Estimation of smart home thermophysical parameters using dynamic series of temperature and energy data

Oleh Sinkevych
Radioelectronic and Computer Systems
Ivan Franko National University of Lviv
Lviv, Ukraine
oleh.sinkevych@lnu.edu.ua

Liubomyr Monastyrskyi
Radioelectronic and Computer Systems
Ivan Franko National University of Lviv
Lviv, Ukraine
lyubomur.monastyrskyy@lnu.edu.ua

Bohdan Sokolovskyi
Radioelectronic and Computer Systems
Ivan Franko National University of Lviv
Lviv, Ukraine
bohdan.sokolovskyy@lnu.edu.ua

Yaroslav Boyko
Radioelectronic and Computer Systems
Ivan Franko National University of Lviv
Lviv, Ukraine
yaroslav.boyko@lnu.edu.ua

Zenyk Matchyshyn
Advanced Technologies
Altran/Lohika
Lviv, Ukraine
zenyk.matchyshyn@gmail.com

***Abstract*—A development of the modern methods dedicated to the energy optimization systems for a smart home is being widely studied by community of academical researches and commercial R&D facilities. To contribute a new method to the existing approaches designed for the understanding of smart home energy losses caused by environmental conditions, an algorithm for estimation of the thermal parameters of the smart home is presented in this paper. Proposed method is based on solving an inverse thermophysical problem in variational formulation, where the temperature and energy data are used as input parameters. This type of problem has been obtained by a reduction of the differential form of a dynamic heat transfer equation to discrete representation in order to use sensor time series data. The realization and testing of developed method has been carried out by using an open-access dataset provided by REFIT SMART Home project. These smart home sensor data have been undergone a cleansing with the subsequent matching with discrete form of the heat transfer problem. Finally, as a result of this process, the effective thermophysical parameters - heat capacity and conductivity have been calculated with the use of solving the inverse problem by one of the state-of-art minimization algorithms.**

***Index Terms*—smart home, energy optimization, inverse problem, heat capacity, heat conductivity**

I. Introduction

Due to the quick development of Industry 4.0 smart devices by various companies, e.g., Nest Labs and Ecobee, modern statistical and machine learning methods allow engineers to solve problems in different areas of applied science. Among the variety of challenging problems which are faced by researchers, of special interest is an identification of the thermal (thermophysical) parameters of the smart home. By this term we mean a set of such parameters as effective heat capacity and conductivity of the residential house. An understanding of these values in combination with energy and temperature data, thermostat settings and resident activity patterns provide more intelligent approach for the development of highly effective energy management system for the smart home.

Among the variety of existing approaches to model the thermal behavior of a building (”white box”, ”black box” and ”grey box” methods), the latter combines either strengths of others: statistical and physical essence of the building’s thermodynamic process and tackles, in some sense, weaknesses of both. Hence, the aim of the above method is to describe heat influx and loss of the building as well. For instance, in [1] measured temperature and energy data were used with the stochastic differential equations to select suitable modeling approach. During that, a lumped parameter model was built and investigated. This method is considered as an initial step in exploration of using the ”grey box” modeling for real smart home data. In [2] heat loss coefficient of a single-family household based on short term monitoring data was calculated using two approaches: ”grey box” method and autoregressive modeling, but with the temperature limitations on boundary zone.

Next contribution to the mentioned approaches was done in [3], where authors used simple representation of dynamic heat transfer equation in order to bring together the relations between the consumed and the lost energy as well as amounts of the former for increasing the indoor temperature using a degree days concept.

To improve existing techniques, in [4] new metrics to measure the effect of thermal mass on the energy required to heat and cool buildings was developed and studied.

In this paper, we propose a new approach for the identification of thermophysical effective characteristics of a smart home: heat capacity and conductivity values within ”grey box” approach. Practically, to implement this technique, we assume that a building is equipped with indoor and outdoor temperature sensors as well as the ability to receive the corresponding energy (depends on heating source type, e.g., electrical or gas) readings at the same point of time.

II. Data description and mathematical model

The aim of current work is to develop improved method for the determination of effective smart home thermophysical coefficients on the basis of given sensor data and mathematical background. Thereby, for this research we did not focus on acquiring own sensor data from equipped residential house, but used open-access REFIT Smart Home database [5]. The reason of choosing these data is that dataset consists of sufficiently vast temperature (rooms, radiators, outdoor) and energy readings from 20 buildings situated in temperate climate zone with distinctive cold periods, which helps to identify heating and cooling cycles. For the convenient use of the data, we have created a sqlite database designed for a quick access to required tables using Python.

A. Data preprocessing

The integral (averaged) thermophysical parameters of a building can be estimated not only via known physical properties of the walls and insulation level, but also from the solution of dynamic heat transfer equation. Data needed to conduct modeling have been chosen as follows: all indoor/outdoor temperature readings from selected annual period (March, 2014 –March, 2015) and corresponding energy (gas) usage.

Data preprocessing steps consist of:

- *Data visualization*. It provides the a priori understanding of data and optimal choice for the further steps of data cleansing.
- *Dealing with missing values*. Here we have used the interpolation (linear) technique [6] to fill values into the missing gaps in time series.
- *Anomaly detection*. Due to errors and noise availability during data accumulation, the outliers filter has been used to detect abnormal behavior of the data. We have chosen the proximity-based algorithm (k-nearest neighbors) for the energy and temperature time series, which has shown good performance for detecting the anomaly values during considered periods.
- *Resampling*. Half and quarter hour time series have been converted to hourly data by summing all correspondent values in particular interval.

Since the energy data were measured in cubic meters, a conversion to kWh has been done [5]. Hence, the resulting time series have been prepared as input data for the developed model.

B. Mathematical Model

We have started our investigation from a point of defining the dynamic differential heat transfer equation applied to building's heating process. Indeed, this equation has taken the following form [7]:

$$c(\mathbf{r})\rho(\mathbf{r})\frac{\partial T_i(\mathbf{r},t)}{\partial t} = div(k(\mathbf{r}) \bigtriangledown T_i(\mathbf{r},t)) + Q(\mathbf{r},t), \quad (1)$$

where $c(\mathbf{r})$ is the heat capacity coefficient, $\rho(\mathbf{r})$ is the substance density, $\mathbf{r}$ is the vector of spatial coordinates, $T_i(\mathbf{r},t)$ is the indoor temperature distribution, $k(\mathbf{r})$ is the heat conductivity, t is the time variable, $div(\cdot)$ is the divergence operator, $\bigtriangledown(\cdot)$ is the del operator, $Q(\mathbf{r}, t)$ is the heating power, consuming in the particular building.

Then, after the integration of previous equation upon the building's volume V and applying the divergence theorem to the first term on right side of the equation, we have reached the next form:

$$c^*\frac{\partial \bar{T}_i(t)}{\partial t} = -k^*(\bar{T}_i(t) - T_e(t)) + \bar{Q}(t) \quad (2)$$

where $c^* = c\rho\ [\frac{J}{m^3 * K}]$ is the effective (averaged over the building's volume) heat capacity, $\bar{T}_i(t)$ is the averaged indoor temperature for the time t, $T_e(t)$ is the outdoor (environment) temperature, $k^* = k\frac{S}{V*w}\ [\frac{W}{m^3}]$ is the effective (averaged) heat conductivity, S is the total walls area, w is the average walls width, $\bar{Q}(t) = \frac{Q(t)}{V}\ [\frac{W}{m^3}]$ is the total heating power being consumed in the building averaged by its volume. Here we assume that the process of dynamic heat transfer occurs through the area S. To simplify this equation, we have replaced the temperature difference $\bar{T}_i(t) - T_e(t)$ by $T^*(t) = \bar{T}_i(t) - T_e(t)$.

The dynamic differential equation (2) reflects our main mathematical model that is used for the optimization and identification purposes.

The next stage of our work after obtaining necessary mathematical inferences is the time discretization of equation (2). This has been done according to consideration of using discrete (time series) sensor data to investigate the range of effective heat capacity and conductivity parameters.

The discrete dynamic problem is illustrated via next representation:

$$c^*\frac{T^*(t_i) - T^*(t_{i-1})}{\triangle t} = -k^*T^*(t_i) + \bar{Q}(t_i) - c^*\frac{\triangle T_e}{\triangle t}, \quad (3)$$

where t_i and t_{i-1} determine the discrete moments of time and the simplified notations in $\triangle(\cdot)$ are $\triangle t = t_i - t_{i-1}$, $\triangle T_e = T_e(t_i) - T_e(t_{i-1})$. To complete this problem for the calculation purpose, we have added the initial condition for the known indoor temperature function:

$$T^*(t_0) = t_0^*, \quad (4)$$

where t_0 is the initial time value. Here we assume that $\bar{Q}(t)$ and $T_e(t)$ are the known functions determined for $t \geq t_0$.

Model (3)-(4) with the corresponding assumptions defines a direct problem, which will be used further to study the calculation of effective heat capacity c^* and conductivity k^* parameters in next section.

III. Investigation of direct problem

This section describes the method of solving direct problem of calculation building's averaged indoor temperature for the known effective heat capacity c^* and conductivity k^* parameters. Recursive differences are applied here to express the relation between two temperature values, calculated on current and previous steps.

The unknown temperature $T^*(t_i)$, derived from equations (3)-(4) is determined via the recurrence relation as follows:

$$T^*(t_i) = \frac{\bar{Q}(t_i)\triangle t - c^*\triangle T_e + c^* T^*(t_{i-1})}{c^* + k^*\triangle t}. \qquad (5)$$

A. Step selection

In equation (5), $\bar{Q}(t_i)$ is the piecewise-continuous function of time which represents the real schema of heating modes (switching on/off) of the gas consumers. During the data visualization, we have noticed that behavior of this function leads to the necessity of using modified finite difference methods [8] in order to smooth the function and increase an accuracy. We has used an information about the minimal duration of heating phases for each building in order to choose the best step for our recursive schema.

To select the optimal step here, we propose a normalized metrics:

$$\sum_{i=1}^{N_p} \frac{|T^*(t_i)_{\triangle t_Q^{(j)}} - T^*(t_i)_{\triangle t_Q^{(j-1)}}|}{|T^*(t_i)|_{\triangle t_Q^{(j)}}} \leq \varepsilon, \qquad (6)$$

where $T^*(t_i)_{\triangle t_Q^{(j)}}$ and $T^*(t_i)_{\triangle t_Q^{(j-1)}}$ are the calculated indoor temperatures at time point i, done for specific steps $\triangle t_Q^{(j)}$ and $\triangle t_Q^{(j-1)}$, $\triangle t_Q^{(j)} \in [0.1, 0.25, 0.5, 1]$ according to 10, 15, 30 and 60 minutes as time steps for the discretization, N_p is the total number of hours, $\varepsilon > 0$ is the stopping criteria.

Numerical experiments have shown that solution of the direct problem, based on particular dataset and implemented step selection algorithm has not been crucially sensitive to the step $\triangle t_Q^{(j)}$, though, in general, we recommend to apply its for a such kind of data. Here, we have chosen $\triangle t_Q^{(j)} = 0.1$ on the basis of equation (6). Also, to reduce a computational time required to analyze that sort of problems, adaptive step selection algorithm can be developed. It uses the procedure of a step decrease if only the energy (gas) switching is detected in the data.

B. Numerical experiments

Before start solving the direct problem defined by discrete equation (5), we have conducted the analysis of energy time series, which belong to winter/autumn and summer periods. The best understanding of building's thermal behavior (e.g., the ability to keep heat) is provided by the investigation of energy consumption and its association with thermophysical parameters during the cold seasons. Although the energy consumption data are known for all the periods, the information about an exact amount of energy used only for heating process has not been available.

To resolve this issue, we have applied the next heuristic algorithm: 1) during warm period the function representing energy data has been integrated and corresponding average value has been set as the portion of energy used for non-heating process; 2) calculated value has been subtracted from cold period data and the corresponding recalculated data have been prepared for the use in numerical experiments.

Figure 1 shows the actual indoor temperature distributions calculated for the cold period (October, 2014 - February, 2015). These calculations have been done for several pairs of the parameters $\{c^*, k^*\}$: $\{5 \cdot 10^3, 0.5\}$, $\{6 \cdot 10^3, 0.4\}$, $\{8 \cdot 10^3, 0.8\}$, $\{4 \cdot 10^3, 0.7\}$, what results in T_1, T_2, T_3, T_4 temperature distributions respectively. Input data for the calculations (initial indoor/outdoor temperatures, energy data) were taken from the dataset corresponded to the arbitrary chosen household. Here we demonstrate the sufficient sensitivity of the solution due to different values of parameters c^* and k^*.

In turn, Figure 2 exemplify the solutions of equation (5), namely $T^*(t_i)$, calculated for different sets of c^* and k^* given as grid of values: $c^* = [1.2 \cdot 10^3, 10^4]$, $k^* = [10^{-1}, 1.0]$. Here, for the combination of parameters c^* and k^* the solutions $T^*(t_i = 100)$ and $T^*(t_i = 500)$ have been calculated for specific time t_i within cold period range. As it is seen, each of built surfaces has global minimum, thus a minimization algorithm in the scope of solving the inverse problem can be applied to identify the optimal parameters c^* and k^*.

IV. Inverse Problem

Here we propose an approach for calculation of the effective parameters c^* and k^*, optimal for defined problem in section III, which is based on variational formulation. The results of direct problem investigation (sensitivity to the required parameters and existence of the optimal points) have approved the possibility of using them as a part of inverse problem. We have formulated the process of seeking parameters c^* and k^* as follows:

$$\sum_{i=1}^{N_p} (T^*(t_i; (c^*, k^*)) - T^*_{actual}(t_i))^2 \to min_{(c^*,k^*)}, \qquad (7)$$

where $T^*(t_i; (c^*, k^*))$ is the calculated temperature as the solution of corresponding direct problem for the given parameters c^* and k^*, $T^*_{actual}(t_i)$ is the actual value of temperature.

This problem requires application of the minimization algorithm to search for the global minimum. We have used two different techniques as brute force algorithm [9] to obtain good initial point and quasi-Newton limited-memory BFGS method [10] to get more accurate solution.

Hence, solution of the inverse problem is rather sensitive to chosen heating period, the range of calculated values c^* and k^*, especially the former one, can be sufficiently wide. For example, the numerical experiments, performed for one chosen building, show that c^* value varies from 1200 to 4200 $\frac{J}{m^3 * K}$ and k^* from 0.6 to 0.65 $\frac{W}{m^3}$ for time periods December, 2014 - February, 2015 and November, 2014 - February, 2015 respectively.

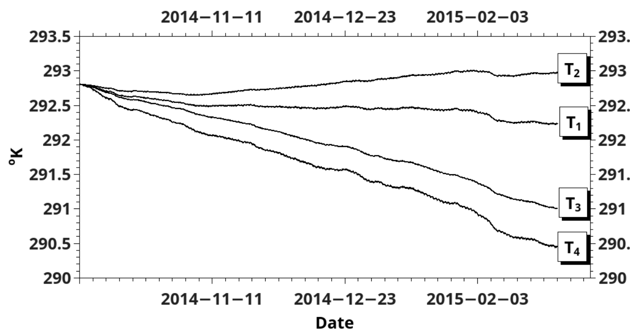

Fig. 1. Calculated indoor temperatures for the set of parameters c^* and k^*

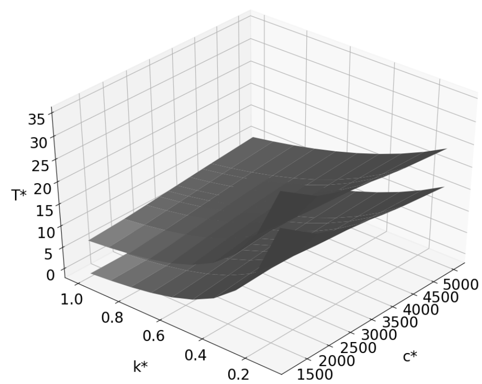

Fig. 2. Calculated temperatures $T^*(t_i = 100)$ (upper surface) $T^*(t_i = 500)$ (lower surface) and for the set of parameters c^* and k^*

V. Discussion And Conclusions

We have conducted the thermophysical modeling using open-access time series data consisting of temperature and consumed aggregated energy dependencies. Due to the fact that available data contain varies artifacts (anomalies, missing values, etc.), there is a necessity of thorough preprocessing with taking into account specific type of the data. The developed model, based on finite-difference method, allows to obtain averaged indoor temperature distribution for the given outdoor temperature data and consumed energy. It should be noted that for the adequate modeling we have to exclude the portion of energy used not for the non-heating purposes from the aggregated energy data. Solving the direct problem is an important stage for formulation of the inverse problem, i.e., estimation of the thermophysical parameters (heat capacity and conductivity) of a building.

Solution of the inverse problem is of importance for the understanding of indoor thermal behavior for the forecasting aims: energy consumption and its reducing, predictive analysis, etc. Because of sensitive of the inverse problem to time series behavior, it needs to pay a lot of attention to the preprocessing algorithms improvement. We are going to consider some aspects of the last task in the frame of our inverse problem in next papers.

Acknowledgment

We thank the research team from Loughborough University, UK: Dr. Steven Firth, Prof. Tarek Hassan, Dr. Tom Kane, Dr. Michael Coleman and PhD researcher Vanda Dimitrou for sharing open access sensor data of smart homes under REFIT project and answering our questions.

References

[1] Dimitriou, V., Firth, S., Hassan, T., Kane, T. and Fouchal, F. (2019). Developing suitable thermal models for domestic buildings with Smart Home equipment. [online] Dspace.lboro.ac.uk. Available at: https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/15930 [Accessed 30 Jan. 2019].

[2] E. Himpe and A. Janssens, "Characterisation of the Thermal Performance of a Test House Based on Dynamic Measurements," Energy Procedia, vol. 78, pp. 3294-3299, 2015.

[3] S. Tabatabaei, W. van der Ham, M. C. A. Klein and J. Treur, "A Data Analysis Technique to Estimate the Thermal Characteristics of a House," Energies, vol. 10, no. 9, p. 1358, 2017.

[4] A. Reilly and O. Kinnane, "The impact of thermal mass on building energy consumption," Applied Energy, vol. 198, pp. 108-121, 2017.

[5] T. Kane, S. Firth, T. Hassan and V. Dimitriou, "Heating behaviour in English homes: An assessment of indirect calculation methods," Energy and Buildings, vol. 148, pp. 89-105, 2017.

[6] G. Milovanovic, Interpolation Processes. Dordrecht: Springer, 2008.

[7] F. Incropera, Principles of heat and mass transfer. Hoboken, NJ: Wiley, 2013.

[8] G. Evans, J. Blackledge and P. Yardley, Numerical Methods for Partial Differential Equations. London: Springer London, 2000.

[9] R. Stephens, Essential algorithms. Indianapolis, IN: J. Wiley & Sons, 2013.

[10] Y. Xiao, Z. Wei and Z. Wang, "A limited memory BFGS-type method for large-scale unconstrained optimization," Computers & Mathematics with Applications, vol. 56, no. 4, pp. 1001-1009, 2008.