Combining Data from Fitness Trackers with Meteorological Sensor Measurements for Enhanced Monitoring of Sports Performance[⋆]

Anna Wachowicz[1][0000−0003−4762−4171], Bożena Małysiak-Mrozek[1][0000−0003−4977−4915], and Dariusz Mrozek[1][0000−0001−6764−6656]

Institute of Informatics, Silesian University of Technology
ul. Akademicka 16, 44-100 Gliwice, Poland
anna.m.wachowicz@gmail.com; bozena.malysiak-mrozek@polsl.pl;
dariusz.mrozek@polsl.pl

Abstract. Systematic analysis of training data has become an inherent element of building a sports condition and preparation for sports competitions. Today's progress in the development of various fitness trackers, smart wearables and IoT devices allows monitoring the level of development of athletic abilities not only for professional athletes but also for sports enthusiasts and sports amateurs. Meteorological conditions prevailing on a given day can significantly affect the effectiveness of training and abilities of a person during sports competitions. However, in order to properly analyze particular sports achievements and the effectiveness of sports efforts, the training data from body sensors should be appropriately combined with weather sensor data. In this paper, we show that, due to approximate nature, this process can be implemented by using the fuzzy join technique.

Keywords: Fuzzy sets, fuzzy logic, data analysis, smart devices, human activity monitoring, IoT, sensors

1 Introduction

In recent years, monitoring of physical activity gained popularity among both sports enthusiasts and cutting-edge technologies fans. As the IoT devices market extends, smartwatches and smart bands are no more out of reach for an average person. They allow monitoring a wide range of activities and parameters – from simply walking and counting the number of steps to controlling advanced

workout with tracking position with GPS signal, recognizing the type of activity and measuring the level of body hydration. People who do sport professionally use wearables as one of many work tools. Sport amateurs also want to monitor and improve their training results, however, they do not have as many devices as professionals nor advice from a sports coach. Wearables, along with corresponding mobile web applications, help amateurs to control the progress of training quality. These applications also provide many useful reports and allow to share the results of the training with other people in social media or online forums for sports enthusiasts.

As online communities gain more and more popularity, people use them for exchanging experiences. They describe the observations and ask various questions. The popular topic is how the weather affects the training performance. Although this issue is covered in many scientific papers, the results usually apply to the professionals. Moreover, from the technical point of view, the answer to this question necessitates joining the training data from wearable devices with weather parameters. Weather parameters can be retrieved from online Web services. However, they are usually provided hourly, while the training may start at any time. This requires the implementation of some flexibility while joining both types of data. This flexibility can be modeled by fuzzy sets.

In this paper, we apply a fuzzy join technique while flexibly combining training data with weather conditions. The fuzzy join is performed as a preparation step for further data analysis allowing to investigate the impact of weather conditions on the training performance. The training data are gathered with the use of smartwatch used for monitoring the running training. With the use of the fuzzy join algorithm, data are combined with the most accurate weather measurements retrieved from an online weather service. The output is then loaded to the Microsoft Azure cloud data center where the analysis can be conducted. In this paper, we present the outcome of the analysis of the impact of weather features such as temperature, air humidity, and pressure or wind speed on the effectiveness of amateurs' running. We also show the duration of particular operations related to data pre-processing performed by the fuzzy join module.

2   Related Works

IoT technologies and wearable devices are frequently applied in monitoring various physiological parameters in terms of people's health state and the activities they perform. World literature gives many examples of such applications. For example, in [14] Yamato shows the platform for the analysis of posture and fatigue on the basis of data gathered from electrocardiograph and accelerator. The analysis can be performed in a smartphone or in the Cloud. Lara and Labrador [8] presented the survey on the usage areas of the human activity recognition (HAR) systems. The authors mention the types of activities recognized by *state-of-the-art* HAR systems which may cover areas such as daily activities, fitness, military, ambulation or transportation. They widely explain the feature extraction step distinguishing the techniques used for measured attributes that are grouped into

three sets: *time-domain, frequency-domain, and others.* The most popular methods for the time-domain group are mean, standard deviation, variance, mean absolute deviation and correlation between axes, whereas the techniques for the frequency-domain set are Fourier Transform and Discrete Cosine Transform. A review of various types of wearable devices and inner sensors used for observing human activities is shown in [10]. In the paper, Mukhopadhyay describes also the architecture of monitoring systems and networks containing various sensors and devices. In [12], Toh et al. draw attention to the design issues for wearable devices. As rightly pointed out by authors, in order to meet all users' requirements smart wearables should be light and low-energy consuming. Devices should be able to send data to the paired application, so that users can access the history of performed activity, like training, or be informed that there is something wrong with a monitored person, regardless of whether it is a sports amateur or an elderly that wears the smart band. In [11], authors describe wearable and implantable sensors for distributed mobile computing. They also present the difficulties and complications that may happen while using wearables. Similarly to the system that we have built and present in sect. 4, the system presented in [11] aims to analyze the impact of weather conditions on the running training parameters.

The relationship between weather conditions and the training performance is not only intuitive but also confirmed by published research. In [4], [13] and [5], the authors study the impact of the temperature and other atmospheric measurements on the performance of marathon runners. All of these articles describe the optimal temperatures for the best running performance among men and women which are between 10℃-15℃. One of the most recent studies that investigate the relationship between various physical exercises (including the running) and weather conditions are presented in [3]. The author analyzes seasonal upper-body strength resistance and running endurance performance, and studies if there are any relationships between the efficiency of these activities and weather conditions. Comparing to previous papers, the running distance is 5 km and the research shows that participants of the conducted experiment gained better results in summer and spring, which are hotter seasons. In our paper, we not only investigate these relationships in the data analysis step, but also focus on the preparation of data with the use of fuzzy join technique before the main analysis begins.

The term *fuzzy join* is widely used in scientific literature but may have different meanings and applications. Many of published papers, including [1], [2], [7], [15], use the term while combining data sets on the basis of flexible character data matching and string similarity with the use of various distance functions, like the Hamming distance. Meanwhile, articles [6] by Khorasani et al. and [9] by Małysiak-Mrozek et al. show how to flexibly combine big data sets with the use of fuzzy join operation by applying the fuzzy sets-based techniques on the numerical attributes. In our paper, we show how we utilize this idea on the numerical values of the time attribute while combining sensor data from a wearable device (a smartwatch) with meteorological data from weather sensors.

3   Fuzzy Sets for Flexibility

Fuzzification of selected attributes of sensor readings may introduce some flexibility while joining various data collections. Fuzzy sets can play a particular role here. The fuzzy sets theory assumes that the membership degree $\mu(x)$ of an object to the set A may be represented with countless values within the unit interval [0,1] [16]. This stays in contrast to the classical set theory that assumes that membership of an object to a set is bivalent – the object either belongs to the set or does not belong to it.

Assuming that X is the universe of points (objects) and x is an element of X, the fuzzy set A in X is defined as an ordered collection of pairs:

$$A = \big\{(x, \mu_A(x)) | x \in X \big\}, \tag{1}$$

where μ_A is the *membership function* defining the set A, and $\mu_A(x)$ is the *membership degree* of the element x to the set A, which takes a value from 0 to 1.

Graphically, the *membership function* is usually represented as a triangular or trapezoidal function. Triangular function is used when there is only one situation such as the value of membership is equal to 1. Fig. 1a shows sample fuzzy set *training time around 9:00 AM* defined with the use of the triangular membership function. This type of characteristic function is defined by three parameters a, b and c, where $a \leq b \leq c$, as follows:

$$\mu_A(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & c < x \end{cases} \tag{2}$$

On the other hand, the trapezoidal characteristic function is described by four parameters a, b, c and d, where $a \leq b \leq c \leq d$, and is defined as follows:

$$\mu_A(x; a, b, c, d) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & b < x \leq c \\ \frac{d-x}{d-c}, & c < x \leq d \\ 0, & d < x \end{cases} \tag{3}$$

Fig. 1b shows a sample fuzzy set *morning time* defined with the trapezoidal membership function, together with the calculation of the membership degree for the beginning of sample sports trainings.

4   Cloud-based Monitoring and Data Analysis System

Training data processing and further analysis are performed in the Cloud-based system presented in Fig. 2. Runners are equipped with the Garmin smartwatch,
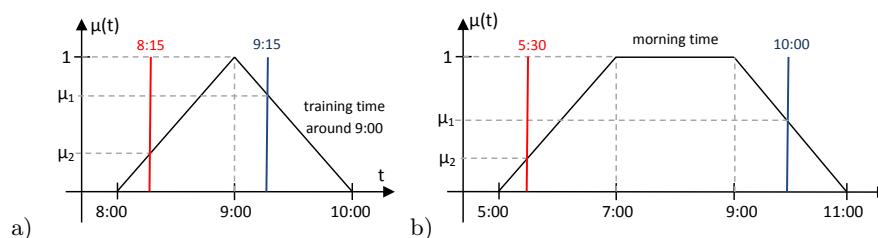
Fig. 1: Sample fuzzy sets defined for the training time: a) *training time around 9:00 AM* defined with the use of the triangular membership function, b) *morning time* defined with the trapezoidal membership function, and calculation of the membership degree for the beginning of sample trainings.

which gathers training parameters with the use of various sensors. In our research, we focused on the effectiveness of running/jogging for a selected group of 15 sports amateurs jogging systematically within a period of one year.

The Garmin smartwatch allowed to collect the following data on the performed training:

- date and time of the training
- training duration,
- heart rate,
- distance,
- calories burned,
- raw GPS data for determining the route,
- running cadence (number of steps a runner takes per minute),
- average speed,
- training type.

Data produced by the smartwatch during the training are collected as training data files stored in the *.tcx* format. *.tcx* is the acronym for Training Center XML introduced by Garmin Company. The format enables the exchange of GPS tracks as an activity with parameters of monitored training, including running, biking, and other forms of activity. The data produced by the smartwatch create a collection of data *at rest*, i.e., these are not constantly monitored data streams, but historical, offline data that are sent to the Cloud after the training is finished.

The Edge gateway module is responsible for transmitting the data to the Cloud. However, before it happens, the training data are combined with weather conditions for the day of performed training. To this purpose, we invoke a URL request to the Dark Sky Web service by using appropriate API (Application Programming Interface). The Web service accepts the date, time and coordinates on the input and returns the following parameters describing the weather conditions on the output:
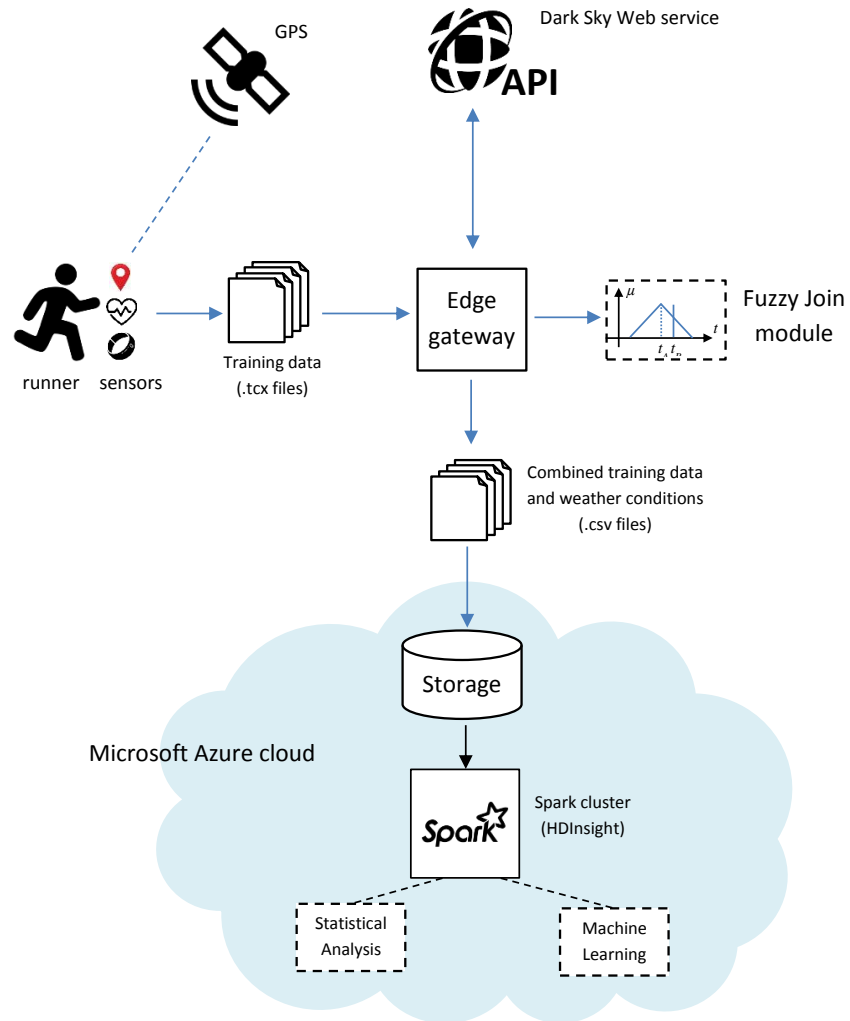
- temperature – real and apparent,

Fig. 2: Architecture of the Cloud-based system for training data analysis.

- air humidity,
- air pressure,
- wind speed,
- dew point,
- cloud cover,
- UV index.

Data collected by the smartwatch are supplemented by the most appropriate meteorological data from weather sensors provided by the Dark Sky Web service in the Edge gateway. The gateway is a device, like an electronic unit, mobile phone or a field computer, which pre-process the data and transmits the data to the Cloud data center. The training data and meteorological sensor measurements are combined by the Fuzzy Join module. The Fuzzy Join module is responsible for data preparation, supplementation and combining before sending the data for further analysis. This phase consists of merging data collected by wearable sensors and atmospheric conditions measurements provided by an online weather Web service (available through appropriate API). The module uses the idea of the fuzzy join with the *fuzzy umbrella* presented in Section 5, to flexibly combine the data from various data sources on the basis of the sensor reading times. The combined data are stored as *.csv (comma-separated) values* files. This format enables to store tabular data in plain text.

Due to large volumes of the training data that can be analyzed and wide scaling capabilities, the analysis phase is performed with use of the Apache Spark engine in the HDInsight cluster in the Microsoft Azure cloud platform. We used Apache Spark 2.3.0 on the HDInsight cluster 3.6.0. Within the Spark-based data analysis we can perform statistical analysis of the influence of the weather conditions on the performance of the training. For this purpose, we calculate Pearson's correlation coefficient. The training efficiency is measured by average running speed and the number of calories burnt during exercises. With the Machine Learning models, like Linear Regression, we can predict the impact of weather conditions on the quality and efficiency during the training (running/jogging).

5 Joining Data Collections through Fuzzification

In its operational lifecycle, the Edge gateway transmits data to the data center located in the Cloud. Training data stored in *.tcx* files and meteorological conditions retrieved from the Dark Sky Web service API are joined within the Fuzzy Join module, which extends the capabilities of the Edge gateway. The data describing particular training collected with the use of a wearable device (a smartwatch) contain much information about training parameters, including the time stamp and coordinates of the location where the training begins. This information is used while retrieving data from the Web service providing weather conditions. Weather conditions are retrieved by specifying the date, time, and coordinates of the training. They are delivered as *.json* objects containing hour-by-hour daily measurements (air temperature (real and apparent),

the percentage level of air humidity, dew point, wind speed, air pressure, and others). The aim of the fuzzy join is to find the most accurate weather conditions based on the time the training begins.

The fuzzy join algorithm (Algorithm 1) implemented in the Fuzzy Join module calculates the value of a membership degree of the time of training to the fuzzy sets U created for each full hour of weather measurement t_{m_j}:

$$\forall_{t_{m_j}} \quad U_j = \{(t, \mu_U(t))|t \in T\}, \tag{4}$$

where $t \in T$ represents all possible time points on the timeline T, and:

$$Supp(U_j) = (t_{m_j} - 1hour, t_{m_j} + 1hour), \tag{5}$$
$$Core(U_j) = \{t_{m_j}\}, \tag{6}$$

where $Supp(U_j) = \{t \in T, \mu_U(t) > 0\}$ is the support of the fuzzy set U_j, and $Core(U_j) = \{t \in T, \mu_U(t) = 1\}$ is the core of the fuzzy set U_j. Such defined fuzzy sets U_j are called *fuzzy umbrellas* and they may cover various training times $t_i \in T$. The fuzzy join algorithm for combining training data with weather conditions is presented Algorithm 1.

Algorithm 1 Combining training data with weather conditions with the use of Fuzzy join algorithm

---

1: **procedure** FUZZYJOIN(t_i, t_{m_j}) ▷ Meteorological conditions are read at full hours
2:     **for each** training time t_i **do**
3:         Convert the beginning of training (t_i) into seconds
4:         Convert the time of weather conditions measurement t_{m_j} into seconds
5:         Find neighbouring times of meteorological measurements $t_{m_{j-1}}, t_{m_{j+1}}$ for the time t_{m_j} and convert them into seconds
6:         $\mu_{U_{j-1}} \leftarrow 1 - \frac{|t_{m_{j-1}} - t_i|}{3600}$
7:         $\mu_{U_j} \leftarrow 1 - \frac{|t_{m_j} - t_i|}{3600}$
8:         $\mu_{U_{j+1}} \leftarrow 1 - \frac{|t_{m_{j+1}} - t_i|}{3600}$
9:         **for each** $\mu \in \{\mu_{U_{j-1}}, \mu_{U_j}, \mu_{U_{j+1}}\}$ **do**
10:             **if** $\mu \notin \langle 0; 1 \rangle$ **then**
11:                 Skip μ
12:             **end if**
13:         **end for**
14:         Join training data for time t_i with meteorological sensor readings from the time $t_m \in \{t_{m_{j-1}}, t_{m_j}, t_{m_{j+1}}\}$ for which $\mu_{t_m} = \max\{\mu_{U_{j-1}}, \mu_{U_j}, \mu_{U_{j+1}}\}$
15:     **end for**
16: **end procedure**

For each training (starting at time t_i), the algorithm converts the time of weather conditions measurement t_{m_j} (Fig. 3) and the beginning of training t_i into seconds (e.g., 10 AM for weather conditions measurement is converted to

36000 s). In the next step, it finds times of meteorological measurements neighbouring to the t_{m_j} $(t_{m_{j-1}}, t_{m_{j+1}})$ and convert them into seconds. Then, it computes the differences between the weather conditions measurement hours $(t_{m_{j-1}}, t_{m_j}, t_{m_{j+1}})$ and the beginning of the training (t_i) (values in seconds). It takes the absolute values of the results and computes the ratio of the values to the number of seconds in one hour (3600 s). For each of the times of weather conditions measurement $(t_{m_{j-1}}, t_{m_j}, t_{m_{j+1}})$ the value of membership function (μ_U) is calculated as the difference of 1 and the calculated ratio. If the calculated value is not between 0 and 1 it has to be rejected, as it is out of the range of the membership degree. Finally, the algorithm takes the maximum of all values of the calculated membership degree and combines training data with those meteorological conditions for which the membership degree is the highest.

The use of fuzzy join allows finding the best matching of the weather conditions to the training as they are chosen by the nearest hour of measurement. The concept of the fuzzy umbrella is presented in Fig. 3. On the timeline, there are hours of weather conditions measurements from sensors in meteorological stations. On the μ axis, there are placed values of membership function computed for each hour of weather measurement.
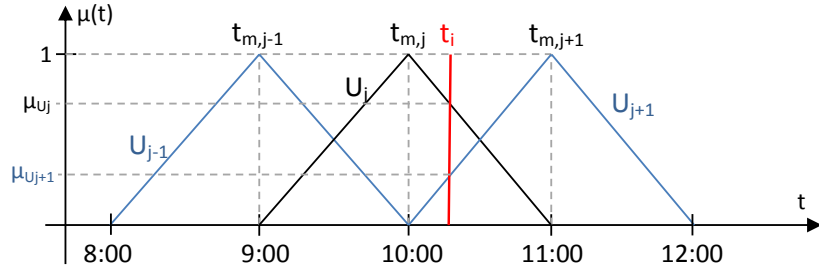

Fig. 3: Fuzzy umbrellas defined for the time of weather sensors measurements $(t_{m_{j-1}}, t_{m_j}, t_{m_{j+1}})$ and calculation of membership degree for training time t_i.

6   Experimental Results

Our experiments covered (1) the analysis of correlations between training parameters like distance, time, average speed, etc. and meteorological measurements, and (2) verification of the performance of the fuzzy join operation. We decided to conduct the experiments for each person independently as everyone had different running habits and experiences. The training data that we collected concerned running on an average distance of 10 km. The runners were amateurs with different level of running experience, various running habits and frequency. Among them, there were men and women, at the age between 20 and 55. The standard deviation of the distance was approximately 1 km whereas for the about 1-hour

training standard deviation value was near to 7 minutes. The running parameters had normal distribution. Weather conditions, in which the trainings were performed also varied. For example, the range of measured temperatures was from -15℃ up to 32℃, while the air humidity was between 26% and 100%.

The results for each of the analyzed set of data (workout) varied. The average values of correlations are shown in Table 1. The correlation measure was presented with the absolute value (without direction) of the Pearson correlation coefficient defined as:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{7}$$

where X, Y are random variables, $cov(X,Y)$ is the covariance, and σ_X, σ_Y are the values of standard deviations of X and Y.

Table 1: Ranges of linear correlations between running parameters and meteorological conditions

ρ	Duration	Distance	Calories	Average speed
Temperature	$0.022 - 0.224$	$0.053 - 0.224$	$0 - 0.196$	$0.046 - 0.286$
Air humidity	$0.045 - 0.386$	$0.046 - 0.205$	$0.014 - 0.174$	$0.023 - 0.313$
Air pressure	$0.006 - 0.132$	$0.021 - 0.228$	$0.014 - 0.350$	$0.026 - 0.255$
Wind speed	$0.019 - 0.500$	$0.021 - 0.364$	$0.019 - 0.292$	$0.012 - 0.646$

The results of the analyzed correlations are diversified for each of the runners. The widest range may be observed for the interactions between training parameters and the wind speed – from only 1%, which means no correlation at all, up to 65%. Large values of the Pearson coefficient were noticed for the air humidity – between this feature and the duration of the training, the correlation for one of the persons reached nearly 39%. For the temperature, opposite to what could be expected, the results were not so satisfying, however, the maximum value of the correlation between the temperature and average speed exceeded 28%. The summary of the average results for all of the runners is presented in Table 2.

On average, the results of our experiments show that there are correlations between trainings parameters and weather conditions. Usually, the average value of the Pearson coefficient is between 10% and 20%. Although the value of the correlation seems to be low, particular values for individual athletes are strongly differentiated. This shows that some of them are able to perform the training assumptions regardless of the prevailing weather conditions. The strongest correlations may be noticed between running parameters and the wind speed while the weakest correlation is for the air pressure. Despite the results seem satisfying, it is important to remember that for every runner the correlations were different – for ones they were stronger, for others – negligible. Moreover, there

Table 2: Average correlations between running parameters and meteorological conditions

ρ	Duration	Distance	Calories	Avg speed
Temperature	0.099	0.144	0.115	0.138
Air humidity	0.153	0.112	0.073	0.144
Air pressure	0.063	0.088	0.117	0.106
Wind speed	0.205	0.119	0.125	0.170

are also many other factors that may have an impact on the results but were not analyzed during our experiments – like the quality and time of the meal before training, the overall condition of the runner or even their health state.

In the second series of experiments, we tested how the fuzzy join operation affects the performance of the data processing on the Edge gateway (Fig. 2). The Edge gateway and the Fuzzy Join module combine data before sending the data to the Cloud for further analysis. The performance tests were conducted on PC station with 8GB RAM and processor Intel(R) Core(TM) i7-3537U CPU @ 2.00 GHz, controlled by the 64-bit Windows operating system. To obtain the most reliable results, during experiments no other applications were running on the machine.

We tested the performance of the fuzzy join on the data set consisting of 850 files (*.tcx*) with data from real training (running/jogging). Results of our experiments presented in Figs. 4 prove that the execution time of the Fuzzy Join algorithm (Fig. 4d) is negligible in comparison to other operations. On average, it takes less than 0.20 milliseconds which is 0.0015% of overall time spent on processing a particular (one out of 850) data file.

During the operation named *Load file* (Fig. 4a) the *.tcx* training data file is opened and all of the needed data are retrieved and prepared for further analysis. The time of this operation is related to the size of the file which depends on the number of measurements. Duration of the URL request (Fig. 4b) to the Web service is determined by the network efficiency and traffic. However, the average time of this step is much shorter than the time of the previous operation. In Fig. 4c we also presented the duration of saving combined training and weather data to an output file performed by the Fuzzy Join module. The time of this operation is on average about 5 milliseconds, which is about 0.44% of the overall time needed for processing a single file. Although the maximum time attracts attention due to its high value, the values of other measures point that it is rather a singular outlying result than a frequent issue.
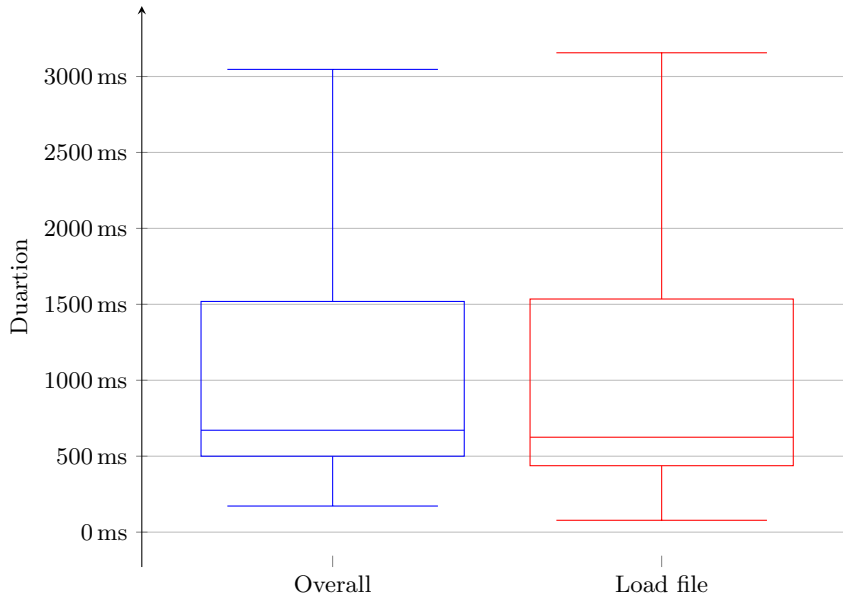
(a) Overall processing and loading file operations
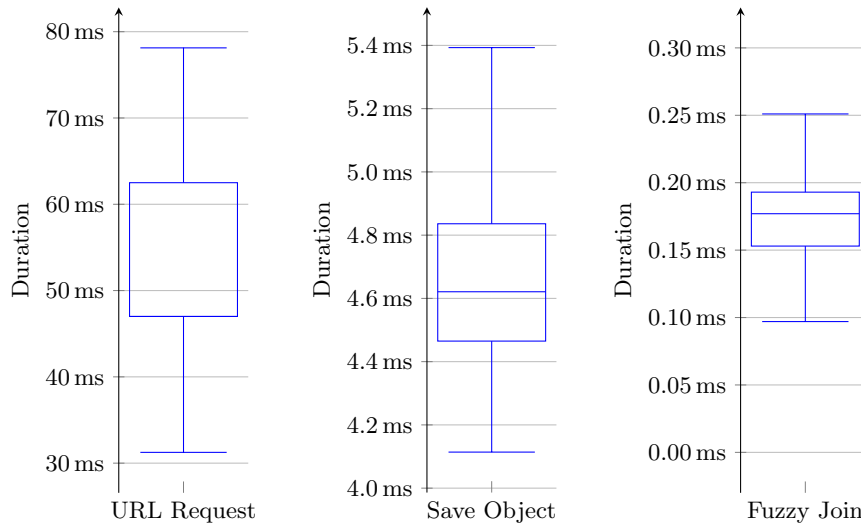
(b) URL Request   (c) Save Object   (d) Fuzzy Join

Fig. 4: Duration of particular operations on one set of the training data.

7   Conclusion

Application of fuzzy sets while joining sensor data on the Edge gateway allows not only to flexibly combine training data collected with a smartwatch with weather conditions but also delivers information on the compatibility of combined time moments for both data sets. This can be an important factor while analyzing the correlation between the performance of the training and the weather parameters, and planning tactics for future training activities.

The fuzzy join algorithm operates on numerical representation of time stamps, similar to the solution presented in [9] that operates on numbers in Big Data cloud environments, and in contrast to works [1], [2], [15] that operate on strings. Moreover, likewise it was presented in works by Yamato [14], Revathi Pulichintha Harshitha et al. [11], Małysiak-Mrozek et al. [9] the whole solution is built upon the cloud infrastructure, which allows for scalable data analysis. However, the fuzzy join is performed on the Edge device, which reduces the amount of the processing work performed in the Cloud.

The data processing performed on the Edge gateway consists of many operations, out of which the fuzzy join takes the least time. This shows that the operation will not introduce any significant delays in data pre-processing. Meanwhile, it is very important, since it supplements the existing training data with additional information that may shed new light on the analyzed data.

During our preliminary data analysis on the Spark cluster in the Cloud, we could notice the existence of correlations between some weather measurements and running efficiency. Some of them were strong and some of them were weak – they largely depend on a person who does the activity. Other factors like talent or psychical strength, which are also mentioned in [13] and [5] but were out of our analysis, could contribute to the results. These elements are difficult to determine. An important issue, also hard to classify, is the level of runner's experience and sports form during a particular workout. Still, we believe that the presence of correlations between meteorological measurements and training parameters is an interesting matter and worth further studies in our future works.

References

1. Afrati, F.N., Sarma, A.D., Menestrina, D., Parameswaran, A., Ullman, J.D.: Fuzzy joins using MapReduce. In: 2012 IEEE 28th International Conference on Data Engineering. pp. 498–509 (04 2012). https://doi.org/10.1109/ICDE.2012.66
2. Deng, D., Li, G., Hao, S., Wang, J., Feng, J.: MassJoin: A MapReduce - based method for scalable string similarity joins. In: 2014 IEEE 30th International Conference on Data Engineering. pp. 340–351 (03 2014)
3. Dhahbi, W.: Seasonal weather conditions affect training program efficiency and physical performance among special forces trainees: A long-term follow-up study. PLoS ONE **13**(10) (10 2018). https://doi.org/10.1371/journal.pone.0206088
4. El Helou, N., Tafflet, M., Berthelot, G., Tolaini, J., Marc, A., Guillaume, M., Hausswirth, C., Toussaint, J.F.: Impact of environmental parameters on marathon running performance. PLoS ONE (05 2012). https://doi.org/10.1371/journal.pone.0037407

5. Ely, M.R., Cheuvront, S.N., O., R.W., J., M.S.: Impact of weather on marathon-running performance. Medicine & Science in Sports & Exercise **39**(3), 487–493 (2007)

6. Khorasani, E.S., Cremeens, M., Zhao, Z.: Implementation of scalable fuzzy relational operations in MapReduce. Soft Computing **22**(9), 3061–3075 (May 2018). https://doi.org/10.1007/s00500-017-2561-3, `https://doi.org/10.1007/s00500-017-2561-3`

7. Kimmett, B., Srinivasan, V., Thomo, A.: Fuzzy joins in MapReduce: An Experimental Study. In: Proc. VLDB Endow. pp. 1514–1517 (08 2015). https://doi.org/10.14778/2824032.2824049

8. Lara, O.D., Labrador, M.A.: A survey on human activity recognition using wearable sensors. IEEE Communications Surveys and Tutorials **15**(3), 1192–1209 (2013)

9. Małysiak-Mrozek, B., Lipińska, A., Mrozek, D.: Fuzzy join for flexible combining big data lakes in cyber-physical systems. IEEE Access **6**, 69545–69558 (2018). https://doi.org/10.1109/ACCESS.2018.2879829

10. Mukhopadhyay, S.C.: Wearable sensors for human activity monitoring : A review. IEEE Sensors Journal **15(3)**, 1321–1327 (2015)

11. Revathi Pulichintha Harshitha, S., Narramneni, P., Raghavee, N.S.: Body sensor using internet of things (iot). ARPN Journal of Engineering and Applied Sciences **13**(8) (04 2018)

12. Toh, W.Y., Tan, Y.K., Koh, W.S., Siek, L.: Autonomous wearable sensor nodes with flexible energy harvesting. IEEE Sensors Journal **14**, 2299–2306 (2014)

13. Vihma, T.: Effects of weather on the performance of marathon runners. Int J Biometeorol **54**(3), 297–306 (05 2010). https://doi.org/10.1007/s00484-009-0280-x

14. Yamato, Y.: Proposal of vital data analysis platform using wearable sensor. In: Proceedings of the 5th IIAE International Conference on Industrial Application Engineering (2017)

15. Yan, C., Zhao, X., Zhang, Q., Huang, Y.: Efficient string similarity join in multi-core and distributed systems. PLoS ONE **12**(3), 1–16 (2017)

16. Zadeh, L.A.: Fuzzy sets. Information and Control **8**, 338–353 (1965)