Water supply telemetry data processing in apartment buildings

Elena Ignatova [1,*]

[1]Moscow State University of Civil Engineering, 26, Yaroslavskoye shosse, Moscow, 129337, Russia

Abstract. The digital transformation of construction and housing and utilities sector involves data-based management. The most accessible data is telemetry of electricity, gas, heat and water consumption. Despite the fact that not all apartment buildings are equipped with meters yet, it is necessary to think about methods for processing measurement data. Intelligent data processing methods are gaining wide popularity. The purpose of this study is the processing and analysis of water consumption telemetry data. The objective of the study is to determine the periods of nighttime water consumption using the clustering method. The data of apartment meters of hourly consumption of hot and cold water are investigated. The measurement period is 1 month. Intelligent cluster analysis was conducted based on the DBSCAN machine learning model (Density-based spatial clustering of applications with noise). Clustering objects are the hours of the day. As a result of the study, the hours of night consumption of cold and hot water were allocated, both in the whole month and separately on weekdays and Sundays. The conclusion is made about the benefits of using intelligent cluster analysis of water consumption telemetry data for effective management of water resources and equipment.

Keywords telemetry, water consumption, apartment building, data mining, clustering

1 Introduction

The stages of digital transformation of construction and housing and utilities sector include data accumulation, creation of platforms for data processing and analysis, management transformation, including data-based decision-making. Data accumulation technologies are the most advanced when accounting for resource consumption. In modern apartment buildings, electricity, cold and hot water, gas and heat consumption meters are used. The meters provide both building communal and apartments resource consumption. According to the Russian Geo Information System of housing and utilities sector, communal cold water consumption meters are installed in 36% of the apartment buildings, individual cold water consumption meters are installed in 63% of the apartments. Communal hot water

*Corresponding author: ignatova@mgsu.ru

consumption meters are installed in 31% of the apartment buildings, individual hot water meters are installed in 58% of the apartments [1].

Not all water consumption meters have the property of automatically transmitting data on resource consumption. Smart meters are usually installed in new apartment buildings. Such water consumption meters can record not only water consumption, but also its temperature. It is planned that by 2025, both apartment and general apartment building and section apartment building metering devices will be intelligent and digital in most new apartment buildings [2]. In the old apartment buildings, it is planned to gradually replace meters with new types of equipment and switch to new technologies.

Modern water consumption metering equipment makes it possible to organize a telemetry system, which can be part of the general telemetry system of a "smart" apartment building [3,4]. The water consumption telemetry system includes water meters that record water consumption volumes in real time, sensors for the status of devices and equipment, devices for transmitting digital data over a distance, as well as data storage.

Data accumulation, processing and analysis can be carried out on digital platforms for "smart" apartment buildings, such as IoT platforms Ujin, Iridi Server, SmartUnity. Information and analytical SCADA systems can be used for building dispatching [5].

As a rule, data processing and analysis is performed in order to obtain statistical data, as well as to control payments for resource consumption [6]. However, there are several tasks that can also be solved based on the analysis of water consumption data: improving water supply safety, preventing water losses, optimizing the operating mode of water supply system equipment, clarifying regulatory indicators for the design of water supply and sanitation systems, smart management of water resources consumption [7,8,9,10,11].

Hourly water consumption data in an apartment building is analyzed to optimize the operation of the water supply system and save resources. In the article [12], the distribution of consumption of cold (HW) and hot water (HW) by apartments during the daytime and night hours of weekdays and Sundays, as well as the distribution of consumption by the entire building of cold and hot water during the daytime and night hours of weekdays and Sundays, are investigated. Analyzing these data, it is possible to determine peak hours of water consumption, areas of the system with increased water consumption, places of possible leaks or accidents, the efficiency of water use by residents, the required parameters of pumping equipment.

Modern pumps can be equipped with automation that regulates their operation depending on the needs of the system. Automation maintains constant pressure in the system by adjusting the pump performance depending on the load change. During peak hours (morning, evening), the load on the pump may increase, which may require more intensive operation of the equipment. Automation can be programmed to reduce pump performance during periods of low water intake, which saves energy and the life of pumping equipment. Periods of increased and decreased load are determined based on the analysis of data on water consumption received from sensors and water meters [13].

The main characteristics of water consumption in the house, such as average, median, minimum, maximum and standard deviation, can be obtained using descriptive statistics [14]. Along with statistical methods of data analysis, artificial intelligence methods are increasingly being used. Artificial intelligence technologies allow solving problems of classification, clustering, optimization, forecasting, and management [15]. The article [16] uses deep learning models of neural networks. Intelligent time series analysis can be used to predict water consumption [17,18]. Artificial intelligence can become a digital assistant for employees of management and resource supply organizations, as well as residents of an apartment building.

The purpose of the study is to analyze the telemetry data of the water supply of an apartment building to determine the periods of night water consumption.

Methods

In this study, the analysis of telemetry data on water consumption of an apartment building in Moscow was carried out. The data of the meters of hourly consumption of hot and cold water for 32 days in the period from January 10 to February 10 are discussed. During the research, the operations of problem statement, accumulation, purification and transformation of data, data research and visualization, preparation of a data mining model, interpretation of the results were performed.

The research problem is defined as the search for periods of night consumption of hot and cold water based on data from hourly measurements of water consumption.

Data accumulation is carried out by recording the consumption of cold and hot water in a multiple of 10 liters. A data table (matrix) has been compiled, the width of which is equal to the number of apartments, and the length is equal to the number of hours in the monthly measurement period.

Data cleaning was performed in apartments in which either meters were not installed, or there were no residents during the study period, or one of the water consumption meters was not working. Thus, the data sample decreased by 11%.

The data conversion is carried out to convert the data into a form and format suitable for a certain type of analysis. In addition, we use data generalization. For example, the total values of water consumption for each apartment, the values of hourly water consumption throughout the apartment building, the values of water consumption of apartments were calculated as a cumulative total over the measurement period.

Data analysis includes visual graph analysis and data mining by solving the clustering problem.

Cluster analysis refers to intelligent machine learning methods without a teacher (Unsupervised Learning). The main idea of cluster analysis (clustering, cluster analysis) is to divide objects into groups or clusters in such a way that within a group these objects are more like each other than to objects of another cluster. Most often, we do not know in advance which clusters our objects need to be divided into. 24 hours of the day were selected as the objects. In this way, we try to identify groups of watches with similar water consumption. Generalized values of hourly water consumption throughout the apartment building (separately for cold and hot water) are used as water consumption data.

Results

The visual analysis of the data was carried out based on graphs of consumption of cold and hot water (Fig.1,2). The graph shows the dependence of the total water consumption for all apartments on the time of day for the entire measurement period. An approximate volume of water consumption of 450 liters/hour has been established, which is accepted as the boundary for separating night consumption. When combining graphs (Fig.3) it can be noticed that the period of night consumption of cold water is shifted to the left, that is, it begins earlier. In general, visual analysis gives only an approximate idea of the process of water consumption in the apartment building.

Intelligent cluster analysis was carried out based on the DBSCAN machine learning model ((Density-based spatial clustering of applications with noise, a density algorithm for spatial clustering with the presence of noise). To calculate the model, you need to set two

parameters: the radius of the neighborhood and the minimum number of neighbors in the cluster. Depending on these parameters, a different number of clusters and a different amount of noise (data that is not included in any cluster) are obtained. The parameters are set so that the number of clusters is maximum, and the amount of noise is minimal.

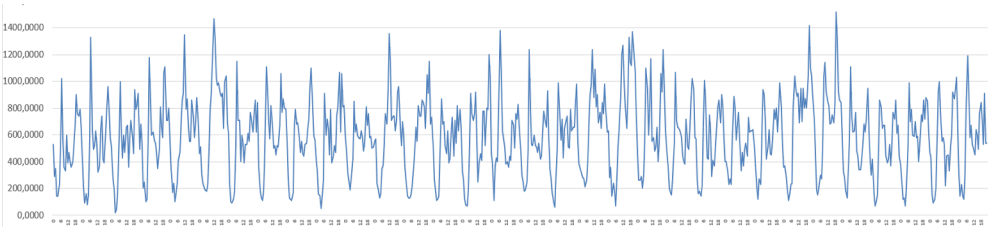

Fig.1. Hourly consumption of cold water in the apartment building during the measurement period.

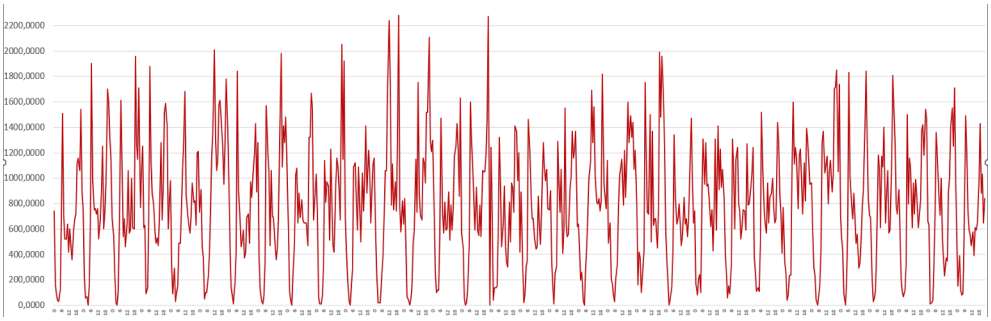

Fig.2. Hourly consumption of hot water in the apartment building during the measurement period.

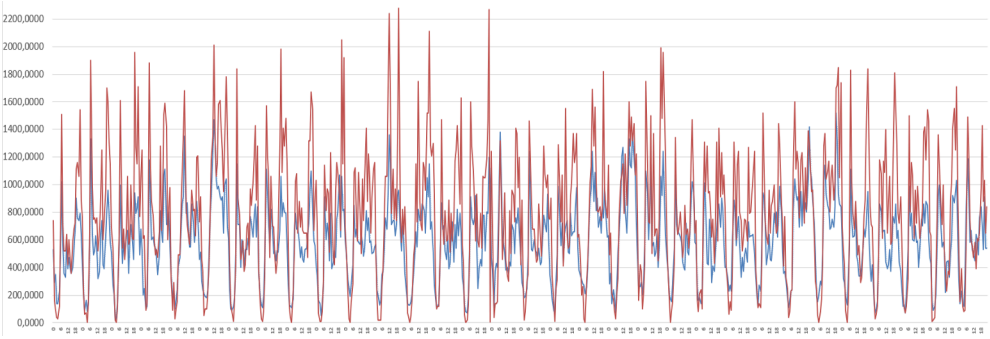

Fig.3 Hourly water consumption in the apartment building during the measurement period.

It is assumed that the minimum number of neighbors in a cluster is 2. The maximum number of clusters, the corresponding values of the radius of the neighborhood and the corresponding noise volumes can be determined in an automated way. According to the calculation results, the maximum number of clusters for cold water is 4, for hot water it is 2.

Figure 4 shows the initial data for the cluster analysis of cold water consumption. The dot graph shows the consumption of cold water in the apartment building depending on the hour of the day in all days of the measurement period. The hours are numbered from 0 to 23. Zones of dense distribution of water consumption values are visible, as well as zones of sparse location of points.

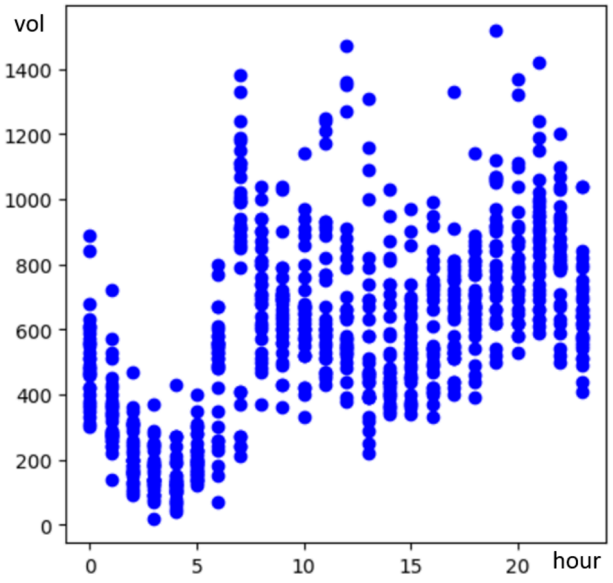

Fig.4. Consumption of cold water in the apartment building depending on the hour of the day during the measurement period.

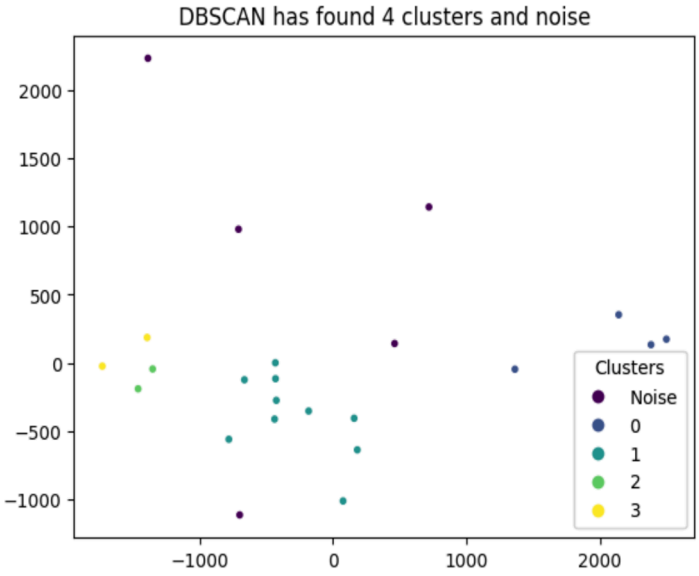

Fig.5. DBSCAN clusters in the analysis of cold water consumption.

Figure 5 shows the result of clustering cold water consumption data. Clusters and noise are highlighted in color. Cluster numbering starts from 0. Noise has a value of -1. Noise means a large variation in water consumption values at the same hours of different days.

Figure 6 shows the initial data for clustering hot water consumption. Data organization is like cold water data organization. Figure 7 shows the result of clustering hot water consumption data.

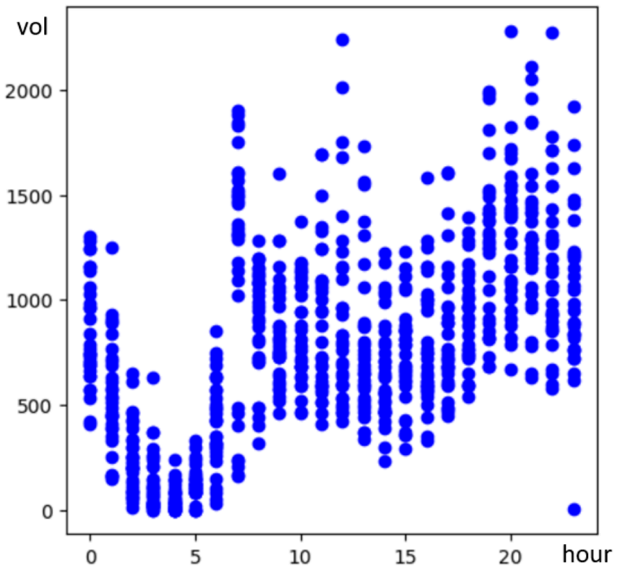

Fig.6. Consumption of hot water in the apartment building depending on the hour of the day during the measurement period.

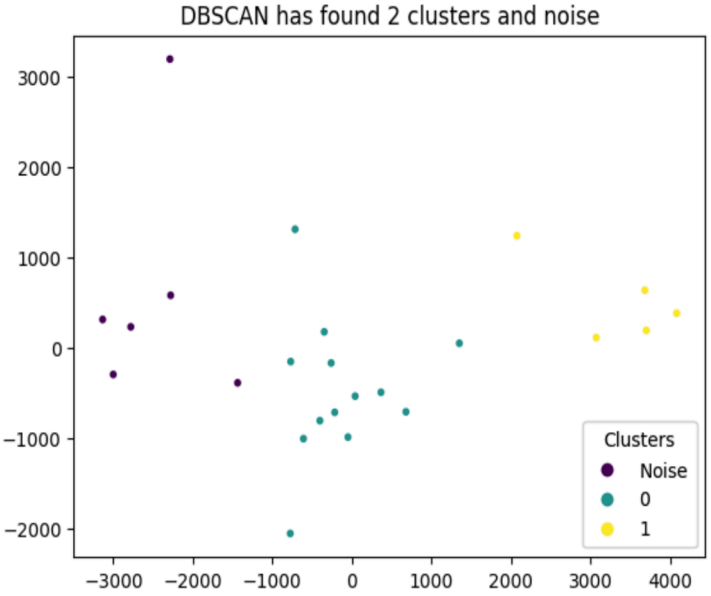

Fig.7. DBSCAN clusters in the analysis of hot water consumption.

In a similar way, intelligent clustering was used to identify periods of reduced nighttime water consumption on working days and weekends. To do this, the data was converted into two separate sets: only working days and only weekends. The results of the study are presented in Table 1. The following research options are shown: 1 – clusters of all cold water, 2 – clusters of cold water on weekdays, 3 – clusters of cold water on Saturday and Sunday, 4 – clusters of all hot water, 5 - clusters of hot water on weekdays, 6 – clusters of

hot water on Saturday and Sunday. The cluster number does not matter. What matters is which clocks are clustered.

Table 1. Distribution of water consumption into clusters.

hour	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
1	-1	**0**	**0**	**0**	**0**	**0**	-1	-1	-1	1	1	1	-1	1	1	1	1	1	1	2	2	3	3	1
2	0	**1**	**1**	**1**	**1**	**1**	0	-1	2	0	0	0	0	0	0	0	0	0	0	2	2	2	2	0
3	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	-1	1	1	2	-1	2	1	1	1	1	1	3	3	-1	1	1
4	0	0	**1**	**1**	**1**	**1**	**1**	-1	0	0	0	0	0	0	0	0	0	0	0	-1	-1	-1	-1	-1
5	0	-1	**1**	**1**	**1**	**1**	-1	-1	0	2	2	-1	2	-1	2	2	2	2	0	-1	-1	-1	-1	-1
6	-1	0	**1**	**1**	**1**	**1**	**1**	**1**	0	2	2	3	-1	3	2	2	2	2	2	2	-1	-1	2	2

Discussion

The data used for the analysis has two advantages. Firstly, they are real, and secondly, these data have already been subjected to visual and statistical analysis by another researcher [12]. Accordingly, it is possible to compare the results. A comparison of the results of determining the night period of water consumption is presented in Table 2.

Таблица 2. Сравнение результатов анализа данных.

Var	**DBSCAN**	**[12]**
1	1-6 h	-
2	1-6 h	2-6 h
3	0-8 h	3-8 h
4	2-7 h	-
5	2-6 h	2-6 h
6	2-8 h	3-8 h

Some problem in the DBSCAN method is the selection of the parameter of the radius of the neighborhood. However, if there is a clustering program, for example, written in Python, it is easy to organize an iterative process of searching for the maximum number of clusters and the minimum amount of noise.

Cluster analysis is used to investigate the nature of the data. It is not used to solve the problem of data forecasting.

Conclusions

Visual data analysis is crude and inefficient. It is especially difficult to analyze data over long-time intervals. Intelligent cluster analysis allows you to identify periods of night water consumption more confidently and reasonably.

Among intelligent clustering methods, the DBSCAN method has several advantages. The method does not require setting the number of clusters, as is done in other methods. The method highlights noise in the data, which can be cleared (if necessary).

Cluster analysis can be applied to other tasks of processing and analyzing water consumption data. For example, you can divide apartments by water consumption level and compare it with the number of residents, the number of water collection points and other data (if available).

The ability to detect noise allows you to look for anomalies in the data to analyze their causes later.

In further research, attention should be paid to the task of restoring and improving the original data.

In the case of mass digitalization of apartment buildings and the mass use of smart meters, cluster analysis can be included in the methodology for processing and analyzing water consumption telemetry data to optimize water resources management.

References

1. Geo Information System of housing and utilities sector. URL: https://dom.gosuslugi.ru/#!/individual-meters
2. The concept of "Smart apartment building". URL: https://lab-cifra.ru/concept
3. S.Yu. Shuvalov, Smart energy saving as part of smart housing and communal services concept. J. Energosberegeniye. **2**. pp. 25-29. (2019)
4. T. E. Akulinushkina, The significance of the use of "Smart Home" technology for the development of housing and communal services in the region. Young Scientist. **18** (256). Pp. 105-109. (2019). URL: https://moluch.ru/archive/256/58586/
5. E. Ignatova, L. Kruglova, Information systems for accounting and data analysis of housing and communal services of apartment buildings. J. Construction and Architecture**. 4**. (2023) DOI: 10.29039/2308-0191-2023-11-4-38-37
6. S. Antsiferov, I. Lushkin, E. Usmanova, Problems of operating the hot water supply system during transfering the hostel to the apartment house status. in Proceedings of the IOP Conf. Series: Materials Science and Engineering. **698**. 066024. (2019)
7. M.L. Croucha, H.E. Jacobs. Defining domestic water consumption based on personal water use activities. J. AQUA — Water Infrastructure. Ecosystems and Society, **70** (7), 1002. (2021) DOI: 10.2166/aqua.2021.056
8. Z. Krippelová, J. Peráčková, Measurement of hot water consumption in apartment building. Urban Water Journal, (2021)
9. J. Nejranowski, W. Szaflik, Hot water consumption time in multi-apartment buildings. Journal of Ecological Engineering. **21**(4), pp.199–202. (2020) https://doi.org/10.12911/22998993/119906
10. A. Chmielewska, Monthly model of hot water consumption. in Proceedings of the E3S Web of Conferences **100**, 1005 0000 (2019) https://doi.org/10.1051/e3sconf /20191000000
11. Kairo Pereira Teodoro da Silva, A. Kalbusch, E. Henning, Guilherme Araújo Lira Menezes, Modeling water consumption in multifamily buildings: a case study in Southern Brazil. (2021) https://doi.org/10.1080/1573062X.2021.1934040

12. Kitaitseva,E. Numerical analysis of hourly consumption of hot and cold water. International Journal for Computational Civil and Structural Engineering. **8** (4). pp. 78-84. (2012)
13. V.S. Ignatchik, S.V. Sarkisov, V.A. Obvintsev, Investigation of coefficients of hourly unevenness of water consumption. J. Water and ecology: problems and solutions. **2** (70). pp. 27-39/ (2017)
14. J. Kim, H. Lee, M. Lee, H. Han, D. Kim, H.S. Kim, Development of a Deep Learning-Based Prediction Model for Water Consumption at the Household Level. J. Water **14**, 1512. (2022) https://doi.org/10.3390/w14091512
15. P. Surendra, H. Deka, N. Rajakumara, Application of Mamdani model-based fuzzy inference system in water consumption estimation using time series. J. Soft Computing - A Fusion of Foundations, Methodologies and Applications. **5**. (2022)
16. L.C.P. Velasco, A.R.B. Granados, J.M.A. Ortega, K.V.D. Pagtalunan, Performance analysis of artificial neural networks training algorithms and transfer functions for medium-term water consumption forecasting. Int. J. Adv. Sci. Appl., **9**, 109–116. (2018)
17. E.V. Odintsova, Analysis and forecasting of time series using computer technologies. Collection of scientific papers of Russian universities "Problems of economics, finance and production management". **52**. pp. 184-186. (2023)
18. A. Tripathi, S. Kaur, S. Sankaranarayanan, L.K. Narayanan, R.J. Tom, Water demand prediction for housing apartments using time series analysis. Int. J. Intell. Inf. Technol. **15**, 57–75. (2019)