T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

AIR TRAFFIC CONTROLLER ALERTNESS LEVEL DETECTION

AZİZ PARMAKSIZ

A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF COMPUTER ENGINEERING

GEBZE
2017

T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

AIR TRAFFIC CONTROLLER ALERTNESS LEVEL DETECTION

AZİZ PARMAKSIZ

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER ENGINEERING

THESIS SUPERVISOR

ASSOC. PROF. DR. MEHMET GÖKTÜRK

GEBZE

2017

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HAVA TRAFİK KONTROLÖRLERİNİN DİNGİNLİK SEVİYESİ TESPİTİ

AZİZ PARMAKSIZ
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMANI
DOÇ. DR. MEHMET GÖKTÜRK

GEBZE
2017

GEBZE TEKNİK ÜNİVERSİTESİ	**YÜKSEK LİSANS JÜRİ ONAY FORMU**

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 31/05/2017 tarih ve 2017/29 sayılı kararıyla oluşturulan jüri tarafından 05/09/2017 tarihinde tez savunma sınavı yapılan Aziz Parmaksız'ın tez çalışması Bilgisayar Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : Doç. Dr. Mehmet GÖKTÜRK

ÜYE : Yrd. Doç. Dr. Burcu YILMAZ

ÜYE : Yrd. Doç. Dr. Ediz ŞAYKOL

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
....../....../......... tarih ve ..../........ sayılı kararı.

İMZA/MÜHÜR

Doç. Dr. Arif Çağdaş AYDINOĞLU
Gebze Teknik Üniversitesi
Fen Bilimleri Enstitüsü Müdürü

SUMMARY

In this work, alertness level of Air Traffic Controller (ATC) was detected by detecting stress of ATC on his/her real work environment and on pre-designed simulation environment. Alertness is state of people having high attention to event and watchful to danger. Stress is an emotion state of people which accused from autonomic nerves system. High-level stress adversely affects the person and cause cognition loose. So detecting stress level of ATC would also mean detecting alertness level of ATC. Galvanic Skin Response (GSR) and Heart Rate (HR) sensors were used to detect stress of ATC. For this experiment on the real controller working area, on simulation environment and while they were resting sensor values were collected. With weka data mining analysis tool those values criticized and 91 percent of success gained to classify stress of ATC by Naive Bayes Classification method, and it has been seen GSR sensor more sensitive and fast response to stress than HR.

Key Words: ATC, GSR, HR, STRESS.

ÖZET

Bu çalışmada Hava Trafik Kontrolörlerinin (HTK) gerçek çalışma ortamındaki ve simülasyon ortamındaki dinçlik seviyesi tespit edilmiştir. Dinçlik insanların durumlara karşı dikkatinin yoğun ve tehlikelere karşı algılarının açık olmasıdır. Stres insanların otonom sinir sistemi tarafından tetiklenen bir duygu durumudur ve stresin yüksek olması insanları olumsuz etkileyerek karar verme mekanizmasını sekteye uğratır. HTK'ların stres tespiti bir başka deyişle HTK'ların dinçliği tespiti anlamına gelmektedir. Galvanik Deri Tepkisi (GSR) ve Kalp Ritim (HR) sensörleri kullanılarak stress tespiti yapılmaya çalışılmıştır. Bu çalışma için kontrolörlerin gerçek çalışma ortamında, simülasyon ortamında ve dinlenme durumundaki sensor değerleri toplanmıştır. Weka veri madeni analiz aracı kullanılarak yapılan sınıflandırmada Naive Bayes sınıflandırma yöntemi ile %91 oranında HTK'ların stres tespiti sınıflandırması başarısı elde edilmiştir ve GSR sensörünün HR'ye göre strese daha duyarlı ve hızlı tepki verdiğini görüldü.

Anahtar Kelimeler: Hava Trafik Kontrolörü, GSR, HR, STRES.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Assoc. Prof. Dr. Mehmet Göktürk for supporting me and lightning my way when I was totally lost.

I would like to thank, Turkish Air Traffic Control Centre and controllers for supporting my thesis.

TABLE of CONTENTS

LIST of ABBREVIATIONS and ACRONYMS

Abbreviations and Acronyms		**Explanations**
ATC	:	Air Traffic Controller
CWP	:	Controller Work Position
GSR	:	Galvanic Skin Response
GTU	:	Gebze Technical University
HR	:	Heart Rate
HRV	:	Heart Rate Variability
HTKM	:	Hava Trafik Control Merkezi
PWP	:	Psudo Pilot Work Position
VCS	:	Voice Communication System
STCA	:	Short Term Conflict Alarm

LIST of FIGURES

LIST of TABLES

1. INTRODUCTION

Human myth creator of the world, adapted to live in any different place, handle many difficulties, learned how to survive, how to evolve when the time came learned the way of the fly and broke the boundaries pass across the sky. However, of course, when doing that, they learned how to work together, they need each other. Into this evolution some people chose to take responsibilities for others being safe, some people chose to avoid responsibility being only him/her being safe.

There are several hardest jobs in this world. Being an engineer on a nuclear factory, general of military, firefighter or air traffic controller, these jobs share one special thing in common, they undertake the responsibility of human life, which is the most important thing in this world. Making a mistake can be accepted for some work field, but those fields of work don't tolerate any single mistake because that can end with a person's life. Also, some people exist to decrease risks, find some solution to alert person before the danger exists, that we call them the scientist. They don't have to have a name or have a reputation. That all they need is to do make researches and prove some danger can be avoided in advance For this reason, with some research and experimentation, people who choose to work in the world hardest profession can be prevented from making mistakes or lost in most. As a result, those who chose the hardest occupation of the world would be alert; Air Traffic Controller would be alert.

1.1. Air Traffic Controller

Air traffic controllers (ATC) are people trained to maintain the safe, orderly, and expeditious flow of air traffic in the global air traffic control system [1]. Pilots must obtain permission and instructions from the air traffic controllers. Also, other ground vehicles like ambulance, fire truck, fuel tanker, follow me… etc. have to obtain permission from ATC. ATC are not only responsible for the safety of aircraft in the air, but also responsible for the safety of traffic on the grand (airport).

The Air Traffic Controller will provide air traffic control services to dozens of aircraft in its control area by using radio to transmit advice, information and instructions, working with many auxiliary units and taking advantage of the innovations brought by the technology to ensure that they fly safely and regularly, it provides. [2]

For example; when an aircraft will fly from Istanbul to Ankara, pilot cannot get up and fly as he wants. It is necessary to fly through designated air corridors and comply with the instructions issued by air traffic controllers. That's why the pilot has to get permission and instructions from the air traffic controllers up to the parking lot from the engine start-up, the flight from the air corridor to the flight level to climb, the descent from the speed to keep it, the landing at the airport to go. An air traffic controller can also be responsible for 10 or sometimes 35-40 flights at the same time.

The duties of air traffic controllers vary depending on whether they are working as tower, approach and airfield or grand controller. Their main tasks are;

- Protect radio and / or radar contact with aircraft;
- Directing the motion of the aircraft at the roadway or at the airport;
- Give instructions on climbing or descending aircraft and allocate the final cruise level;
- Providing information on weather conditions to aircraft;
- To ensure that minimum distances are maintained between aircraft;
- Handling and managing unexpected situations and situations, emergencies and non-plan traffic.
- Control of movements to be performed from the track or the track;
- Transportation and management of aircraft movements around the terminal;
- Transportation and management of vehicles' movements around the airport.

1.2. ATC and Stress

ATC have to make countless sudden decisions within a short period of time to ensure flight safety and flow of traffic on a daily basis.

With increased air traffic, the workload of the controllers has also increased. This responsibility and workloads make ATC naturally stressful [3].

ATC, responsible for air traffic and human safety, must be alert to all situations that can occur at any moment. This can be achieved with their alertness levels being high.

Identifying emotional states of the controllers are not enough in terms of the continuity and safety of air traffic management. Determining alertness level of the controller can be much more sufficient to increase air traffic management performance.

Being able to safely manage aircraft in different speeds, altitudes and directions in different types of airplanes without changing the possibility of emergency breakdowns and collisions and safely under varying meteorological conditions in a restricted airfield requires that this professional group be knowledgeable, careful and skillful as an orchestra conductor. While evaluating the flow of intensive information from electronic devices such as radar, radio and telephone, it is very difficult to combine visual information with them, and to manage airplanes in the nighttime conditions. In this process, communication in a special terminology and symbology in English is another problem. The possibility that even small mistakes are transformed into plane crashes is a fearful dream of ATC [4].

This group of professors is in a consensus on the fact that the workload stress is too much, the work must make quick and decisive decisions without making any mistakes, and the controller is the brains and key personnel of air traffic. Hypertension and ECG anomalies in the ATC profession group [5], which is such a stressful task, are twice as many as the control groups; psychosomatic complaints such as stomach ulcers, indigestion, headache, and chest pain were also found to be significantly higher [6], [37].

1.3. Work Schedule of ATC

There are 4 teams in an air traffic control center. In day duty 11 hours between 08:30 - 19:30 hours, at night time 13 hours between 19:30-08:30. In teams where 4 teams are working, one day the team working in the day shift comes to watch the night the next day. 2 days after the night shift (the day after the night shift and the next day). During the seizure period, he usually works for 2 hours and rests for 2 hours. [8] The controller works continuously for about 6 hours in the 11 hour day shift and the remaining 5 hours to help his teammates who are resting and intensive traffic. Operating hours in seizures may increase with the number of controllers on the base and the traffic intensity.

- Stress

The peripheral nervous system (PNS) consists of nerves and ganglia, except the brain and spinal cord. The main function of the PNS is to communicate (connect) the organs and limbs with the central nervous system (CNS. The peripheral nervous system is separated by the somatic nervous system and the autonomic nervous system. PNS constitutes a large part of the nervous system [39].

You can think nerves as telephone wires. The nerves carry messages from our bodies to the brain. The chemical or biological effects that occur in the body or around the body and cause a certain reaction in the body are called the warning. Alerts are picked up by special cells in our senses. Receive warning, nervous transport to the CNS. Alerts cause chemical or electrical changes in nerve cells. This format is called the excitation message. The stimulus message is transmitted to the CNS. Responses to messages are transmitted to the muscles, organs, and glands by nerves. The stimulus message is evaluated in the head and an answer is generated against the warning. The resulting response is transmitted to the organ or construct by nerves again, and reacts to the stimulus.

Stress is a physiological response of the body to the mental, emotional, and physical challenges people faced [9]. High level of stress adversely affects person's mental and physical activities [10], [11].

Autonomic nervous system (ANS) is responsible for body’s unconscious activities [12]. Blood pressure, sweat secretion, electrical changes of heart rate those all triggered by ANS.

ANS has two branches, Sympathetic Nervous System and Parasympathetic Nervous System (PNS) [13]. While SNS produces “fight or flight” response, PNS relax the body and normalize effect of SNS.

Measuring SNS activities and the intensity of these activities can give clues about the emotional state of a person [17].

- GSR and Stress

GSR, also known as Skin Conductance (SC) or Electro Dermal Response (EDR), that skin resistance is measured by electrodes placed on two fingers [36], [37]. When a person meets with a stress factor, SNS activates the sweat glands and person starts producing more sweat than normal [14]. Sweat makes a salty layer on the skin and salty layer known as electrically conductive, therefore skin conductivity increase. When skin electrical conductivity increases, conversely skin electrical resistance decreases [17], [18], [21].

- HR and Stress

Changes in Heart Rate (HR) helps to determine stress [17], [20], [23]. On fight or flight response high adrenalin being produced and heart pulse rhythm goes fast. Within an optical HR sensor, an infrared light touches to skin and catches the changes in the amount of reflection of light that reflected from blood flowing under the skin. Those reflection gives information about heart pulse. On each pulse, different intensity level of blood pass from the body and when light passes the skin, the reflection of light from blood dynamically change through blood intensity. By measuring the time interval between each pulse can give information about HR and HR Variability (HRV).

2. LITERATURE WORK

Coghlan [15], claim that there are 3 core problems to judge emotion of human by using indicator from ANS that are Baselining Problem, Timing of Data Assessment Problem and Intensity of Emotion. He said by Baselining Problem which changes of ANS will be measured, by Timing of Data Assessment Problem on which time of data will be considered as emotion change and by Intensity of Emotion how to addresses emotional changes with the physiological response must before be decided and defined. When evaluating his studies with his claim, It is seen that measuring effect of sweat secretion which triggered by ANS simply give answer to "Baselining Problem", giving an challenging task and degree of difficulty help us to solve "Intensity of Emotion" and monitoring response of client on task, solve the "Timing of Data Assessment Problem."

2.1. Effect of Stress

Stress has effect on perception, attention, memory, decision making, problem solving and response execution, all stages that have been identified and studied by cognitive psychologists, are candidates for degradation, and triggered by some stressor; for humans they might be extreme temperatures, loud or noxious noises, infectious diseases, sleep deprivation, extreme heavy or prolonged workloads, time pressures, social pressures, and intense negatively-toned emotions [7].

2.1.1. Effect of Stress on Perception and Attention

Stress, has important effects on attention and perception. But these effects are not always same. It depends in serious ways on the qualitative features of the stressor. Different stressors have different effects on performance. One stressor might cause shifts in attention or a failure to inhibit irrelevant stimuli, while another stressor causes a lapse of attention or attentional narrowing in the same task.

Callaway and Dembo [47] studied about the effect of stress on attention. They tested the visual judgment of sizes of objects by subjects exposed to one of two different conditions. One group was asked to inhale amyl nitrite, which is a drug and effect on person in seconds and last long to several minutes as decreased blood pressure, followed by an increased heart rate, flushed face and neck, dizziness, and headache. The purpose of that exposure was to test the effects of different physiological changes in visual perception. The second group of subjects was requested to put their foot into a bucket of ice water. Bucket of ice water would make a stressful condition to make them feel thermal discomfort. The authors found that both groups tended to judge the objects as larger than they are.

Diaz, Hancock, and Sims [46] reported that a speed/accuracy trade-off, reducing the effectiveness of visual search under noise conditions. Lavine, Sibert, Göktürk and Dickens [30] examined simultaneous eye movements and human performance during vigilance tasks that require frequent visual scanning. A four-digit rectangular array of warnings changed at a 4-second rate for a 30-minute duty cycle. Participants were asked to respond by tapping into specific, rare signal sequences under 50 dBA white noise and 90 dBA intervals and unpredictable sound burst conditions (SBC). With time-tasking, the subjective fatigue ratings increased, the total number of fixations in the target cells decreased, the number of fixations decreased, and fixtures appeared in the target cells in both conditions. Thus, the usual precautionary performance decline was repeated. The fixation period did not show any significant change with time or condition. Out-of-sight visual crawls followed less than the target crawl paths in both conditions. Corrections with SBCs suggest that target markers may help to balance precautionary degradation of background stimuli with a higher stimulus function, increasing the likelihood of burns and hits.

2.1.2. Effect of Stress on Memory

It was proved by so many experiments noise, threat, thermal conditions, and some others stressor reduce working memory performance [44], [45]. For example, it has been proved with several investigations workload increase the number of errors on working memory tasks.

Burrows [43] examined memory performance under the conditions of workload (increasing word list) and time pressure (quick presentation) that measure recall correctness and speed in the recognition test. Burrows' first experiment was designed to present a six-word list; each word was kept visible for two seconds and then changed to the next word (on the computer screen). The list of the words was presented twice in each case. During the next recognition test, people were asked to make the decision as soon as possible about whether or not the visible sheen was on the previous list. Burrows conducted 12 trials of new words, half and half of the target words presented earlier. After completing these experiments, each participant was asked to evaluate subjective stress levels and perceived performance levels like they were feeling. Burrows then offered a second list of six words; he gave instructions to add these additional words to the original list. After 12 similar tests, after 50 self-assessments, these topics were presented as an additional list of six words. This process continued until each person was exposed to 60 words and 120 recognition attempts. The self-rating of the stress and the perception of success in the task were also assessed throughout these re- peations. After three days, each subject was retested with a similar series of memory tests. The results of this experiment showed that as the number of items to remember increases, the memory accuracy decreases gradually. For example, after exposure to only six words, 95% of the averages upon recognition are correct. This rate gradually increased to three quarters after the presentation of all 60 calves. The increase in similarly remembered words caused a strong correlation with self-rated stress, so stress ratings were matched at an exceptional rate with workload measures and reductions in memory performance. The writer noted that the memory pens gradually diminished during their accumulation, if the decline in memory accuracy was significant and significant. This model was also shown between the response time measurements. Subjective evaluation performance correctly represented objective measures, which led to the idea that subjects perceived their own diminished performance and correctly perceived it.

2.1.3. Effect of Stress on Decision Making

Ruffell Smith [48] investigated the effects of increased workload on flight crew performance. They made experiment with a three-person flight team of 20 people from a large civilian freight forwarder. They gave subject a series of flight simulator scenarios under various workload conditions like ordinary to abnormal flight scenarios. Flight teams wore physiological monitoring equipment during their observations. Data were recorded in five dimensions, including the observer's comments, the crew's heart rate, documentation of each flight, flight deck communications, and aircraft parameters. The crew initially flew in uneventful leg then unexpected conditions were created like mechanical failures, fuel loads and weight problems, weather conditions, etc. At first period, the crew had a relative calm resulting then after it ended with a more complex landing sequence, including a hand-held ILS (Instrument Landing System) approach and three-motor landing. During experiments, the writer noted that a series of mistakes were made. These are categorized into several main types including navigation, communication, system faults, flying faults, tactical decision errors, and crew integration or resource management challenges. This study shows that periods of increased workload, which are believed to be part of normal flight operations, increase the frequency and volume of errors. In addition to a list of faults and fault frequencies, Ruffell Smith also presented a narrative account of how mistakes occurred in different phases of flights. He noted that procedural mistakes tend to be linked to poor coordination between crew members under stressful conditions. The greatest deviations from heart rate appeared immediately following the take-off and landing procedures. Moreover, heart rate capability was also associated with the pilot's flight, which pilot in each case did not fly-pilots tended to experience an increase in heart rate during flight control. It was also observed that the response times for the captains significantly increased when they were faced with the decision of shutting down a motor during flight. However, this was usually only during their control. He noted that numerous human factor problems remain unaddressed for airline operations, highlighting issues ranging from instrumentation and ergonomics to control list design.

2.2. Stress Detecting

There are many studies in the literature about identifying whether people are stressed or not [13], [17], [18], [21], [22], [24], [29], [31], and [32]. Several physiological measures have been used as an index of user mental effort and mental workload. PNS measures include cardiovascular measures such as Heart Rate Variability (HRV), measures of pupil diameter [40], [41], respiratory measures [42], and measures of electrodermal activity (EDA) or Galvanic Skin Response (GSR).

Each of these measures has its strengths and weaknesses in terms of the relative difficulty in collecting, analyzing, and interpreting their meaning; each also differs along the lines of sensitivity, diagnosticity, intrusiveness, reliability, and generality of application [23].

2.2.1. Stress Detecting By GSR Measures

The GSR sensor is the most commonly used sensor to identify the stress state. Conglahm [15], in his experiment he made clients listen to an experimental music while clients sit on a chair that a GSR sensor connected to it. He observed that when music volume and rhythm changed GSR value of client also changed.

María Viqueira et al. built and designed GSR based sensor to detect stress of person and send stress data over wireless communication [18]. They used Jennic JN-5148 boards to acquiring data, and send data with Zigbea board. They chose Jennic board, because those two boards use same protocol to communicate. While detecting GSR, they fed board with 1.8 V input and measure output voltage. To detecting different emotional state of subject they implemented several testes. Relaxing, reading book, taking deep breath, thinking something that make them nerves, solving mathematical operation which are testes they implemented. On each test they saw subject's GSR value respond different. They used weka data analysis tool to implement several classifier algorithm to find best matched classifier. With BeyesNet algorithm they observed 93% rate of success to classify stress state.

2.2.2. Stress Detection By HRV Measures

Changes in the heart rhythm are another method used to determine stress. In stress state SNS activities cause changes on HRV. HRV is simply the beat-to-beat variation in the R-R interval (Figure 2.1). Count of beat in specific unit of time called Heart Rate (HR). Generally it is calculated per minute. Analysis on HRV features collected under two major fields, Time Domain Analysis and Spectral Domain Analysis.

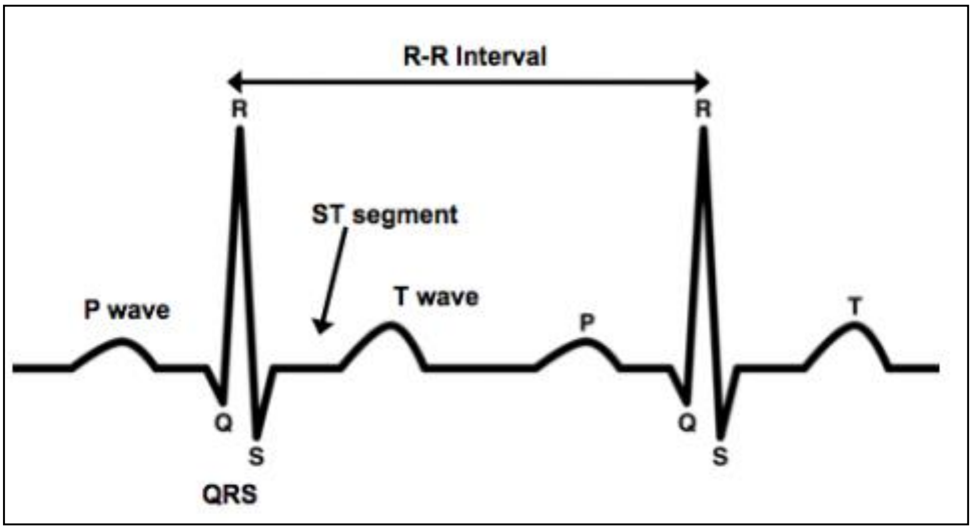

Figure 2.1: R-R interval of Heart Beat.

2.2.2.1. Time Domain Analysis

Time domain measures are done with measuring mean of HR (meanHR), mean of all RR intervals (meanRR), the percentage of heartbeat intervals greater than 50 ms with a difference in successive heartbeat intervals (pNN50), standard deviation of all RR intervals (SDNN) and, root mean square of the difference between successive RR-intervals (rMSSD). Low meanRR shows person in stress state, and this is also seen in high meanHR.

Salahuddin et al. point that RR and HR intervals measured within 10 second, rMSSD and pNN50 within 30 second can be trusted to monitor mental stress in mobile setting [52]. He made an investigation on ultra-short term analysis of HRV. He stated that long time analysis of HRV like 5 min is not applicable for conventional product. He made experiment with 24 subjects. For each subject he made them rest on confortable chair for a 10 second without any physical activity. That was for measuring normal level of HRV. Then he implemented a Stroop Test to the subject in time range of 2.5 minute and 5 minute. After test phase, he randomly sampled collected RR data to time ranges which are from 10 second to 150 second in step of 10 second increment. He segmented meanRR, meanHR, rMSSD and pNN50 of result with both stroop test and baseline data in 10 to 150 time range. He saw that RR and HR within 10 second, rMSSD and pNN50 within 30 second can be trusted to monitor mental stress.

2.2.2.2. Frequency Domain Analysis

Frequency domain HRV measures based on power spectral density (PSD) analysis of the HRV data. It includes power spectral density computation of RR interval by means of Fourier transform. They are very low frequency (VLF), low frequency (LF) and high frequency (HF), normalized very low frequency spectrum (nVLF), normalized low frequency spectrum (nLF), normalized high frequency spectrum (nHF), difference of normalized low frequency spectrum and normalized high frequency spectrum (dLFHF).

The experiment in [20] authors give a coin-stacking task to stimulus and observe HR frequency changes. By using Fast Fourier Transform (FFT) transformation they transform HRV data to frequency domain data to get LF HF, and LF/HF ratio of HRV. While subject was on stressed task they observed the changes on HF, LF of HRV data. When subject in stress state they saw LF ration of HRV was bigger and HF is smaller, but when in not stress state HF ratio was bigger and LF ratio was smaller. After all, they improved changes in LF/HF ratio affected the mental status of the subjects and increased their sympathetic nerves activity.

Fhen Tso et al. on their resource use GSR, Electrocardiogram and Accelerometer to detect stress of clients on three states while walking, sitting and standing [17]. Accelerometer placed on hip significantly helped classify activities such as sitting, running, crawling, and lying down. They used Accelerometer to recognize whether subject, walking, sitting or standing and maximize the difference of signal while sitting, walking or standing. They requested participants to respond to color which asked on Stroop test [25] and solves the simple mathematical question by mental calculation in a limited time while they are resting, sitting and walking. They made an experiment with 20 participants. They collected data on three 10 minute phases. Before experiment they made participants listen to meditation music for 10 minutes and labeled those data as “not stressed”. Then they requested each participant to solve stress test which is stroop test and mental calculation for 10 minutes. After that step, they requested participants to rest for 10 minutes. To analyze data they separate data as a 60-second window for two reasons. First, the algorithm that they used in that study can distinguish between stressed and baseline segments using 60-second windows. Second, they claimed 60-second feature window reduces the impact of misclassified R-peaks by averaging HRV features within the window. Then, they used FFT to convert the time-domain RR-interval series to the power spectrum. So they got HF, LF and, LF/HF. Within all, they gained 80% success classifying stress by decision tree classifier.

One other previous work belongs to Mozos [22]. Authors measured Electro dermal activity and Heart Pulse activities of patient on Trier Social Stress Test, which include public speaking and cognitive tasks that make participants under cognitive stress. In public speaking tasks they asked some question to participants while camera on records. They used camera for the purpose of make participants also feeling under stress. In cognitive task, they made experiment by using time pressure. In one task they request participant sit on desk and prepare a presentation for job interview in 3 minute. Then request participant to speak continuously during 5 minute. After then, they requested participant to count down in step of 13 from 1022 with 5 minute time limit. They completed their experiment with 18 volunteering students from the School of Psychology at the University of Lincoln. To analyzing result they used AdaBoost, SVM, and KNN classifier. When comparing precision and recall of those classifiers, AdaBoost was most efficient classifiers with 0.94 accuracy, 0.94 precision and 0.96 recall values.

2.2.3. Stress Detection By Eye Pupil Measures

As a result of SNS activities, eyes also give information about human emotional state. With technologic development it is possible to extract eye gazing, fixation, blinking, and even changes in pupil diameter. Eye movements have been rapidly growing research interest in psychology, ergonomics, and computer science etc. Resources being done with eye tracker in Human Machine Interacting domain generally focus on usability of computer interface [49]. Some researches showed that eye movement behavior can used to predict human performance [50].

Wang [51], made a statistical analysis on ATC' eye movement behavior on ATC environment, to investigate relation between eye fixation and ATC experience. In order to investigate eye movement activity on complex task they used ATC simulator which have exactly same interface and functionality with real ATC system of Air Traffic Management Bureau (ATMB) of Zhejiang Province. They prepared three levels of traffic scenarios that based on the real schedule as easy, normal and hard. 3 female and 22 male has vary controller experience participated that experiment. All participants were requested to do their normal job with their normal equipment and clothes. Also wearing glasses permitted to feeling as they are on real environment. Experiment showed that, eye movement of participants has more 12 years experience was quite different from other controllers (who had less than 12 year controller experience). While average of time duration to find Area of interest (AOI) is less for experienced ATC, average of fixation time is much longer than other. So that tells experienced ATC more quickly find AOI and less saccade on AOI.

Povilas Treigys [53] make an investigation to automatically identify the iris and pupil of the eye in the video stream and to parameterize the identified structures in order to make assumptions if the subjected is stressed or not. Acording to Treiygs eye tracking can be done in tree way:

- Eyes are covered with special mirror lenses or magnetic field detectors. This allows you to measure eye movements very precisely in the form of eye tracking. However, this approach is limited because it is contact-dependent.
- When a light reflected from the eyes is recorded in the video stream. In this case, because the pupil size of the eye is sensitive to the density of the light, the reflected light illumination and video recording are performed in the near infrared (NIR) light range. Furthermore, in order to detect eye movements, image analysis techniques should be applied in the recorded video frame. This approach is useful when it is non-contact.
- Eye tracking is performed by electrical potential measuring devices applied around the eye area. This approach is useful because eye movements can also be viewed in the dark or when the envelope is closed. However, just like the first approach, it is contact dependent.

Among three options they chose second option to measure eye pupil diameter in stress task. Because contacting any device to eye by alone is enough to make subject stressed. Also in this work it was described as methods based on image analysis can be classified into three classes: patterns, properties and sample-based

[54]. In the pattern-based method, the global template defining the eye shape must first be defined. The regions of the image are then compared to the global template, and the best match indicates the region of the eye region. Property -based methods use the image intensity characteristics of the eye region, such as the contrast differences between iris, pupil and sclera and the color of the eye anatomic structures. Sample based methods can perceive eye structures according to the eye photometric definition. These methods require large amounts of data where different people have eye samples at different light, color, or angle conditions. In this experiment authors make test with 4 students in university. During the investigation they record the eye pupil reaction of stimuli from the beginning of test till the end. Various questions were submitted to the students during the test, while they had to choose only one correct answer. After the test was finished, they processed the recorded video material frame-by-frame. They used Haar classifier to find the eye region. Then, the detected region was passed to the algorithms the detected region was passed to the algorithms for image edge detection and pattern recognition. After, parametric estimates of the identified structures were obtained using the least-squares algorithm. They represent test result on periodograms. Point represents the correct test answer and circle represents the incorrect test answer. They obtained result of the points tend to group together shows that the periodograms are similar, so the students are stressed.

3. METHOD

3.1. Time-domain HRV measures

Time domain measures are done with measuring mean of HR (meanHR), mean of all RR intervals (meanRR), the percentage of heartbeat intervals greater than 50 ms with a difference in successive heartbeat intervals (pNN50), standard deviation of all RR intervals (SDNN) and, root mean square of the difference between successive RR-intervals (rMSSD). In this study meanHR was used to investigate stress of controller.

3.2. Frequency-domain HRV measures

Frequency domain HRV measures based on power spectral density (PSD) analysis of the HRV data. PSD can be calculated by Fast Fourier Transform (FFT) or Autoregressive (AR) based measures which includes energy and variance of HRV. The spectral measures have the advantage of relating the power of variation in different frequency bands to different ANS activities. The common frequency bands of HRV are the very low frequency (VLF, 0–0.04 Hz), low frequency (LF, 0.04–0.15 Hz), and high frequency (HF, 0.15–0.4 Hz). VL and LF frequency is related with SNS activities and HF is related with PNS activities of Heart [20], [33].

$$X(f) = \int_{-\infty}^{\infty} x(t).\, e^{-2j\pi ft}\mathrm{dt} \qquad (3.2)$$

3.3. Classification

In this study Naive Bayes, KStar, Random Forest, and Decision Table classifier algorithms were used. Because of low data Leave One Out Cross Validation method was applied to obtain best accuracy result.

3.3.1. Naïve Bayes

Bayes method mostly used algorithm for medical data classification Naive Bayes model assumes that each of the features it uses are conditionally independent of one another given some class. Figure 3.3 shows the Bayes theorem where P(C|X) is the posterior probability of class (c, target) given predictor (x, attributes), P(C) is the prior probability of class, P(X|C) is the likelihood which is the probability of predictor given class, P(X) is the prior probability of predictor.

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p(x)} \tag{3.3}$$

3.3.2. KStar (K*)

KStar algorithm is an Instant Based (IB) learned, and IB learners are Lazy learner algorithm. Lazy Learner also called less time in training and most time in predicting Classification by KStar is done by summing the probabilities of the new sample into all the members in a category.

3.3.3. Random Forest Classifier

The goal of Random Forest is to combine the decisions of many multivariate trees, each of which has been trained in different training clusters, instead of producing a single decision tree. Pre-loading and random property selections are used when creating different training clusters. The CART algorithm is used when multiple decision trees are created. Before determining the attributes at each level, the attributes are determined by calculating all the trees, then the attributes used in all the trees are combined and the most used attribute is selected. The selected attribute is included in the hierarchy and the same operations are repeated at other levels

.

3.3.4. Decision Table Classifier

The decision table is a compact and precise way to model complex logic. Decision tables take on the task of associating actions with conditions such as flow chart, if-then-else, and switch-case statements, but it does so more efficiently. In test techniques such as end value analysis and equivalence class, there is usually a data input and tests are performed based on these data inputs. However, if different values of the entered values lead to different results from the system, it is difficult to test such a situation with an extreme value analysis or equivalence class. The decision table test technique is more a test technique based on business logic and business rules and is preferred in tests where there are a large number of combinations (eg data entries).

3.4. Experiment Tools

On Literature search it was seen that there are many method to detect stress. Measuring Eye pupil diameter, cardiac activity, Electro Dermal Activity or Restoration are some of them. This investigation is about Air Traffic Controller, so it is not permitted connect some devices to ATC environment, or some stress sensor which may distract their attention while on work schedule. Thus, for better performance accuracy, and to not distract ATC attention, a GSR sensor which connected to user by wearing coat on two fingers (Figure 3.1) and an ear clipped HR sensor (Figure 3.2) were used. These two sensors work with Arduino Uno R3 development cards (Figure 3.3). Arduino Uno is programmable microcontroller board. It is connected to the computer with USB jack and fed with 5V electrical voltage.

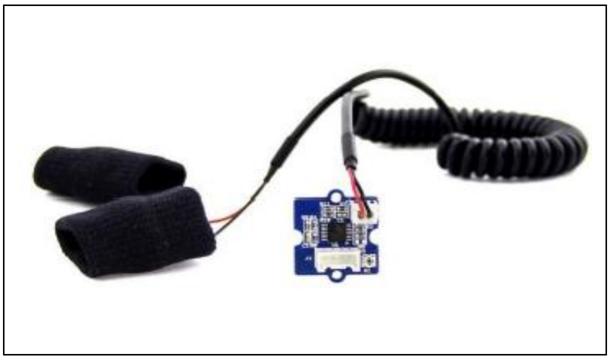

Figure 3.1: GSR Sensor.

Figure 3.2: Ear clipped HR Sensor.

Figure 3.3: Arduino Uno Microcontroller Board.

Arduino UNO needs the program called sketch to run. An ide specialized for Arduino called as Arduino Sketch was used to write the program and upload it to the microcontroller.

To connect sensors to board by avoiding messy wires a Grove Base Shield (Figure 3.4) was plugged into the Arduino. Figure 3.5 shows the detailed hardware block diagram of the experiment system.

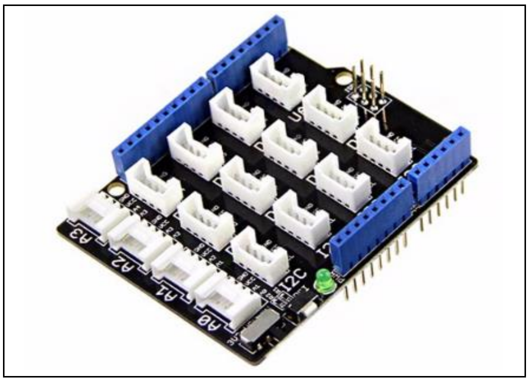

Figure 3.4: Grove Base Shield.

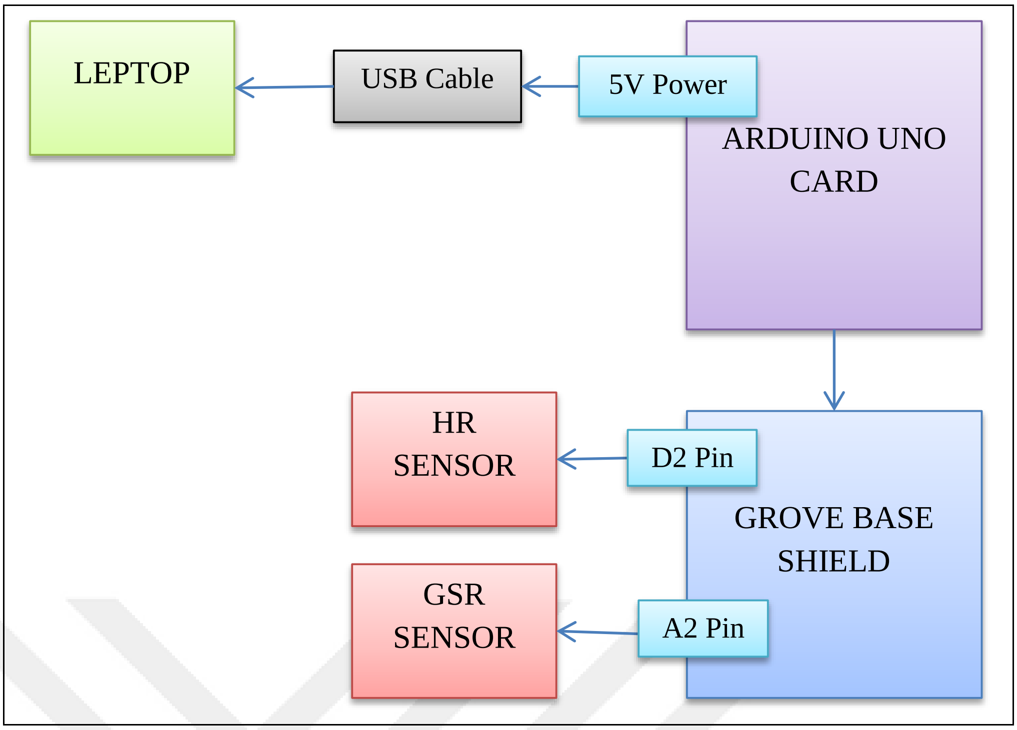

Figure 3.5: Hardware Block Diagram.

3.5. Collection Data

Two separate experiments were done with those sensors. At first experiment GSR and HR data collected from ATC. In second experiment instead of HR data RR data collected. So with RR data HRV measurement and HF, LF transforms could be done. Figure 3.6 shows the block diagram of experiment.

The GSR value obtained from sensor shows the galvanic skin resistance. The skin electrical conductivity value is Siemens (S) value and the skin electrical resistance value is read as ohm (R). When the sensor was free 600 GSR value is read. 100 ms time-frequency was chosen for GSR. Sensor value logged with the time stamp by gobetwino which is free exe tool. Gobetwino catches the Arduino print messages. Figure 3.7 shows the block diagram of the software used.

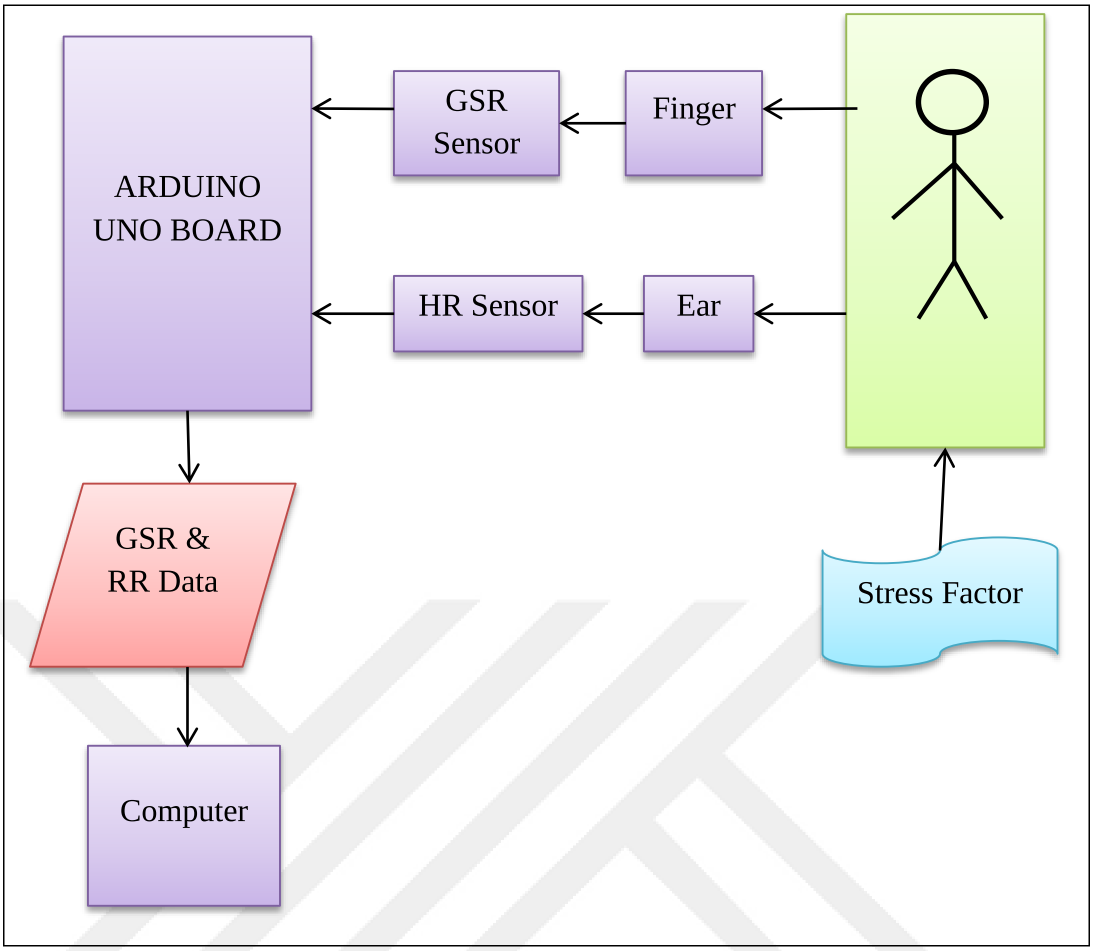

Figure 3.6: Archictecture Block Diagram.

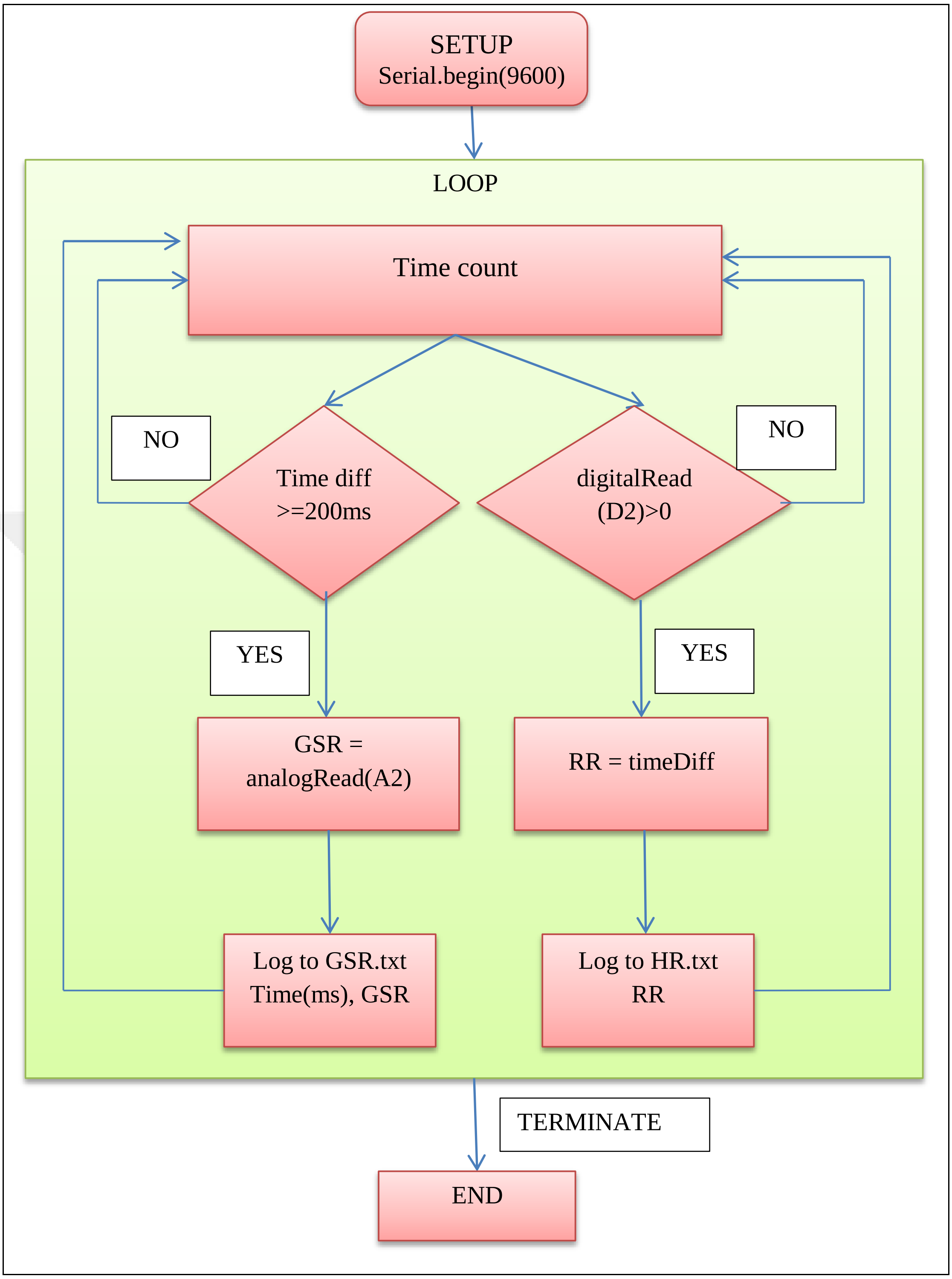

Figure 3.7: Software Block Diagram.

The HR value collected in 20-second time-frequency by an ear clipped HR sensor connected to the microcontroller.

Data collection was done while HR and GSR sensors both connected to the person. For each experiment one log file created. Figure 3.8 shows the contexts of the log file for experiment-1. At second experiment two separated log files created for GSR and HRV values. So that made analyzing process much faster.

```
10.02.2017 15:50:17;_10635-GSR: 47
10.02.2017 15:50:17;_10735-GSR: 47
10.02.2017 15:50:17;_10835-GSR: 47
10.02.2017 15:50:17;_10937-GSR: 47
10.02.2017 15:50:17;_11037-GSR: 46
10.02.2017 15:50:17;_11138-GSR: 47
10.02.2017 15:50:17;_11239-GSR: 47
10.02.2017 15:50:17;_11339-GSR: 47
10.02.2017 15:50:17;_11392-Hert_Rate_is:    93
10.02.2017 15:50:17;_11440-GSR: 47
10.02.2017 15:50:17;_11540-GSR: 47
```

Figure 3.8: Content of Log file created for each experiment.

3.6. Test Protocol

3.6.1. Experiment-1

Experiments were done with 20 (9 female, 11 male) controllers worked at the Turkish Air Traffic Control Center. Before the experiment, each controller was informed about the experiment and agreed with using their results. The working experience of the controllers varies from 6 months to 19 years. The age range of males is between 26 and 35 and females are between 26 and 42. 7 of participants use cigarettes. In this study, the responses to the workload of the controllers and the possible traffic conflict situations were measured in the real environment and in the CWP simulation environment. Both environments have an exactly same user interface and HMI.

The sensor values were collected while controllers on CWP board, out of CWP board and on CWP simulation environment.

In the simulation environment, each of 10 participants has a 6-month controller experience was asked to control traffics in the approach sector on the CWP board for 35 minutes. In a different room, 4 controllers take pseudo pilot positions and implement instructions ordered from voice communication system. At the beginning, the controller has 1 aircraft in the field of responsibility. As time progressed number of traffic increased to 36. During the experiments conflicts were consciously created at 10th and 20th minutes, to monitor the stress level of the controller at conflict time.

In the real CWP environment, 10 minutes sensor values were collected from the subjects with the experience of controlling between 1 and 19 years. During this time, participants were asked to do their routine work. At the moment of operation, the number of traffic within the responsibility of the controllers varies between 12 and 35. Participants were checking transit flights and generally making traffic separation. In this environment, the events that can cause the stresses and time were recorded by monitoring the operation of the controllers.

4-minute sensor values were collected from the subjects out of real and simulation CWP board, to obtain the normal GSR and HR values of the subjects.

3.6.1.1. Data Analysis

As a result of experiment 33 log file collected where each log file includes GSR and HR values of ATC during the time. Before analysis GSR and HR values separated from the log file and 2 new files created for each of sensor data. For each participant and HR and GSR, a graphic was drawn through those data.

Table 3.1 shows the min, max, and average values of GSR and HR data of participant of being tested on CWP environment, Table 3.2 shows the values while participants were on Simulation Environment and Table 3.3 show the values of while participants were relaxing. When take a look on these three tables it is obvious to see that GSR values are bigger on Table 3.3 and lesser on Table 3.2. Therefore, by only GSR values, it can have already been told that, participants of Table 3.2 might be more stressed than participants of Table 3.3. The average values of GSR 335 and HR 85 were obtained while controllers were relaxing. In the simulation environment, participants' GSR values were observed to fall below the initial level due to increased traffic and intensity. For simulation environment these values measured as 100 GSR and 98 HR, respectively. In the real CWP environment average of 230 GSR and 80 HR were measured.

Table 3.1: GSR and HR value of CWP Environment.

User	Min GSR	Max GSR	Avr. GSR	Min HR	Max HR	Avr. HR
User1	295	321	307	80	110	96
User2	319	379	346	57	73	64
User3	224	342	259	84	129	104
User4	157	258	185	59	109	84
User5 t1	51	132	76	76	102	86
User5 t2	114	188	152	78	115	96
User6	186	248	221	63	90	69
User7	113	171	133	72	100	82
User8	193	250	222	62	106	89
User9	188	264	224	67	93	83

Table 3.2: GSR and HR value of Simulation Environment.

User	Min GSR	Max GSR	Avr. GSR	Min HR	Max HR	Avr. HR
User11	24	276	116	94	128	104
User12	127	241	183	72	92	82
User13	0	211	28	82	175	100
User14	170	310	202	70	93	79
User15 t1	7	163	92	69	93	81
User15 t2	0	50	15	73	90	81
User16	0	68	19	73	106	89
User19	177	313	252	87	128	99

Table 3.3: GSR and HR values while Relaxing.

User	Min GSR	Max GSR	Avr. GSR	Min HR	Max HR	Avr. HR
User1	401	431	418	66	82	73
User5	203	292	257	83	103	93
User6	274	471	355	64	82	69
User7	248	339	273	66	88	76
User9	257	297	278	92	127	98
User11	309	403	342	79	91	85
User13	103	168	138	72	103	86
User15	84	382	211	63	98	77
User16	121	161	142	77	95	83
User17	329	382	356	86	107	96
User18	234	449	334	81	99	92

3.6.1.2. GSR Analyze by Plot Graph

The GSR values collected from simulation environment show a falling line for all users, however, while relaxing these values did not show any significant changes. On real CWP environment except conflict occurrence GSR values also follow in the same level but it was under the level of relaxing. In Figure 3.9 GSR values of User-11 while relaxing initially was measured about on 400 levels. Then it gradually drops to 340 levels and with little changes it flows on 340 lines. This shows that in first 20 second subject might be a little stressed. Then him stress level didn't changed. It can be explained as subject felt a bit uncomfortable with coat of GSR sensor which put on his finger. This initial drop of GSR value was seen on all GSR plots of participants. Figure 3.10 shows the GSR plot of User-11 while he was on Simulation Environment. It can be seen on this plot, participant GSR value always decreased. This give information about subject stressed and his stress level increased during Simulation test. On simulation environment subjects started with controlling 1 aircrafts, when time passed traffic count increased to 36. Intense traffic and workload make pressure on subject. So seeing such a result like this plot did not surprised.

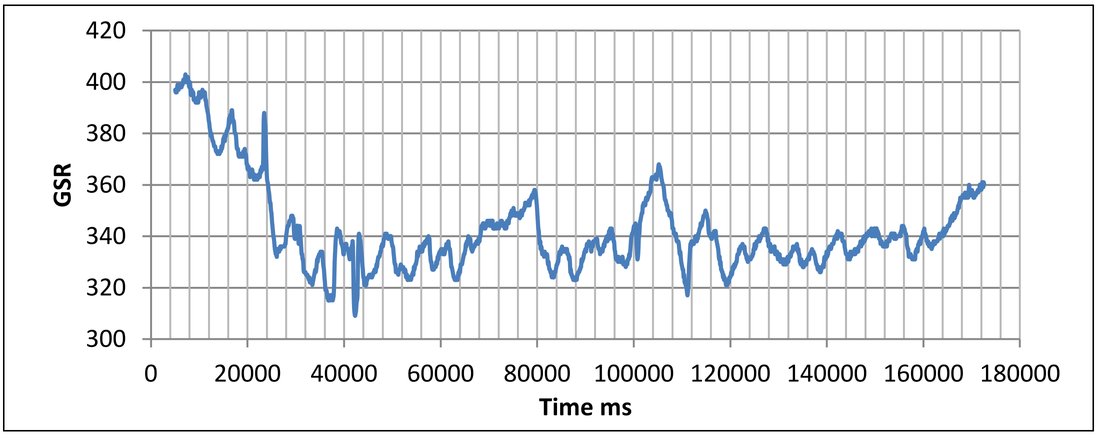

Figure 3.9: User-11 GSR values of while relaxing.

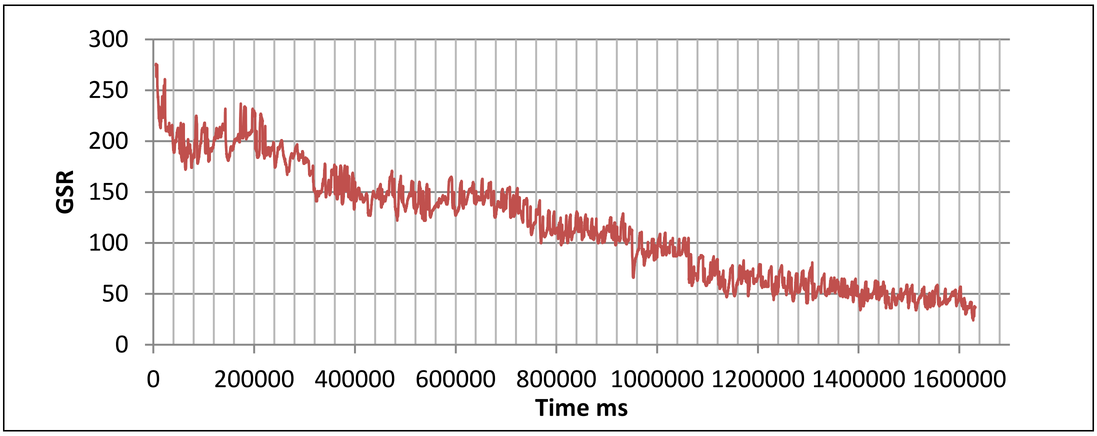

Figure 3.10: User-11 GSR values of while on Simulation Env.

Figure 3.11 shows the plot of GSR values of participant-1 while she was relaxing. As participants-11, GSR values of participiant-1 don't show any significant changes while relaxing. This is expected result, because there is no stress factor to make participant stressed. Most of Participant of on CWP environment has more than 10 years experience. Experience of such those years helps controllers to overcome high work load and unpredicted events. Participant-1 has 19 years experience in controller profession. Figure 3.12 shows the plot of GSR values of Participiant-1 on CWP environment. GSR values show lower level than while relaxing, but flow on same level. This result can be explained as participiant-1 has a stress level while she was working, and this has not been triggered by any other event than normal operation.

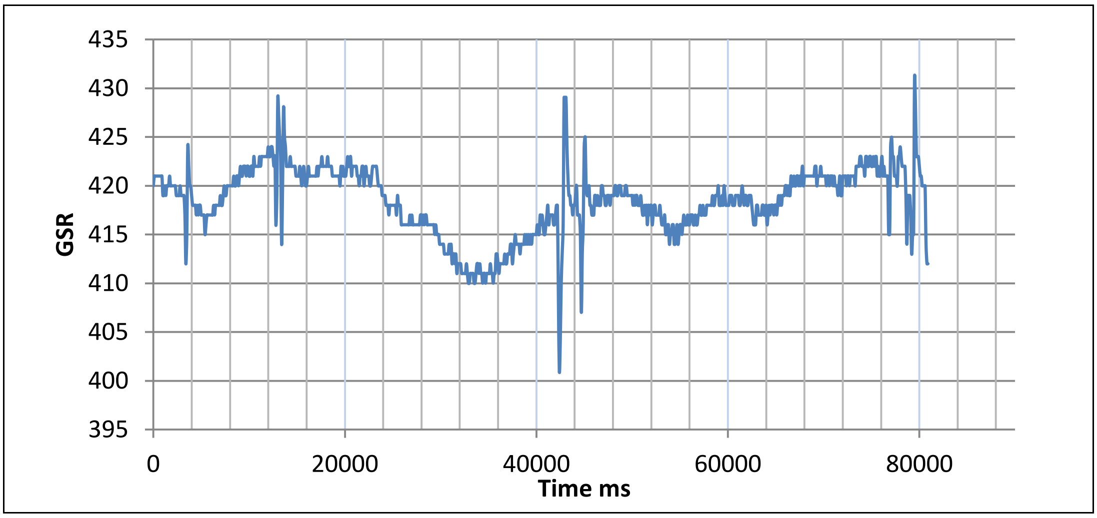

Figure 3.11: User-1 GSR values of while relaxing.

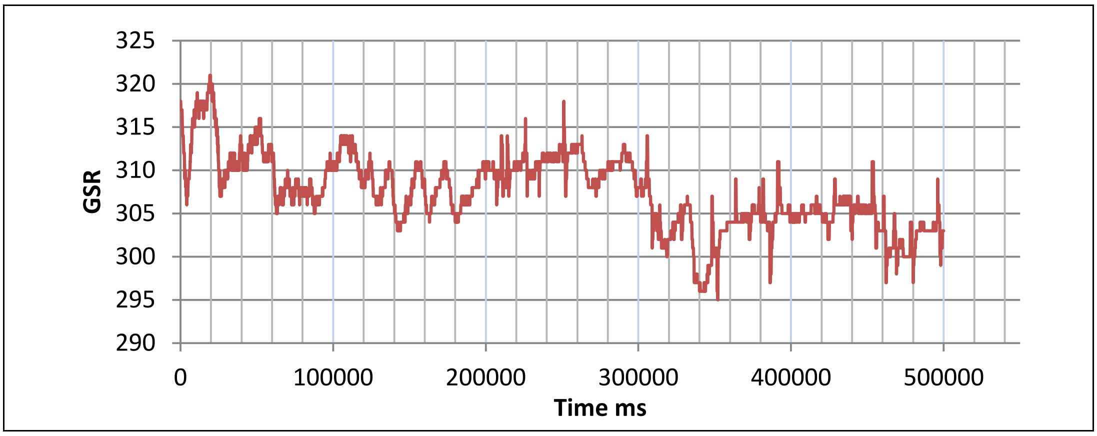

Figure 3.12: User-1 GSR values of while relaxing, of while on Real CWP Env.

Figure 3.13 shows the plot of GSR values of participant-4 while she was working CWP environment. Initially her GSR value was on 230 levels. At 100 second she came up with Short Term Conflict Alert (STCA) which is most important alert type controller faced. STCA warns controller that there are two aircraft about to be conflicted. If ATC don't separate them, those aircrafts might be ended with crashing.

Arrow on Figure 3.13 points the time of conflict occur and GSR changes of participiant-4 on that time. Sharp decrease shows that participant stress level increased. When the time progressed, her GSR value didn't increased. This can be explained as she was still under effect of STCA.

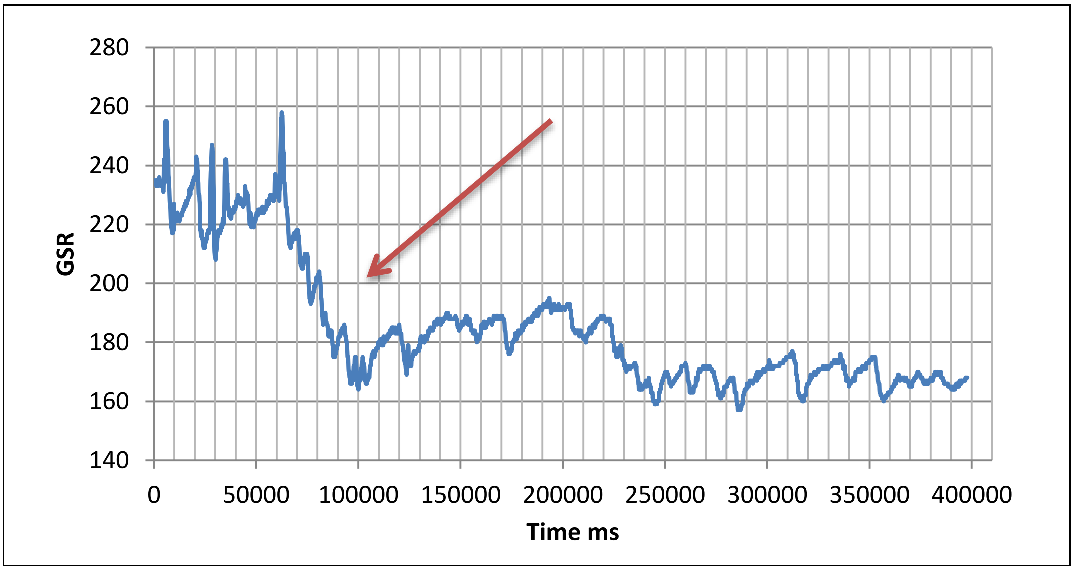

Figure 3.13: User-4 GSR values when conflict occured at real CWP Environment.

3.6.1.3. HR Analyze by Plot Graph

When comparing HR values, it was seen that average of HR values was much bigger when the controller on CWP and on simulation environment. By itself, HR result didn’t show the logical result as GSR changes did. While GSR values show a decreasing line for User11 on simulation environment which is seen in Figure 3.10, the HR value seen on Figure 3.15 showed rising values at some interval, while on other intervals showed declining values. This showed that HR graphics is not enough to extract meaningful information about the stress of person. When take a look plot of HR values of participant-4 (Figure 3.14), same explanation can be done.

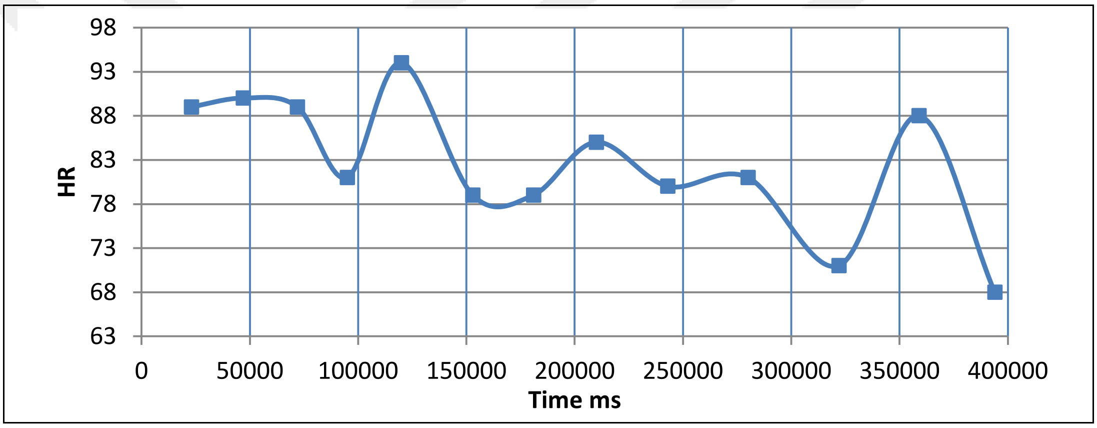

Figure 3.14: User-4 HR values at real CWP Environment.

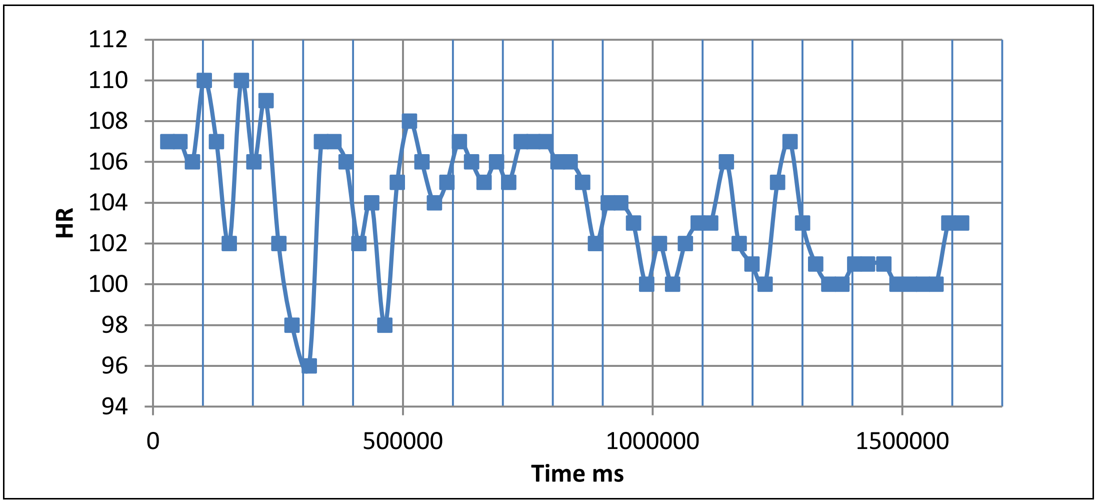

Figure 3.15: User-11 HR values on Simulation Environment.

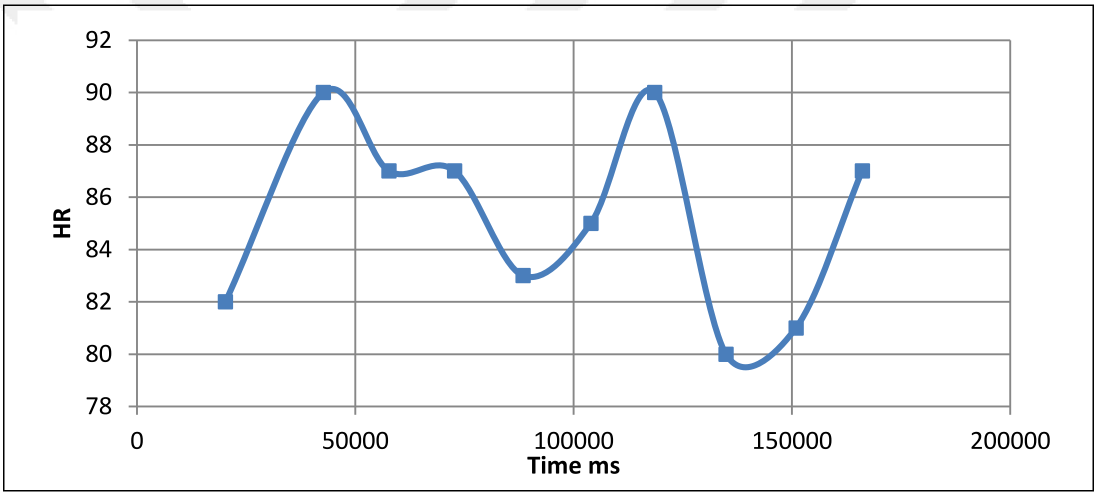

Figure 3.16: User-11 HR values while relaxing.

3.6.1.4. Stress Classification

Datasets were created with values selected from the time ranges of the conflict occurs while experiment was done in Real CWP Environment, from the time ranges of the traffic intensity arises in the Simulation environment, and from the time ranges of the while ATC are relaxing. Those time intervals previously noted while observing controller deal with conflict and on simulation. Simulation data and conflict data labeled as stressed, relaxing data and real CWP data which during experiment flow in the same level labeled as not stressed as shown in Table 3.4.

To verify the different tests, WEKA learning machine was used for testing. Because of having low sample Leave One Out Cross Validation (LOOCV) method should have been applied. So k fold number was chosen as 33 for LOOCV. Naïve Bayes method mostly used algorithm for medical data classification. Naive Bayes model assumes that each of the features it uses are conditionally independent of one another given some class. For this experiment Navie Bayes method applied to data, a good result was obtained from Naive Bayes algorithm and Naïve Bayes Multinomial algorithm as 91% success. Naive Bayes Multinomial classifier is a specific instance of a Naive Bayes classifier which uses a multinomial distribution for each of the features. Secondly, KStar algorithm was applied to data. KStar algorithm is an Instant Based (IB) learned, and IB learners are Lazy learner algorithm. Lazy Learner also called less time in training and most time in predicting Classification by KStar is done by summing the probabilities of the new sample into all the members in a category. 87% success of classifying stress state was obtained from KStar algorithm. While Naïve Bayes Multinomial Algorithm was much sufficient to find not-stressed state, Naïve Bayes and KStar algorithm were very effective in detecting the stress situation as can be seen in Table 3.5, 3.6, 3.7 showed the confusion matrix of algorithms.

Table 3.4: Classification Table.

GSR	HR	CLASS
353	64	notstressed
336	68	stressed
259	103	notstressed
226	83	notstressed
195	88	stressed
307	95	notstressed
418	73	notstressed
355	68	notstressed
221	69	notstressed
224	94	notstressed
206	87	stressed
224	82	notstressed
257	91	notstressed
152	95	stressed
76	85	stressed
206	79	notstressed
144	84	stressed
273	82	notstressed
133	76	stressed
334	92	notstressed
28	100	stressed
138	86	stressed
278	98	notstressed
92	81	stressed
15	81	stressed
211	77	notstressed
19	89	stressed
142	83	stressed
356	96	notstressed
116	104	stressed
342	85	notstressed
252	99	stressed
399	88	notstressed

Table 3.5: Confusion Matrix of Naïve Bayes.

Stressed	Not-Stressed	
13	2	Stressed
1	17	Not-Stressed

Table 3.6: Confusion Matrix of Naïve Bayes Multinomial.

Stressed	Not-Stressed	
12	3	Stressed
0	18	Not-Stressed

Table 3.7: Confusion Matrix of KStar.

Stressed	Not-Stressed	
13	2	Stressed
2	16	Not-Stressed

Table 3.8: Naïve Bayes Classifier Detailed Accuracy.

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,867	0,056	0,929	0,867	0,897	0,817	0,900	0,935	stressed
	0,944	0,133	0,895	0,944	0,919	0,817	0,900	0,83	notstressed
Weigh Avg.	0,909	0,098	0,910	0,909	0,909	0,817	0,900	0,878	

Table 3.9: Naïve Bayes Multinomial Classifier Detailed Accuracy.

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,800	0,000	1,000	0,800	0,889	0,828	0,919	0,945	stressed
	1,000	0,200	0,857	1,000	0,923	0,828	0,922	0,872	notstressed
Weigh Avg.	0,909	0,109	0,922	0,909	0,908	0,828	0,921	0,905	

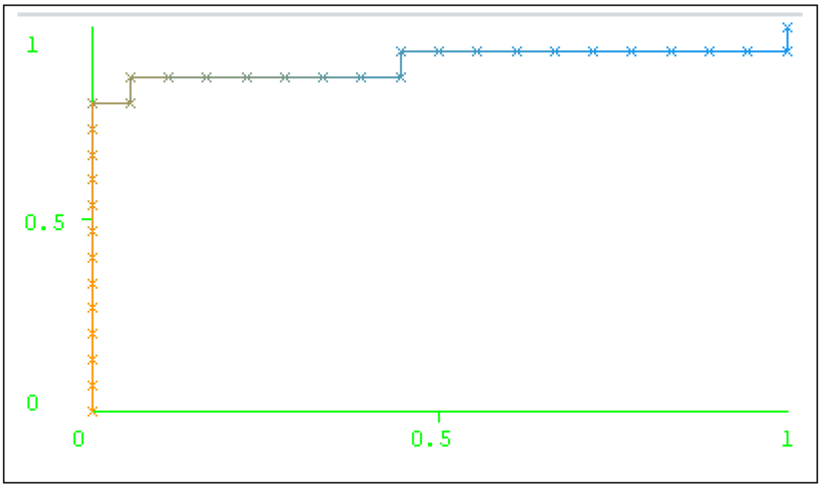

Figure 3.17: Naïve Bayes ROC curve of Stress State.

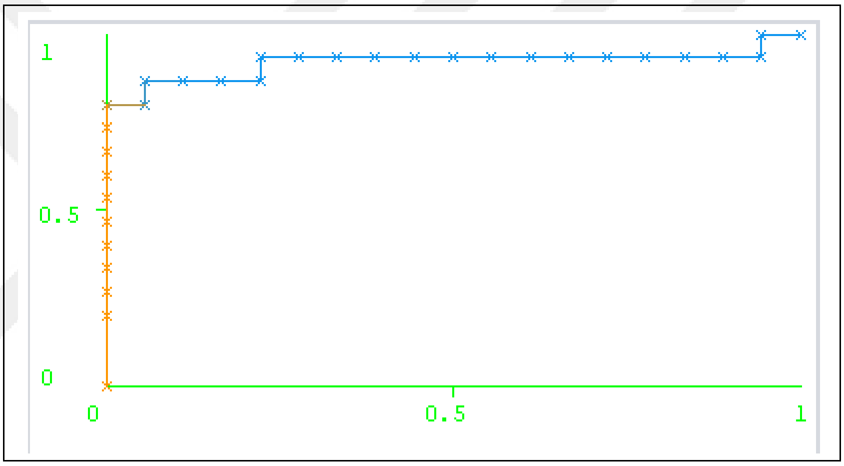

Figure 3.18: Naïve Bayes Multinomial ROC curve of Stress State.

3.6.2. Experiment-2

On experiment-1 HR value failed to help detect stress of ATC by alone. Because collecting data was done in 20 second period. Obtained HR result after 20 seconds is not enough to tell about a person stress, and that time period is enough for an aircraft to collapse or lost. Instead HR value, HRV value might help better to detect stress. So in second experiment, some experiment was done by focusing on GSR and RR data.

For some extra ordinary reasons, it is not possible to repeat the experiment with same participants or on the same environment. Experiment-2 was done on a simple fruit ninja online game. That there some fruit jumped on screen and user cut those fruits before they drop. If gamer miss cut fruit or cut bomb which seen on the screen instead of fruit he/she punished with one life. This game was chosen for 3 reasons. First of all, patient will be in a situation to decide correct decision among countless. Second, he/she would have time obligation for a correct decision where ATC also have. Third, this game was easy to train and it reaches its difficulty level in short time period, so the stress of patient would be measured in short time period task.

Experiments were done with 10 participants 2 female 8 male. The age range of participants is between 26 and 35. Before the experiment, 2 minutes data were collected while they were in relaxing position and those data labeled as not stressed. Then 6-minute fruit ninja task was given to the subject. If one participant lost his/her all 4 lives before task end, he/ she continued on task by restarting the game.

3.6.2.1. Data Analysis

HRV value analyzed with Kubios HRV, which is advanced and easy to use software for heart rate variability (HRV) analysis. Kubios provide some features to analyze RR data. It provides Time Domain method, Nonlinear method and, Frequency Domain Method. By frequency domain method, RR time domain interval transformed to the frequency domain in order to get HF, LF and LF/HF values. Here, Fast Fourier Transform (FFT) and Autoregressive (AR) based method applied to RR data to calculate HRV.

The common frequency bands of HRV are the very low frequency (VLF, 0–0.04 Hz), low frequency (LF, 0.04–0.15 Hz), and high frequency (HF, 0.15–0.4 Hz).

The algorithm used here to analyze HRV values provides 30-second window frames. Therefore, each 6-minute RR data was split into 12 windows frames. Table 5-2 shows the Frequency domain value of HRV. On each frame, FFT and AR were applied. LF value shows the activity of the SNS, and HF shows the PNS activities [7].

After Experiment-2 60 dataset were collected. Each dataset contains the GSR and LF/HF ratio of participations. High LF ratio in LF/HF indicated the subject in a stress state. FFT and AR methods both provide LF/HF ratio. In experiment FFT transformation was chosen for HRV analysis because it showed much close behavior with GSR value than AR. Table 3.10 and 3.11 show the frequency result of HRV of user-6 and Figure 3.19 shows GSR graphic of user-6. During 90 seconds GSR value shows that user is not in a stress state and LF/HF ratio of FFT spectrum also tells HF value is bigger which is related with PNS activities, so subject is not stressed. Between 90-120 seconds GSR tends to decrease which means that subject in a stress state. LF/HF ratio of FFT also tells SNS activities in charge, therefore subject is in stress state.

Table 3.10: User 6 Frequency Domain HRV results (Frame 1, 2, 3).

	Frame 1		Frame 2		Frame 3		
Frequency-Domain Results	FFT spectrum,	AR spectrum,	FFT spectrum,	AR spectrum,	FFT spectrum,	AR spectrum	…
Peak frq.							…
VLF (Hz):	0.033333	0.040000,	0.040000,	0.040000,	0.040000,	0.040000,	…
LF (Hz):	0.150000	0.100000,	0.060000,	0.150000,	0.050000,	0.150000,	…
HF (Hz):	0.173333	0.150000,	0.366667,	0.350000,	0.383333,	0.326667,	…
Absolute powers							…
VLF (ms^2):	705.247	13.572.662,00	5.109.686,00	5.390.749,00	1.873.402,00	4.635.672,00	…
LF (ms^2):	7.204.395	54.458.012,00	30.734.981,00	18.456.432,00	4.983.836,00	15.280.329,00	…
HF (ms^2):	15.395.025	26.002.498,00	88.798.983,00	50.772.768,00	35.397.958,00	42.393.467,00	…
VLF (log):	42.560,00	72.132,00	62.363,00	62.899,00	52.329,00	61.390,00	…
LF (log):	65.799,00	86.026,00	80.306,00	75.206,00	62.114,00	73.317,00	…
HF (log):	73.392,00	78.634,00	90.915,00	85.325,00	81.718,00	83.522,00	…
Relative powers							…
VLF (%):	30.233,00	144.311,00	40.786,00	71.869,00	43.943,00	74.082,00	…
LF (%):	308.847,00	579.023,00	245.328,00	246.060,00	116.902,00	244.193,00	…
HF (%):	659.972,00	276.471,00	708.798,00	676.899,00	830.300,00	677.486,00	…
Normalzd powers							…
LF (n.u.):	318.475,00	676.675,00	255.759,00	265.113,00	122.275,00	263.731,00	…
HF (n.u.):	680.548,00	323.097,00	738.936,00	729.315,00	868.463,00	731.691,00	…
Total power (ms^2):	23.326.771,00	94.051.508,00	125.281.120,00	75.007.854,00	42.632.728,00	62.574.700,00	…
LF/HF ratio:	0.4680,	2.0943,00	0.3461,	0.3635,	0.1408,	0.3604,	…

Table 3.11: User 6 Frequency Domain HRV results (Frame 4, 5, 6).

	Frame 4		Frame 5		Frame 6		
Frequency-Domain Results	FFT spectrum ,	AR spectrum,	FFT spectrum,	AR spectrum,	FFT spectrum ,	AR spectrum	...
Peak frq.							...
VLF (Hz):	0.040000 ,	0.040000,	0.040000,	0.040000,	0.040000,	0.040000 ,	...
LF (Hz):	0.123333 ,	0.110000,	0.123333,	0.126667,	0.046667,	0.090000 ,	...
HF (Hz):	0.150000 ,	0.360000,	0.150000,	0.150000,	0.373333,	0.366667 ,	...
Absolute powers							...
VLF (ms^2):	1.560.130,00	2.148.740,00	95.786,00	236.752,00	255.895,00	152.327,00	...
LF (ms^2):	9.126.458,00	14.274.942,00	16.236.608,00	9.270.000,00	872.284,00	1.381.920,00	...
HF (ms^2):	7.120.136,00	9.886.411,00	5.780.831,00	5.514.678,00	1.728.022,00	1.266.112,00	...
VLF (log):	50.499,00	53.701,00	22.595,00	31.644,00	32.422,00	27.234,00	...
LF (log):	68.163,00	72.637,00	73.924,00	68.320,00	44.685,00	49.286,00	...
HF (log):	65.681,00	68.963,00	63.597,00	63.126,00	51.521,00	48.411,00	...
Relative powers							...
VLF (%):	87.436,00	81.575,00	0.4331,	15.748,00	89.080,00	54.157,00	...
LF (%):	511.487,00	541.938,00	734.119,00	616.593,00	303.653,00	491.319,00	...
HF (%):	399.044,00	375.331,00	261.373,00	366.808,00	601.546,00	450.145,00	...
Normalzd powers							...
LF (n.u.):	560.494,00	590.074,00	737.312,00	626.458,00	333.347,00	519.451,00	...
HF (n.u.):	437.278,00	408.668,00	262.510,00	372.677,00	660.372,00	475.920,00	...
Total power (ms^2):	17.843.006,00	26.340.537,00	22.117.131,00	15.034.224,00	2.872.637,00	2.812.673,00	...
LF/HF ratio:	1.2818,00	1.4439,00	2.8087,00	1.6810,00	0.5048,	10.915,00	...

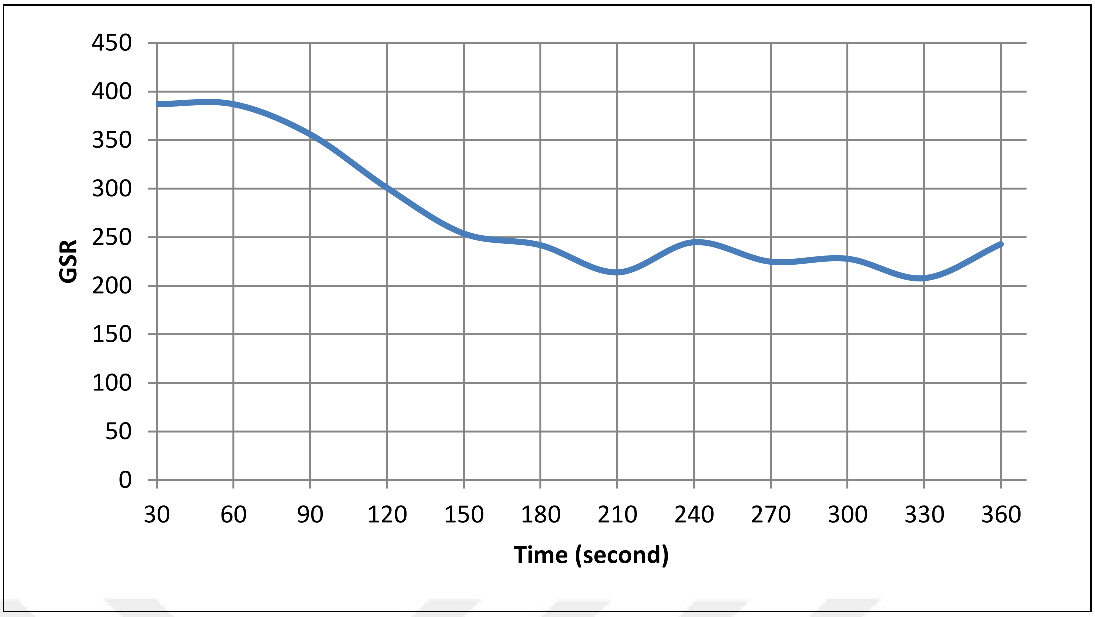

Figure 3.19: User-6 GSR values.

3.6.2.2. Stress Classification

To verify the different tests, WEKA learning machine was used for testing. For the better accurate result, LOOCV method was chosen. Random Forest, Bayes Net and Decision table algorithms applied to data. A good result was obtained with Random Forest as 86% success of classifying of stress, secondly Bayes Net algorithm and Decision Table gave 85% success of classifying of stress. Among three algorithms, Random Forest was much efficient to find stress state as it correctly labeled 30 dataset as stressed where 35 dataset were initially labeled as stressed, and correctly labeled 22 dataset as notstressed where 25 dataset were initially labeled as not stressed. The results of Decision Table and Bayes Net differ by a number than Random Forest result. Bayes Net and Decision Table were able to correctly classify 29 dataset as stressed and 22 dataset as not stressed. Table 3.12, Table 3.13, Table 3.14 shows the confusion matrix of algorithms. It is possible tell classification result was successful by looking Roc Curve of Random Forest algorithm which is seen on Figure 3.20 and Decision Table algorithm seen on Figure 3.21.

Table 3.12: Confusion Matrix of Random Forest Classifier.

Stressed	Not-Stressed	
30	5	Stressed
3	22	Not-Stressed

Table 3.13: Confusion Matrix of Bayes Net.

Stressed	Not-Stressed	
29	6	Stressed
3	22	Not-Stressed

Table 3.14: Confusion Matrix of Decision Table.

Stressed	Not-Stressed	
29	6	Stressed
3	22	Not-Stressed

Table 3.15: Random Forest Detailed Accuracy.

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,857	0,120	0,909	0,857	0,882	0,730	0,906	0,941	stressed
	0,880	0,143	0,815	0,880	0,846	0,730	0,906	0,852	notstressed
Weig. Avg.	0,867	0,130	0,870	0,867	0,867	0,730	0,906	0,904	

Table 3.16: Bayes Net Detailed Accuracy.

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,829	0,120	0,906	0,829	0,866	0,700	0,729	0,850	stressed
	0,880	0,171	0,786	0,880	0,830	0,700	0,729	0,741	notstressed
Weig. Avg.	0,850	0,141	0,856	0,850	0,851	0,700	0,729	0,805	

Table 3.17: Decision Table Detailed Accuracy.

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,829	0,120	0,906	0,829	0,866	0,700	0,746	0,816	stressed
	0,880	0,171	0,786	0,880	0,830	0,700	0,746	0,692	notstressed
Wg. Avg.	0,850	0,141	0,856	0,850	0,851	0,700	0,746	0,764	

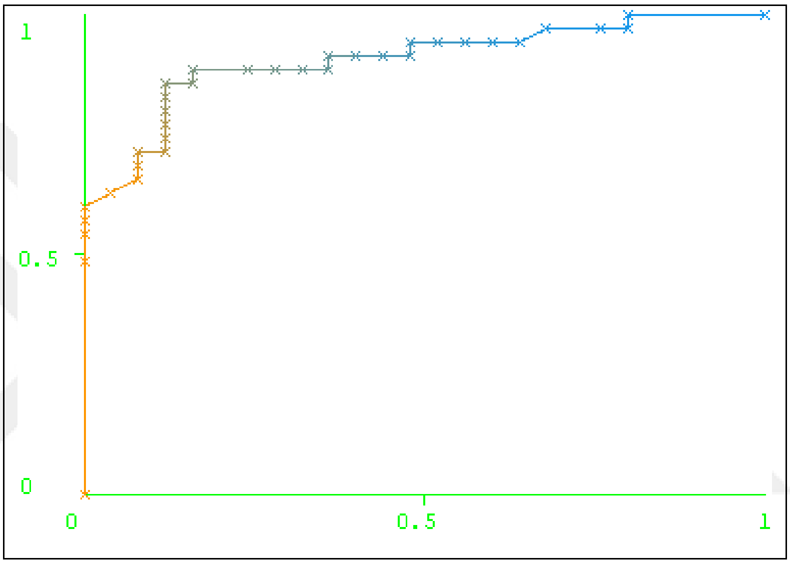

Figure 3.20: Random Forest ROC curve of Stress State.

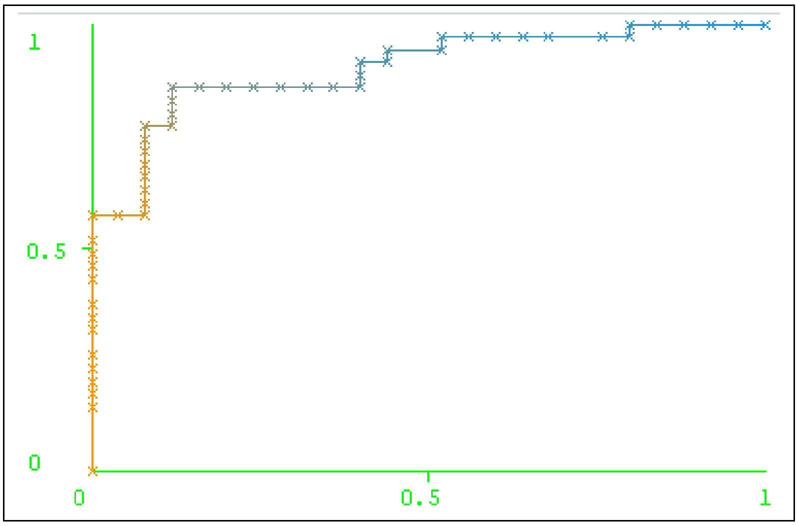

Figure 3.21: Decision Table ROC curve of Stress State.

4. RESULT

Two separate experiments have been done with GSR and HR sensors. Both experiment-1 and experiment-2 showed that the GSR values are more sensitive to events in the stress detection of ATC. That makes it possible to use the GSR sensor as a priority sensor for stress definition.

On experiment-1 the HR values were obtained in 20 seconds period. These values were analyzed in Time Domain method. Mean of HR for subject shows bigger value while on stress task. Thus, it is obvious to say subjects get stressed. But to specify in which time interval they get stressed Time Domain method did not help. GSR values were collected in 200ms interval. The plots drown with GSR data shows the changes in time interval. By looking GSR plot of subject it is possible to tell estimated stressed time. But by looking the time plot of HR data, it was almost impossible to estimate at which time interval subject stressed. HR data was sampled in 20 second interval. By alone HR result didn't enough to detect stress. To interpret result more HR data was need. In ATC domain every single second is important. The need for more time to perform stress detection with the HR value while already 20 seconds long is indicative that this method is not a suitable method for ATCs. Because that 20 second time period is enough for an aircraft to collapsed or lost. So with GSR sensor stress changes can obtain faster. The variety of GSR value in 2 seconds show us how GSR sensor sensitive to the event. That is the second reason why GSR sensor must be used as priority sensor for detecting stress of ATC.

In the real environment, the GSR values at the CWP board from beginning to end were at the same level unless a conflict occurs, but this value tended to decrease in the simulation environment. When the two situations are compared, it is clear to see that in the simulation environment the stress level of the participant increases with the increasing workload. Work experience was another factor. It was seen on experiment-1, GSR values of controllers who have over 10 years of experience and control the same amount of traffic flow at the same level.

On experiment-2 RR data was used instead of HR data to investigate what if RR data used in experiment-1. For some extra ordinary reasons, it is not possible to repeat the experiment with same participants or on the same environment. Experiment-2 was done with 10 partcipiant on a simple fruit ninja online game. HRV features were analyzed in PSD. While on test phase GSR and RR data show changes. Experiment-2 showed that HRV value and GSR value both are successful to detect stress of subjects. With choosing RR feature and PSD analysis it was become possible to detect stress of subject in 30 second window. While subject stressed power of LF showed bigger value, and while relaxing HF values shows bigger value. To get the stress of ATC with hybrid sensor GSR and HRV sensor both can be used together. While in extraordinary condition or specific illness like "Hand Sweeting", result of one another can be used as confirmatory. So alertness of ATC would be detected by two different effects of ANS activity and that would be the more accurate result when decision made based on one ANS activity.

REFERENCES

[1] Web 1, (2017), https://en.wikipedia.org/wiki/Air_traffic_controller, (Date of Access: 01/04/2017).

[2] Whitfield D., Jackson A., (1982), "The air traffic controller's picture as an example of a mental model", IFAC, 15 (6), 37-44.

[3] Zeier H., (1994), "Workload and psychophysiological stress reactions in air traffic controllers", Ergonomics, 37 (3), 525-539.

[4] Joulia A., Le Tallec J. C., (2007), "UAVs: What and how to sense? To avoid what?.", 2nd AIAA Infotech@ Aerospace Conference and Exhibit, Rohnert Park, California, USA, 7-10 May.

[5] Ming E. E., Adler G. K., Kessler R. C., Fogg L. F., Matthews K. A., Herd J. A., Rose R. M., (2004), “Cardiovascular reactivity to work stress predicts subsequent onset of hypertension: the Air Traffic Controller Health Change Study”, Psychosomatic Medicine, 66 (4), 459-465.

[6] Costa G., (1996), “Occupational stress and stress prevention in air traffic control”, Technical Report No: ISBN 92-2-110070-7, Institute of Occupational Medicine, University of Verona, Italy.

[7] Bourne L. E., Yaroush R. A., (2003), "Stress and cognition: A cognitive psychological perspective", Technical Report No: NASA 2003-2012282, CA: NASA Ames Research Center, NASA, USA.

[8] Web 2, (2017), http://tatcaankara.org/, (Date of Access: 05/03/2017).

[9] Cooper C. L., Dewe P. J., (2004), ”Stress: A Brief History”, 1st Edition Blackwell Publishing.

[10] Dickerson S. S., Kemeny M. E., (2004), “Acute stressors and cortisol responses”, Psychological Bulletin, 130 (3), 355-391.

[11] Rai D., Kosidou K., Lundberg M., Araya R., Lewis G., Magnusson C., (2011), “Psychological distress and risk of long-term disability”, Journal of Epidemiology and Community Health, 66 (7), 586-592.

[12] Schmidt A., Thews G., (1989), "Autonomic Nervous System", 2nd Edition, Springer-Verlag.

[13] Gillian P., (2006), Human Physiology, 3rd Edition, Oxford University Press.

[14] Darrow C. W., (1964), “The rationale for treating the change in galvanic skin response as a change in conductance”, Psychophysiology, 1 (1), 31–38.

[15] Niall C., Knapp R. B., (2008), "Sensory Chairs: A system for biosignal research and performance", 8th International Conference on New Interfaces for Musical Expression NIME08, 233-237, Genova, Italy, 5-7 June.

[16] Levenson R. W., Ekman P., Heider K., Friesen W. V., (1992), "Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra", Journal of personality and social psychology, 62 (6), 972.

[17] Sun F. T., Kuo C., Cheng H. T., Buthpitiya S., Collins P., Griss M., (2010), "Activity-aware mental stress detection using physiological sensors", Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering , 76, 282-301.

[18] Villarejo M. V., Zapirain B. G., Zorrilla A. M., (2012), "A stress sensor based on Galvanic Skin Response (GSR) controlled by ZigBee" Sensors, 12 (5), 6075-6101.

[19] Papathanasiou G., Georgakopoulos D., Papageorgiou E., Zerva E., Michalis L., Kalfakakou V., Evangelou A., (2013), "Effects of smoking on heart rate at rest and during exercise, and on heart rate recovery, in young adults", Hellenic J Cardio, l54 (3), 168-177.

[20] Seong H. M., Lee J. S., Kim W. S., Lee H. S., Youn Y. R., Shin T. M., (2004), “The analysis of mental stress using time-frequency analysis of heart rate variability signal”, Journal of Biomedical Engineering Research, 25 (6), 581-587.

[21] Sandulescu, V., Andrews S., Ellis D., Bellotto N., Mozos O. M., (2015), “Stress detection using wearable physiological sensors”, Lecture Notes in Computer Science, 9107, 526-532.

[22] Mozos O. M., Sandulescu V., Andrews S., Ellis D., Bellotto N., Dobrescu R., Ferrandez J. M., (2017), “Stress detection using wearable physiological and sociometric sensors”, International journal of neural systems, 27 (2), 1650041.

[23] Rowe D. W., Sibert J., Irwin D., (1998), “ Heart rate variability: Indicator of user state as an aid to human-computer interaction”, Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 480-487, Los Angeles, California, USA, 18-23 April.

[24] Vrijkotte T. G., Van D. L. J., De Geus E. J., (2000), “Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability”, Hypertension, 35 (4), 880-886.

[25] Stroop J., (1935), “Studies of interference in serial verbal reactions”, Journal of Experimental Psychology, 18, 643-661.

[26] Clifford G. D., Tarassenko L., (2005), "Quantifying errors in spectral estimates of HRV due to beat replacement and resampling", IEEE transactions on biomedical engineering, 52 (4), 630-638.

[27] Bakker J., Mykola P., Natalia S., (2011), "What's your current stress level? Detection of stress patterns from GSR sensor data", IEEE 11th International Conference on Data Mining Workshops, 573-580, Vancouver, Canada, 11 December.

[28] Jongyoon C., Gutierrez O., (2011), "R. Removal of Respiratory Influences from Heart Rate Variability in Stress Monitoring", IEEE Sensor Journal, 11 (11), 2649-2656.

[29] Beyaz A., Beyaz R., (2015), "Traktörde Güvenlik AmaçlıI Galvanik Deri Tepkisi Sensörü Kullanım Olanaklarının Belirlenmesi", Mühendislik Bilimleri ve Tasarım Dergisi, 3 (3), 121-125.

[30] Lavine R. A., Sibert J. L., Gokturk M., Dickens B., (2002), "Eye-tracking measures and human performance in a vigilance task", Aviation, space, and environmental medicine, 73 (4), 367-372.

[31] Karakaya O., Barutcu I., Kaya D., Esen A. M., Saglam M., Melek M., Ersel O., Muhsin T., Özlem B. E. & Kaymaz C., (2007), "Acute effect of cigarette smoking on heart rate variability", Angiology, 58 (5), 620-624.

[32] Tarvainen M.P., Karj P.A., Koistinen A.S., Valkonen-Korhonen M.V., (2000), "Principal Component Analysis of Galvanic Skin Responses", In Proceedings of 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 6075-6101, Piscataway, NJ, USA 23-28 July.

[33] Tarvainen M. P., Niskanen J. P., Lipponen J. A., Ranta-Aho P. O., Karjalainen P. A., (2014), "Kubios HRV–heart rate variability analysis software", Computer methods and programs in biomedicine, 113 (1), 210-220.

[34] Web 3, (2017), https://en.wikipedia.org/wiki/Naive_Bayes_classifier, (Date of Access: 10/07/2017).

[35] Perkins, A., (1994), "Saving money by reducing stress", Harvard Business Review, 72 (6), 12.

[36] Zhai J., Barreto A., (2006), "Stress detection in computer users based on digital signal processing of noninvasive physiological variables", In Proceedings of IEEE 28th Annual International Conference on Engineering in Medicine and Biology Society, 1355-1358, New York, NY, USA, 30 August-3 September.

[37] Massot B., Baltenneck N., Gehin C., Dittmarm A., McAdams E., (2011), "EmoSense: An Ambulatory Device for the Assessment of ANS Activity Application in the Objective User Evaluation of Stress with the Blind", IEE Sensors Journal, 12 (3), 543–551.

[38] Welch K. M. A., Bowyer S. M., Aurora S. K., Moran J. E., Tepley N., (2001), "Visual stress induced migraine aura compared to spontaneous aura studied by magnetoencephalography", The journal of headache and pain, 2, 131-136.

[39] Matic A. I., (2014), "Introduction to the Nervous System: The autonomic nervous system and the central nervous system", American Medical Writers Association Journal, 29 (2), 51-55.

[40] Pedrotti M., Mirzaei M. A., Tedesco A., Chardonnet J. R., Mérienne F., Benedetto S., & Baccino T., (2014), "Automatic stress classification with pupil diameter analysis", International Journal of Human-Computer Interaction 30 (3), 220-236.

[41] Henckens M. J., Hermans E. J., Pu Z., Joëls M., Fernández G., (2009), "Stressed memories: how acute stress affects memory formation in humans", Journal of Neuroscience, 29 (32), 10111-10119.

[42] Ritz T., Steptoe A., Dewilde S., Costa M., (2000), "Emotions and stress increase respiratory resistance in asthma", Psychosomatic Medicine, 62 (3), 401-412.

[43] Burrows D., (2002), “Memory overload: The effects of amount of information, rate of presentation, and reorganization”, Cognitive Technology, 7, 20-30.

[44] Gomes L. M. P., Martinho-Pimenta A. J. F., Castelo-Branco N. A. A., (1999), “Effects of occupational exposure to low frequency noise on cognition”, Aviation ,Space & Environmental Medicine, 70 (3), 115-118.

[45] Fowler B., Prlic H., Brabant M., (1994), “Acute hypoxia fails to influence two aspects of short-term memory: Implications for the source of cognitive deficits”, Aviation, Space. & Environmental Medicine, 65, 641-645.

[46] Diaz D., Hancock P. A., Sims V. K., (2002), “Ecological visual search under noise stress”, Present & future research, 23rd Army Science Conference, 98-101, Orlando, Florida, USA, 2-5 December.

[47] Callaway E., Thompson S.V, (1953), “Sympathetic activity and perception”, Psychosomatic Medicine, 15, 443-455.

[48] Smith H. P. R., (1979), “A simulator study of the interaction of pilot workload with errors, vigilance, and decisions”, Technical Report No: NASA Technical Memorandum 78482, Moffett Field, CA: NASA Ames Research Center, USA.

[49] Jacob R. J., Keith S. K., (2003), "Eye tracking in human-computer interaction and usability research”, Ready to deliver the promises, Mind, 2 (3), 573-605.

[50] Ha J. S., Seong P. H., (2014), “Experimental investigation between attentional-resource effectiveness and perception and diagnosis in nuclear power plants”, Nuclear Engineering and Design, 278, 758-772.

[51] Wang Y., Cong W., Dong B., Wu F., Hu M., (2015), "Statistical analysis of air traffic controllers’ eye movements", The 11th USA/Europe ATM R&D Seminar, 453-461, Lisbon, Portugal, 22-26 June.

[52] Salahuddin L., Cho J., Jeong M. G., Kim D., (2007), “Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings”, 29th Annual International Conference of the IEEE, 4656-4659, 23-26 Augst.

[53] Treigys P., Virginijus M., Kaklauskas A., (2012), "Analysis of iris and pupil parameters for stress recognition", Information Technology and Control, 41 (1), 7-14.

[54] Daugman J. G., (1993), “High confidence visual recognition of persons by a test of statistical independence”, IEEE Trans. Pattern Analysis and Machine Intelligence, 1148-1161.

BIOGRAPHY

Aziz Parmaksız was born in Şanlıurfa in 1988. Graduated from Şanlıurfa Atatürk Highschool in 2005, and then in 2006, he started university and took a bachelor degree in Computer Engineering from Izmir Institute of High Technology in 2012, and in the same year, he started work as a researcher for Tubitak Bilgem company. He has been working with a great team and within them, he made first Turkish Air Traffic Controller system for DHMI. While working on Tubitak he started his master degree in Computer Engineering at Gebze Technical University in 2013.

APPENDICES

Appendix A: Publications under the scope of thesis work

Parmaksız A., Göktürk M. (2017), "Air Traffic Controller Alertness Level Detection", 17nd International Conference on Engineering Research and Applications, 356-359, Istanbul, Turkey, 17-18 May.