Simulation of Municipal Heating Network Based on Days with Similar Temperature

Dolinay V., Vasek L.

Abstract—This article describes preparation and verification of particular part of the model for heat distribution. Presented improvement focuses on heat supplies proposal which is based on seeking and identification of days with similar outdoor temperature behavior. The model of distribution system for heat consumption was prepared but the main questions, for real system management as well as for simulation, still remain: When, how much and in what heat condition to deliver into the urban agglomeration? The first and simplest answer could be the similar conditions as we used yesterday or better, day alike the one we want to control now. The main task of heat supply systems is to maintain all needs associated with heat consumption. From the information we had got from previous (similar) day, the new control could starts. The idea is to make the model more precise and offer resources to improve existing control, for more accurate function.

Keywords—Heat, consumption, model, temperature, simulation.

I. INTRODUCTION

This paper deals with approaches to improving Municipal heating network simulation model [12]. Model offers possibility to describe distribution and consumption of heat energy in the municipal heating systems. Behavior of water temperature in return line, mass flow and consumption can be trace. There are many different approaches for simulation models and operational optimization of heating networks [1]; [2] and heat-load modeling [8]. Our approach is to use data mining combined with simple model of heating network [11]. The adaptive parts of the model utilize real data measured in distribution systems to set up its internal structures for subsequent use in prediction and regulation.

For our model, the chosen city was simplified and model can be trained on the real measured data [13]. The main aim of all experiments face the question: When, how much and what temperature to set-up for hot water supply. Several experiments, presented in this article lately, have confirmed importance of finding days whose outside temperature is close to one we are just need to control. Such, for “tomorrow” we need to know weather forecast to choose right interference on heating water temperature. And based on it also the database can be scan to select day with likely the same values. The found day becomes the base for “tomorrow” heat supply proposal. The expectations are that the data from that day are telling us consumption needs and also provide information about trace in time.

Simulation and control can be described in these steps:

- seek and choose best matching day from the past (Looking for days of ancient history has no sense, e.g. previous years, heating season, because the system is constantly changing. Days from the surrounding area should be preferred. It is also advisable to monitor the previous days, because if it is such a day following a significant change in the weather, the behavior of consumers is considerably in an unstable state)
- train the model (Behavior of consumers is such nonlinear that better than trying to find a general function describing its behavior is to identify a particular period and model optimization for a given situation)
- predict behavior for proposing day (System trained on a similar day can learn from any mistakes and optimize the management of individual variables to the optimal operation)

II. MODEL IMPLEMENTATION

The training experiments described below are based on real data measured by the heat producer and distributor company. The city about eighty seven square kilometers with about sixty seven thousand citizens has been chosen for setting up model [1] and identification of its parameters [12]; [13]. Location has been split into four parts to embrace the whole area, shown on Fig. 1.

The simulation model basically contains two types of parameters:

- static, e.g. length and diameter of the pipes,
- variable.

![Fig. 1 Location split (4 parts with supply and return pipe lines)](image)

The variable parameters are covering variability of the system and are adapted by the evolution algorithms during model training. Those variables held information about amount of heat mass needed for particular time. The evolution algorithm used in this model is described in detail in [11].
Lately, the location split was extended into twelve consumers. The principle of distance equal parts embracing the whole area has been kept. The virtual pipelines system was put more precisely by recalculating the pipes capacity and pipes diameters [13]. See fig. 2 for pipe hierarchy.

![Fig. 2 Pipe hierarchy with 12 consumers](image)

The key benefits of new pipe model were in its accuracy, the transport delay which varies with each consumer (exchange station) seems to be more close to reality. The exact values are unknown, they are directly depending on current mass flow, which is not measured for each consumer or net node and therefore can not be accurately compared.

The main disadvantage of the new pipe module is higher time-consuming calculation for a single simulation step, the number of discrete quantum of flow of fluid (water) DFQ [12];[13] increased by a greater number of nodes, where the mass is divided.

For the fast calculation, the pipe model only with one consumer was prepared. The pipe capacity was held. Pipe diameter was recalculated to use average distance and keep pipes capacity. Speed of model identification greatly increased, but this simple model reflects the changes in heating water temperature to the temperature of water in the return line, which is not usually the behavior of the real system. Through the distribution of consumers, the real system smoothes this peaks. One consumer system is however suitable for fast verification of new algorithms. See fig. 3 for schematic diagram.

![Fig. 3 Simple, one consumer model](image)

### A. Flow modeling

Compressibility of water in the pipe is insignificant and does not need to be included in the model.

In each simulation step the flow quantum, denoted $iDFQ$, in the network is monitored [12]. Shown in the following picture, we can apply simple terms:

$$jV_i = \pi * D^2 / 4 * L \quad (1)$$

and

$$jV_0 = M_j * \Delta t_j \quad (2)$$

where:

- $jV_i$ is volume of $iDFQ$ on input to distribution network
- $M_j$ mass flow in simulation step $j$.  

To monitor the flow quantum passing through the distribution network, it is of course necessary to respect the fundamental physical laws applicable to the fluid flow and heat energy transfer - conservation of mass and energy and the law of continuity.

In accordance with the law of continuity we have to follow several rules for flow quantum:

1. While passing through the section, the DFQ don’t alter its volume, its length $L$ in a pipe varies depending on the diameter $D$ of the current pipe line.
2. Each section is divided into pipe lines, as noted above. While $DFQ_j$ is passing two consecutive pipe lines $p$ and $q$ of the section, $DFQ_j$ is split into the two new $DFQ - DFQ_p$ and $DFQ_q$. $DFQ_p$ is the part of $DFQ_j$, which remains in the pipe line $p$ - does not reach the border between pipe lines $p$ and $q$. $DFQ_q$ is contrary of $DFQ_p$ - part of $DFQ_j$ which passed transition of $p$ and $q$ in current simulation step and switched into $q$ pipe line.
3. The above mentioned $DFQ_j$’s splitting rules are valid also for each $DFQ_k$, which is entering node $k$. Each $DFQ_k$ is split into two parts – $DFQ_p$ and $DFQ_q$. $DFQ_p$ is the part of $DFQ_k$, which reached the node in particular simulation step and $DFQ_q$ is the part which does not reach it.  

For each output section $j$, which is outgoing for node $k$, $DFQ_j$ is created. From the law of continuity we can use equation:
\[ \sum \text{Vol}(DFQ_l) = \sum \text{Vol}(DFQ_j) \] (3)

where:
- \(\text{Vol}(DFQ)\) describe function, which presents particular part of flow quantum,
- \(\sum\) on the left is processed for all section, which allows flow to come into the node,
- \(\sum\) on the right is processed for all sections, which allows flow to come out from the node.

1) Heat transfer modeling

For each flow quantum, which is at a given time in the distribution network, its heat balance is calculated in each simulation step. The heat balance is based on law of the preservation of the heat energy. The heat energy change in single simulation step, is described by the equation [12]:

\[ \Delta Q_i = \sum (\text{DFQ}) \] (4)

where:
- \(\Delta Q_i\) and \(\sum (\text{DFQ})\) describe the amount of heat energy contained in the \(\text{DFQ}_i\) at the beginning of the simulation step \(j\) and simulation step \(j+1\).

The next equation is still valid:

\[ iQ_i = V_i \rho c_i \Delta T_i \] (5)

where:
- \(c_i\) is the heat capacity constant for the fluid (water),
- \(\rho\) is water density,
- \(V_i\) is volume of \(\text{DFQ}_i\) and \(\Delta T_i\) its temperature.

Presented decrease of the heat \(\Delta Q_i\) in \(\text{DFQ}_i\) during the time interval \(\Delta t\) arises from the fact that this heat is transferred to the surroundings, either in the form of losses (supply line) or in the form of consumption (consumers) [12].

Pipeline losses can be determined by the relationship

\[ iQ_i \text{loss} = k_p \Delta t \] (6)

where:
- \(k_p\) is the heat transfer coefficient in the current pipe line \(p\),
- \(\Delta t\) is water temperature for the \(\text{DFQ}_i\),
- \(\Delta T_{p,\text{ext}}\) is the outside temperature for the pipe line \(p\), both in simulation step \(j\).

Coefficient \(k_p\) is based on pipe structure - pipe material, style and material of insulation, pipe seating, etc.

For the heat consumption at consumer \(r\) at time interval \(\Delta t\), the following equation can be used:

\[ iQ_i \text{cons} = s_i (T_i, T_{\text{ext}}, \ldots) \Delta t \] (7)

where:
- \(s_i(\ldots)\) is the function describing heat consumption for the consumer \(r\).

Determination of this function is obviously very difficult, but for the final solution of this task, especially in terms of its accuracy for those particular parts "consumers", it is very important. There may be applied many different important factors such as:
- type of the day - workday, weekend, holiday etc.,
- part of the day - morning, afternoon, evening, night,
- type of the consumers in the particular part of the network - flats, schools, industrial companies etc.,
- other weather conditions - sun intensity, wind, air humidity.

To determine the functional dependences of heat consumption on these factors it is also possible to successfully use the proposed simulation model. This usage of the model will be included in the identification of model parameters for given conditions.

2) Applicability

It is expected that the proposed simulation model will be used in the control system SHDC for the following purposes:


II. Prediction of appropriate timing of the supplied amount of heat energy for the next period [11].

I. Identification of model parameters for the selected time period

As mentioned, essential for the modeling approach to SHDC is to determine the function \(s_i(\ldots)\) used in equations (7). This means that it is necessary to choose the appropriate form of parametric functions and find values of the parameters for the given conditions.

The procedure will be described in detail on simple example where the function \(s_i(\ldots)\) shall only affect consumption fluctuations during the day. We will therefore assume that the function \(s_i(\ldots)\) will have the form

\[ s_i(\ldots) = \lambda_r \Delta T_i \text{ext} * k_h \] (8)

where:
- \(\lambda_r\) is the coefficient of heat transfer in segment \(r\) (here we suppose that the segments are pipe sections as well as consumer units, depending on the value of the coefficient \(\lambda\)),
- \(\Delta T_i\) is the current temperature \(\text{DFQ}_i\) for the particular simulation step \(j\),
- \(T_i\text{ext}\) is the current outside temperature and
- \(k_h\) is coefficient which corrects heat consumption oscillations during a day.
The coefficient \( k_h \) is considered as a discrete function of time, which changes its value at one hour intervals. Assuming the daily frequency of consumption of heat, this is probably a reasonable assumption, so it is necessary to determine 24 coefficients \( k_h \). When we use function \( s, (...) \) in form of equation (8) then for each \( k_h \) it is possible, in the time between hours, to use the value from the previous hours or use interpolated values. It is possible to use linear or other more complex interpolation. For our purpose of use \( k_h \) and needful accuracy, the linear interpolation is fully suitable.

To determine function \( k_h \) we can use the following procedure, which uses the simulation model in combination with genetic algorithm:

1. To select the appropriate length of time period, longer than one day. For the simplicity at the beginning of each simulation experiment we assume that the distribution network is empty. Therefore, for the beginning of the simulation, we must simulate the preceding interval, to let distribution network fills with hot water and let the state of the network become stabilized. This period is given at least as time delay for transportation, which means that all water from the source must have enough time to pass the system and come back to source again. Then we must follow at least one day period to obtain discrete values of \( k_h \) for each hour of the day.

2. For the chosen time interval to select and verify necessary historical data. For this task the following data are required (measured in time period \( \Delta t \)):
   - Temperature of heating water \( T_{vv} \), measured on source output. This is also considered as an input for distribution network. For our purpose only one source is expected, but model is generally capable to use various combinations of sources and consumers.
   - Temperature of returned water \( T_{vv} \) – output of distribution network ergo reentry into source. It should be noted that this value is implicitly influenced by the size of the time delay for distribution network but this is already covered in the structure of simulation model.
   - Total mass flow \( M \) in distributive network, again measured on the network input (output of the source).
   - Air temperature measured for particular points of distributive network. Quantities of these values will depend on the density of the network, its measuring points for the outside temperature at that location. If measuring points are located at different places than the landmarks of the distribution network, it is also possible to use interpolated values, in this case, interpolation in the area, i.e. two-dimensional. Also, if there are meteorological data measured in a different time period than \( \Delta t \), the value interpolated in the timeline has to be used.
   - Pressure ratios in the major individual network points. These are particularly important in cases where the topology of distribution networks is complicated. According to these, the value of the mass flow in different sections of the distribution network may be determined - without the knowledge of pressure ratios can be mass flow in each section only estimated.

3. To determine searched values \( k_h \) for 24 points of timeline is possible to use several methods based on principles allowing us to find a function(s) which should have the best course approximating analyzed variables. One option is for example to use genetic algorithms. In the presented solution was the method PSO (Particle Swarm Optimization) used - see [7]. This method has been lately compared with other methods, such as SOMA, neural networks [10], and Levenberg-Marquard algorithms for nonlinear methods of least squares. It was found that the results achieved in terms of accuracy and speed of convergence is similar. PSO is therefore comparable for the determination of the correction factors and we use it.

4. The main idea of calculation is that the approximated function is the timing of the \( T_{vv} \) for the output of the distribution network to source.

5. For calculating the approximating function the simulation model will be used. For each simulation experiment, we set sought coefficient \( k_h \) for selected points in the timeline to values generated by the PSO algorithm. With this sets of parameters one simulation experiment is processed for the selected time slot according to point 1. Rate of the quality of approximation is calculated according to equation (9), i.e.

\[
F = \sum (T_{vv\text{ meas}} - T_{vv\text{ calc}})^2
\]

where:
- \( F \) is value of approximation accuracy (fitness),
- \( T_{vv\text{ meas}} \) is measured value of returned water temperature for the time interval \( \Delta t_j \) and
- \( T_{vv\text{ calc}} \) is calculated value of returned water temperature for the simulation step \( j \).

The sum is performed for the entire simulated period in each simulation step. It is thus the sum of squares of deviations between the measured values \( T_{vv} \) (the values measured on the real system) and the calculated values \( T_{vv} \) (the values calculated using the simulation model) in terms of the search parameters - the values of \( k_h \) in 24 points of the timeline.
The calculated F value is transmitted into PSO algorithm, which will use it to generate the next set of values of the coefficients of \(k_h\), with which the process repeats. After a certain number of iterations (from practical results, there is a few thousand iterations necessary), the calculated values \(T_{vv}\) are close to measured values \(T_{vv}\) with sufficient accuracy. So designated coefficient \(k_h\) can then be used in the simulation model as the parameters for its further use - see section II in chapter IV.

An example result of the calculation of the coefficient \(k_h\) was shown in Fig. 5.

Later, the sequence of 24 points for one day was replaced by a smaller amount. Tested was point spacing of six, four and three hours. The current implementation uses a spacing of three hours. Individual points are connected using spline functions (before line – see fig.5).

An example for comparison of measured and predicted values of \(T_{vv}\) in identified model, based on calculated coefficients, is shown in fig. 6.

![Fig. 5 Calculated \(k_h\) coefficient](image)

![Fig. 6 Identification of SHDC - calculated and measured temperature of \(T_{vv}\)](image)

It is obvious that the described procedure can be reused if the typical pattern of consumption of heat will vary in different types of days - a working day, weekend day, holiday, etc. In this case, repeat the procedure for each typical day, with the use of appropriate choice of time period in accordance with point 1. Then, for each type of day we obtain own set of values \(k_h\) that the model "switch" according to the type of the day which is being modeled [11].

These procedures are thus used to identify the model, i.e. to find such values of model parameters that provide the most accurate approximation of the function characterizing the system behavior - in this case of SHDC it is the timing of \(T_{vv}\).

It is also apparent that this procedure can be similarly used to seek the values of correction coefficients for other variables that affect the consumption of heat, like any other variable weather - sunshine, wind direction and magnitude, relative humidity, etc. [4]. Also, the procedure may be easily modified, if we assume a different shape and form of the function \(s(\ldots)\).

II. Prediction of appropriate timing of the supplied amount of heat energy for the next period

One of the practically applicable results tied to the use of the simulation model SHDC is to design a predictor of heat supply, useful for its production, as well as for its distribution.

Such a predictor is a very convenient tool that enables efficient management of SHDC and meets the objectives stated at the beginning of this article. It makes it possible to perform dynamic calculations, which take into account both the characteristics of the SHDC and its temporal changes in response to external conditions (especially climate), which at the particular moment SHDC influence. Without such tools, the whole procedure is based only on experience and intuition of management and controllers, which can lead, especially when there are rapid changes in the state and not fully achieving the normal state system, to not fully effective control of production and distribution of heat.

Based on the results of this work with the simulation model, a procedure for the design, processing and use of predictors in the heat of urban agglomeration consumption was proposed. The procedure is described in the next sections.

It should be noted that the idea of building a predictor in the project gradually evolved and changed. At the beginning of the project, the opinion prevailed that it will be enough to identify a model for a longer period and once identified, the model could be used for this period without any modifications. But it turned out that this approach does not lead to sufficiently accurate results, the characteristics of the system are not sufficiently "stable". It is probably due to both the details of the processed model and major stochastic nature of the whole system. Therefore, the idea was abandoned, and the work moved in the direction of the procedure described below. Given the fact, that the work on the project is not yet finished, described process reflects the present state of the solution and may not be definitive.

The proposed procedure is based on several fundamental ideas:

- SHDC will behave similarly under similar conditions.
- Predicted period is suitable to choose short and necessary calculations (including simulation) to perform for the shorter period repeatedly. This is depending, as already mentioned, on speed changes inside the SHDC and size sampling period \(\Delta t\). If this period is few minutes, it is easy to perform calculations repeatedly.

To control the production and distribution of heat, there are two control variables - temperature \(T_r\) and mass flow \(M\). It will
be necessary to find and use an appropriate cost function, which allows the required amount of heat supplied to optimally divide to the parts obtained mass flow \( M \) and the temperature \( T_v \). Search for this objective function is the task for other parts of the project and is not in this article further discussed [11].

III. TRAINING SAMPLES

The above mentioned common model was subjected to a series of tests to determine the best period for the identification and subsequent prediction that best suit reality. Tested variant are shown below.

A. Particular tests for time period identification

Several periods were tested to verify various assumptions.

1) Month and more

The main idea of those tests was to cover variety of weather during a winter season and model identified on such long period will enclose the most situations. As shown on Fig.7, the prediction coming from such trained model has many inaccuracies. Even the identification was not able to adapt to all abnormalities.

2) Month and more

Subsequent alteration comes up with the same ideas as Month experiment – to adapt model parameters for long period but train model for short prediction, just after trained samples. The results are quite better compared to month samples but disadvantage of this approach is high sensitivity to outside temperature changes.

For example, if temperature changed into values which were not included in training samples, the subsequent prediction for these cases become more inaccurate.

3) Week

Another cut in training samples length brought another improvement. The model was adapted more accurate but problem with incoming days which were significantly different still remains. See Fig. 8.

4) Day

As can be seen from previous experiment the best result could be obtained if the model is adapted for a small time period but subsequent conditions must be met. This approach will be described and developed further in the chapter “Similar day”. Difference between Original (measured) and Calculated (predicted) course, which can be seen in Fig.7 and 8 is reduced when shorter time period is used for training. Output values, shown in those figures are returned water temperatures (water in the return line).

IV. SIMILAR DAY

Unlike previous experiments, the examination of the period just prior to the desired section of prediction is not required. This improvement is based on the principle of finding similar days (periods), and application of its relevant model parameters on the stretch of the same nature.

The advantage of this approach should be also that, having regard to the accuracy of weather forecasts for the period length of 24 hours. So there should be no unexpected fluctuations and thus should be removed error caused by previous procedures inappropriateness of the samples used to identify the model.

To determine whether the day (period) is "similar" to another, the minimum variations of outdoor temperature were used. The day called “similar day” to day we just want to predict is considered such a day when outdoor temperature forecast deviation from the measured temperature will be minimal. The periods is called “similar day” but similarity is sought in wide range, surrounding area is also taken in account. Fig 9 shows three similar periods. System used weather forecast for 7th to 9th of March 2011. Anyway, prediction was required just from 8th of March 12 PM to 9th 6 PM but behavior is more obvious if we look cross all period.
The best matching period (green) were found in previous year (March 2010) but, as mentioned earlier, does not mean best conditions for subsequent use. This is well observable on curves below – consumed heat. The values in the previous year differ from the current year. Similar outside temperature in current year give us similar curves for other variables as well as for consumed heat.

The water temperature in the return line is another of the values on which similarity can be observed system behavior. One may consider the temperature dependence of water in the return line to the outdoor temperature, but can not ignore day time. I.e. if we determine the dependence by regression of these two temperatures, it will be only the big generalization, because even the outdoor temperature is the same, the water temperature in return line typically vary in the range to five degrees of Celsius. We are still considering our selected location (system). In this system, the maximum variance of temperature in the return line is approximately fifteen degrees Celsius – considering whole heating season. Fig 12 shows a good resemblance of these values for similar days. Day from the last year is again different, which confirms the previously presented inappropriateness for prediction.

Fig 13 shows a similar behavior in the consumption of energy, which is determined by practice and monitoring system in the previous days. The values would of course be best to set up a system that will operate at optimal levels at low cost.
similar days. The difference of the previous year is not only in quantity but especially in time. On the contrary, two similar days in the same year show a similar distribution of consumption over the time. However, there is variation in the observable quantities. Variation in the amount brings uncertainty into the system, which works only on the basis of similarity of outdoor temperatures. To improve the system, remains the task of identifying the cause of deviations and incorporate it into the model.

V. CONCLUSION

The experiments focused on adaptation for longer periods (week, month, etc.) suggest the increased need for finding the parameters reflecting the diversity of the system behavior on a different course with the outside temperature. Tests show that the most appropriate method of identification and prediction of returned water temperatures into the model can be considered an application of “similar day” method. In subsequent experiments, it would be useful to examine the possibility of establishing different ranges of similarity search. Described results show merely “day” meant in common view from midnight to midnight of the following day.

Experiments have shown that it is appropriate to seek similar days in the distant past. Day with a minimum deviation from the predictions but found in last year’s heating season is not suitable for further use. However, this restriction greatly reduces the ability to use an algorithm of “similar days” to the early days of the heating season. The number of days in which we can find similarity is minimal at the beginning and so the first cold days of winter can not have similar day in current heating season yet. To improve the system, it is important to find a compensation algorithm, which can be used even for days, which have similar course, but not an exact temperature.

However, the model takes into account most important parameters affecting the amount of the heat and the experiments carried out so far are showing good performance.

ACKNOWLEDGMENT

The work was performed with financial support of research project NPVII-2C06007, by the Ministry of Education of the Czech Republic and is also supported by the European Regional Development Fund under the Project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089.

REFERENCES


