Computers in Human Behavior 53 (2015) 427–434

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full Length Article

A comprehensive study of parameters in physical environment that impact students' focus during lecture using Internet of Things

Ana Uzelac *, Nenad Gligoric, Srdjan Krco

University of Belgrade, Faculty of Organizational Sciences, Jove Ilica 154, Belgrade, Serbia

ARTICLE INFO

Article history:
Received 19 September 2014
Revised 9 April 2015
Accepted 14 July 2015
Available online 23 July 2015

Keywords:
Smart classroom
Internet of Things
Classification
Machine learning
Pattern recognition
Student's focus

ABSTRACT

We describe and analyze the impact of several parameters of the physical environment in a classroom on students' focus, where the term "focus" refers to the students' subjective feeling of their ability to concentrate on a lecture at a given moment. The primary goal is to identify those parameters that significantly affect students' focus during the lectures. We had measured several parameters in a real classroom environment using different low-cost smart devices. The research is based on the dataset collected from 14 recorded lectures attended by 197 students. We had measured five parameters of the physical environment and extracted 22 features from the lecturer's voice. After analyzing collected measurements, we had identified eight parameters that have shown to have statistically different values for "focused" and "not focused" segments. We used obtained dataset to test different classifiers and their ability to correctly classify "focused" against "not focused" segments of the lectures. We found out that AdaBoost M1 classifier had the best overall recognition accuracy (86.78%). After performing additional series of trials we identified three parameters that could be removed from the original dataset without changing classifier's accuracy, which left us five uncorrelated parameters that have shown to have significant impact on students' focus.

1. Introduction

Previous studies (Felder & Brent, 1999) have shown that students cannot stay fully focused throughout the lecture. It has been proved that student's attention begins to decrease approximately 10 min after the beginning of a lecture. At the end of a lecture, students remember 70% of the information presented in the first ten, and only 20% of the information presented during the last ten minutes of a lecture (Hartley & Davies, 1978). Therefore, detecting parts of the lecture where students' focus is decreased is important as some actions can be performed in order to stimulate their focus. If students are focused on the lecture most of the time, they would remember more information presented, and their benefit from the lecture would be maximized.

There are many studies that investigated the influence of different parameters on students' performance and achievements by comparing their results received after lecture (Bako-Biro, Clements-Croome, Kochhar, Awbi, & Williams, 2012; Bronzaft & McCarthy, 1975; Coley, Greeves, & Saxby, 2007; Crook & Langdon, 1974; Downs & Crum, 1978; Evans & Maxwell, 2007; Ito, Murakami, Kaneko, & Fukao, 2006; Johnson, 2001; Kyzar, 1977; Molhave, Bach, & Federsen, 1986; Murakami, Kaneko, Ito, & Fukao, 2006; Otto, Hudnell, House, & Molhave, 1992; Shaughnessy, Haverinen-Shaughnessy, Nevalainen, & Moschandreas, 2006; Wargocki & Wyon, 2007). It has been shown that there is a significant relationship between students' ability to concentrate and their academic performance (Egong, 2014), which indirectly indicates that the same parameters may have a high impact on students' focus as well. However, none of the previously conducted studies have considered the direct effect of these parameters on students' focus, as it requires their instant feedback. Additionally, studies are rarely investigating the influence of more than one parameter at the same time, and most experiments were conducted in the laboratory environments. To the best of our knowledge, there is no comprehensive study that has tried to simultaneously identify and analyze parameters that have a significant impact on students' focus in the real classrooms.

The development of different technologies results in changes and enhancements of the educational process as well. For example, education has largely been influenced by the ICT development, resulting in the emergence of different e-learning platforms, virtual learning environments, tele-education systems, etc. Therefore, it is expected that the recent emerge of Internet of Things (IoT) will

* Corresponding author.
E-mail address: ana.uzelac@sf.bg.ac.rs (A. Uzelac).

http://dx.doi.org/10.1016/j.chb.2015.07.023
0747-5632/

change the teaching and learning process as well. As this concept is new, many standards for its key components are still missing. One of the organizations that promotes a unified approach to the development of technical standards defines Internet of Things "as a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies" (ITU-T, 2015). Everything is based on a "thing" which can be defined as "object of the physical world (physical things) or the information world (virtual things), which is capable of being identified and integrated into communication networks" (ITU-T, 2015). General device is defined as a "device that has embedded processing and communication capabilities and may communicate with the communication networks via wired or wireless technologies", including "equipment and appliances for different IoT application domains, such as industrial machines, home electrical appliances, and smart phones" (ITU-T, 2015). The range of new applications based on the IoT technology is broad and diverse, i.e. e-health, traffic, environmental monitoring, smart homes, smart classrooms, etc. This paper focuses on using IoT in smart classrooms. Smart classrooms can be defined as intelligent environments equipped with an assembly of many different kinds "of hardware and software modules such as projectors, cameras, sensors, face recognition module", and many more (Xie, Shi, Xu, & Xie, 2001). In our case, a smart classroom is equipped with a set of sensors able to monitor parameters of the physical environment (for example CO_2, temperature, humidity, noise) and a Bluetooth headset used to capture lecturer's voice. The aim of this study is to identify parameters of the physical environment in a classroom and evaluate their influence on students' focus. Selected parameters will be later used to implement smart classroom system that would be able to determine in real-time if the classroom environment is optimized to maximize student's ability to concentrate on a lecture at a given moment.

The main contributions of this manuscript are: (1) An innovative approach to analyze the impact of different parameters in the physical environment on students' focus, (2) Identification and the comprehensive analysis of the parameters in the physical environment that influence students' focus, (3) to the best of our knowledge this is the first attempt to measure, analyze and correlate features extracted from the lecturer's voice with the students' focus.

1.1. Literature review

Nowadays learning is becoming more interactive and modern classrooms are expected to be more student-centric. Learning Management System (LMS) is continually being improved by applying innovations from ICT field, such as integrating m-learning (Bogdanovic, Barac, Jovanic, Popovic, & Radenkovic, 2014), cloud computing (Despotovic-Zrakic, Simic, Labus, Milic, & Jovanic, 2013), or gLearning (Lytras & Ordoñez de Pablos, 2011). LMS is opening to Personal Learning Environment (PLE) (García-Peñalvo, Conde, Alier, & Casany, 2011), where PLE represents rather a new approach to the use of new technology in learning than a piece of software (Attwell, 2007). PLE is learner-centric and enables learners to have the control over the learning environment. Proposed PLE frameworks uses different technologies, such as mobile phones (Attwell, Cook, & Ravenscroft, 2009), Web 2.0 tools (Kompen, Edirisingha, & Monguet, 2009; Rahimi, Van den Berg, & Veen, 2015), distributed Web 2.0 tools (Juarros, Ibáñez, & Crosetti, 2014), social semantic web technologies (Halimi, Seridi-Bouchelaghem, & Faron-Zucker, 2014), and cloud services (Rizzardini, Linares, Mikroyannidis, & Schmitz, 2013). Furthermore, some researchers tried to blend personalized and conversational learning methods in classroom contexts (Atif, 2013) while others proposed a service-based approach to define mobile personal learning environments that facilitate communication with institutional learning platforms (Conde, García-Peñalvo, Alier, & Piguillem, 2013).

There are still very few studies that use IoT in the learning environments. Applications are mostly related to using technologies such as RFID or NFC for locating students and calculating their attendances (Chang, 2011; Shen, Wu, & Lee, 2014). In another application, IoT is used in synergy with crowdsourcing to create a model for smart e-learning environment, where students can provide preferred values of environmental variables that can later be used for creating optimal learning environment (Simic, Stavenovic, & Djuric, 2014).

Another smart classroom environment that is based on IoT technology presents a system that is capable to detect the level of students' interest in near real-time with the accuracy of 80% (Gligoric, Uzelac, Krco, Kovacevic, & Nikodijevic, 2015). During the experiment, the behavior of the students was monitored using a camera and a broadband microphone while lecturer's activity was measured by an accelerometer (built in a smartphone placed in his/her pocket). The stress in this study was on monitoring students and their activities while in the current work we have focused on monitoring environmental parameters. In addition, the present study is oriented to determine the impact of different environmental parameters on students' focus that will altogether with the previously determined level of students' interest enable us to better assess the lecture quality.

Another smart classroom environment related to this study is a classroom equipped with emotion monitoring system which is able to detect students' attention and emotion in real time (Luo, Zhou, Wang, & Shen, 2009). Student's attention is recognized by detecting and analyzing student's eye movement while student's emotion is recognized by short and long term features of speech. The system is able to give the lecturer an instant feedback if students are actively involved in the presentation. It is strictly designed for distance learning and is not intended to be used in "face-to-face" teaching.

There are few studies that investigate or review influence of more than one parameter on student's concentration, performance and/or achievements (Howarth & Hoffman, 1984; Mendell & Heath, 2005; Wargocki & Wyon, 2007). One such study investigated the influence of different weather variables on concentration; it was concluded that three predictor variables for concentration, in order of importance, were: humidity, temperature, and hours of sunshine (Howarth & Hoffman, 1984).

Different studies have been conducted to find the relationship between one of the parameters of the physical environment and students' performance or achievements. Parameters that have been explored so far include temperature (Pepler & Warner, 1968; Pilman, 2001; Schoer & Shaffran, 1973; Wargocki & Wyon 2007; Wyon, 1970), air quality (Bako-Biro et al., 2012; Coley et al., 2007; Ito et al., 2006; Molhave et al., 1986; Murakami et al., 2006; Otto et al., 1992; Shaughnessy et al., 2006; Wargocki & Wyon, 2007), and environment noise (Bronzaft & McCarthy, 1975; Crook & Langdon, 1974; Downs & Crum, 1978; Evans & Maxwell, 2007; Johnson, 2001; Kyzar, 1977).

Numerous studies confirmed the negative impact of inadequate temperature on student's performance (Pepler & Warner, 1968; Pilman, 2001; Schoer & Shaffran, 1973; Wargocki & Wyon, 2007; Wyon, 1970). Other studies are oriented to air quality, where the term "air quality" refers to the existence of specific gases or volatile organic compounds (VOC), amount of CO_2 as well as ventilation rates that supply a classroom with the outdoor air. A great number of studies support the statement that either low ventilation rate or high level of CO_2 has negative impact on student's performance (Bako-Biro et al., 2012; Coley et al., 2007; Ito et al., 2006;

Murakami et al., 2006; Shaughnessy et al., 2006; Wargocki & Wyon, 2007). Some works have investigated the influence of VOC on performance (Molhave et al., 1986; Otto et al., 1992) and concluded that exposure to the higher values of VOC impaired subjects' ability to concentrate (Molhave et al., 1986), while lower values of VOC had no influence (Otto et al., 1992).

Noise is one of the most investigated parameters that affect the learning process, and many experiments have been conducted to find the relationship between the noise and the student's ability to perform various cognitive tasks. It is generally accepted that any kind of noise has a negative impact on academic performance (Bronzaft & McCarthy, 1975; Crook & Langdon, 1974; Downs & Crum, 1978; Evans & Maxwell, 2007; Johnson, 2001; Kyzar, 1977). When noise levels are high, a lecturer is frequently interrupted and forced to repeat some parts of a lesson that leads to time lost (Crook & Langdon, 1974; Kyzar, 1977). In addition, noise disables transfer of information between a teacher and students. It has been confirmed that the level of noise, more specifically – echo – has a significant impact on students' concentration and could lead to the lesson misinterpretation (Johnson, 2001). Additionally, it has been shown that the student needs more effort to process teacher's voice when the noise level is high (Downs & Crum, 1978).

From the previous work, it can be concluded that there are similar researches with the aim to assess the influence of the psychical environmental parameters to students' achievements and performances. The correlation between students' focus during the lecture and their performances and/or achievements exists (Egong, 2014), but our aim was to directly quantify parameters that influence students' focus during the lecture. Beside the parameters that were already analyzed in the previous studies, we suspected that there are additional ones that we have examined in this study.

1.2. Selecting parameters

Previous studies provide a range of potential environmental parameters that have been confirmed to influence students' performance and achievements such as temperature, humidity, air pressure, noise level, and level of CO_2. As previously mentioned, lower levels of VOC do not affect performance (Otto et al., 1992), and concentration of VOCs generally found in indoor environments are even lower than the studied concentration, therefore we conclude that levels of VOC commonly found in classroom environments cannot significantly affect students' focus. All previously mentioned parameters can be monitored using different sensors; we decided to track all of them in order to identify and select those with the most significant impact.

Previously conducted researches shown that the lecturers have a great impact on students' ability to concentrate. It has been confirmed that a lecturer can affect students achievement and satisfaction through her/his expressive behaviors (Murray, 1997), which can be expressed through different behavioral channels such as the face, speech, the body, and tone of the voice (Ambady & Rosenthal, 1992). Many studies reported that analyzing non-linguistic vocal features such as pitch, rhythm, energy, speech rate, intonation, perceptual loudness and voice quality can give us a great insight of speaker's emotional state (Batliner, Fischer, Huber, Spilker, & Noth, 2003; Fairbanks & Pronovost, 1939; Fernandez, 2004; Huang, Chen, & Tao, 1998; Huber et al., 2000; Lee & Narayanan, 2002; Williams & Stevens, 1981), which may influence students' engagement during lecture and their success in the classroom (Zembylas & Schutz, 2009). The number of features that can be extracted from a speech signal is nearly unlimited. We decided to use social signaling measures proposed by Pentland (2004), which have proved to have great success in very different situations: predicting outcomes in interactions such as negotiations, speed dating, friendships. He proposed the usage of 22 features of the sound that can be easily extracted and have shown to be important for measuring social signals (Pentland, 2004).

Table 1
List of measured parameters.

Parameter	Device used for measurement/recording
CO_2	CO_2 sensor
Temperature	Temperature sensor
Air pressure	Air pressure sensor
Humidity	Humidity sensor
Noise level	Noise level sensor on smartphone
Lecturer's voice	Bluetooth headset connected to a netbook

Due to the previous discussion, a teacher's voice seems to be a carrier of useful information in terms of measuring expressive behavior; therefore we decided to record it using Bluetooth headset.

The list of parameters that we decided to measure is presented in Table 1.

2. Method

2.1. Participants

For the purposes of the present study, we recorded 14 lectures (each 90 min long) attended by 197 students (69 female and 128 male students). Students' ages ranged from 18 to 20. Some of the students attended more than one recorded lecture. The experiment was conducted at the Belgrade University over the period of five months. The lectures were recorded in the same classroom, and every group had between 15 and 21 students.

2.2. The recording setup

Recording setup was prepared with respect to the following requirements: (1) not to use equipment that cannot be categorized as either IoT thing or IoT device using definitions described in Recommendation Y.2060 (ITU-T, 2015), (2) do not attach anything to the student, and (3) the location of sensor should not jeopardize the validity of measurements.

Recordings were done using the following devices (see Fig. 1):

- Samsung Galaxy SII smartphone – located in the middle of the classroom and used for tracking environment noise.
- eb700 device – originally it was used in the Smart Santander FP7 project (SmartSantander, 2014). The eb700 device is equipped with several gas sensors including CO_2 weather sensors (temperature, air pressure, and humidity), location (GPS – which we did not use) and a mobile network interface (GPRS). We placed it in the classroom far from the windows and heating sources and used it for tracking air temperature, humidity, air pressure, and concentration of CO_2. Our aim was to receive results in natural settings without our interference, and therefore we let the lecturer and/or students open windows when they wanted to.
- Jabra EasyGo Bluetooth Headset – connected to a local computer; it was used to record lecturer's voice.

2.3. Procedure

2.3.1. Parameters of the physical environment

We had simultaneously measured five different parameters of the physical environment: CO_2, temperature, air pressure, humidity, and noise. Values of temperature and humidity were combined

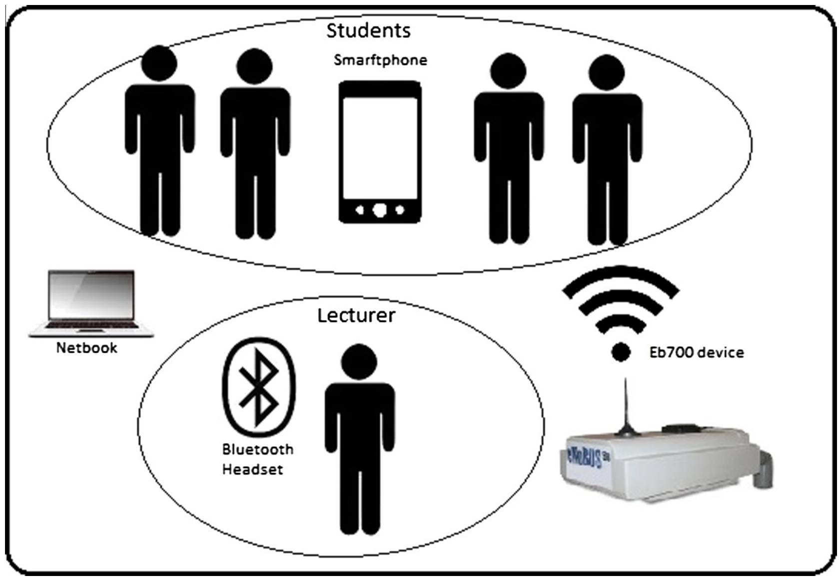

Fig. 1. The recording setup.

using humidex formula (Masterton & Richardson, 1979). The humidex combines temperature and humidity into one number to reflect the temperature perceived and represents a measure of thermal discomfort. As it considers the two most important factors that affect thermal comfort, it is a better method of measuring how hot a person really feels; more than either temperature or humidity alone (Orosa, Costa, Rodríguez-Fernández, & Roshan, 2014). To the best of our knowledge, this is the first study where the influence of the humidex on students' focus is examined.

Sampled noise levels on each time-frame were analyzed using the following statistical methods: average, average of the absolute deviations, sum of squares of deviations, median, and standard deviation.

The sound was recorded at 8 kHz using a Bluetooth headset connected to a netbook. The feature extraction was done using a toolbox developed by the Human Dynamics group at Media Lab (Pentland, 2006). This resulted in the following 22 acoustic features: mean of formant frequency, mean of confidence in formant frequency, mean of spectral entropy, mean of largest autocorrelation peak, mean of location of largest autocorrelation peak, mean of number of autocorrelation peaks, means of energy in frame, mean of time derivative of energy in frame, standard deviation of formant frequency, standard deviation of confidence in formant frequency, standard deviation of spectral entropy, standard deviation of value of largest autocorrelation peak, standard deviation of location of largest autocorrelation peak, standard deviation of number of autocorrelation peaks, standard deviation of energy in frame, standard deviation of time derivative of energy in frame, average length of voiced segment, average length of speaking segment, fraction of time speaking, voicing rate, fraction speaking over, and average number of short speaking segments.

2.3.2. Annotating data

The aim of the data annotation was to differentiate segments of the lectures on which students were "focused" against segments on which they were "not focused" on a lecture. As we have defined focus as the students' subjective feeling of their ability to concentrate on a lecture at a given moment, techniques that determine focus by analyzing someone's eyes movement or EEG could not be applied here as we were more interested in students' self-evaluation of their focus. Furthermore, both mentioned methods require additional devices to be attached to the student that contradicts our initial requirements. The only way to obtain this information was by asking students to rate if they feel able to concentrate or not at a given moment. This was indicated during a lecture by pressing one of the two buttons on a web page. As students were asked to perform only one click, the obstruction to students' activity was minimal. After a time, we assume that students got used to this action and that the disturbance became minimal. The total recorded material lasted 21 h, and the segmentation was done based on the number of votes in 30 s intervals. The selection of the time frame length was based on the preliminary pilot experiment done with different time frames and achieved accuracies.

While preparing the training dataset we annotated segments of a lecture as "focused" and "not focused" with respect to the following restrictions:

- Segments with more than 90% negative votes were annotated as "not-focused".
- Segments with more than 90% positive votes were annotated as "focused".
- Segments with less than 90% either negative or positive votes were discarded.
- Multiple votes from the same students for one segment were discarded.
- Segments with excessive hum were discarded.

After applying these restrictions, we received a dataset with 121 30-s-long segments, where every segment represents a dataset instance with 31 attributes.

3. Results

The null hypothesis states that there is no difference between the attribute values on the segments labeled as "focused" and their values on the segments labeled as "not focused". After performing *t*-test using significance value of 0.05 for each attribute, the null hypothesis was rejected for eight parameters. Therefore, we segregated these eight attributes to investigate their impact on students' focus.

These eight attributes that were rejected by the null hypothesis have values that significantly differ on the segments labeled as "focused" from the segments labeled as "not focused". Therefore, we have segregated them to investigate if we can use these values to conclude if a particular segment of a lecture belongs to the "focused" or "not focused" class.

From all the features extracted from the lecturer's voice, the null hypothesis was rejected for the following attributes: mean of formant frequency, standard deviation of formant frequency, standard deviation of confidence in formant frequency, and standard deviation of the number of autocorrelation peaks. From all the statistical analyzes we performed on the noise level, the null hypothesis was rejected for the following two attributes: the average of the absolute deviations and standard deviation. From all the parameters of the physical environment, the null hypothesis was rejected for the humidex and the level of CO_2.

Detailed list of the attributes that were rejected by the null hypothesis is presented in Table 2.

As it was expected from the previous researches (Bako-Biro et al., 2012; Coley et al., 2007; Ito et al., 2006; Molhave et al., 1986; Murakami et al., 2006; Otto et al., 1992; Shaughnessy et al., 2006; Wargocki & Wyon, 2007), the mean value of the amount of CO_2 for the segments on which students were focused was lower than for the segments on which students were not focused. This confirms previous findings that greater values of CO_2 have an adverse impact on students' focus. Our experiment shows that students are more focused when the humidex is lower, which means that students feel less hot.

Along with the previously described expected findings, there are several new contributions. When students are quiet, lecturer's voice is more distinguished and produces more variations in noise levels, and that explains why the average of the absolute deviations on the noise level is higher when students are focused. Similarly, when lecturer's voice is more dominant than the noise, values of noise levels are more spread out, which leads to having the mean of sum of squares of deviations and the mean of standard deviation of noise levels to be higher when students are focused.

We found out that formant frequency values for the lecturer's voice were higher when students were focused while the standard deviation is lower. Standard deviations of formant frequency and standard deviations of confidence in formant frequency play a role in determining the emphasis that indicates the strength of the speaker's motivation (Lepri, 2009). Higher values of formant frequency imply that the lecturer has strong motivation while little variations in standard deviation represent the strength of lecturer's mental focus (Lepri, 2009) which leads to the better presentation and causes better focus. Autocorrelation is used to find periodic components in the signal. Voiced frames are characterized by small number of autocorrelation peaks while unvoiced frames result in a large number of small peaks. Therefore, high values for standard deviation of number of autocorrelation peaks represent frames with a high mixture of voiced and unvoiced segments which can be interpreted as frames with higher student's activity, and yield to better focus.

3.1. Machine algorithm performance

As a result of data annotation and analysis, we received the dataset with eight attributes that have significantly different values on the segments labeled as "focused" from the segments labeled as "not focused". Our next step was to examine if we can use these values to train a classifier to recognize if a particular segment of a lecture belongs to the "focused" or "not focused" class. We have used the obtained dataset to train ten different classifiers with the aim to find the one with the highest recognition accuracy.

Weka toolkit (Hall et al., 2009) was used to evaluate their classification performance using 10-fold cross-validation (Refaeilzadeh, Thang, & Liu, 2008). The performance was measured using the following indicators: Accuracy, True positive rate, False positive rate, Precision, Recall time and Cohen's Kappa coefficient (Vieira & Kaymak, 2010). Every classifier was tested using the same dataset as the input to the Weka toolkit. The AdaBoost M1 classifier (Freund & Schapire, 1996) showed the best recognition accuracy (86.78%) with the highest Kappa value (0.74), while kNN had the worst recognition accuracy (79.34%) with the lowest Kappa value (0.59). This can be explained by the size of our dataset: if the training dataset is small, then high bias/low variance classifiers (such as AdaBoost) have the advantage over the low bias/high variance classifiers (kNN) (Manning, Raghavan, & Schutze, 2008). The complete list of evaluated classifiers, with their recognition accuracies are presented in Table 3.

In order to evaluate the impact of individual attributes on the classifier accuracy, we conducted a series of trials by removing one of the attributes in each trial. Kappa statistics was used as an indicator of attribute significance (Vieira & Kaymak, 2010). The recognition accuracy of the classifier after removing each one of the attributes can be seen in Table 4.

Table 2
Analysis of the attributes rejected by the null hypothesis.

Parameter name	Device used for measuring/recording	How it was sampled or analyzed	Mean value for focused segments	Mean value for not focused segments	Significance (p)
CO_2	CO_2 sensor	One value for 30 s interval	468.5593	469.4396	0.001
Humidex	Temperature and humidity sensor	Temperature combined with humidity using humidex formula	14.49529	15.83215178	0.001
Mean of formant frequency	Bluetooth connected to a netbook	Extracted from 8 kHz audio sequence	255.2373333	235.5921541	7.60525E-14
Standard deviation of formant frequency	Bluetooth connected to a netbook	Extracted from 8 kHz audio sequence	0.259177	0.276481967	0.006
Standard deviation of confidence in formant frequency	Bluetooth connected to a netbook	Extracted from the recorded 8 kHz audio sequence	0.681755	0.66376557	0.014
Standard deviation of number of autocorrelation peaks	Bluetooth connected to a netbook	Extracted from the recorded 8 kHz audio sequence	2.947512	2.7924623	0.002
Average of the absolute deviations	Noise level sensor	Sampled every second (30 values per sequence) on which we performed average of the absolute deviations	3.17759	2.845405	0.014
Standard deviation	Noise level sensor	Sampled every second (30 values per sequence) on which we performed standard deviation	3.9661557	3.50103017	0.006

Table 3
Results for different classifier using 10-fold cross-validation.

Algorithm	Accuracy (%)	TP	FP	Precision	Recall	Kappa
AdaBoost M1	86.78	0.87	0.13	0.87	0.82	0.74
Random forest Breiman (2001)	84.30	0.84	0.16	0.84	0.84	0.69
MultiBoosting Webb (2000)	82.64	0.83	0.17	0.83	0.83	0.65
Nearest neighbor using non-nested generalized exemplars Martin (1995)	82.64	0.83	0.17	0.83	0.83	0.65
Decision table Kohavi (1995)	81.82	0.82	0.18	0.82	0.82	0.64
Alternating decision tree Freund and Mason (1999)	81.82	0.82	0.18	0.82	0.82	0.64
Bayes Net	80.99	0.81	0.19	0.81	0.81	0.62
Best first decision tree Shi (2007)	80.99	0.81	0.19	0.81	0.81	0.62
NaïveBayes John and Langley (1995)	79.34	079	0.21	0.79	0.79	0.59
kNN	79.34	0.79	0.21	0.79	0.79	0.59

Table 4
Accuracy of the classifier after attribute removal.

Feature	Accuracy (%)	Kappa
Original	86.78	0.74
Mean formant frequency (from sound)	64.46	0.29
Standard deviation of formant frequency (from sound)	86.78	0.74
Standard deviation of confidence in formant frequency (from sound)	86.78	0.74
Standard deviation of number of autocorrelation peaks (from sound)	85.95	0.72
Average of the absolute deviations (from noise)	84.30	0.69
Standard deviation (from noise)	86.78	0.74
CO_2 (from air)	85.95	0.72
Humidex (combination of temperature and humidity)	85.95	0.72

Results suggest that some attributes can be removed without affecting classifier's recognition accuracy. We removed three of them (the standard deviation of formant frequency, the standard deviation of confidence in formant frequency, and the standard deviation from noise) and repeated the cross-validation using the AdaBoost M1 algorithm. Our final dataset was reduced to only five attributes: the mean formant frequency, the standard deviation of number of autocorrelation peaks, the average of the absolute deviations of noise levels, the level of CO_2, and humidex. This reduction has not changed the classifier's recognition accuracy. Our results suggest that AdaBoostM1 algorithm is able to effectively determine if students are focused on a lecture at a given moment only by analyzing values of these five attributes.

4. Discussion

In this study, we have used IoT devices to measure numerous parameters of the physical environment in order to distinguish those that significantly affect students' focus. From five measured parameters of the physical environment, we segregated three parameters that have shown to be significant for determining students' focus: the level of CO_2, the average value of the absolute deviations received from the noise, and the combination of temperature and humidity (humidex). Some other measured parameters (such as air pressure) appeared irrelevant as their levels were not significantly changing over time. It is left unclear if some of them would affect students' focus under different conditions. Beside this limitation, the advantage of this study is that it was conducted in uncontrolled natural settings opposite to the other researches that were mostly conducted in strictly controlled laboratory environments.

Although it has been confirmed that a lecturer can affect students' achievement through his/her expressive behavior (Murray, 1997), and that the voice is one of the channels for distributing expressive behavior (Ambady & Rosenthal, 1992), none of the studies have identified voice features that have influence on students' focus. In the present study, we have investigated the influence of 22 different voice features and segregated two that have shown to be significant in determining students' focus: means formant frequency, and standard deviation of number of autocorrelation peaks.

Additionally, this paper provides performance evaluation of ten different machine learning algorithms and their ability to correctly recognize "focused" and "not focused" segments of the lecture using only values of these five parameters. It was revealed that AdaBoost M1 classifier had the best recognition accuracy among them (86.78%).

Determining "focused" and "not focused" segments of the lectures was based on the votes received from students in real time during the lecture, which differs from other studies that mostly rely on the data received after the lecture.

During the experiments, we have used devices that were available to our team, but some different device combinations are also possible. Alternative for the eb700 device is a cheap Raspberry Pi device (raspberrypi.org, 2014) equipped with necessary sensors, or Galileo board (Intel, 2014). Instead of using a smartphone for measuring noise levels, any other smart device (such as a tablet) could had been used. All mentioned devices represent an IoT device as defined by ITU Recommendation (ITU-T, 2015).

4.1. Future work

One of the main limitations is that study was carried out for a short term in similar learning environments with similar number of students, so it would be useful to conduct a long-term study in different environments with different number of students. Therefore, the first future step would be to record more lectures in a diverse classroom settings with dissimilar number of students in order to extend the existing dataset and to evaluate trained classifier under different conditions. The next step will be directed towards implementing the system that will be able to automatically determine students' focus. We plan to use the same devices that we used in this research (smartphone, eb700 device, and a Bluetooth headset). These devices will be measuring parameters in real time and sending them to the server for further analysis. In order to determine if the students are focused on a lecture or not, we will use the AdaBoost M1 classifier trained on the extended dataset.

4.2. Conclusions

This research represents an innovative approach to analyzing the impact of different parameters in the physical environment on students' focus based on the Internet of Things concept. Although prior researches agreed that certain parameters of the physical environment such as temperature, environment noise and level of CO_2 impact students' performance and/or achievements, to the best of our knowledge, there is no study that has tried to examine their direct influence on students' focus. In this research, we have identified and thoroughly analyzed parameters of the physical environment that influence students' focus. Additionally, to the best of our knowledge, this is the first study where the correlation between lecturer's voice features and students' focus was interpreted.

Our results suggest that gathering of the following five parameters: the mean formant frequency (extracted from the lecturer's voice), the standard deviation of number of autocorrelation peaks (extracted from the lecturer's voice), the average of the absolute deviations of noise levels, the level of CO_2, and humidex is a viable strategy for recognizing the students' focus during lecture. The experiment was conducted outside the laboratory – in the natural settings which implies that the results can be used in the real classrooms during actual lectures.

Additionally, the received results ensure us that it is possible to implement a smart classroom system that would be able to determine in real-time if a classroom environment is optimized to maximize student's ability to concentrate on a lecture at a given moment. Also, received results can be used as a basis for more general system that would be able to automatically optimize learning environment in real-time, to inform the lecturer if students' are focused or not, and give him/her useful suggestions related to his/her presentation.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*(2), 256–274.

Atif, Y. (2013). Pedagogy-based technology enhancement in tertiary education. In P. Ordóñez de Pablos & R. Tennyson (Eds.), *Strategic role of tertiary education and technologies for sustainable competitive advantage* (pp. 295–311). IGI-Global.

Attwell, G. (2007). The personal learning environments – The future of eLearning? *eLearning Papers, 2*(1). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.3011&rep=rep1&type=pdf> Retrieved 28.02.15.

Attwell, G., Cook, J., & Ravenscroft, A. (2009). Appropriating technologies for contextual knowledge: Mobile personal learning environments. In *Proceedings of the 2nd world summit on knowledge society* (pp. 15–25).

Bako-Biro, Z., Clements-Croome, D. J., Kochhar, N., Awbi, H. B., & Williams, M. J. (2012). Ventilation rates in schools and pupils' performance. *Building and Environment, 48*, 215–223.

Batliner, A., Fischer, K., Huber, R., Spilker, J., & Noth, E. (2003). How to find trouble in communication. *Speech Communication, 40*(1), 117–143.

Bogdanovic, Z., Barac, D., Jovanic, B., Popovic, S., & Radenkovic, B. (2014). Evaluation of mobile assessment in a learning management system. *British Journal of Educational Technology, 45*(2), 231–244.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Bronzaft, A. L., & McCarthy, D. P. (1975). The effect of elevated train noise on reading ability. *Environment and Behaviour, 7*(4), 517–527.

Chang, C. H. (2011). Smart classroom roll caller system with IOT architecture. In *Proceedings of the 2nd international conference on innovations ni bio-inspired computing and applications* (pp. 356–360).

Coley, D. A., Greeves, R., & Saxby, B. K. (2007). The effect of low ventilation rates on the cognitive function of a primary school class. *The International Journal of Ventilation, 6*(2), 107–112.

Conde, M. A., García-Peñalvo, F. J., Alier, M., & Piguillem, J. (2013). The implementation, deployment and evaluation of a mobile personal learning environment. *Journal of Universal Computer Science, 19*(7), 854–872.

Crook, M. A., & Langdon, F. J. (1974). The effects of aircraft noise in schools around London Airport. *Journal of Sound and Vibration, 34*(2), 221–232.

Despotovic-Zrakic, M., Simic, K., Labus, A., Milic, A., & Jovanic, B. (2013). Scaffolding environment for adaptive e-learning through cloud computing. *Educational Technology & Society, 16*(3), 301–314.

Downs, D., & Crum, M. (1978). Processing demands during auditory learning under degraded listening conditions. *Journal of Speech and Hearing Research, 21*, 702–714.

Egong, A. I. (2014). Reading culture and academic achievement among secondary school students. *Journal of Education and Practice, 5*(3), 132–136.

Evans, G. W., & Maxwell, L. (2007). Chronic noise exposure and reading deficits: The mediating effects of language acquisition. *Environment and Behaviour, 29*(5), 638–657.

Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotions. *Speech Monographs, 6*, 87–104.

Felder, R. M., & Brent, R. (1999). How to improve teaching quality. *Quality Management Journal, 6*(2), 9–21.

Fernandez, R. A. (2004). *Computational model for the automatic recognition of affect in speech*. Doctoral dissertation. <http://affect.media.mit.edu/pdfs/04.fernandez-phd.pdf> Retrieved 17.09.14.

Freund, Y., & Mason, L. (1999). The alternating decision tree algorithm. In *Proceedings of the 16th international conference on machine learning* (pp. 124–133).

Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. In *Proceedings of the 13th international conference on machine learning* (pp. 148–156).

García-Peñalvo, F. J., Conde, M. Á., Alier, M., & Casany, M. J. (2011). Opening learning management systems to personal learning environments. *Journal of Universal Computer Science, 17*(9), 1222–1240.

Gligoric, N., Uzelac, A., Krco, S., Kovacevic, I., & Nikodijevic, A. (2015). Smart classroom system for detecting level of interest a lecture creates in a classroom. *Journal of Ambient Intelligence and Smart Environments, 7*(2), 271–284.

Halimi, K., Seridi-Bouchelaghem, H., & Faron-Zucker, C. (2014). An enhanced personal learning environment using social semantic web technologies. *Interactive Learning Environments, 22*(2), 165–187.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10–18.

Hartley, J., & Davies, I. K. (1978). Note-taking: A critical review. *Programmed Learning and Educational Technology, 15*, 207–224.

Howarth, E., & Hoffman, M. S. (1984). A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology, 75*(1), 15–23.

Huang, T. S., Chen, L., & Tao, H. (1998). Bimodal emotion recognition by man and machine. In *Proceedings of ATR workshop on virtual communication environments, Japan*.

Huber, R., Batliner, A., Buckow, J., Noth, E., Warnke, V., & Niemann, H. (2000). Recognition of Emotion in realistic dialog scenario. In *Proceedings of the international conference spoken language processing* (Vol. 1, pp. 665–668).

Intel (2014). Intel Galileo Gen 2 Development Board. <http://www.intel.com/content/www/us/en/do-it-yourself/galileo-maker-quark-board.html> Retrieved 11.09.14.

Ito, K., Murakami, S., Kaneko, T., & Fukao, H. (2006). Study on the productivity in the classroom (Part 2): Realistic simulation experiment on effects of air quality/thermal environment on learning performance. *Healthy Buildings*, 207–212.

ITU-T (2015). Telecommunication standardization sector of international telecommunication union. Overview of the internet of things, Recommendation ITU-T Y.2060. <http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=y.2060> Retrieved 28.2.15.

John, G. H., & Langley, P. (1995). Estimating continuous distributions in Bayesian classifiers. In *Proceedings of the 11th conference on uncertainty in artificial intelligence* (pp. 338–345).

Johnson, E. (2001). Let's hear it for learning: Improving classroom acoustics can maximize student productivity at a minimal cost. *American School and University, 73*(11), 28–30.

Juarros, M. V., Ibáñez, S. J., & de Benito Crosetti, B. (2014). Research results of two personal learning environments experiments in a higher education institution. *Interactive Learning Environments, 22*(2), 205–220.

Kohavi, R. (1995). The power of decision tables. In *Proceedings of the 8th European conference on machine learning* (pp. 174–189).

Kompen, R. T., Edirisingha, P., Monguet, & J. M. (2009). Using web 2.0 applications as supporting tools for personal learning environments. In *Proceedings of the 2nd world summit on knowledge society* (pp. 33–40).

Kyzar, B. L. (1977). Noise pollutions and schools: How much is too much? *Council of Educational Facilities Planners Journal, 4*, 10–11.

Lee, C. M., & Narayanan, S. S. (2002). Combining acoustic and language information for emotion recognition. In *Proceedings of the international conference on speech and language processing* (pp. 873–876).

Lepri, B. (2009). *Multimodal recognition of social behaviors and personality traits in small group interaction*. Ph.D. Thesis, DIT University of Trento.

Luo, Q., Zhou, J., Wang, F., & Shen, L. (2009). Context aware multimodal interaction model in standard natural classroom. *Lecture Notes in Computer Science, 5685*, 13–23.

Lytras, M. D., & Ordóñez de Pablos, P. (2011). Software technologies in knowledge society. *Journal of Universal Computer Science, 17*(9), 1219–1221.

Manning, C., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. <http://nlp.stanford.edu/IR-book/html/htmledition/choosing-what-kind-of-classifier-to-use-1.html> Retrieved 17.09.14.

Martin, B. (1995). *Instance-based learning: Nearest neighbor with generalization*. Master's thesis. <http://www.cs.sun.ac.za/~kroon/courses/machine_learning/lecture2/martin_instance.pdf> Retrieved 17.09.14.

Masterton, J. M., & Richardson, F. A. (1979). Humidex: A method of quantifying human discomfort due to excessive heat and humidity. *Atmospheric Environment Service, Environment Canada*, 1–79.

Mendell, M. J., & Heath, G. A. (2005). Do indoor pollutants and thermal conditions in school influence student performance? A critical review of the literature. *Indoor Air, 15*(1), 27–52.

Molhave, L., Bach, B., & Feddersen, O. F. (1986). Human reactions to low concentrations of volatile organic compounds. *Environment International, 12*, 167–175.

Murakami, S., Kaneko, T., Ito, K., & Fukao, H. (2006). Study on the productivity in the classroom (Part 1) field survey of the effects of air quality/thermal environment on learning performance. *Healthy Buildings*, 271–276.

Murray, H. G. (1997). Effective teaching behavior in the college classroom. In Raymond P. Perry & John C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 171–204). New York: Agathon Press.

Orosa, J. A., Costa, A. M., Rodríguez-Fernández, A., & Roshan, G. (2014). Effect of climate change on outdoor thermal comfort in humid climates. *Journal of Environmental Health Science and Engineering, 12*, 46.

Otto, D. A., Hudnell, H. K., House, D. E., & Molhave, L. (1992). Exposure of humans to a volatile organic mixture. I: Behavioral assessment. *Archives of Environmental Health, 47*(1), 23–30.

Pentland, A. (2004). Social dynamics: Signals and behavior. In *Proceedings of IEEE int'l conf. developmental learning*. <http://web.media.mit.edu/~sandy/TR-579.pdf> Retrieved 13.09.14.

Pentland, A. (2006). A computational model of social signaling. In *Proceedings of the 18th international conference on pattern recognition* (Vol. 1, pp. 1080–1083).

Pepler, R. D., & Warner, R. E. (1968). Temperature and learning: An experimental study. *ASHRAE Transactions, 74*(2), 211–219.

Pilman, M. S. (2001). The effects of air temperature variance on memory ability. National Undergraduate Research Clearinghouse, 4. <http://www.webclearinghouse.net/volume/> Retreived 21.01.13.

Rahimi, E., Van den Berg, J., & Veen, W. (2015). Facilitating student-driven constructing of learning environments using Web 2.0 personal learning environments. *Computers & Education, 81*, 235–246.

Raspberry Pi (2014). What is raspberry Pi? <http://www.raspberrypi.org/> Retrieved 11.09.14.

Refaeilzadeh, P., Thang, L., Liu, H. (2008). Cross Validation. Arizona State University. http://leitang.net/papers/ency-cross-validation.pdf Retreived 14.09.2014.

Rizzardini, R. H., Linares, B. H., Mikroyannidis, A., & Schmitz, H. C. (2013). Cloud services, interoperability and analytics within a ROLE-enabled personal learning environment. *Journal of Universal Computer Science, 19*(14), 2054–2074.

Schoer, L., & Shaffran, J. (1973). A combined evaluation of three separate research projects on the effects of thermal environment on learning and performance. *ASHRAE Transactions, 79*, 97–108.

Shaughnessy, R. J., Haverinen-Shaughnessy, U., Nevalainen, A., & Moschandreas, D. (2006). A preliminary study on the association between ventilation rates in classrooms and student performance. *Indoor Air, 16*(6), 465–468.

Shen, C. W., Wu, Y. J., & Lee, T. C. (2014). Developing a NFC-equipped smart classroom: Effects on attitudes toward computer science. *Computers in Human Behavior, 30*, 731–738.

Shi, H. (2007). *Best-first decision tree learning*. Master's thesis. <http://researchcommons.waikato.ac.nz/bitstream/handle/10289/2317/thesis.pdf?sequence=1&isAllowed=y> Retrieved 17.09.14.

Simic, K., Stojanovic, M., & Djuric, I. (2014). A model for smart e-learning environment. In *Proceedings of the 14th international symposium SymOrg* (pp. 377–382).

SmartSantander (2014). *SmartSantander*. <http://www.smartsantander.eu/> Retrieved 13.09.14.

Vieira, S., & Kaymak, U. (2010). Cohen's kappa coefficient as a performance measure for feature selection. In *Proceedings of IEEE international conference on fuzzy systems* (pp. 1–8).

Wargocki, P., & Wyon, D. P. (2007). The effects of moderately raised classroom temperatures and classroom ventilation rate on the performance of schoolwork by children. *HVAC&R Research, 13*(2), 193–220.

Webb, G. I. (2000). MultiBoosting: A technique for combining boosting and wagging. *Machine Learning, 40*(2), 159–196.

Williams, C. E., & Stevens, K. N. (1981). Vocal correlates of emotional states. In J. K. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 221–240). New York, NY: Grune & Stratton.

Wyon, D. P. (1970). Studies of children under imposed noise and heat stress. *Ergonomics, 13*(5), 598–612.

Xie, W., Shi, Y., Xu, G., & Xie, D. (2001). Smart Classroom – an Intelligent Environment for Tele-education. In *Proceeding of the 2nd IEEE pacific rim conference on multimedia* (pp. 662–668).

Zembylas, M., & Schutz, P. A. (2009). Research on teachers' emotions in education: Findings, practical implications and future agenda. In P. A. Schutz & M. Zembylas (Eds.), *Advances in teacher emotion research: The impact on teachers' lives* (pp. 367–377). New York, NY: Springer.