

p.m. for a period of 7 months. Moreover, we develop a smartphone application to obtain the rooms' occupancy and to collect all the users' preferences.

In the following, Section II discusses the state of the art of smart HVAC systems, introduces the concept of the SIoT and illustrates an overview of the personal comfort models. Section III describes the real scenario, whereas Section IV depicts the system overview, the used architecture, and the implemented control algorithm. Finally, Section V presents the performance evaluation while Section VI draws final conclusions.

II. BACKGROUND

A. Energy Management Models

Developing innovative energy management methods in several scenarios, such as smart homes or smart buildings, is critical and has the potential to help save energy. In the literature, many algorithms and frameworks have been proposed for smart home scenarios, e.g., to eliminate the power consumption of unattended appliances and devices or to set the scheduling for any device [6]. Moreover, more complex scenarios are represented by smart buildings, in which automated processes and control have to consider a great number of devices and an increased amount of data [7]. In such a scenario, algorithms based on triggers and simple methods of control are no longer effective, and models based on forecasting and other complex techniques are required [8].

Among the works considering a general scenario of a smart home/building, a technique for energy efficiency is represented in [9]. The authors conduct experiments for efficient energy and water management through social networks and interacting technology. In these terms, a centralized architecture is depicted, in which interactive public displays, personalized dashboards, and notifications assume particular importance.

Moreover, several energy management models are illustrated for more complex scenarios, where one of them is represented by the smart building. Two well-known approaches are illustrated in [10] and [11]. In the first work, the authors illustrate a centralized approach to address the cost-optimal operation of smart buildings integrating a centralized HVAC system, photovoltaic generation and electrical storage devices. The model, based on a Model predictive control (MPC) strategy, proposes to reduce the energy consumption of the HVAC by the prediction of internal and external temperature. In the second work, the authors propose an intelligent energy management solution for smart buildings in an off-grid mode considering renewable energy systems, thermal and electrical equilibrium and lighting loads. An objective centralized function is introduced to optimize the electrical storage and energy cost, providing a compromise time interval for using home appliances. Other two smart building models based on mathematical optimisation are described in [12] and [13]. In the first paper, the authors propose a price-sensitive operational model for an HVAC system in a smart building scenario. A centralized controller is responsible for optimizing electricity usage by leveraging people's occupancy information and considering renewable energy and battery storage. In the second one, the authors depict a decentralized framework

TABLE I
ANALYSIS OF THE EXISTING ENERGY MANAGEMENT MODELS.

Ref	Approach	Technique	RES	Comfort	Users
[9]	Trigger	Centralized	✗	✗	✗
[10]	Linear	Centralized	✓	✓	✗
[11]	MPC	Centralized	✓	✗	✗
[12]	Optimization	Centralized	✓	✗	✗
[13]	Optimization	Decentralized	✓	✗	✓
[14]	DL	Centralized	✗	✗	✗
[15]	SAX	Centralized	✗	✗	✗
[16]	GOA, BFA	Centralized	✗	✗	✗
[17]	Trigger	Centralized	✗	✗	✗
Our	Genetic	Hybrid	✓	✓	✓

to manage the energy exchanges within a community of Smart Buildings. The proposed algorithm takes care of the power profiles of the users and, through a shared objective function, attempts to optimize consumption based on a set of renewable energy sources. Moreover, a complex centralized approach for smart buildings is depicted in [14], in which the authors introduce a DL and IoT-based approach to control the operation of a HVAC system in order to reduce energy consumption. The proposed system controls the operation of the air conditioners based on the presence of people in the specific area and sets the scheduling accordingly. Furthermore, another centralized approach is described in [15], where a method for discovering abnormal patterns in electricity loads and detecting anomalies in building energy management is depicted. Based on a time series representation method, i.e. the Symbolic Aggregate appRoXimation (SAX), the model is able to detect abnormal energy consumption in specific time windows of the day considering the building occupancy and systems operation schedule, e.g. for HVAC.

The last group of papers considers a case of intelligent building in a specific working environment, namely smart offices. Two recent works are illustrated in [16] and [17]. In the first work, the authors present a centralized technique for appliances' scheduling in an office. Based on two objective functions, i.e. the Grasshopper Optimization Algorithm (GOA) and Bacterial Foraging Algorithm (BFA), the proposed approach manages the power scheduling of several appliances in order to reduce energy consumption and pollution. In the second work, the authors illustrate a centralized lighting system that focuses on improving the control of lights in a smart office and propose to satisfy the goal of saving energy. The proposed approach controls the office lights based on several trigger signals, such as people's presence and command sent by users, and manage the scheduling accordingly.

However, even though such advanced energy management models depict acceptable results, to the best of our knowledge there are still some issues that they are failing to address. The first issue is related to the lack of RESs, which could improve consumption and so reduce cost and pollution. Furthermore, very few approaches consider the users' thermal comfort. The works mainly focus on energy consumption and overlook an important factor, i.e. the users' satisfaction: almost no one

of the analyzed papers evaluates and gives importance to the choices of the users, which represent the main actors of the whole system. The analysis of the thermal profiles of the users could help to adapt the system's behaviour to user needs and provide them with a trade-off between energy consumption and comfort. To sum it up, Table I summarizes the representative contributions described above and the classification of the analyzed works.

Indeed, this work aims at improving energy consumption in a smart building scenario by developing a “user-centric” solution, which includes the presence of a RES, i.e. a solar photovoltaic farm, and the users' thermal comfort.

B. SIoT Paradigm and Architecture

The idea of using social networking notions within the IoT to allow objects to autonomously establish social relationships is recently gaining fast popularity. The driving motivation is that a social-oriented approach is expected to support the discovery, selection, and composition of services and information provided by distributed objects and networks [18], [19]. In this paper, we refer to the SIoT model proposed in [5], where a set of forms of socialization among objects exists and the relationship between objects can be defined as:

- *Parental Object Relationship (POR)* among similar objects built in the same period by the same manufacturer and the production batch;
- *Co-location or Co-work Object Relationship (CLOR and CWOR)* like humans do when they share personal (e.g., cohabitation) or public (e.g., work) experiences;
- *Ownership Object Relationship (OOR)* for objects owned by the same user;
- *Social Object Relationship (SOR)* when objects come into contact, sporadically or continuously, due to relations among their owners (e.g., devices belonging to friends).

All these kinds of relationships are created and updated within the SIoT platform taking into account the object's features (i.e., brand, type, computational capabilities, characteristics, etc.) and social life, that is related to their meeting other objects [20].

In this work, the authors propose a cloud-based SIoT platform to implement the proposed system, based on a well-known architecture named Lysis [21]. The platform foresees a four-layer architecture, as depicted in Figure 1: Application, Aggregation, Virtualization and Real World. The lower layer is made up of the devices, which have the role to sense the physical environment and provide data to the higher layers. The Virtualization layer is made of Social Digital Twins (SDTs), which represent the digital counterparts of any entity of the real world enhanced with social capabilities, describing their characteristics and the services they are able to provide. The Micro Engine (ME), i.e. the main entity of the Aggregation layer, is a mash-up of one or more SDTs and other MEs, and it is responsible for getting and processing information from SDTs into high-level services requested by applications at the higher level. Finally, the Application layer is installed in the Cloud and partially in the devices, so that applications can be deployed and executed by exploiting one or more MEs.

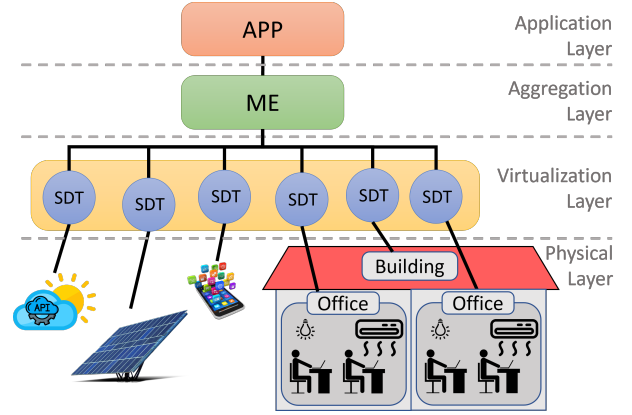


Fig. 1. SIoT Architecture.

To implement the proposed system, we had to extend the Lysis architecture: to virtualize the PV and the open weather services, we defined new SDT templates, while we implemented new MEs responsible for the predictions and the optimization. Moreover, new interfaces have been developed in order to allow communication between the platform and the new implemented SDTs and MEs. Below are provided examples of the main primitives used by the objects:

- *CREATE FRIENDSHIP $type_friendship$* : primitive used by SDTs to create a friendship with another SDT. This can be called automatically when a new object with the same owner is registered in the platform or sent manually by objects when specific conditions are satisfied, e.g. after a specific number of meetings at work for the CWOR.
- *SET SCHEDULING $scheduling_vector$* : called by the ME when the optimization is calculated. Then the ME directly contacts each SDT of the offices and sends them the new scheduling of the HVACs for the evaluated day.
- *NOTIFY USER PRESENCE*: as soon as a user's smartphone enters an office, each SDT will notify the ME in order to evaluate if a new optimization is required. Different users can have different preferred temperatures, and the ME considers all the users trying to provide adequate thermal comfort.
- *NOTIFY ERROR $variable$* : each SDT evaluates the proposed scheduling based on the predictions of several variables, such as the internal temperature or the power produced. If it detects errors or variations concerning the predictions, it notifies the ME, so a new optimization is provided.

C. User Comfort

Personal comfort models have the important purpose of analyzing and predicting thermal comfort requirements, increasing users' acceptability and associated energy benefits in both shared and single occupant built environments [22]. Thermal comfort refers to the satisfactory human perception of the thermal environment and describes the condition in which people feel comfortable. Several international standards, such as [23], adopt the PMV (Predicted Mean Vote) model, originally developed in the second half of the 1960s by

Fanger. The author proposes an index to represent the mean value of the thermal sensation votes of a group of people occupying a specific internal area. Based on several studies on the perception of healthy adults, the model calculates the thermal comfort between the body and the environment on a 7-point thermal sensation scale from -3 (cold) to 3 (hot). Furthermore, the PDD (Predicted Percentage of Dissatisfied) index, calculated as a function of the PMV index, quantifies the percentage of thermally dissatisfied people in an environment. The standards recommend a PMV index between -0.5 and 0.5, which corresponds to an optimal indoor temperature when PDD is lower than 10%.

Several works in the IoT scenario have been proposed to address the issue of thermal comfort by integrating the PMV-based model [24]. Among them, in [25] the authors revisit Fanger's thermal model using a Deep Learning (DL) algorithm. The proposed architecture collects all the comfort factors discussed in the PMV model from a set of IoT devices and evaluates a new score through a deep neural network. The proposed approach is then compared with the original model and tested under different machine learning algorithms. Another IoT application that makes use of the PMV and PDD indices is presented in [26]. The authors depict an architecture for live calculation of the indices and IoT monitoring. The environment is represented through thermal images, and real-time sensory data maintain a comfortable temperature. Two other approaches based on the PMV are described in [27] and [28]. In [27] the authors propose a study of the existing thermal PMV sensations and target different types of people's disabilities. The thermal comfort data are collected through an IoT architecture, and a revisited PMV model is illustrated starting from a prediction model based on a DL algorithm. Whereas, in [28] the authors propose an IoT smart system that can control the air conditioning to provide a thermally comfortable environment. The sensors' data are integrated with the PMV-based model, and the internal temperature is set accordingly in order to increase the users' satisfaction.

However, because of individual differences, one significant challenge is represented by providing a thermal environment that makes every occupant satisfied [29]. In these terms, the PMV index does not investigate how people perceive differently even if they are exposed to the same thermal environment, and no physiological factors or cultural differences are considered. All the analyzed models are designed and tested to work with the well-known PMV-based model, if anything, with customization for some category of users. The PMV score of an individual can be influenced by levels of physical activity, clothing insulation, as well as the parameters of the thermal environment. However, thermal comfort owns several personal characteristics, and so it is also influenced by user preferences. To address the issue of individual differences, we propose an improved PMV factor, in which each user can adapt the score based on his choices, and the relations between all the parameters described in the PMV model are preserved.

III. REAL SCENARIO

This paper proposes a real system for energy management in a smart building scenario. The innovative part stands in

involving the rooms, and the corresponding HVAC systems, the people working in those rooms to collaborate through social relations among their digital counterparts in order to minimize the energy cost without discarding the users' comfort. Nowadays, almost every aspect of the real world can be observed through the use of devices, so that people's behaviour can be inferred by their smartphones, while data of interest in a room, such as the internal temperature or the energy consumption of an appliance, can be monitored through a control unit. Let us call the set of devices in the smart building as $\mathcal{N} = \{n_1, \dots, n_i, \dots, n_M\}$, where the smart building itself can be seen as a device and thus identified with n_1 . Every device can sense and collect information regarding the environment even if not directly interested in using the data; through the Internet, this information is made available to potentially every other object. Among all the devices in the smart building, we can identify two subsets: the set of rooms, i.e. the set of devices representing the functionalities of a room, that can be expressed as $\mathcal{R} \subseteq \mathcal{N}$ and the set of users' smartphone, represented as $\mathcal{U} \subseteq \mathcal{N}$.

In our scenario, we are considering several parameters to be monitored, from the energy, both consumed and produced, to the thermal comfort of the workers in the offices, their occupancy in the rooms and the temperature, both internal and external to the building. All the rooms are connected through OOR relations among themselves and with the building as well as with the RES to obtain the produced energy since they all belong to the same owner, i.e. the owner of the smart building. Each room is tied to the relevant HVAC inside through a CLOR, while workers' smartphones create CWOR relations with the room they work in.

As depicted in Figure 1, each device is associated with a virtual counterpart (i.e. SDT) on the SIoT platform. All the information sensed by the devices, such as temperature or users' presence, are sent to their virtual counterpart, so that if an object needs them, it can query its social network and access its friends' data. This means that information among friends is exchanged at the Virtualization Layer, i.e. among Web Services used to implement the SDTs on the Cloud, which could also be running on the same physical machine.

The whole process starts whenever an application installed on a smart building needs to optimize its energy consumption. Figure 2 illustrates the steps of the overall process. The proposed system makes use of the friends of the building n_1 to assist the building in minimizing energy consumption while maximizing the users' comfort. This is achieved through an optimal scheduling for the HVAC units; to this, the SDT of the building crawls its social network to obtain the profiles of all the entities, which include the open-data regarding external temperature and its forecast and of the historical information stored in the SDTs n_1 's friends. All this information is used at the Aggregation level so that it is possible to make predictions regarding the SDTs' behaviours. To this, a ME is created for each of the entities depicted in Figure 2: the power absorbed from different HVAC based on the functioning mode and their model, as a building can have installed different HVAC at different times; the temperature preferences of the users in each room depending on the time of the year and the related

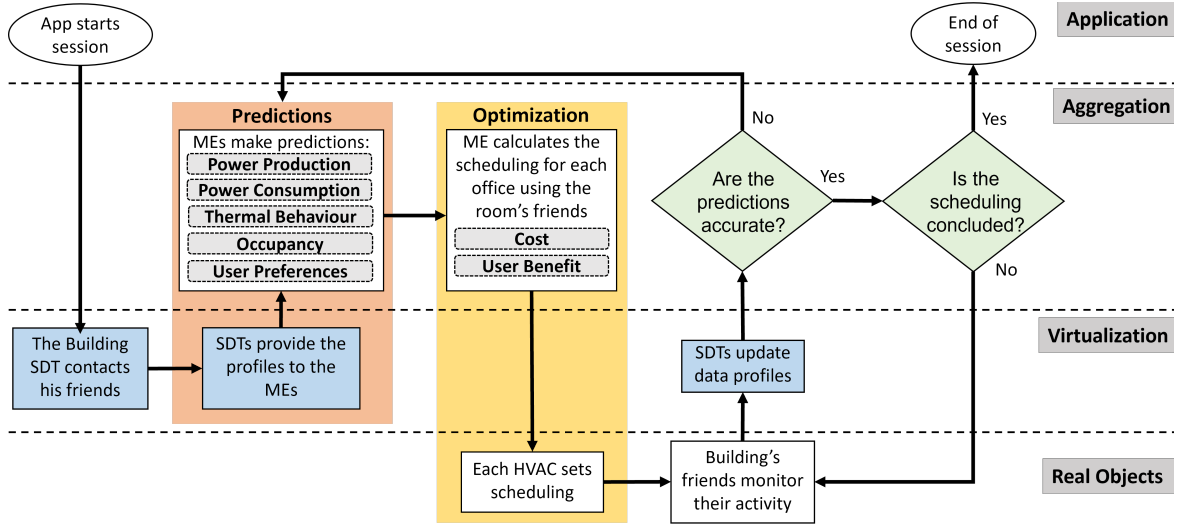


Fig. 2. Flow chart of the system.

temperature; the dynamic thermal behaviour of the offices; the room occupancy based on employees' usual working hours (arrival and departure); the power production of the RES based on the external forecast weather. All these predictions are learned, computed and updated autonomously based on the information collected during the normal functioning of the system: they are described in greater detail in Section IV-A.

These predictions are given as input to the ME in charge of finding the optimal scheduling for each office, which will weigh the cost related to the electricity consumption of the building and the average user benefit during the day, expressed as thermal comfort, through a genetic algorithm as described in Section IV-B. Once the scheduling has been set for each HVAC, it is sent to the corresponding real objects for its implementation. During the day, all the SDTs in the building verify that their own status is coherent with MEs' predictions. If all the predictions are accurate, the scheduling can continue till the end of the day. However, if one of the predictions fails, the related ME is updated with the new data and a new prediction is made. In this case, the optimization ME updates the scheduling to take into account the variation and communicate the new scheduling to all the HVAC. This is done because a wrong prediction on an office, e.g. related to its occupancy, can change the scheduling of the entire building.

IV. SYSTEM DESIGN

This section proposes the system design and illustrates the functioning of the aggregation layer, the core of the energy management process. The first aspect revises the predictions that are based on the information collected during the normal functioning of the system, while the second one evaluates the optimization algorithm that computes the best scheduling for each HVAC based on a genetic algorithm.

A. Predictions

The optimization control algorithm relies on several prediction problems. In particular, the scheduling of each HVAC is

formulated for the entire day and based on the predictions of the upcoming variables.

1) *Power Production:* The first prediction corresponds to the power production of the RES, the photovoltaic (PV) farm, based on the external forecast weather. In general, solar irradiation and the energy produced by PV systems should be measured precisely over a long period, with complicated and expensive measures. For this reason, various works have been proposed in the literature to simplify them by introducing different ways to predict, such as stochastic models. In a recent study, described in [30], authors design and install a PV in order to implement some prediction algorithms. Many methods are used for this purpose, such as linear and non-linear regression models and artificial neural networks. A comparison between the different modelling techniques shows that regression techniques are slightly less accurate. Nevertheless, the algorithm's complexity is low and fast in terms of time consumption. According to these studies, we have selected a regression model for our prediction algorithms, which do not represent our project's main focus. The approach proposed in this prediction consists of a multiple linear regression [31], in which the predicted value y is calculated based on the following regression equation:

$$y = \sum_i \beta_i x_i + \epsilon_i \quad (1)$$

where β_i depicts the regression coefficient for the feature x_i and ϵ_i represents the error term. After every working day, the prediction model is updated and trained so that it can learn from its past experience and provide a more accurate evaluation. In specific, the algorithm is trained by three features, i.e. the solar irradiance, the external temperature and humidity, and the regression coefficients β_i^{RES} for the power production prediction are obtained in order to satisfy the following equation:

$$\arg \min_{\beta \in \mathbb{R}} \sum_n^N (P_n^{RES} - \sum_i^3 x_{i,n}^{RES} \beta_i^{RES})^2 \quad (2)$$

where N represents the number of historical data used to train the regression model, and P_n^{RES} refers to the power production historical data.

The features considered for the power production prediction are related to the weather forecasts, which is not the focus of the paper. To this, we rely on open-data¹, which allows us to save resources and make the system faster.

2) *Power Consumption*: The second prediction is represented by the power absorbed from all the different HVACs in the smart building; this consists of an essential factor for energy efficiency optimization. Each room owns an HVAC system as described in Section III. However, two leading technologies concern them: the inverter technology and the ON-OFF one. Both systems offer similar functions but differ in terms of what type of compressor motor is running. The traditional non-inverter air conditioner switches the compressor motor from ON to OFF and vice versa. Whereas the inverter Air Conditioner (AC) compressor is always in the ON state: after reaching the desired temperature, it runs the motor with the minimal speed in order to maintain the room at the desired temperature. With a focus on power consumption, the inverter needs more power at the start but guarantees less consumption in the long term. Therefore, the two technologies correspond to two different load profiles. The profile of the non-invert AC derives from the specifications supplied by the vendor. So when the HVAC is in the ON state, the system considers a constant power consumption; differently, the consumption is zero. However, the load profile of inverter AC needs more detailed studies [32]. As for the power production prediction, our approach consists of a regression model, the Lasso [33], which correlates the energy consumption with several features. Developed from historical performance data, this model permits the creation of an energy usage profile thanks to five features, i.e. the room temperature, the humidity and the configuration parameters of the HVAC (temperature, mode and fan), and generates the power consumed P_r^{HVAC} in the room r by the HVAC. The system is continuously trained during the days thus allowing it to improve prediction and adapt to HVAC system degradation. As illustrated for the power produced, the power consumption is evaluated with the regression equation illustrated in (1), in which the Lasso regression coefficients β_i^{HVAC} are calculated in order to achieve the following equation:

$$\arg \min_{\beta \in \mathbb{R}} \sum_n^N (P_n^{HVAC} - \sum_i^5 x_{i,n}^{HVAC} \beta_i^{HVAC})^2 + \lambda \sum_i^5 |\beta_i^{HVAC}| \quad (3)$$

where λ represents the tuning parameter set for the Lasso regression.

3) *Thermal Behaviour*: A further prediction is depicted by the dynamic thermal behaviour of the offices, essential for planning the HVACs' scheduling. Many studies have tried to explain how indoor conditions are related to outdoor ambient weather. Among the works are proposed several techniques, such as [34] and [35], that make use of machine learning techniques in order to disregard the building thermal parameters. In

the first work, the authors examine how indoor conditions are strictly related to outdoor weather. The correlation is examined through a linear regression model, with several historical features such as indoor and outdoor temperatures. Instead, in the second paper, the authors design an artificial neural network for representing the room's thermal behaviour, using as inputs only the measures of outdoor air relative humidity, indoor air temperature and relative humidity at the past 12 hours. Our approach, as for the two previous predictions, is based on a regression model, namely Ridge [36], that permits reducing computational time and complexity. Each new day, the model is trained through the past 15 days and makes use of six features, i.e. the internal temperature and humidity, the configuration parameters of the HVAC (temperature, mode and fan), and the number of people inside the office, to generate T^{in} , that represents the value of indoor temperature. In specific, the prediction of the internal temperature is evaluated with the regression equation illustrated in (1), and the Ridge regression coefficients β_i^{in} are calculated in order to achieve the following equation:

$$\arg \min_{\beta \in \mathbb{R}} \sum_n^N (T_n^{in} - \sum_i^6 x_{i,n}^{in} \beta_i^{in})^2 + \lambda' \sum_i^6 (\beta_i^{in})^2 \quad (4)$$

where λ' represents the tuning parameter set for the regression.

4) *Occupancy*: One of the most important features of thermal behaviour prediction is depicted by the presence of people. In recent years many studies have been proposed to represent and predict the occupants' behaviour and their presence in indoor environments [37]. A common classification considers the representation of the occupancy data in two different ways: binary vector and probability distribution. For the first class, in [38] authors create a presence dataset based on historical data, in which each occupant is associated with a binary vector for a whole day. Its elements represent the occupancy status per minute: 1 means that the user is inside the room, instead 0 if she is outside. The prediction process consists of a supervised machine learning algorithm, namely the k-nearest neighbour, that considers as outputs the time of the occupancy and the total duration of presence. The second class is illustrated in [39], in which the occupancy model is scheduled in the format of daily probability distributions. Based on the predictor attributes, such as the day of the week, a decision tree predicts the occupancy state of an office, vacant or occupied. However, even though such techniques depict acceptable results, these exhibit several gaps. The first approach needs a large amount of data to create the dataset and it gives the same weight to all days, failing to distinguish the week-end and other vacations. Moreover, the second work does not evaluate different users, considering only the presence of a generic occupant in an office. The approach introduced in this paper tries to overcome these gaps by providing historical data of occupancy over the two weeks preceding and evaluating each user autonomously. In specific, for each time interval t , a value of occupancy O_u is associated with each user u as follows:

$$O_{u,t} = \frac{1}{D} \sum_d O_{u,t}^d \cdot W_d \quad (5)$$

¹<https://api.weatherbit.io/>

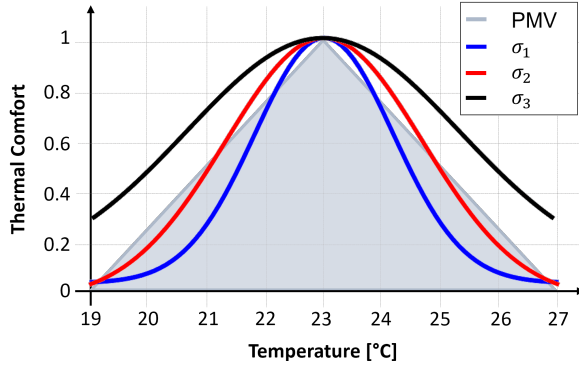


Fig. 3. PMV and adaptive model curve.

where the weights W_d are selected to give more importance to a specific day of the week and D and $O_{u,t}^d$ represent the total number of analyzed days and the value of occupancy in the specific day d respectively.

5) *User Preferences*: The user comfort relies on many factors, based on the building, thermal parameters such as capacity or resistance, and on the users, such as the metabolic rate or their worn clothes [40]. As explained in Section II-C the majority of works are developed in accordance with Fanger’s model, standardized in EN ISO 7730 [41]. The model represents thermal comfort in terms of PMV, which is determined by the heat balance of the human being with his/her environment. The coefficient is based on many parameters and is unsuitable for the users’ diversity. Several research topics tried to overcome this issue by implying personalized conditioning systems and the influence of personal factors, such as age or thermal history. However, these models have higher computational complexity and are not usable for real-time applications. We propose an improved PMV factor in which each user can adapt the score based on her choices, preserving the relations between all the parameters described in the PMV model. The proposed function, that it will be explained in detail in the next section, is obtained from the analysis of the PMV model function under specific conditions. In specific, Figure 3 illustrates the comparison between the two functions. The PMV model curve is depicted considering constant parameters, such as relative air velocity and value of clothes insulation, and it is normalized in the interval $[0, 1]$. The best thermal comfort value corresponds to a temperature of 23°C ; however, the value can not be changed based on user preferences. To solve the issue, a thermal comfort curve is represented in order to follow user preferences [42]. In specific, the temperature set by the user is used as the mean value of the Gaussian, and different values of variance are based on ON-OFF interactions by the users with the AC, i.e. the initial Gaussian curve is set with $\sigma_2 = 2.8$, while command ON or OFF determine the width: the command ON corresponds to the highest value of variance $\sigma_3 = 5.6$ and OFF to the lowest $\sigma_1 = 1.4$. If the users’ thermal comfort falls below 0.5, this is identified as an error in predictions, and so a new optimisation process is started.

The system communicates with the users, and vice versa,

through an app for Android and IOS systems and so adapts the thermal comfort to their habits, creating a profile for each user. The app allows the user u to change his/her desired temperature T_u^{pref} , switch off and on the HVAC for the related office and modify other parameters, such as the swing or the fan. These commands are used to set the thermal comfort model, as we will explain in the next section, and to collect and save all the preferences and so used for the optimization function. Moreover, the application is used to detect the presence of users in the rooms.

B. Genetic Algorithm Application: Optimal HVAC Scheduling

The proposed approach, as mentioned in the previous sections, includes an optimization algorithm that computes the best scheduling for each HVAC based on a Genetic Algorithm (GA). The GA is adopted to overcome the problems of classical optimization methods, such as nonlinear programming, that have limitations in searching for absolutely optimal solutions and are sometimes trapped in local minimums. The GAs fundamentals were first introduced by John Holland [43]. The GA is a global search technique that has arisen with inspiration from Darwins natural principle. This method, which can be utilized to solve optimization problems, is based on the theory of natural selection and the biological evolution process. Fundamentally, the GA creates a population that develops through time using reproduction and mutation methods. The GA general idea can be summarized with the following steps:

- Objective function (OF) definition, according to the specific problem under investigation;
- Definition of GA suitable coding, which consists in the characterization of a string useful to operate with the genetic operators. The decision variables in the GA are usually encoded into a binary string as a set of genes corresponding to chromosomes in biological systems. A group of chromosomes is called “population”. In the proposed approach, each solution (referred to as a single HVAC/room) can be coded by using a vector V . A binary coding is sufficient to define the daily pattern for the HVAC utilization (1 for the ON state and 0 for the OFF state of an HVAC). In other words, for the HVAC operation has been adopted an electric model that assumes a power absorption equal to the predicted value when the unit is on. In fact, some HVAC units do not include high-efficiency devices (e.g. HVAC inverter equipped) with variable power according to the temperature variation requested but a simple on/off model has been assumed;
- GA operators definition;
- Set of the initial population: the creation of a suitable initial population is important to have a good evolution of population during the different generation steps;
- By applying genetic operators on selected people, the children population is produced.

Each individual in the population is evaluated by the fitness function, linked to the problem OF; the adopted fitness value is defined as a ratio between cost and benefit due to the HVAC

pattern definition. In particular, the fitness function is obtained according to the following equation:

$$FF_t = -\log\left(\frac{c_t}{b_t}\right) \quad (6)$$

where c_t and b_t are the cost and benefit, respectively. More details about their calculation will be provided in the next subsection. The evolution starts from a random population.

Only individuals with the best OF/fitness value are selected, using a random procedure, to form a mating pool set, that each chromosome has a different probability to survive, correlated with the OF/fitness value. The selection method adopted is the roulette wheel selection. In the roulette wheel selection, the probability of choosing an individual for breeding of the next generation is proportional to its fitness; the better the fitness is, the higher is the chance for that individual to be chosen. In particular, the ratio between the OF/fitness and its average value returns the expected number of sure reproductions; the integer part is the number of sure reproductions of the individual, while the residual to the near integer is equal to the probability of another reproduction of the same individual. With the aim of guaranteeing elitism, a small portion of the best individuals (10%) from the last generation is carried over (without any changes) to the next one.

The crossover operator is applied to a subset of the mating pool by taking a pair of chromosomes (parents that will yield a pair of offspring chromosomes). This operation is performed by choosing a random position in the string and then swapping either the left or right substrings of this position (one-point crossover) with its chromosome mate. After the crossover, the mutation operator can be applied. For the chromosome to be mutated, the values of a few positions in the string are randomly modified. A very low mutation probability is assumed to avoid complete loss of the genetic information carried through the selection and crossover methods. The new population is then used in the next GA iteration. The procedure ends when a maximum number of generations is reached.

1) *Costs and Benefits*: As discussed before, the adopted fitness value is defined as a ratio between cost and benefit due to the HVAC pattern definition. In other words, the aim of the proposed approach consists in finding the best compromise between the cost needed to use the HVAC (energy costs) and the benefits, that are essentially related to the thermal comfort of the users. The evaluation of cost and benefit is performed for every single solution obtained by genetic operators.

In the following, the cost c_t and benefit $b_{u,t}$ adopted are presented.

$$c_t = \left(\sum_r P_{r,t}^{HVAC} - P_t^{RES} + P_t^{loads} \right) \cdot C_t^{electr} \quad (7)$$

where $P_{r,t}^{HVAC}$ represents the power consumed in the room r by the HVAC at time t , P_t^{RES} depicts the power produced by the photovoltaic farm, P_t^{loads} the remaining used power consumption, such as light and socket for PCs, and C_t^{electr} expresses the electricity tariff.

The benefit $b_{u,t}$, due to the optimal HVAC scheduling, expresses the thermal comfort for the user u at time t and

it is computed as follows:

$$b_{u,t} = O_{u,t} \cdot \exp \left[-\frac{(T_{t+1}^{in} - T_{u,t}^{pref})^2}{2(\sigma_{T_{u,t}^{pref}})^2} \right] \quad (8)$$

It is defined from the contribution of two terms: the occupancy $O_{u,t}$ and the Gaussian function, illustrated in Section IV-A5, with the mean value that corresponds to the user's favourite temperature $T_{u,t}^{pref}$ and the variance $\sigma_{T_{u,t}^{pref}}$ based on the user preferences. Moreover, T_{t+1}^{in} represents the value of the internal temperature at time $(t+1)$ and corresponds to its prediction. Finally, the total benefit value is then calculated as the mean of all the user benefits as follows:

$$b_t = \frac{1}{U} \sum_u b_{u,t} \quad (9)$$

C. Complexity

The evaluation of the system's complexity considers the two main phases of the aggregation layer: the predictions and the optimization. In the first phase, all the predictions are evaluated simultaneously; for this reason, the time complexity for this phase can be approximated to the complexity of the regression algorithm with the most significant number of features, i.e. the thermal behaviour prediction. Considering k the number of features, the time complexity can be described as $O(k^3)$. In the second phase, the genetic algorithm is responsible for the optimization and so for the scheduling for each office. The complexity depends on the genetic operators, their implementation, the representation of the individuals and the population, and the fitness function. Considering the proposed genetic algorithm with g the number of generations, n the population size, m the size of the individuals, and $O(1)$ the approximated complexity of the fitness function, the complexity of the genetic algorithm is on the order of $O(gnm)$. Hence, the time complexity of the proposed system is $O(k^3) + O(gnm)$.

V. EXPERIMENTAL EVALUATION

A. Set-Up in a Smart Building

We implement the whole system in 3 offices at the University of Cagliari, which usually follows a working time from 8.00 a.m. to 8:00 p.m., over a long period of 7 months, from February to August, for a total of 9 users, distributed as 5 in the first office, and 2 for each of the other offices. For each room, we create a small Wireless Sensor and Actuator Network (WSAN) in order to monitor and control the status of the HVAC. In specific, we test three HVACs belonging to two different technologies, i.e. two Inverter and one ON-OFF. The controller for the WSAN is represented by a Raspberry Pi 3 Model B+ and it is connected to a DHT22 for temperature and humidity monitoring and to an InfraRed Light Emitting Diode (IR LED) to send commands to the HVAC. Furthermore, the HVAC's power consumption information is obtained from a Smartplug, the Edimax SP-2101W. Moreover, a small WSAN is created for the photovoltaic farm, in which two non-invasive current sensors are connected to an additional Raspberry used

to measure the produced power. Finally, room occupancy is obtained by the smartphone app, which communicates the presence to the Raspberry through a beacon Bluetooth and collects all the user preferences to adapt the proposed thermal comfort function.

Moreover, we create an instance of every object involved in the scenario in the SIoT platform, so one SDT for each room is developed in order to obtain the same interfaces to communicate with the main application through the ME. Each SDT has the important role of supplying the room's thermal data and controlling the offices' HVACs. Moreover, one SDT for each of the users' smartphones is associated with the system and can collect their preferences and presence. Moreover, two other SDTs are created, one for the PV to provide the produced power and another one for the external weather forecast data, respectively. At this point, all the offices create an OOR relationship since they have the same owner, while each user smartphone SDT creates a CWOR relationship with its office. When people stay at work, they meet the offices and establish a friendship after a specific number and duration of meetings. In this way, social relationships can help the system to create the users' profiles, adapt to their choices and temperatures, and detect the presence inside the offices. The ME can evaluate all the data discovering the social relationships between the SDTs through encrypted communication. In specific, it implements all the prediction functions and the optimization and, as a result, sets each HVAC scheduling. Moreover, it is responsible for verifying if the status of predictions is accurate; the system checks the status every 15 minutes, and if one of the predictions fails, the considered ME is updated with the new data, and new predictions and optimization are made. In specific, the scheduling time is set equal to 15 minutes considering the time needed to modify the internal temperature in the offices and consecutively to the time needed for the calculation of the genetic algorithm. The genetic algorithm is tested with a machine with the following specifics: *Processor* Intel(R) Xeon(R) E5-1620 CPU @3.60GHz, *Memory* 32GB Samsung 1600 MHz, *Storage* SanDisk SSD PLUS 240GB and *Operating System* Ubuntu 22.04 LTS. Moreover, the genetic algorithm is set with a size of individuals equal to 144 (4 intervals times every 12 hours for 3 offices), a population size equal to 25 and a number of generations equal to 5. Under all these specifics, the complexity of the system is mostly related to the genetic algorithm as $gnm \gg k^3$; for this reason, the genetic algorithm calculates the scheduling with an average of 6 minutes. Finally, for the occupancys weight $O_{u,t}^d$, the system considers the unitary value for the working days, from Monday to Saturday, and 0 for all the non-working days on the calendar, i.e. both Sundays and bank holidays. At the same time, the system's cost is expressed in terms of power and no electricity tariff is considered for the following analysis.

B. Analysis of Performance

This section shows and analyses the performance of each module and of the whole system. The main goal is to decrease the energy consumption of the HVACs and at the same time increase the users comfort. The performances are evaluated by taking into consideration one or all the three offices.

TABLE II
 R^2 OF THE MACHINE LEARNING REGRESSION MODELS.

Prediction	Regression Model R^2			
	Linear	Lasso	Ridge	Polynomial
Power Production	0.94	0.92	0.91	0.94
Power Consumption	0.98	0.99	0.98	0.99
Thermal Behaviour	0.70	0.80	0.85	0.79

The focus of the first set of simulations is the validation of the different regression models used in three prediction problems, i.e. the power production, the power consumption and the thermal behaviour predictions. In order to validate the performance of the machine learning algorithms and compare them with other regression models, we have computed the accuracy in terms of R-squared (R^2) as the metric. R^2 is a metric that measures the variation in the dependent variable that can be predicted by the independent variable(s) in a regression analysis. It ranges from 0 to 1, where a score of 1 means the model fits the data perfectly, and a score close to 0 implies poor fit. R^2 is employed to assess the accuracy of the model and compare it with other models. We compare the performance of four well-known regression models, the Linear, the Lasso, the Ridge and the Polynomial (with a degree equal to 3) regressions. To test their validity, we select the prediction of one week, and all the algorithms are trained with the same data from the previous month. The results are shown in Table II: as expected, all the regression models present good results in terms of errors, except for the linear regression in the thermal behaviour prediction. In specific, consequently to these performances, as explained in Section IV-A, we make use of a liner regression for the power production prediction, the Lasso model for the power consumption, which provides an advantage also in terms of computational load with respect to Polynomial, and the Ridge model for the thermal behaviour prediction. Figure 4 shows an example of these predictions for a single working day, from 8:00 to 19:00. The Figure illustrates how accurate the regression models are for the first two predictions, i.e. P^{RES} with R^2 equal to 0.96 and P^{HVAC} equal to 0.99. Moreover, the last graph depicts good results for the thermal behaviour prediction (dotted line), with a minimum value of benefit equal to 0.88, and for the internal temperature T^{in} (dashed line) in which R^2 is equal to 0.86.

The next set of results examines another prediction algorithm, the occupancy, used to predict the presence of people inside the offices. Figure 5 illustrates the prediction in a working day for a single user compared to the real presence. The graph exhibits how accurate is the prediction algorithm, in which the value of occupancy is greater than 0.5 in interval times of presence, e.g. from around 9:30 to 11:30, while it is lower than 0.5 for interval times in which the user is not inside the office, e.g. during launch time from 13:30 to 14:30. The system then uses the value of occupancy in order to weigh the user benefit, giving more importance to people truly inside the offices. Moreover, regarding the latest prediction algorithm, i.e. the user preferences, this is based on the user preferences set by the smartphone applications. The app allows

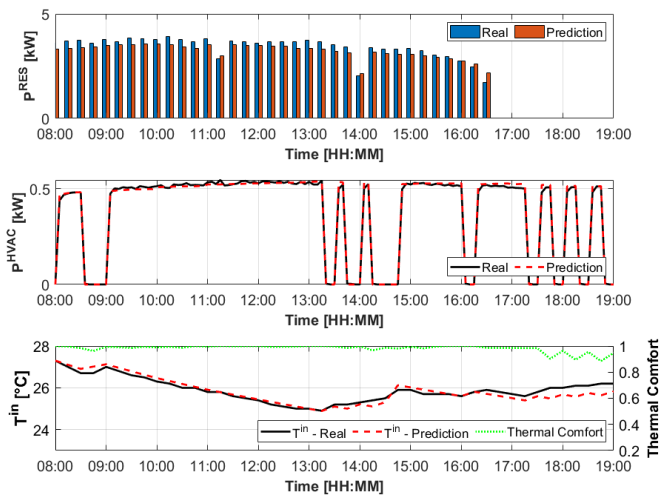


Fig. 4. Power data and thermal behaviour prediction - working day (June)

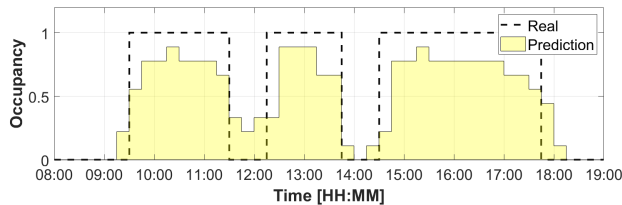


Fig. 5. Occupancy prediction for a single user in an example day.

users to change their favourite temperature, switch off and on the HVAC and modify other parameters, such as the swing or the fan. So no specific set of simulations is carried out; however, these parameters are included in all the following experimentations.

It is important to show how the entire system works in a single office. To test the performance, Figure 6 depicts the real presence of users, the behaviour of the indoor temperature and the HVAC power during a normal working day (June 29th). The HVAC autonomously decides to work in cooling mode. The first graph shows the number of users inside the office between the working time, from 8:00 to 18:00, analyzed through the users' smartphones, while the second graph illustrates the real temperature (solid line), the preferred one (dashed line) and the resulting benefit (dotted line), and finally, the last graph shows the power consumption of the HVAC inside the office. The graph illustrates how an optimal benefit is guaranteed during the working day, with a minimum value of it equal to 0.86 in the presence of people. The favourite temperature for the users varies in accordance with the specific favourite temperature of each user and based on their presence in the room. In order to increase the users' benefit, the system turns on the HVAC just before they would enter the office. Moreover, this choice is allowed by an accurate balance of the consumption of the HVACs and the power produced by the PV generator. As for the internal temperature and the HVAC power, the system makes a prediction for the presence. The system is able to predict the presence of people and then schedule the functioning of the HVAC in order to maintain an

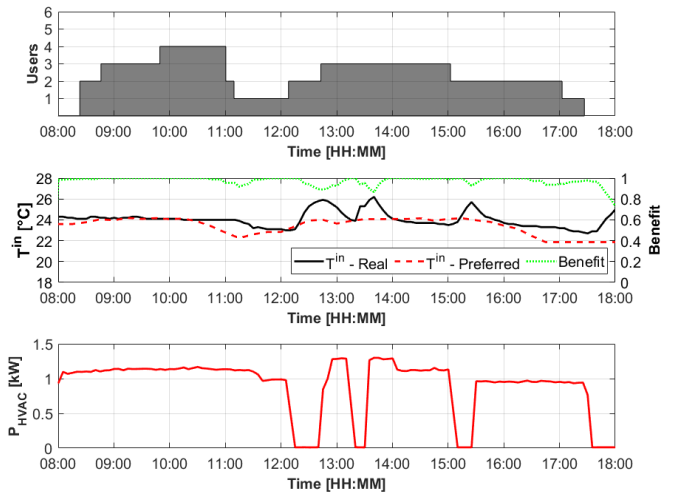


Fig. 6. Presence, temperature and power consumption for a single office in an example day.

adequate indoor temperature and reduce power consumption. Taking into account the previous week, it creates a curve of presence for each office for the working day. This curve allows an accurate prediction of the indoor thermal values and better management of the optimization. After 17:30, the system, in accordance with the absence of users, does not turn on the HVAC and overlooks the preferred temperature. As the systems' goal is to maximize the trade-off between the users' comfort and the cost, it also happens that the HVAC switches off even if people are inside, causing some peaks in the temperature or it stays on even if only one person is inside to keep the room cold making use of the energy from the PV.

The next set of results illustrates the users' interactions over a long period of 7 months. Figure 7 illustrates the number of total interactions for 5 users: each user interacts with the system through a smartphone application, where it is possible to set the preferred parameters. In a traditional approach, i.e. without the proposed system and considering only two different commands (ON and OFF), the system would count at least 120 interactions per month, which represent 2 commands per day for each office over 20 days per month, while the maximum registered number of interactions in our approach corresponds to 20. The graph exhibits how the system could not adapt its behaviour to the users' comfort perfectly in the first month because it is working only on a generic user's information and still needs to learn each user's preferences. However, in the following months, it can learn from users' feedback and requires fewer number of users' improvements. Nevertheless, in the month of May, due to the new season in Cagliari, from the cold season to the warm one, the system needs more interactions. After this first month with the new operation mode, the system is able to adapt again to the users and to the new season, requiring fewer interactions.

Finally, the last set of results is aimed at understanding how our system reacts in terms of benefit and cost. System performances are compared with a "traditional" approach, where users can decide the behaviour of the HVAC through remote control, and there is no energy management system.

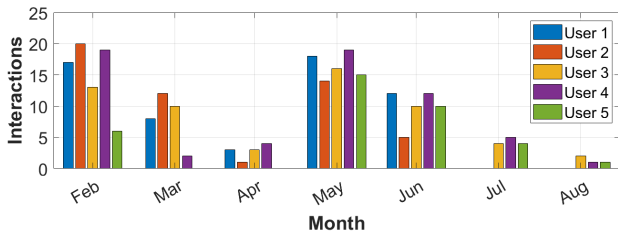


Fig. 7. Number of users' interactions during the real scenario test.

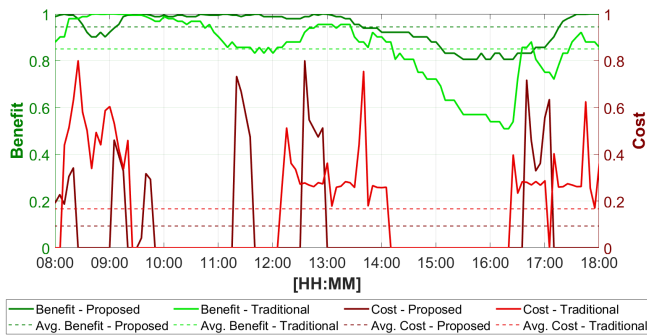


Fig. 8. Cost and Benefit for an office for a single day.

In order to compare the two approaches, two days with a similar pattern of power produced have been selected, and the same power consumed by loads has been considered. Figure 9 shows, according to equations (2) and (3), the benefit and cost normalised on the maximum cost for the selected office. Figure 8 shows the benefit and cost normalised on the maximum cost for the selected office. The graph illustrates how the system is able to reach a high level of benefit with intelligent scheduling and to follow the occupancy profiles better than the traditional use, as can be seen for the period of absence in the room from around 14:00 to 17:00. This is due to the fact that frequently the users do not turn off the HVAC during breaks. The average of benefit for the day is then equal to 0.94, as described by the green dotted line, while for the traditional approach is 0.85. Regarding the cost, the system avoids high consumption and ensures a low cost. The HVAC systems only stay operative in order to guarantee that the benefit does not drop too much, which usually is achieved in half an hour of functioning; otherwise, the system is in off mode to save energy. The average cost for the working day is 0.09, while for the traditional approach is 0.17. Furthermore, Figure 9 illustrates the comparison of the consumption for the three offices and the relative benefit for users. The graphs show how the scheduling of the HVACs is generally provided not simultaneously, with the exception of times in which the benefit for users could considerably drop, e.g. at the start of the day, when the desired temperature is different. The system provides benefits more significant than 0.8 for all the offices.

VI. CONCLUSIONS AND FUTURE WORK

This paper presents a user-centric approach in a smart building scenario towards a cognitive Smart HVAC system.

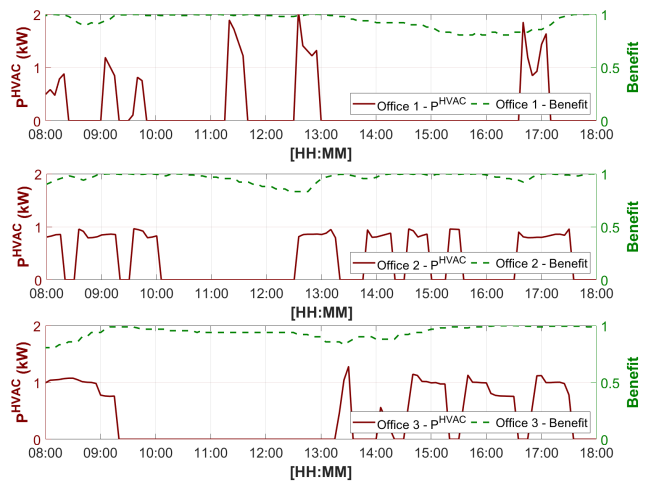


Fig. 9. p_{HVAC} and benefit for all the offices for a single day.

The proposed system implements a scheduling mechanism taking into account not only energy cost savings but also the thermal comfort of users inside the building. Moreover, the system takes advantage of the SIoT paradigm to create users profiles and thermal characterizations of the building and so adapt the HVACs functioning to them. The algorithm and the SIoT architecture have been implemented and tested in a real indoor environment, taking into consideration offices and users, a RES, various device classes, and relationships between the involved nodes. In particular, the whole system has been tested in 3 offices at the University of Cagliari, which usually follows a working time from 8.00 a.m. to 7:00 p.m. for a period of 7 months. Experimental results prove that the implemented system is able to adapt to users needs and ensure an acceptable comfort level while at the same time reducing energy costs with reference to traditional usage.

ACKNOWLEDGEMENTS

This work has been developed within the project "Design and Implementation of a Novel Hybrid Energy Storage System for Microgrids", which is funded by the Sardinian Regional Government (Regional Law no. 7, August 7, 2007) under the Grant Agreement no. 68 (Annuity 2015).

REFERENCES

- [1] K. Song *et al.*, "Energy efficiency of end-user groups for personalized hvac control in multi-zone buildings," *Energy*, vol. 206, p. 118116, 2020.
- [2] United Nations Environment Programme (UNEP) - Division of Technology, Industry and Economics (DTIE), "Cities and buildings," <https://goo.gl/7V6YIX>.
- [3] C. D. Korkas *et al.*, "Intelligent energy and thermal comfort management in grid-connected microgrids with heterogeneous occupancy schedule," *Applied Energy*, vol. 149, pp. 194–203, 2015.
- [4] S. Baldi *et al.*, "Model-based and model-free plug-and-play building energy efficient control," *Applied Energy*, vol. 154, pp. 829–841, 2015.
- [5] L. Atzori *et al.*, "The social internet of things (siot)—when social networks meet the internet of things: Concept, architecture and network characterization," *Computer Networks*, vol. 56, no. 16, pp. 3594–3608, 2012.
- [6] F. E. Aliabadi *et al.*, "Coordination of smart home energy management systems in neighborhood areas: A systematic review," *IEEE Access*, vol. 9, pp. 36417–36443, 2021.

- [7] D. Minoli *et al.*, "Iot considerations, requirements, and architectures for smart buildings energy optimization and next-generation building management systems," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 269–283, 2017.
- [8] I. T. Michailidis *et al.*, "Proactive control for solar energy exploitation: A german high-inertia building case study," *Applied Energy*, vol. 155, pp. 409–420, 2015.
- [9] E. Curry *et al.*, "Internet of things enhanced user experience for smart water and energy management," *IEEE Internet Computing*, vol. 22, no. 1, pp. 18–28, 2018.
- [10] G. Bianchini *et al.*, "An integrated model predictive control approach for optimal hvac and energy storage operation in large-scale buildings," *Applied Energy*, vol. 240, pp. 327–340, 2019.
- [11] S. M. Hakimi and A. Hasankhani, "Intelligent energy management in off-grid smart buildings with energy interaction," *Journal of Cleaner Production*, vol. 244, p. 118906, 2020.
- [12] M. Ostadjafari *et al.*, "Smart building energy management using nonlinear economic model predictive control," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.
- [13] O. Van Cutsem *et al.*, "Cooperative energy management of a community of smart-buildings: A blockchain approach," *International Journal of electrical power & energy systems*, vol. 117, p. 105643, 2020.
- [14] M. Elisisi, M.-Q. Tran, K. Mahmoud, M. Lehtonen, and M. M. Darwish, "Deep learning-based industry 4.0 and internet of things towards effective energy management for smart buildings," *Sensors*, vol. 21, no. 4, p. 1038, 2021.
- [15] A. Capozzoli *et al.*, "Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings," *Energy*, vol. 157, pp. 336–352, 2018.
- [16] I. Ullah *et al.*, "An efficient energy management in office using bio-inspired energy optimization algorithms," *Processes*, vol. 7, no. 3, p. 142, 2019.
- [17] C.-T. Lee *et al.*, "Design and implementation of a leader-follower smart office lighting control system based on iot technology," *IEEE Access*, vol. 10, pp. 28 066–28 079, 2022.
- [18] A. Ortiz *et al.*, "The cluster between internet of things and social networks: Review and research challenges," *Internet of Things Journal*, IEEE, vol. 1, no. 3, pp. 206–215, June 2014.
- [19] P. Mendes, "Social-driven internet of connected objects," in *Proc. of the Interconn. Smart Objects with the Internet Workshop*, 2011.
- [20] C. Marche *et al.*, "How to exploit the social internet of things: Query generation model and device profiles dataset," *Computer Networks*, vol. 174, p. 107248, 2020.
- [21] R. Girau *et al.*, "Lysis: A platform for iot distributed applications over socially connected objects," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 40–51, 2016.
- [22] L. A. Martins *et al.*, "A systematic review of personal thermal comfort models," *Building and Environment*, vol. 207, p. 108502, 2022.
- [23] ANSI/ASHRAE Standard 55-2020, *Thermal environmental conditions for human occupancy*. American Society for Heating, Refrigerating and Air-Conditioning Engineers, 2020.
- [24] P. O. Fanger *et al.*, "Thermal comfort. analysis and applications in environmental engineering." *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.
- [25] W. Zhang *et al.*, "Thermal comfort modeling for smart buildings: A fine-grained deep learning approach," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2540–2549, 2018.
- [26] M. Shahinmoghdam *et al.*, "Bim-and iot-based virtual reality tool for real-time thermal comfort assessment in building enclosures," *Building and Environment*, vol. 199, p. 107905, 2021.
- [27] B. Brik *et al.*, "An iot-based deep learning approach to analyse indoor thermal comfort of disabled people," *Building and Environment*, vol. 203, p. 108056, 2021.
- [28] A. Ali *et al.*, "Iot-based smart air conditioning control for thermal comfort," in *2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*. IEEE, 2019, pp. 289–294.
- [29] Z. Wang *et al.*, "Individual difference in thermal comfort: A literature review," *Building and Environment*, vol. 138, pp. 181–193, 2018.
- [30] H. A. Kazem and J. H. Yousif, "Comparison of prediction methods of photovoltaic power system production using a measured dataset," *Energy Conversion and Management*, vol. 148, pp. 1070–1081, 2017.
- [31] G. A. Seber and A. J. Lee, *Linear regression analysis*. John Wiley & Sons, 2012, vol. 329.
- [32] S. Baldi *et al.*, "Joint energy demand and thermal comfort optimization in photovoltaic-equipped interconnected microgrids," *Energy Conversion and Management*, vol. 101, pp. 352–363, 2015.
- [33] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [34] J. L. Nguyen *et al.*, "The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity," *Indoor air*, vol. 24, no. 1, pp. 103–112, 2014.
- [35] L. Mba *et al.*, "Application of artificial neural network for predicting hourly indoor air temperature and relative humidity in modern building in humid region," *Energy and Buildings*, vol. 121, pp. 32–42, 2016.
- [36] C. Hans, "Bayesian lasso regression," *Biometrika*, vol. 96, no. 4, pp. 835–845, 2009.
- [37] C. D. Korkas *et al.*, "Occupancy-based demand response and thermal comfort optimization in microgrids with renewable energy sources and energy storage," *Applied Energy*, vol. 163, pp. 93–104, 2016.
- [38] Y. Peng *et al.*, "Using machine learning techniques for occupancy-prediction-based cooling control in office buildings," *Applied energy*, vol. 211, pp. 1343–1358, 2018.
- [39] S. Doca and T. Hong, "Occupancy schedules learning process through a data mining framework," *Energy and Buildings*, vol. 88, pp. 395–408, 2015.
- [40] R. F. Rupp *et al.*, "A review of human thermal comfort in the built environment," *Energy and Buildings*, vol. 105, pp. 178–205, 2015.
- [41] I. ISO, "7730: Ergonomics of the thermal environment analytical determination and interpretation of thermal comfort using calculation of the pmv and ppd indices and local thermal comfort criteria," *Management*, vol. 3, no. 605, p. e615, 2005.
- [42] C. Marche and M. Nitti, "Iot for the users: Thermal comfort and cost saving," in *Proceedings of the ACM MobiHoc Workshop on Pervasive Systems in the IoT Era*, 2019, pp. 55–60.
- [43] D. E. Golberg, "Genetic algorithms in search, optimization, and machine learning," *Addion wesley*, vol. 1989, no. 102, p. 36, 1989.



Claudio Marche Claudio Marche received his PhD in Electronic and Computer Engineering in 2023 and the M.Sc. degree in Telecommunication Engineering with full marks in 2018 at the University of Cagliari. Since graduation, he has been working as a Researcher in the Department of Electrical and Electronic Engineering at the University of Cagliari in the Net4U research group. He is currently a Researcher, and his current research interests include the Internet of Things (IoT) and Machine Learning.



Gian Giuseppe Soma is Research Fellow at the University of Cagliari since 2009. He graduated in Electrical Engineering at the University of Cagliari in 2005 (magna cum laude); he was awarded with the Ph.D. degree by the same university in 2009. He co-authored more than 103 papers published in international journals and conference proceedings. The main fields of interest are distribution planning and operation with smart grids, microgrids, and multi-objective methodologies for network planning.



Michele Nitti is an Assistant Professor at the University of Cagliari, Italy since 2015. Currently, he is a member of the editorial board for the IEEE IoT Journal, Elsevier Computer Networks and MDPI IoT and co-founder of an academic spin-off (GreenShare s.r.l.) which works in the mobility sector. His main research interests are in architecture and services for the Internet of Things (IoT).