A Probabilistic Approach for Detection and Analysis of Cognitive Flow

Debatri Chatterjee, Aniruddha Sinha, Meghamala Sinha
TCS Research, Tata Consultancy Services Ltd.
Kolkata, India
Email: {debatri.chatterjee, aniruddha.s}@tcs.com,
meghamala.sinha@gmail.com

Sanjoy Kumar Saha
Computer Science & Engineering,
Jadavpur University,
Kolkata, India
Email: sks_ju@yahoo.co.in

Abstract

A performer may undergo a task with varying difficulty level. It is important to know the mental state in order to maintain the optimum level of performance. The mental state of an individual varies according to their IQ levels, task difficulties or other psychological or environmental reasons. We have tried to measure the cognitive state of individuals, while they are performing tasks of various complexity levels, using physiological responses like brain activation, heart rate variability and galvanic skin response. In this paper we have proposed a Bayesian network based model to probabilistically evaluate the cognitive state of an individual from the difficulty levels of the tasks, IQ level of the individual and observations made using the physiological sensing. Twenty subjects with various IQ levels are asked to play a modified Tower of London (TOL) game having three complexity levels: low, medium and high. The sensor data collected have been used to train the Bayesian model for generating the conditional probability distribution for the desired cognitive state. Results show that it can be used as a tool to determine the current cognitive state of any individual, provided we know their IQ score. In case of any contradiction between the desired cognitive state (obtained from prior knowledge) and the observed cognitive state (obtained during testing), the personal insights of a performer is analyzed.

1. INTRODUCTION

Cognitive flow is defined as a state of mind which is achieved while a person is performing a task with complete concentration and engagement. This state is closely related to the balance between the challenge of any task and the skill of the person executing the task. If the task difficulty is low compared to the held skill level of an individual, the person tends to be in the bored state (Csíkszentmihályi, Mihály, 1990). Alternatively, if the task difficulty is high, and the skill level is low, the person is supposed to be in the anxiety state. However, if the skill and task difficulty level matches, the person enter *Flow states* i.e. a state of focused concentration with a sense of enjoyment (Csíkszentmihályi, Mihály, 1997), (Csíkszentmihályi, Mihály, 1999). In our case, the skill level of an individual is directly related to his or her IQ level and is treated as the prior knowledge of the individual. On the other hand, the challenge of the task is synonymous to the task difficulty level. Hence the flow-boredom state can be represented as shown in Figure 1.

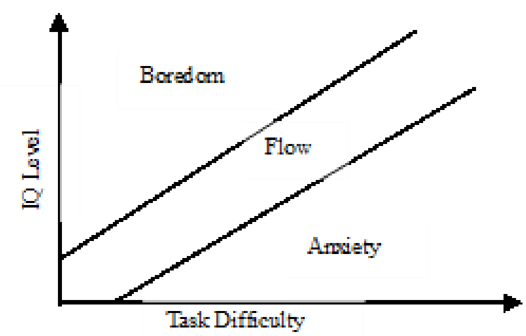


Figure 1: Flow state for IQ level and Task Difficulty

For a better learning experience, it is necessary for students to remain in the flow state throughout the session. It is a challenging task to create an environment which is enjoyable to a student and attracts complete attention as well as keep them motivated. To achieve this it is important to detect whether the student is in their flow state based on which we can induce a tailored learning environment. Present state of the art literature suggests using standard flow state questionnaires for measuring flow state. This method is indirect and might be biased. A more reliable approach would be to estimate the flow state based on physiological changes in an individual (Sinha, Aniruddha, et al, 2015).

In this paper, our main purpose is to detect the cognitive flow with the help of physiological signals and probabilistic reasoning models. In order to do this, a modified Tower of London game (Schnirman, Geoffrey M., Marilyn C. Welsh, and Paul D. Retzlaff, 1998) is implemented for various difficulty levels in PEBL: The Psychology Experiment Building Language (Mueller, Shane T., and Brian J. Piper, 2014).

Our work mainly consists of following phases –

(i) Extraction and analysis of data from various physiological sensors like Electro-encephalogram (EEG), Heart-rate variability (HRV) and Galvanic Skin response (GSR) while a participant is performing a standard psychometric test of various difficulty levels.

(ii) Creating a Bayesian Network (BN) framework with the objective to correlate the nodes based on sensor data as mentioned in previous step with the state nodes by using data obtained in step (i).

(iii) Validation of BN based model for test data

We have used the Bayesian Network based model because it provides a theoretically efficient and consistent mechanism for processing imprecise and uncertain information. Although there has been a growing interest among researchers to use BN in the field of education and student knowledge evaluation (Chrysafiadi, K. and Virvou, M, 2013), very little work has been done to model EEG and other physiological sensor data. We have created a model for our problem domain by representing a high level probability distribution over a set of random variable denoting the different states which have direct dependencies among themselves. Conditional probability tables for each node are updated based on the sensor data.

The paper is organized as follows-. In section 2, we have given a brief summary of existing approaches adopted for analysing flow state and related works. In section 3, we have described the design of the game, experimental setup, methodologies for the analysis of sensor data and construction of Bayesian network. Section 4, presents the analysis of results that we achieved from the experiments. Finally in section 5, we have concluded and have provided the future prospect of this work.

2. RELATED WORK

The 'Flow state' has been described (Csikszentmihályi, Mihály; Harper & Row, 2015) as an experience achieved by a person while performing a task and is dependent on the personality and ability of the person. A Flow state can only be achieved when a person is engaged in an active task. Passive tasks such as watching television, taking a bath etc. do not induce flow experience (Csikszentmihályi, M., Larson, R., & Prescott, S, 1977), (DelleFave, A., & Bassi, M., 2000). The three conditions (Csikszentmihályi, M.; Abuhamdeh, S. & Nakamura, J, 2005) that must be satisfied to achieve flow are: (a) the task must have a clear goal and rate of progress, (b) the task should accompany a continuous feedback process to help the person maintain the flow state and (c) the challenge level of the task and the skill level of the person must be balanced. In the fields of education, detecting the flow state of mind is important to provide appropriate learning environment for students.

Various researches have been conducted for Flow measurement in different domains like piano playing (de Manzano, Örjan, et al, 2010), video-game (Kramer, Daniel, 2007), online games (Hsu, Chin-Lung, and Hsi-Peng Lu, 2004), social networking sites (Mauri, Maurizio, et al, 2011), e-commerce business (Koufaris, Marios, 2002) etc.

Most of these approaches use indirect methods for measuring flow (Csikszentmihalyi, Mihaly, and Isabella Selega Csikszentmihalyi, 1992), (Novak, Thomas P., and Donna L. Hoffman, 1997), (Nakamura, J. and Csíkszentmihályi, 2009). The EEG is recently being used extensively in the fields of education with the help of Brain Computer Interface (BCI) technology (van Schaik, Paul, Stewart Martin, and Michael Vallance, 2012). A greater left temporal alpha activity is noticed (Kramer, Daniel, 2007) indicating the flow state of the performer in comparison to the right temporal lobe. The mid beta and theta activity also have a distinctive effect on performance whereas no significant effect was found in the delta waveforms. Higher alpha activity coupled with lower beta activity to characterize the flow state (Mauri, Maurizio, et al, 2011). Researchers are attempting to measure boredom, anxiety etc. from EEG signals (Chanel Guillaume et. al, 2008 and Berta Riccardo et.al, 2013). Recently low cost devices are being used for analysing the effect of various elementary cognitive tasks (Chatterjee, Debatri et al., 2015). Some of these works also suggested using other physiological responses like GSR and heart rate for assessing the flow state (Chaouachi, Maher and Claude Frasson, 2010). The main problem of using multichannel physiological sensors is that, we have to find out an appropriate mechanism for fusing the results obtained from multiple sensors.

Bayesian network (Pearl, Judea, 1986) is becoming an increasingly popular technique to model uncertain and complex domains. Unlike classical statistical models, BN allow the introduction of prior knowledge into models. This prevents extraneous data to be considered which might alter desired results. BN uses the concept of conditional probability which is proven to be very useful in applications to the real world problem domain, where probability of occurrence of an event is conditionally dependent on the probability of occurrence of a previous event.

Bayesian Network modelling has been used in the areas of medicine (Koufaris, Marios, 2002), document classification, information retrieval (Luis M. de Campos, Juan M. Fernández-Luna and Juan F. Huete, 2004), image processing, decision support system (F.J. Díez, J. Mira, E. Iturralde and S. Zubillaga, 1997*)*, gaming, bioinformatics (Neapolitan, Richard, 2009), gene analysis (Friedman, N.; Linial, M.; Nachman, I.; Pe'er, D, 2000) etc.

In the present work, BN, a desired state is derived based on the static information namely IQ of the individual and difficulty of the task. The estimated state is derived based on the sensor observation while the task is performed. The desired and the estimated state may not match as the prior characterization of individual in an exact manner is not possible. In that case, the model can help to decide the task level to maintain the flow state of the individual.

3. METHODOLOGY

This section explains the game designed and the methodology adopted for measurement and analysis of physiological signals and creation of BN model based on the findings.

3.1 GAME DESIGN

Tower of London (TOL) is a classical puzzle based game used by psychologists for assessment of executive functions and planning capabilities of an individual. We modified the standard TOL-R (Schnirmanetal, 1998) game for three levels of difficulties (Difficult, Moderate and Easy) in PEBL. For a game session, a target configuration is shown at the top of the screen as shown in fig. 2. The goal is to move a pile of disks given at the bottom so that the given assembly matches the target configuration shown on the top of the screen. Participants can only move one disk at a time, and cannot move a disk onto a pile that has no more room (indicated by the size of the grey rectangle). Participant has to click on the pile they want to move a disk off, and it will move up above the piles. Next, they click on another pile, and the disk will move down to that pile. There is a time limit to finish each game. Participants are instructed to finish each game within the allotted time. If the participant fails to finish the game within the session related time, the session ends and a new game starts. Information like start time of the game, duration, number of moves, total number of success etc. per session are stored.

The complexity of each game is determined with the help of the number of different coloured disks, the minimum number of moves needed to finish the game and the number of available empty space among the stacks for the disk to move around. Hence, the game complexity G_{com} is defined as,

$$G_{com} = \frac{N}{N_{disk} + N_{space}} \quad (1)$$

where, N is the total number of moves, N_{disk} is number of disks in the game and N_{space} and is the available number of empty space.

Screen shots for three different levels of games are shown in Figure 2. The minimum number of moves per disk for the low difficulty game is chosen as two and the total number of different disk is two, the complexity for this level of game is 3/(2+4)=0.5. Similarly, the complexities for the medium difficulty game is 4/(3+3)=0.67 and high difficulty game is 8/(4+2)=1.33 according to (1). These calculations show that the complexity for each task increases with respect to increase in difficulty level. The number of moves and number of disks for various complexity levels are chosen based on user feedback collected from 20 participants who also participated in the data collection.

The low difficulty level session consist of 10 set of games with 30 sec duration each. Similarly the medium session consist of 6 set of games each having 30 sec duration. The difficult session consist of 4 set of games with 30 sec duration each. For lower complexity game it is usually finished by most of the participants before 30 seconds, hence the number of games for various complexity levels are varied so that the completion time for easy, medium and hard sessions are comparable. Finally, it is found that for all the participants the minimum time to complete among all types (low, medium and high) of tasks is 90 seconds, hence the corresponding sensor data are considered for further processing.

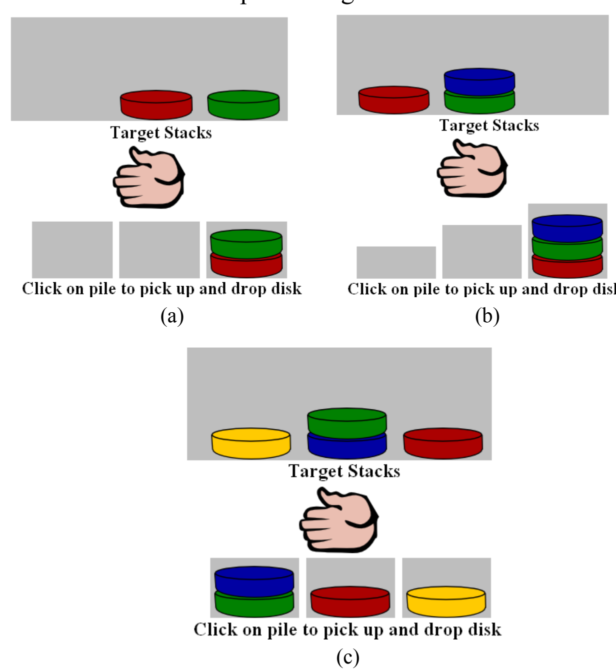


Figure 2: Screenshot of the three Tower of London games (a) Low (b) Medium and (c) High

3.2 EXPERIMENTAL SETUP

3.2.1 Participants

We have selected a group of 20 participants of varying IQ, from our research lab. The IQ scores are found to vary between 80 and 131. The average age of the participants are 22-30 years. They are all right handed male engineers belonging to similar socio-economic background. They had normal or corrected to normal vision. The selection is made to reduce the participant related bias so that the variation is only in their IQ levels.

3.2.2 Data Collection

An in house python based data capture tool is used for the data collection. The capture tool allows participants to play the game on a standard 17 inch computer screen placed at a viewing distance of approximately 25 inch and simultaneously allows to collect the EEG, GSR and Photoplethysmogram (PPG) signals. Participants are asked to play the game while wearing a single lead EEG device from Neurosky[1]. It is a dry sensor with a lead placed in FP1 position and the grounding is done with a clip fixed to left earlobe. We have used a GSR device from eSense[2] to record the variations in skin conductance level. All the participants are right handed and wore the GSR sensors on the middle and index fingers of the left hand. The right hand is kept completely free so that they can play the game comfortably using mouse. The oxygen saturation level and the pulse rate are measured by the pulse oximeter from Contec[3], through the left ring finger. The devices used are shown in Figure 3. The participants are asked to play three session of the game (High, Low and Medium) with a resting period of 5 min between each game. For half of the participants the order of the game followed is high, medium, low and for remaining half low, medium, high sequence is followed. During the game, scores, number of moves, duration etc. are also recorded for further analysis.

Figure 3: Data collection devices: (a) Neurosky EEG device (b) GSR device from eSense (c) Pulse oxymeter from Contec

3.2.3 Participant Feedback Form

The standard Game-flow indicator (GFI) feedback form has been used to assess if a participant experienced boredom and flow experiences during the experiments and the findings are used as the reference. This questionnaire based feedback form described by (Bakker, A. B, 2005) and (Bakker, A. B., 2008) is ideal to measure level of engagement while playing a game. For doing this, the overall scores for both flow and boredom questionnaires are calculated assuming 1= strongly disagree, 2= disagree, 3 = undecided, 4 = agree and 5 = strongly agree. Different questions are asked regarding the experience of participants while playing the game.

The participants are asked to fill up three feedback forms, one for each of the three games, immediately after the end of each session. This feedback is necessary as they provide a ground truth for cognitive flow along with the game data. We have used the feedback forms to analyze the contradictions which we have seen during the Bayesian Network analysis as explained in section 3.4.

3.3 SENSOR DATA ANALYSIS

3.3.1 *Skill-challenge analysis using EEG signals*

In the present work, we have experimented with various frequency band energies described by (W. Klimesch, 1999), (H. Sijuan, 2010) and time domain Hjorth parameters as described by (Gudmundsson, Runarsson, Sigurdsson, Eiriksdottir, Johnsen, 2007), (V. Carmen, et al. 2009) as shown in (2) and (3).

$$F = \{E^{\delta}, E^{\theta}, E^{\alpha}, E^{l\beta}, E^{m\beta}, H^{a}, H^{m}, H^{c}\} \quad (2)$$

The first five features are the energies in various frequency bands namely, delta (E^{δ} as 0.5 - 4 Hz), theta (E^{θ} as 4 – 7.5 Hz), alpha (E^{α} as 7.5 – 12.5 Hz), low-beta ($E^{l\beta}$ as 12.5 – 16 Hz) and mid-beta ($E^{m\beta}$ as 16 – 20 Hz) respectively. The energies in each band are extracted using Welch's power spectral density as defined by (Welch, Peter, 1967). The last three features in (2) are the Hjorth parameters namely Activity (H^{a}), Mobility (H^{m}) and Complexity (H^{c}) respectively as given in (3).

$$H^{a} = \mathrm{var}(x(t)) \quad H^{m} = \sqrt{\frac{H^{a}(\frac{dx(t)}{dt})}{H^{a}(x(t))}} \quad H^{c} = \frac{H^{m}(\frac{dx(t)}{dt})}{H^{m}(x(t))} \quad (3)$$

Here *x(t)* indicates the time domain signal in a window of duration 1 sec and $\frac{dx(t)}{dt}$ is the first order derivative of the signal.

3.3.2 *Analysis of GSR signal*

The galvanic skin response (GSR) is the electro-dermal response where the skin conductance changes with the

[1]http://neurosky.com

[2]https://www.mindfield.de/en/biofeedback/products/esense/esense-skin-response

[3]http://www.coopermedical.com/overnight-pulse-ox/cms-50d-plus-recording-fingertip-pulse-oximeter.html

amount of secretion from the sweat glands in presence of stressful, likeable events. Therefore GSR can be used as a good predictor of concentration, mental workload etc. in flow study (Nourbakhsh, Nargess, et al, 2012). During flow experience, subjects should experience a higher concentration and focus on the task. This can be measured using the GSR.

The GSR device consists of two electrodes which applies a constant voltage to the skin. The current which then flows through the skin can be detected. The GSR signal consists of two components: phasic, which is the fast varying component and tonic, which is the slow component. Both contain important information associated with specific physiological aspect of the mental states. The tonic component (T_c) is calculated only by taking the inverse transform of first few Fourier coefficients (4) and the phasic component (P_c) is calculated by inversing the higher order coefficient of the Fourier coefficients (5).The eSense GSR sensor has 5Hz sampling frequency hence we have used first 4 components ($0<=k<=3$) of the Fourier coefficients which corresponds to 0.5Hz. The IFFT in (4) and (5) corresponds to Inverse Fast Fourier Transform.The tonic component (T_c) of the GSR signal is computed for windows of duration 1 second. Next to investigate the distribution T_c over the task interval, we compute the mean and the kurtosis as given in (6) using N successive windows for each task (low, medium, high). These parameters provide the statistical information on the way the skin conductance changes for an individual over time, for a given task.

$$T_c = IFFT(\sum_{n=0}^{N-1} x(n).e^{-j(\frac{2\pi}{N})nk})\ ,\ k=0,1,2,3 \qquad (4)$$

$$P_c = IFFT(\sum_{n=0}^{N-1} x(n).e^{-j(\frac{2\pi}{N})nk})\ ,k=4,5,...N-1 \qquad (5)$$

$$T_c^{mean} = mean(T_c(n)),\ T_c^{kurt} = kurtosis(T_c(n)),\ 1 \le n \le N \qquad (6)$$

3.3.3 *Calculation of HRV from PPG signals*

Stress level can be evaluated using Heart rate variability (J. Taelman, S. Vandeput, A. Spaepen and S. Van Huffel, 2009), (McDuff, Daniel, Sarah Gontarek, and Rosalind Picard, 2014). When the task challenge is low compared to the subjects' skill level, then the heart rate variability is high compared to the flow state where task challenge matches the skill. We have used the SPO2 device wearable on the ring finger for sensing the Photoplethysmogram (PPG) signal. We have calculated three time domain HRV parameters namely 1) rMSSD (root mean square of successive differences between adjacent NN intervals), 2) SDSD (Standard deviation of successive differences between adjacent NN intervals), 3) SDNN (Successive difference between NN intervals) as explained in (McDuff, 2014). Out of these SDNN is found to give a better indication of the stress level for our experiment.

3.4 BAYESIAN NETWORK CONSTRUCTION

Our problem statements for creating the Bayesian Network are as follows -

- A number of random participants are performing a task in a controlled environment.
- The task can be of three difficulty levels: easy, moderate or hard.
- The participants are categorized based on their intelligence levels i.e. the IQ scores.
- The participants can be in either of three states: flow, boredom or anxiety.
- Their performance and feedback are recorded to determine their mental state and experience.
- During data capture, three physiological sensors namely EEG sensor, HRV sensor and GSR sensor are used.

Now a Bayesian Network (BN) framework is created with the objective to diagnose and investigate the different relationships between the sensors nodes and the state nodes as shown in Figure 4.

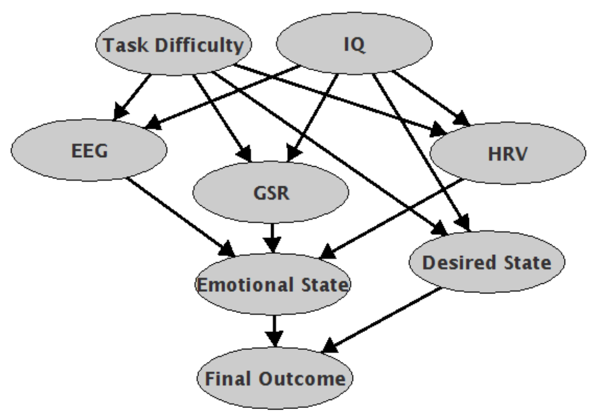


Figure 4: Proposed Bayesian Network

The BN consists of three types of nodes:

Evidence nodes: Task Difficulty (TD), Intelligence Quotient (IQ)

Sensor nodes: GSR, GSR, HRV

State nodes: Emotional State (O), Desired State (D), Final Outcome (F)

Through in the Evidence nodes we can input the static knowledge or evidence which we have in our problem domain. In our case, we know beforehand, the IQ level of the participants and also the difficulty levels of task they are given. The Evidence nodes serve as parents to the Sensor nodes, since the sensor readings are a direct causal effect of the two conditions (IQ & TD). Based on the results obtained from sensors, we can predict the current cognitive state of the participants. The derived state i.e. the Emotional state (O) from the sensor reading is compared with the Desired state (achieved from ground truth during training) and finally a conclusion (F) is drawn based on these two states. We have used a separate desired state in order to compare it to the results derived from the signals about what emotional condition the subject is currently in.

We have created the Bayesian Network in SamIam[4], a comprehensive tool for modelling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche, UCLA.

At the beginning of the experiment, the participants are categorized based on their IQ levels measured by a standardized IQ[5] test prior to playing the game. This IQ test is a free online test from Brainmetrix.

The IQ level and the task difficulty level serve as evidence nodes in the BN. Depending on the state of these two nodes, the sensor readings vary. Hence all the sensor nodes have them as immediate parents. The readings of all three sensors are used to analysis the emotion state of the participants and are compared to the ground truth i.e. the desired state. If two state matches, then we can conclude the actual state via the final outcome and that the Bayesian Network is expected to give the correct outcome. If not, then there is a case of contradiction caused by one of more sensors giving faulty state. In case of a contradiction between observed state and desired state the BN model can be used to resolve this conflict by reasoning between the various nodes using probabilistic queries.

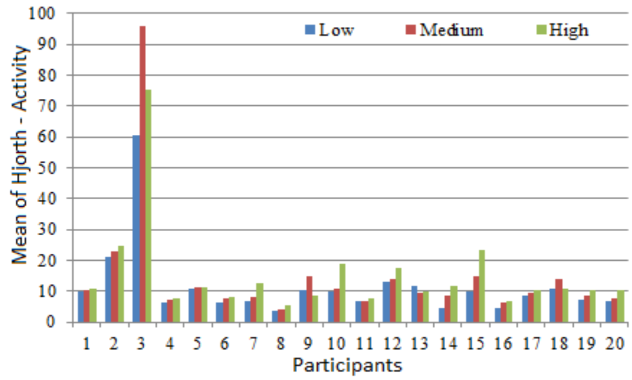
4. RESULTS AND DISCUSSIONS

4.1 EEG SIGNAL ANALYSIS

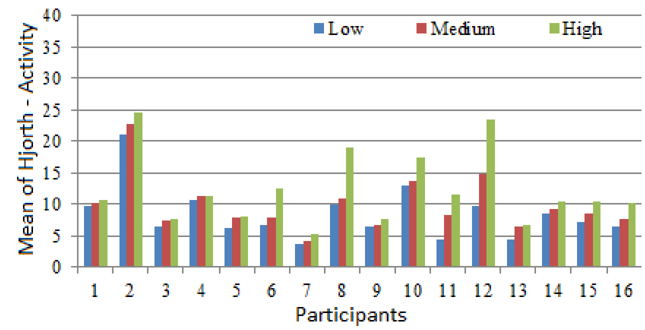
Various features of the EEG signals namely alpha, beta, theta, delta, attention, meditation, Hjorth etc. for all the experiments for all the participants are extracted using (2) and (3). Among all features the *Activity* measure of Hjorth parameter is found to be indicative of variations of brain signals with difficulty level. The raw EEG signal is used for calculating activity for windows of duration 1 sec and the overall mean is taken for each session of the game of all the participants. The results for different difficulty levels (Low, Medium and High) are combined separately for all the 20 participants and compared as shown in Figure 5(a). We have found good amount of separation between three different tasks for 16 out of 20 participants. The comparison for these selected participants are shown in Figure 5(b).

4.1 GSR SIGNAL ANALYSIS

The GSR data is subdivided into a number of windows of duration 1 sec. Next we calculate both tonic and phasic power using (4) and (5).The phasic power does not show sufficient separation between different task levels whereas the tonic gives good separation. The mean and kurtosis of the tonic power is calculated using (6). The plot of mean and kurtosis of tonic component for GSR are shown in Figure 6(a) and Figure 6(b). The line representing the medium difficulty game for all the subjects is found to overlap over the other two; the separation between the medium games across participants is not consistent. Hence for better representation, we have only plotted the GSR for high and low difficulty games.



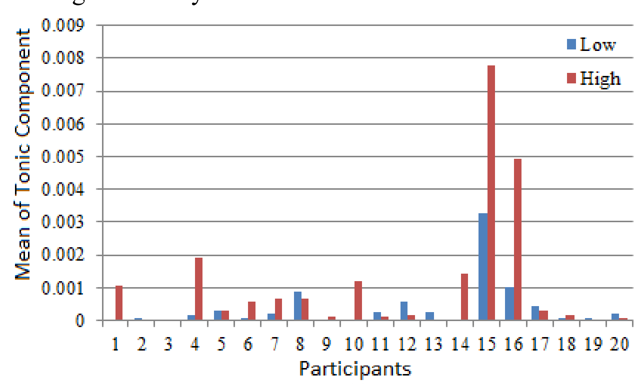
(a)



(b)

Figure 5: Separation between Low, Medium, High difficulty task for the Hjorth calculation of EEG signals for (a) 20 subjects and (b) 16 subjects

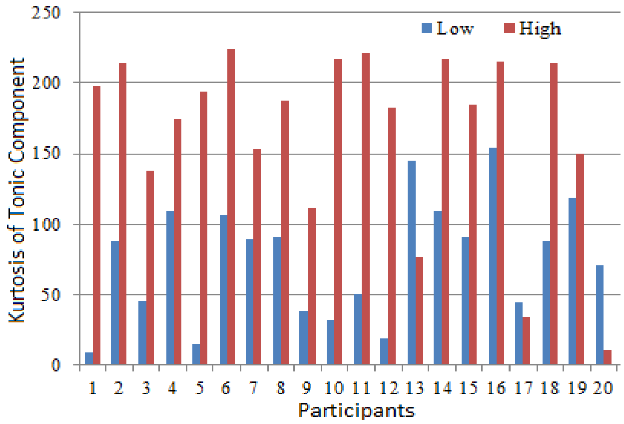
Out of 20 participants 8 participants (6, 7, 8, 9, 10, 11, 12 and 20) played the game in high-medium-low sequence and the remaining participants played in low-medium-high sequence. It is evident from the plots that kurtosis performs better than mean of the tonic power in separating the low and high difficulty tasks.



(a)

[4]http://reasoning.cs.ucla.edu/samiam/

[5]http://www.brainmetrix.com/free-iq-test/



(b)

Figure 6: Separation between the (a) Mean and (b) Kurtosis of the Tonic (μS) component of GSR during the Low and High difficulty games.

4.2 PPG SIGNAL ANALYSIS

The PPG signal analysis does not give a clear, consistent separation between difficulty levels of the task for all the participants. We have plotted the successive difference between NN intervals (SDNN) in Figure 7. For 10 out of 20 subjects there is a separation between the two SDNN values out of which only 5 have the value for the low game lower than the high game. Hence any substantial information cannot be derived from this result of PPG analysis.

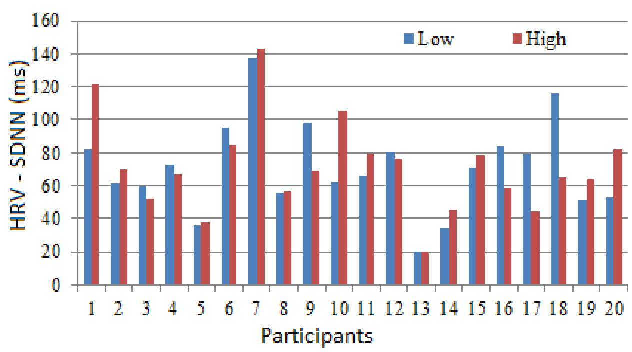


Figure 7: Separation of the HRV(SDNN) in msec for all the 20 subjects between the high and low difficulty game

4.3 BAYESIAN NETWORK ANALYSIS

The Bayesian network framework, given in Figure 4, has been trained based on the results obtained from physiological sensor data collected during the experiments. The conditional probability tables associated with each node have been updated according to the occurrence of their respective parent node.

Given a participant of a particular IQ level and solving a game of a particular difficulty level, we have collected each of the sensor data and have calculated their probability of occurrence over all possible conditions. The scores of IQ less than 89 are treated as Low IQ, scores between 90 and 109 are treated as Medium IQ and scores with 110 is treated as High IQ. In the present study we got 6 participants having high IQ, 9 participants having medium IQ and 5 participants having low IQ.

We have used the data for 18 out of 20 participants to train the BN. The sensor data for all the training participants (18 subjects) are classified into three levels (Low, Medium, High) based on their observations. An example is shown in Table 1 where the levels of the EEG sensor data are shown for 5 training participants with High IQ, playing three levels of games. It can be seen that for the easy game (TD = low), all of them have low EEG feature values, for medium game (TD = medium), one of them has high EEG feature and similarly for difficult game (TD = high). The probability of occurrence of each level is calculated and updated in the respective Conditional Probability Table (CPT) of the BN. After the training the final CPT for the EEG node is shown in Table 2. The same rule is followed for the rest of the sensor nodes as well.

Table 1: The levels for EEG Sensor data (Mean of Hjorth - Activity) for 5 Participants with High IQ and playing three levels of game (TD). Here L-Low, M-Medium, H-High.

TD	**EEG level (Mean of Hjorth - Activity)**				
L	L	L	L	L	L
M	M	H	M	M	M
H	H	M	H	H	H

Table 2: Conditional probabilities of the EEG node is shown for the pair (TD, IQ)

EEG	**(L,H)**	**(M,H)**	**(H,H)**	**(L,M)**	**(M,M)**	**(H,M)**
L	0.98	0.02	0.02	0.98	0.01	0.01
M	0.01	0.79	0.19	0.01	0.98	0.01
H	0.01	0.19	0.79	0.01	0.01	0.98
EEG	**(L,L)**	**(M,L)**	**(H,L)**			
L	0.5	0.25	0.25			
M	0.25	0.5	0.25			
H	0.25	0.25	0.5			

The data for the remaining 2 participants (having high IQ and low IQ) are used to validate whether the BN nodes are providing the correct results. Different combinations of evidences for these two subjects are checked in the BN (Figure 8 through Figure 11) and the reasons for contradictions are explained. The states of the BN are shown as rectangular blocks containing the percentage equivalent of the probability values of the random variables.

Case 1: A participant of high IQ is playing a high difficulty level game

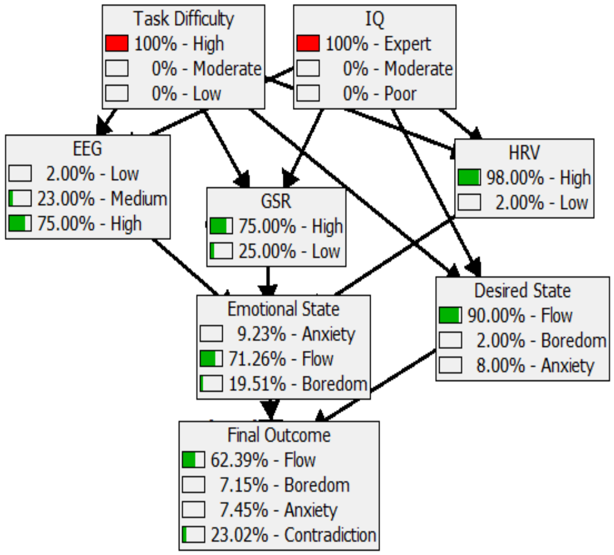


Figure 8: Bayesian Network in Query mode for the evidence Task Difficulty= High and IQ=Expert/High

In this case we can see that the Emotional State (O) matches with the Desired State (D) hence the final outcome is Flow state. Also all the sensors are providing the expected result, except the EEG feature value. The subject used in this experiment has high IQ. From the questionnaire based survey we have found that the participants has reported to be sometimes in anxious state during the high difficulty game. This could be the reason for the EEG feature value to be very high for him. We have found from participant's feedback that he was mostly in the flow state. Hence the output of the Bayesian model seems to be correlating with the participant's feedback.

Case 2: A participant of high IQ is playing a low difficulty level task

In this case we can see that if a participant with high IQ is playing a low difficulty task, the sensor nodes are providing the desired states. This is shown in Figure 9. The EEG feature value and GSR states are in low state as expected, as the participant is expected to be in a relaxed state. However, the HRV indicates a low state as opposed to the expected high state leading to the contradiction. If a highly intelligent participant is playing a very easy task with low concentration (GSR=Low) then he / she might be in a restless state which can lead to anxiety. Hence the true mental state cannot be correctly determined.

Case 3: A participant of low IQ is playing a high difficulty level task

In this case we can see that all the states are providing the expected result except the EEG feature value as shown in Figure 10. The participant in this experiment has low IQ. According to participant's feedback, he was in the flow state without being too anxious and matches with the Emotional State of the BN. However a contradiction is shown in the Final Outcome indicating a deviation of the behaviour of the participant from the expected one.

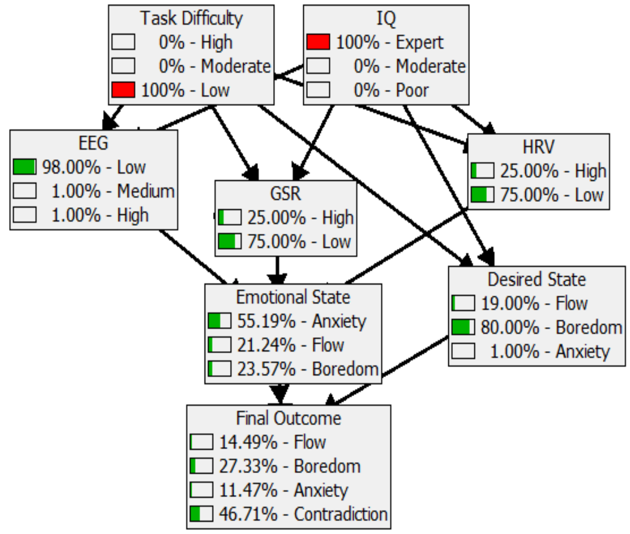


Figure 9: Bayesian Network in Query mode for the evidence Task Difficulty=Low and IQ=Expert/High

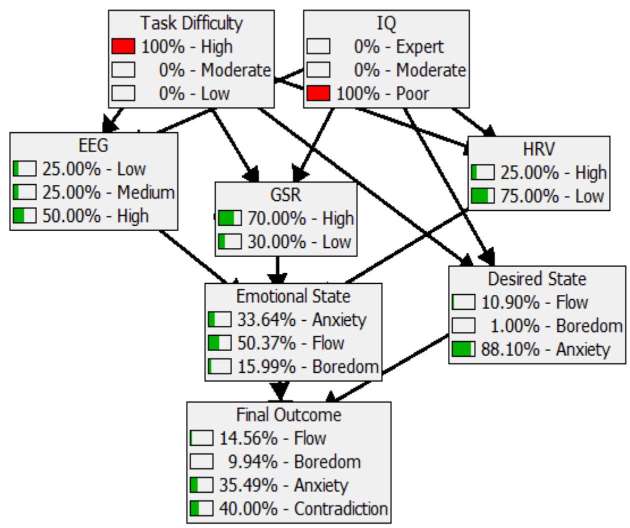


Figure 10: Bayesian Network in Query mode for the evidence Task Difficulty= High and IQ=Poor/Low

Case 4: A participant of low IQ is playing a low difficulty level task

In this case we can see that all the states are providing the expected result except the EEG feature value as shown in Figure 11. The participants used in this experiment has low IQ. We have found from participants' feedback that he was in the relaxed state and did not find the game too difficult. This could be the reason for the EEG feature value to be low but with very low percentage. It can also be observed from all the four cases that if the contradiction option is ignored in the final outcome state, then the next highest probability belongs to the state same as the desired state.

This shows that while the Bayesian Network provides the correct prediction of the actual mental state, given the sensor data and static knowledge, it also hints at the possibility of a faulty sensor data based on the probability of the contradiction.

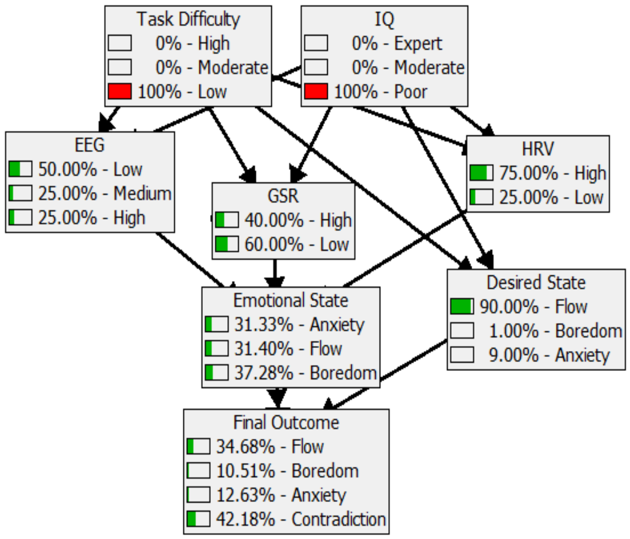


Figure 11: Bayesian Network in Query mode for the evidence Task Difficulty= Low and IQ=Poor/Low

5. CONCLUSION AND FUTURE WORK

In this work we have measured the cognitive state of a participant performing task of varying difficulty level from brain signals and certain physiological parameters like heart rate variability and galvanic skin response. Results show that EEG (Hjorth-Activity) and GSR signals can be successfully used to distinguish the performance for all the participants in three different level tasks. On the other hand, the analysis of the HRV data is not providing consistent information. These results can be fused together and analyzed along with the performance and feedback data to get further insights of the mental state for the participants.

From our proposed Bayesian Network we need to determine whether there is a contradiction between observed and desired state. In case of contradiction we need to find out whether there are any sensor(s) nodes giving faulty data. In future a large number of participants are to be considered for further analysis. Moreover, the workflow of the Bayesian Network should be dynamic for handling a large number of data. Several knowledge under uncertainty can be determined in future by using such probabilistic graphical models. In the area of education, models can be designed for generating a sequence of questions which are intermixed in difficulty level to analyze the variation in cognitive states of an individual during the assessment process.

Acknowledgements

We are thankful to the participants who have provided their time and inputs for the data collection for the experiments.

References

Bakker, A. B. (2005). "Flow among music teachers and their students: The crossover of peak experiences". Journal of Vocational Behavior, 66, 26–44. Retrieved from http://dspace.library.uu.nl/handle/1874/10711.

Bakker, A. B. (2008). "The work-related flow inventory: Construction and initial validation of the WOLF". Journal of Vocational Behavior, 72(3), 400–414.

Berta, Riccardo, et al. (2013). "Electroencephalogram and physiological signal analysis for assessing flow in games." Computational Intelligence and AI in Games, IEEE Transactions on 5.2: 164-175.

V. Carmen, et al. (2009). "Time Domain Parameters as a feature for EEG-based Brain-Computer Interfaces" in Neural Networks 22: 1313-1319

Chanel, Guillaume, et al. (2008). "Boredom, engagement and anxiety as indicators for adaptation to difficulty in games." Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era. ACM.

Chaouachi, Maher, and Claude Frasson. (2010). "Exploring the relationship between learner EEG mental engagement and affect." Intelligent tutoring systems. Springer Berlin Heidelberg.

Chatterjee, Debatri, et al. (2015). "Analyzing elementary cognitive tasks with Bloom's taxonomy using low cost commercial EEG device." Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP). IEEE.

Chrysafiadi, K. and Virvou, M. (2013). "Student modelling approaches: A literature review for the last decade." Expert Systems with Applications, 40(11): 4715-4722.

Csíkszentmihályi, Mihály, (1990). "Flow: The Psychology of Optimal Experience", New York: Harper and Row, ISBN 0-06-092043-2.

Csikszentmihalyi, Mihaly. (1997). "Finding flow.".

Csikszentmihalyi, M. (1999). "If we are so rich, why aren't we happy?" American Psychologist. Vol 54(10), 821-827.

Csikszentmihályi, Mihály; Harper & Row. (2015). "FLOW: The Psychology of Optimal Experience" Retrieved 2 April 2015.

Csikszentmihalyi, Mihaly, and Isabella Selega Csikszentmihalyi, eds. (1992). "Optimal experience: Psychological studies of flow in consciousness." Cambridge University Press.

Csikszentmihályi, M., Larson, R., & Prescott, S. (1977). "The ecology of adolescent activity and experience." Journal of Youth and Adolescence, 6, 281-294.

Csikszentmihályi, M.; Abuhamdeh, S. & Nakamura, J. (2005), "Flow", in Elliot, A., Handbook of Competence and Motivation, New York: The Guilford Press, pp. 598–698.

deManzano, Örjan, et al. (2010). "The psychophysiology of flow during piano playing." Emotion 10.3: 301.

DelleFave, A., & Bassi, M. (2000). "The quality of experience in adolescents' daily lives: Developmental perspectives." Genetic, Social, and General Psychology Monographs, 126, 347-367.

F.J. Díez, J. Mira, E. Iturralde and S. Zubillaga (1997). "DIAVAL, a Bayesian expert system for echocardiography". Artificial Intelligence in Medicine (Elsevier) 10 (1): 59–73.

Friedman, N.; Linial, M.; Nachman, I.; Pe'er, D. (2000). "Using Bayesian Networks to Analyze Expression Data". Journal of Computational Biology 7 (3–4): 601–620.

S. Gudmundsson, P. Runarsson, T. Sigurdsson, S. Eiriksdottir, K. Johnsen, (2007). "Reliability of quantitative EEG features", in Clinical Neurophysiology 118: 21622171, August 2007.

Hsu, Chin-Lung, and Hsi-Peng Lu. (2004). "Why do people play on-line games? An extended TAM with social influences and flow experience." Information & Management 41.7: 853-868.

W. Klimesch, (1999). "EEG alpha and Theta oscillations reflect cognitive and memory performance: a review and analysis" in Brain Research Reviews 29.

Koufaris, Marios. (2002). "Applying the technology acceptance model and flow theory to online consumer behavior." Information systems research 13.2: 205-223

Kramer, Daniel. (2007). "Predictions of performance by EEG and skin conductance." Indiana Undergraduate Journal of Cognitive Science 2: 3-13.

Luis M. de Campos, Juan M. Fernández-Luna and Juan F. Huete (2004). "Bayesian networks and information retrieval: an introduction to the special issue". Information Processing & Management (Elsevier) 40 (5): 727–733.

Mauri, Maurizio, et al. (2011). "Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook." Cyberpsychology, Behavior, and Social Networking 14.12: 723-731

McDuff, Daniel, Sarah Gontarek, and Rosalind Picard. (2014). "Remote measurement of cognitive stress via heart rate variability." Engineering in Medicine and Biology Society (EMBC), 36th Annual International Conference of the IEEE.

Mueller, Shane T., and Brian J. Piper. (2014). "The psychology experiment building language (pebl) and pebl test battery." Journal of neuroscience methods 222: 250-259.

Nakamura, J. and Csíkszentmihályi, M. (2009). "Flow Theory and Research." In Snyder, C. R. and Lopez, S. J. eds. Oxford handbook of Positive Psychology.Oxford University Press, Oxford, 195-206.

Neapolitan, Richard (2009). Probabilistic Methods for Bioinformatics. Burlington, MA: Morgan Kaufmann. p. 406. ISBN 9780123704764.

Nourbakhsh, Nargess, et al. (2012). "Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks." Proceedings of the 24th Australian Computer-Human Interaction Conference. ACM.

Novak, Thomas P., and Donna L. Hoffman. (1997). "Measuring the flow experience among web users." Interval Research Corporation 31.

Pearl, Judea. (1986). "Fusion, propagation, and structuring in belief networks." Artificial intelligence 29.3: 241-288.

Pearl, Judea, and Stuart Russell. (1998). "Bayesian networks." Computer Science Department, University of California.

Schnirman, Geoffrey M., Marilyn C. Welsh, and Paul D. Retzlaff. (1998). "Development of the Towerof London-revised." Assessment 5.4: 355-360.

Schnirman, G.M., Welsh, M.C., Retzlaff, P.D. (1998). "Development of the Tower of London-Revised." Assessment 5(4), 355 –360.

H. Sijuan, (2010). "Feature extraction and classification of EEG for imagery movement based on mu/beta rhythms", 3rd International Conference on Biomedical Engineering and Informatics (BMEI).

Sinha, Aniruddha, et al. (2015) "Dynamic assessment of learners' mental state for an improved learning experience." Frontiers in Education Conference (FIE), 2015. 32614. IEEE.

J. Taelman, S. Vandeput, A. Spaepen and S. Van Huffel, (2009). "Influence of Mental Stress on Heart Rate and Heart Rate Variability", 4th European Conference of the International Federation for Medical and Biological Engineering, vol. 22, pp. 1366-1369.

Van Schaik, Paul, Stewart Martin, and Michael Vallance. (2012). "Measuring flow experience in an immersive virtual environment for collaborative learning." Journal of Computer Assisted Learning 28.4: 350-365.

Welch, Peter. (1967). "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms." IEEE Transactions on audio and electroacoustics: 70-73.