Follow your heart

Heart rate controlled music recommendation for low stress air travel*

Hao Liu, Jun Hu[1] & Matthias Rauterberg[1]
Beijing, China / [1]Department of Industrial Design, Eindhoven University of Technology

Long distance travel is an unusual activity for humans. The economical cabin environment (low air circulation, limited space, low humidity, etc.) during the long haul flights causes discomfort and even stress for many passengers. In-flight video and music systems are commonly available to improve the comfort level of the passengers. However, current in-flight music systems do not explore how the content can be used to reduce passengers stress. Most of these systems are designed and implemented assuming a homogeneous passenger group that has similar tastes and desires. In this paper, we present a heart rate controlled in-flight music recommendation system for reducing the stress during air travel. The system recommends personalized music playlists to the passengers and attempts to keep their heart rate in a normal range with these playlists. Experiments in a simulated long haul flight cabin environment find that the passengers' stress can indeed be significantly reduced through listening to the recommended music playlists.

Keywords: stress; music recommendation; bradycardia; tachycardia; heart rate control; stress reduction; bio feedback; heart rate variability

1. Introduction

Travel by air, especially over a long distance, is not a natural activity for humans. Many passengers experience some degree of discomfort and even stress when flying. Excessive stress may cause passengers to become aggressive, over-reactive and even endanger their health (World Health Organization, 2005). Many airlines have realized the potential of in-flight entertainment (IFE) system for improving

* Accepted on 08 Oct 2014 by *Interaction Studies – Social Behaviour and Communication in Biological and Artificial Systems*

ISSN 1572–0373 / E-ISSN 1572–0381

customers' comfort level. In-flight music system is commonly installed as part of the IFE system on long haul flights for the passengers. However, the current in-flight music systems do not explore how the music can be utilized to reduce stress. Music is simply broadcasted through a limited number of channels, or delivered interactively as audio on demand. This provides mental distraction, however it is unclear whether it leads to stress reduction or not.

In this paper, we present a heart rate controlled music recommendation system for stress free air travel. It incorporates the passenger's heart rate signal as an indicator of stress to be measured non-intrusively and to be fed into a music recommendation process. Music properties, passenger's music preference and heart rate are taken into account to recommend customized music playlists to the passenger. By listening to the recommended music playlists, the system keeps the passengers heart rate within a normal range and reduces the passengers stress level during long haul flights.

The design and implementation of the system went through several iterations, each of which focusing on different issues and problems. The earlier designs focused more on the music preferences and the adaptively (Liu et al., 2008), whereas the later iterations introduced the bio-signals especially the heart rate as an input for the system and as an indicator for stress (Liu et al., 2009a,b). Preliminary empirical research was carried out and the results were used to refine the system and the experiment (Liu, 2010). The refinement of the system and experiment continued based on these experiences and findings, and this paper is to present the final results.

This paper is organized as follows: In Section 2, the state of the art of in-flight music systems and music recommendation systems are investigated. Next, an adaptive framework for heart rate con-trolled in-flight music recommendation systems is introduced. In Section 4, the software architecture of the system is presented. Section 5 reports the user experiments that were conducted to validate the system design, followed by discussions and conclusions.

2. Related work

2.1 Current in-flight music systems

Music service is an important part of the IFE system ever since IFE becomes available to aircraft passengers. Since the middle of the last century, commercial flights became available for daily public transportation. Entertainment was then requested by passengers to help passing the time. It was delivered in the form of centralized movie projection during lengthy flights, in addition to food and drink services. IFE systems were upgraded to CRT (Cathode Ray Tube) based systems in the late 1980s and early 1990s. In 1985, using Philips Type Cassette technology, the first audio player system was offered to the passengers by Avicom. Around the same time, CRT-based displays began to be installed on the ceiling of the aisles of aircrafts. In the mid-1990s, the first in-seat video systems began to appear, and LCD (Liquid Crystal Display) technology started to replace CRT for overhead video. In the late 1990s and early 2000s, the first full-fledge in-seat audio/video on-demand systems began to appear (Siddiqi, 2011).

To configure the aircrafts according to their budget and marketing strategies, airplane producers (e.g. Boeing and Airbus) provide customized IFE solutions. Liu (2010) investigated the installed IFE systems in the aircrafts of Airlines of Lufthansa, Air France, British Airways, American Airlines, Delta Airlines, and Japan Airlines. These airlines are considered to be top airlines in Europe, North America and Asia.

The result of our investigation shows that all the installed in-flight music systems are designed based on a preset concept, assuming a homogeneous passenger group that has similar tastes and desires. These IFE systems present the same user music interface and contents to every passenger.

To get desired music services during air travel, users need to interact with the IFE system, using touch screens or in-seat controllers to browse through extensive lists and selects the desired music tracks from provided options. Users tend to get lost in too many options and too much of work to orient themselves in these options. When this is the case, the system does not improve the situation, but could make the situation even worse (Liu, 2010).

2.2 Current music recommendation methods

A music recommendation system aims to support users in selecting music items while interacting with a large collection of music items. Since the middle nineties, there has been much work done in both academia and industry in developing music recommendation systems. They are usually classified into the following categories based on the recommendation methods used: content-based filtering, collaborative filtering, context-based filtering, and hybrid approaches. These methods are based on the information from users, peers, contexts and music items (Woerndl et al., 2007).

Content-based filtering recommends music to a user based on the description of the music and the user's music preference. Although the application areas differ, content-based music recommendation systems share in common descriptions of music items, user's music preferences, and algorithms to match music items to the user's preferences (Cano et al., 2005; Çataltepe and Altinel, 2007; Chai and Vercoe, 2000). Content based filtering is capable of clustering similar music tracks based on content features. After the user selects an item, similar music items can be recommended. The similarity can be at a higher level of user preference, but also be at a lower level of audio signals (Knees et al., 2006). Unfortunately, the recommendation is limited by the content features that can be extracted.

Collaborative filtering uses the correlation between users with similar tastes and their preferences in the past. Users with common interests recommend music tracks to each other (Basilico and Hofmann, 2004; Sarwar et al., 2001; Su and Khoshgoftaar, 2009). Collaborative filtering has been widely used and developed in both industry and academia. It is capable of recommending music without accurate music content analysis. However, at the system initialization stage, it will face the cold-start problem: the system cannot effectively serve its duty until it has gathered sufficient information about the users and items.

Context-based filtering aims at improving user satisfaction level by tailoring music recommendation according to the context information. For example, Baltrunas et al. (2011) designed a system to recommend music for car drivers based on the traffic condition and the drivers mood. Kaminskas and Ricci (2012) have an overview on such tools and techniques as well as related research challenges in context-aware music retrieval and recommendation. In recent years, there has been some work done towards context-based filtering in serving music on mobile platforms. In these applications, context information such as location, time, activity, is taken into account in order to provide users with context-aware mobile music services (Masuhr et al., 2008; Mokbel and Levandoski, 2009; Suh et al., 2007). However, there is still a lot to be done, to include broader, more precise context information at a higher level (for example mood and emotion) into the recommendation process, and to embed the recommendation unobtrusively into users' daily life.

All these three methods have their own specific advantages. None of them can perform well in every situation. In some applications, they are combined (hybrid recommendation) to improve the recommendation performance (Su et al., 2010; Yoshii et al., 2006), or equipped with active learning (Elahi et al., 2013; Rubens et al., 2011) and active bootstrapping (Golbandi et al., 2011) to satisfy users new to a system. However, none of these approaches are directly applicable in our design challenge. We will present a new way of content and context based filtering approach. The research approach presented is based on triangulation as introduced by Rauterberg (2006); this approach had been already successfully applied to a different design challenge (Abrazhevich et al., 2009).

3. Framework

Next a user scenario is introduced to highlight the desired system behavior, based on which a system framework is designed.

3.1 User scenario

In order to foster the customer loyalty, in addition to offering the frequent travelers the rewards such as bonus miles, speedy services, and free hotels, the airline SEAT offers passengers the rewards of "additional services" such as personalized IFE. To get the access to these extra services, the passenger needs to apply a club member card from SEAT. Joanna Smith is a frequent air traveler. She always flies with SEAT. In order to get the "additional service" from SEAT, she applied to be a club member of SEAT. She had filled out the membership application form which asked for her name, gender, date of birth and music preferences. After SEAT received and approved her application, her personal data was saved into the database of the IFE system. When Joanna buys a flight ticket from SEAT and later selects a seat number online or at the check-in desk, her profile will be linked to her in-seat entertainment system.

During the long haul flight, when Joanna sits in the seat, her heart rate is measured non-intrusively by heart rate sensors embedded in the seat textile. If her heart rate is higher than normal and she wants to listen to music, the system will recommend her a personalized music playlist which can decrease her heart rate back to the normal. If Joanna's heart rate is normal and she wants to listen to music, the system will recommend her a personalized playlist which helps to keep her heart rate at normal; if her heart rate is lower than normal and she wants to listen to music, the system will recommend a personalized playlist that can increase her heart rate back to normal. If Joanna does not like the recommended music, she can decline it and manually reselect the music she likes. During this process, the system logs the interaction. The system can then learn and adapt to her latest music preferences. The result of the recommendation will be influenced by the status of the heart rate as well as her preferences. By using the heart rate controlled music recommendation system during the flight, Joanna gets less stressed and her comfort level is improved.

It can be noticed that this scenario does not take into account the situation some people may sit somewhere else rather than their own seat. The system can adapt to this situation easily by identifying the passenger in one way or another, or if the passenger or the cabin crew would provide the information about the change. Furthermore, the implementation of such a scenario shall also take the privacy issues into account. The passenger shall be aware that her heart rate is measured automatically by a sensor embedded in the seat. These situations and issues shall be taken care of in the implementation of the real system, however they are not the focus of this research hence are not covered in this paper.

3.2 Adaptive framework

The framework starts by observing the passenger's current heart rate and creating an internal representation of the passenger's heart rate state. The information in this representation then must be processed in order to determine whether the passenger's heart rate is normal. If this is the case, the adaptive inference recommends a list of personalized music to keep the passenger's heart rate at normal according to the user profile. If the passenger's heart rate is faster or slower than normal, the adaptive inference recommends a personalized music playlist according to the user profile in order to adjust the passengers heart rate. The user profile includes the passenger's demographic and music preference information. This information can be filled out explicitly by the passenger when applying the membership.

If the adaptive inference recommends the music items that one does not like, one can reject the recommendation and reselect the music item oneself. During this process, the system logs the interaction between the passenger and the system. By mining on the log information, the user preference learning component can learn from it and update the music preference in the user profile. Figure 1 illustrates the design of the system framework. Figure 2 illustrates the working procedure of the framework.

Next the main components in the framework, their working procedures and the coordination among them will be described in detail.

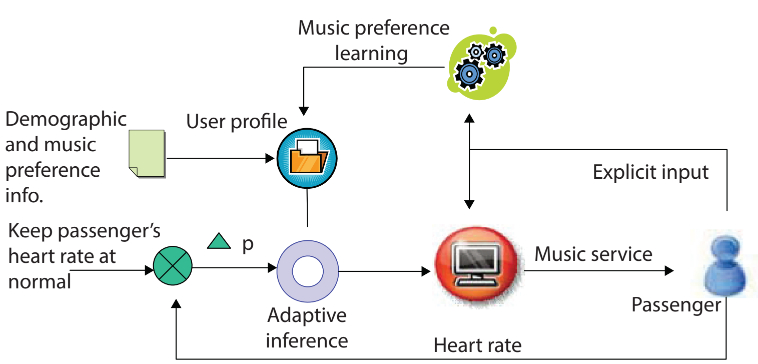

Figure 1. System Framework

3.2.1 *Heart rate as an indicator of stress*

The heart rate is the number of heartbeats per unit of time – typically expressed as Beats Per Minute (BPM). A normal heart rate is the BPM when a person is at

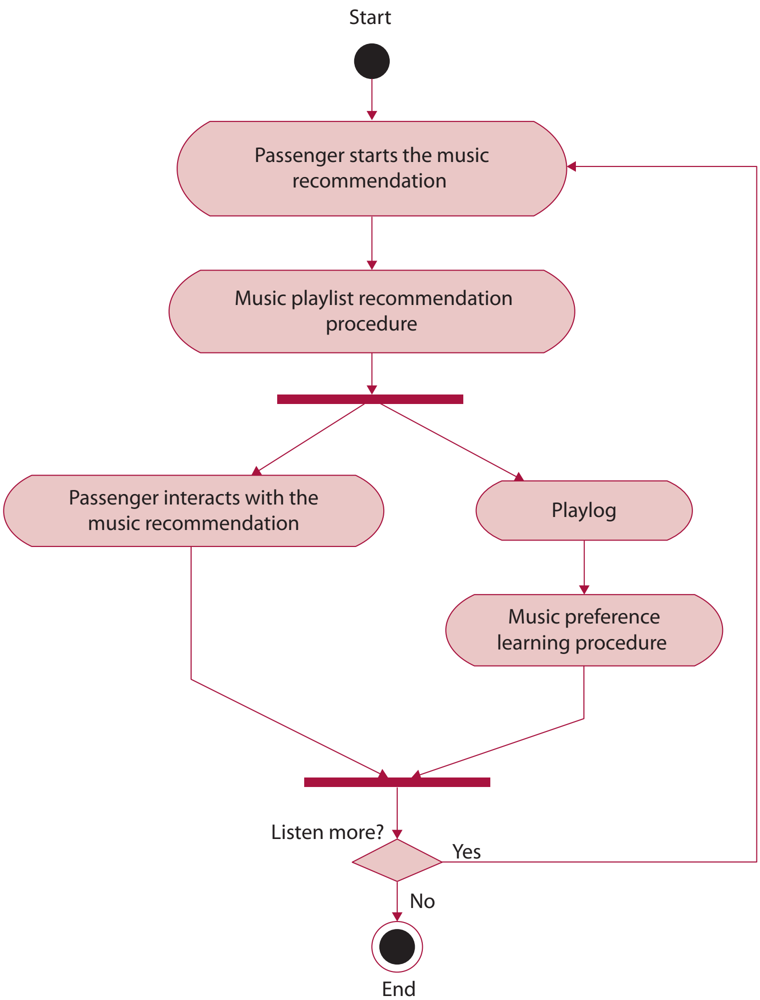

Figure 2. System working procedure

rest. For a young child (age 0–6), the normal heart rate is in the range of 90–110 BPM. For an older child and a teenager (age 6–18), the normal heart rate at rest is in the range of 70–100 BPM, and for an adult (age 18 and older), 60–100 BPM. An individuals normal heart rate range may depend on other factors such as air temperature, body position, and gender and body size, however in the research we focus on unusually high or low heart rate that may related to stress.

If the heartbeat rhythm is disrupted and the heart rate is lower than the normal, it is called bradycardia. Bradycardia may be caused by stress, a heart disease or the natural process of aging. It may cause tiredness, dizziness, breathlessness, blackouts, and if severe enough, death. This occurs because the person with bradycardia may not be pumping enough oxygen to the heart, which may cause heart attack-like symptoms (Kunz, 2008).

If a person's heart rate is higher than normal, it is called tachycardia. When the heart beats too fast, every heart beat pumps less efficiently and provides less blood flow to the rest of the body, including the heart itself. It is a normal response to stress, excitement, anxiety or exercise. A fast heart beat causes shortness of breath, dizziness, fainting, chest pain, severe anxiety, and even a heart attack if it persists. This occurs because the decreased flow of the necessary amount of oxygen to the heart causes myocardial cells to begin to die off (Kunz, 2008).

Research has showed that heart rate can be used as an indicator of stress. Elwess and Vogt (2005) found that during exams and when their graded results were returned to them, the heart rate of the most of the students increased substantially. Taelman et al. (2009) reported that a person's heart rate is significantly faster when the person is in mental stress. If one is in an excessive physical stress state, it may fatigue one's heart and cause bradycardia, or a slow heart rate.

3.2.2 *Music*

Music is an art form consisting of sound. Elements of sound in music include pitch, rhythm, and the sonic qualities of timbre and texture. It is created by artists to express emotions, feelings and thoughts. Music nowadays can be recorded, stored, transferred, and played back in an audio le in a digital format. Currently, there are tens of these formats, such as mp3, wma, wav and ogg. In this paper, mp3 is used. It is the most popular format for downloading and storing music and is supported by most of the digital music players.

Music metadata is usually part of the music file that is used to describe the characteristics of the music. ID3[1] is a metadata container used in conjunction with the mp3 audio file format. It allows information such as the title, artist, album, track number, tempo or other information about the music piece to be stored together in the file.

3.2.3 *User profile*

The user profile includes the passenger's demographic information and music preference. User demographic information includes items such as name, date of birth, gender etc. Date of birth is used by the system to determine the age dependent normal heart rate ranges.

1. http://www.id3.org

User's music preference is the user's long term evolving commitment to certain categories of music. Rentfrow and Gosling (2003) found that "when people discuss their music preferences they tend to do so first at the level of genres. Then, they talk to lesser extent subgenres. Only later, they step up to broader terms (e.g. loud) or down to specific artists (e.g. Van Halen) or songs (e.g. *Running with the Devil*)". In the implementation of this paper, the genre is the level at which the system starts the investigations of music preference. The music preference is modeled by a set of preference items with a weight of likeness, described in more details below.

The passenger's demographic information is input explicitly by the passenger when applying for the membership card. In the membership application form, the passenger needs to fill out name, gender, and date of birth. The passenger can also indicate to which extent she likes the fourteen popular music genres in the form. Those extents are then translated into the weight of the music preference item, on a scale of 1 to 7, where "like" is translated into 1 and "like strongly" is translated into 7. The "dislike" option was not included, since the "like" option towards a music genre can already provide sufficient items for recommendation. If a genre is not marked as any scale between 1 and 7, the weight is set to 0. Hereafter, these music preference items are used for music recommendation to avoid the cold-start problem. If the passenger is not willing to indicate music preference explicitly, the music preference will be learned by the system implicitly during the use of the system.

3.2.4 *Relation between music, heart rate and stress*

There are a few studies in literature that investigate the relation between heart rate and music tempo. Bernardi et al. (2006) found that listening to music with a slow or meditative tempo has a relaxing effect on people, slowing breathing and the heart rate. Listening to faster music with a more upbeat tempo has an opposite effect – speeding up respiration and the heart rate. Atluri (2008) reported similar results. Iwanaga (1995) found that people prefer music with the tempo ranging from 70 to 100 per minute which is similar to that of adults' heart rate in normal daily situations.

There is a long literature list involving the use of music for reducing stress. Miluk-Kolasa et al. (1996) showed that music was one of the relaxing adjuncts in modulating the ascent of autonomic responses to negative stress. Knight and Rickard (2001) reported that relaxing music attenuated blood pressure and heart rate after a stressful task; moreover, the level of anxiety was reduced after listening to relaxing music. The tempo of the music being listened to is an important parameter. Steelman (1991) investigated a number of studies of music effect on relaxation. He concluded that music items with tempos of 60 to 80 BPM reduce the stress and induce relaxation, while music items with tempos between 100 and 120 BPM stimulate the sympathetic nervous system. White and Shaw (1991) reported similar results. They found that tempos (40 to 60 BPM) slower than the average human's heart rate induce suspense, while tempos of 60 BPM are the most soothing.

Stratton and Zalanowski (1984) concluded that there is a significant correlation between degree of relaxation and preference for music. User preference, familiarity or past experiences with the music have an effect on positive behavior change.

3.2.5 *Adaptive inference*

The adaptive inference component is the central part of our framework. It mediates among the user's heart rate, the user profile, and the music to recommend personalized playlists to keep the user's heart rate at normal thus reduce the stress.

The adaptive inference component works with the ECA (Event-Control-Action) mechanism, that is, when the event occurs, if the condition is verified, then the action is executed. It receives the events (e.g. the passenger wants to listen to music) from the user and reacts on the personalized music recommendation controlled by the user's current heart rate.

When the passenger wants to listen to music during the flight, the adaptive inference component acquires the user profile and the current heart rate. According to the user's age it determines which state the user's heart rate is in and decides upon the playlist to be recommended accordingly.

If the user's heart rate is in a tachycardia state, the system recommends a personalized heart-rate-decreasing playlist to transfer it back to normal; if the user's heart rate is in a bradycardia state, the system recommends a personalized heart-rate-increasing playlist to bring it back to normal; otherwise, if the user's heart rate is in a normal state, the system recommends a personalized heart-rate-keeping music playlist to keep it at normal. Figure 3 illustrates the intended heart rate state changes with personalized music playlists.

The hybrid music recommendation method used in our framework combines context-based and content-based filtering recommendation approaches. There are two reasons why the collaborative filtering is not used. First, long haul flight passengers come from highly heterogeneous user groups that have different cultural background and nationalities. We know that culture does matter (Hu and Bartneck, 2008), and the major concern in relation to different cultural background and nationalities is the difference in the collections and styles of the music, and the similar preferences may appear mainly among the similar cultural background and nationalities. Dividing the passengers into cultural groups then applying collaborative filtering may improve the recommendation quality. However

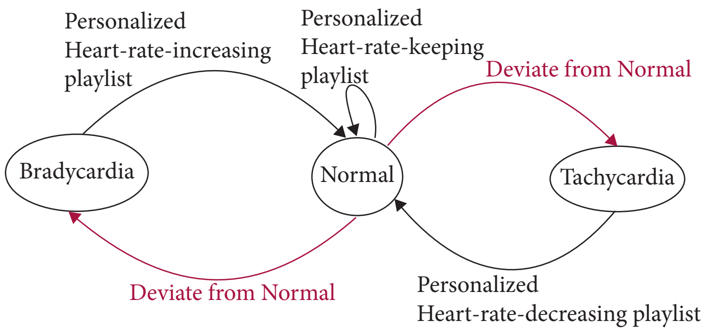

Figure 3. Intended heart rate state transfer with personalized music playlists

if the individual passengers preferences have already been taken into account to solve the cold-start problem in recommendation, we can then focus more on the relation between heart rate controlled recommendation and stress, instead of other recommendation strategies. Second, according to Stratton and Zalanowski (1984), there is a significant correlation between the degree of relaxation, and the music preferences and past experiences of the user; hence content-based filtering approach works better for the purpose of relaxation than the collaborative filtering approach. Content-based filtering approach is used to match the recommendation to the user's music preference, while context-based music filtering recommendation approach is used to adapt the music to the user's current heart rate.

If the passenger declines the recommendation, the system provides sufficient functionalities for the passenger to browse and select preferred music items to compose a music playlist manually.

3.2.6 *Interaction*

One of the important principles of designing user-system interaction is that the user must feel and actually be in control. It is identified by Dumur et al. (2004) as one of the most important principles to be taken into account when designing a more comfortable interaction space for aircraft passengers. It means that the passengers are aware of things that are going on during the interaction and they should not feel lost or confused. They shall be confident that everything is under control. For our system, three measures are identified to ensure the users are in control of the interaction processes: (1) they should be able to use and exit the system at any time; (2) they should be able to accept and decline the recommendations; (3) they should be able to browse and select preferred music items to compose a music playlist manually if they decline the recommendation.

Interaction with the system can be either explicit or implicit. Implicit input is for example the heart rate gathered by sensors in the seat. Passengers express their needs to the system by explicit input such as pressing a button and filling out the profile forms. System output is for example music recommendation and playback. The output is the system's response to the passenger's explicit or implicit input. The interaction is depicted in Figure 4.

During the flight, if the passengers want to listen to music, they can explicitly start the system by pressing a button. Once the system is started, it uses both the passenger's heart rate and the user profile to recommend music. The passengers can accept the recommendation and listen to the music playlist, but can also stop and exit the system at any time by explicitly pressing on the exit button or other buttons that lead to other parts of the system (for example, movies and games).

If the passenger declines the recommendation, the system provides the functionalities of a normal music system. The passenger can browse through the options in terms of artists or albums, and select the desired music items to manually compose a playlist. When the passenger listens to the self-selected music items, the system logs this information for learning the passenger's music preference.

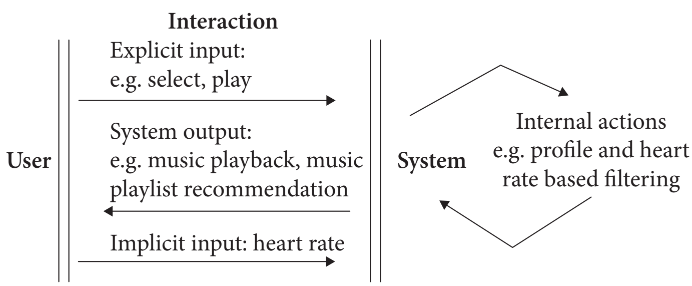

Figure 4. Interaction between the user and the system

3.2.7 *Preference learning*

In the user profile, the music preference items are either given by the passengers when they filled out the membership application forms, or later learned by the system by mining on the interactions between the passenger and the system.

The learning process adopts the method of "relevance feedback" to capture an appropriate snapshot of the user's interests in the music. Relevance feedback has been employed in several types of systems for the purpose of personalization (Haas et al., 2004; Hoi et al., 2006; Koenemann, 1996). Relevance feedback approaches are based on either an explicit or implicit feedback gathering scheme. In the explicit feedback gathering scheme, users provide the ratings with predefined scales. In the implicit feedback gathering scheme, ratings are inferred by mining on user interactions with the system in a transparent manner.

Implicit relevance feedback gathering techniques are proposed as unobtrusive alternative or supplement to explicit ratings. Click-throughs and time spent viewing a document are among the possible sources of implicit feedback. The system can thus monitor the user interaction to estimate the user preference.

Compared to explicit rating, the main benefits of implicit rating include: (1) it removes the cognitive cost of relevance judgment explicitly; (2) and it can be gathered in large quantities, and be aggregated to infer the relevance.

By the implicit relevance feedback, two types of music preference can be learned: (1) the music preference about which the user did not give the answers in the application form. In ID3 there are 126 genres of music in total, which makes it very tedious if the user is asked to rate on every of them; (2) user's latest music preference shown via the interaction with the system during the flight. User feedback on the recommendation can be implicitly captured by logging the interaction between the user and the system. The logged information includes the music items played and declined.

The working procedure of the music preference learning is as follows. Once a music item is played or declined, the system acquires the genre of the music item and check whether there is a music preference item in the user profile with the same genre. If there is a preference item and the music item is played, the weight of the music preference item increases 0.1, if there is a preference item and the music item is declined, the weight of the music preference item reduces 0.1; if there is not a preference item and the music item is played, a new music preference item is created where the weight is one and the genre is the played music item genre; if there is not a preference item and the music item is declined, a new music preference item is created with the weight zero. After learning the system updates the music preference items in the user profile.

4. Implementation

The software architecture was designed and implemented to realize the designed system framework. There are five abstraction layers in the designed software architecture. At the bottom layer is the resource layer. It serves the upper layers with music items, the user profile, and heart rate signals from the sensors.

In the second layer there is the resource manager. It includes the managers of the resources. The music manager is responsible for synchronizing the information about the available music on the physical storage to the corresponding metadata in the database, including managing the process of registration and deregistration. The heart rate manager collects signal readings from sensors, stores and updates heart rate information in database. The user profile manager collects and updates the passenger's demographic information and music preference.

The third layer is the database, decoupling the upper layer components from the lower layer. For example, replacing or updating components in the resource manager layer would not affect the rest of architecture as long as the corresponding data structures used to store data in the database are kept intact.

The fourth layer contains the adaptive control units. It includes the component that logs the passenger's feedback and interaction, the component for learning user preference, as well as the adaptive inference component. The log component is responsible for logging passenger's feedback to the recommended music into the database. The learning component is responsible for learning user's music preference based on the logged feedback and interaction. It forwards the learning results to the database and updates the user music preference. The adaptive inference mediates the user profile, the heart rate data and available music items to recommend personalized and heart rate controlled music playlists.

The fifth layer is the graphical user interface. This is the only layer to which the users have access. Figure 5 shows the software architecture.

The software architecture is implemented using Java, MySQL, Jamon[2] and Javascript with a client/server structure. The implementation is based on open source software Sockso[3] which provides a good foundation for a client/server

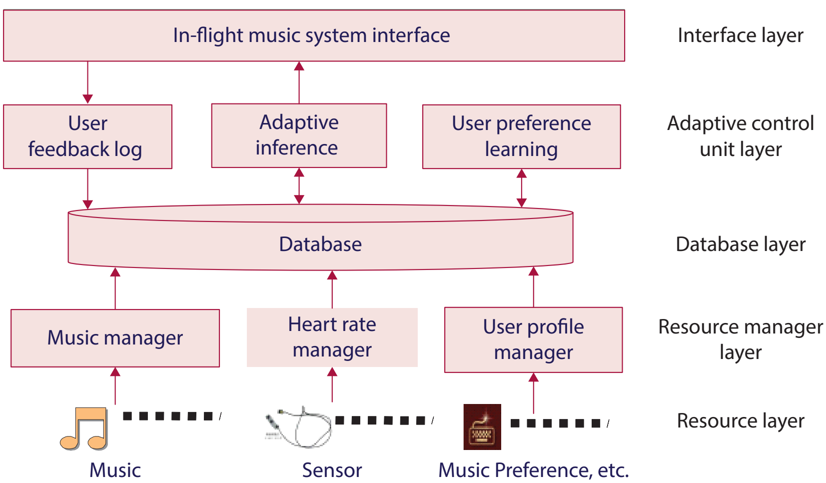

Figure 5. System Architecture

2. http://www.jamon.org
3. http://sockso.pu-gh.com

structured music player. The software is designed to be installed on an on-board server, delivering the content to in-seat computers over an intranet. The in-seat client provides a browser based interface for the passenger to interact with the service.

5. User experiment

5.1 Setup

The user experiments were conducted in the cabin simulator built as a test bed. The test bed is part of the simulation lab at the Department of Industrial design, Eindhoven University of Technology. It consists of a small scale aircraft cabin, a motion platform, a projection, and a control room. The aircraft cabin is divided into an economy class section, a business section, a lavatory and a kitchen. The cabin is supported by a motion platform that simulates the effects of taxing, taking off, landing and turbulence. The projector hung above the aircraft cabin projects the simulated window view on a wall next to the cabin. The control room is equipped with computers to support IFE systems and with observation monitors that are connected to surveillance cameras installed in the cabin. A surround sound system is installed to simulate the sound and noise during the flight. Figure 6 is the top view of the test bed.

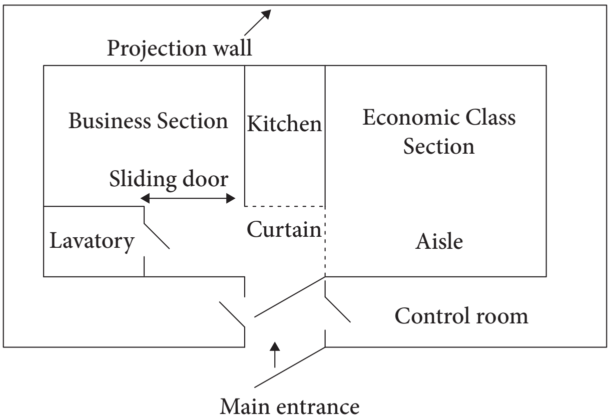

Figure 6. Top view of the test bed

In the economic class section are six economic class seats. Each seat is equipped with a touch screen connected to a personal computer in the control room via an EGA (Enhanced Graphics Adapter) extension cable. At each seat, a

SONY NC40 noise canceling earphone is provided. An extension cable connects the earphone jack from the armrest of the seat to a personal computer in the control room. In each seat, an Emfit[4] sensor film is non-intrusively embedded under the cover of the seat to measure the subjects heart rate. The Emfit processor and the power supply are installed under each seat. A USB extension cable connects the processor to a personal computer in the control room. Under the hand luggage compartment, speakers are installed for standard cabin announcements. Figure 7 shows the economy class section.

Figure 7. Economy class section

Our aircraft cabin in the flight simulator is equipped with a lavatory and kitchen to enable long-haul flight simulations. The lavatory is equipped with a camping toilet and a wash basin. The business class section in this work is used as the rest place for the flight attendant during the user experiments.

The control room is equipped with computer systems which are used to support, control and monitor the cabin simulator. For supporting the IFE system, six personal computers and a server computer are connected to form an on-board intranet providing IFE to the “passengers”. Each personal computer supports one in-seat touch screen and one normal LCD monitor in the control room via a video splitter. For controlling and monitoring the cabin, an additional personal computer is used for observation and announcement, connected to the surveillance cameras and the announcement speakers inside the cabin. Two other computers are used, one for the window view projection, the other for controlling the motion platform. Figure 8 shows the control room section.

4. http://www.emfit.com/en/sensors/productssensors/l-series

Figure 8. Control room section

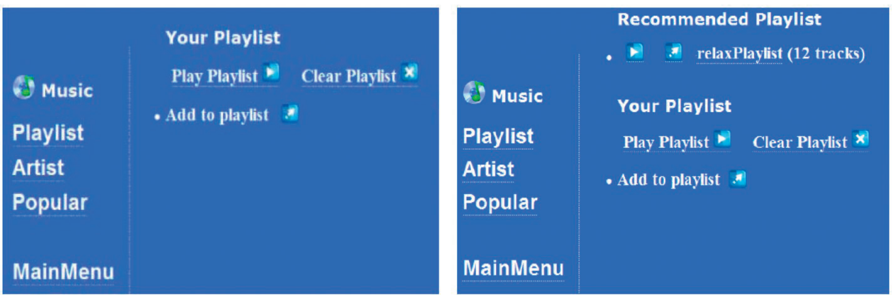

Figure 9. Music interface for the control (left) and experimental (right) groups

The passenger's window view is simulated with a $3m \times 6m$ projection on a wall using a Sanyo PLC-WXU300 wide screen LCD projector. The wall is painted white and it is 1.3m away from the cabin windows. The motion platform is constructed using four compressed air bags controlled by a computer. Each air bag supports one of the four corners of the cabin. The computer controls the air bags to be flatted and deflated in order to move the cabin to simulate the situations of taking off, landing and turbulence.

All test subjects are provided with IFE during the simulated flights. The subjects are divided into two groups, one control group, the other the experimental group. The entertainment content they receive is exactly the same but the experimental group would also have the option of heart rate controlled music recommendation. Hence the graphic interface differs only in terms of this additional recommendation option. Figure 9 (left) shows the interface for the control group, and Figure 9 (right) shows the one for the experimental group. The main difference is that the experimental group has the access to the bio-signal controlled recommended music playlist.

5.2 Test subjects

Twelve subjects were recruited to participate in the user experiments and all got 50 Euros for their effort. Six were allocated to the control group and the other six were allocated to the experimental group. Each group had three males and three females. Half of the subjects in each group were from Asia and the others were from Europe. The ages of the subjects in the control group ranged from 21 to 33. The ages of the subjects in the experimental group ranged from 23 to 32. The professions in the control group included one newspaper reporter, two blue collar workers and three engineers. The professions in the experimental group included one student, two blue collar workers and three engineers. Table 1 shows the details of the test subjects.

Table 1. Test subjects' profile

	Control group	**Experimental group**
Number of subjects	6 (3male, 3 females)	6 (3male, 3 females)
Nationality	3 European, 3 Asian	3 European, 3 Asian
Age	21–33	23–32
Professions	Reporter: 1; Blue collar workers: 2; Engineers: 3	Student: 1; Blue collar workers: 2; Engineers: 3

5.3 Procedure

The KLM flight KL0895 from Amsterdam Schiphol international airport to Shanghai Pudong international airport was simulated in our experiments.

The simulation procedure was a time-driven work flow. It coordinated the static set up with the dynamic services to create a virtual environment to give test subjects long haul flight experiences. In the simulation procedure, the boarding, taxing, take-off, flying, turbulence, landing, un-boarding flying situations are simulated by the synchronization of moving platform service, sky view projection service, and are enhanced by the captain information service and the attendant service. The participants were able to communicate with each other and the flight attendant, request the service and visit the lavatory as if they were in a real flight. Figure 10 illustrates the simulation procedure.

The KLM KL0895 left Amsterdam at 6:20pm and arrived in Shanghai next day at 4:55am (Shanghai local time 10:45am). The control group user experiment

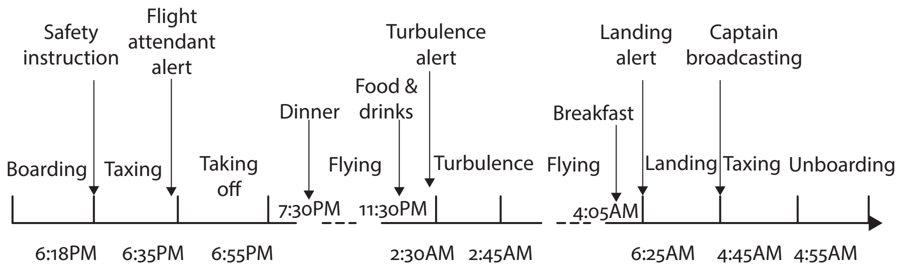

Figure 10. Simulation procedure

Figure 11. User experiment picture taken from the observation camera

was conducted first. The experiment with the experimental group was conducted one week later. Figure 11 shows two pictures taken from the observation camera during the experiment.

5.4 Experiment variables

5.4.1 *Variables*

Experiment variables include controlled, simulation quality, independent and dependent variables (Table 2). The simulation variable measures the test subjects' long haul flight experience. In this paper, the experience is measured by a presence questionnaire. Presence is defined as the subjective experience of being in one place or environment, even when one is physically situated in another (Bartneck and Hu, 2005; Hu, 2006). Presence means the "passenger's" subjective experience of being in the long haul flight; even when the "passenger' is physically sitting in the test bed. The presence questionnaire by Witmer and Singer (1998) was customized to measure the test subjects' presence. Controlled variables include the cabin environment (temperature, noise and humidity) and hearing ability. Hearing ability of the test subject is important for the effect of music listening. It is measured by the hearing ability questionnaire. The independent variable includes two test conditions: the test condition values are the control group and the experimental group. The dependent variables include the passenger's heart rate, stress level (objective measure indicated by the ratio between LF power and HF power of the heart rate variability) and stress scale (subjective measure indicated by self-report stress scale questionnaire (Bartenwerfer, 1969)).

Table 2. Variables

Simulation variables	Controlled variables	Independent variables	Dependent variables
Presence	Temperature Humidity Noise Hearing ability	Test condition (control group, experimental group)	Heart rate Stress

5.4.2 *Data acquisition*

5.4.2.1 Control variables. The cabin environment is measured every hour during user experiments (start from 6:30pm to 4:30am). The cabin temperature (Celsius) and humidity (percentage of water in the air) are measured by a wireless air temperature/humidity monitor and recorded in the control room. The inside cabin noise is measured by the flight attendant with the hand-held noise measurement equipment in decibel (dB). Hearing ability is measured by a questionnaire. It is composed by four questions to investigate the test subjects' hearing abilities. Test subjects need to fill out the questionnaires before the user experiments.

5.4.2.2 Simulation quality. The quality of the simulation is measured by the Presence questionnaire. It consists of five questions each on a 1 to 7 Likert scale where 1 represents 'not at all', 'at no time', etc. and 7 represents 'very much', 'almost all the time', etc. The questionnaires were filled out by the test subjects right after the user experiments before leaving the test bed.

5.4.2.3 Independent variable. Test condition values are the control group and the experimental group.

5.4.2.4 Dependent variables. Heart rate is measured in real time using Emfit bio sensors. It is recorded to MySQL database with BPM values. Additionally the stress was also measured by other two methods. One was the subjective self-report stress scale questionnaire (Bartenwerfer, 1969). It was distributed by the flight attendant to subjects every hour (started from 7:00pm until next day 5:00am). For each subject, she needs to report her stress scale 11 times with this stress scale questionnaire throughout the whole 'flight'. If the test subject was sleeping, her stress scale was reported by the flight attendant. This value was to determine which the middle scale point is between the "deep, dreamless sleep" state and "I am on the green in a forest and dreaming with open eyes" state. The flight attendant collected questionnaires five minutes after the distribution. The stress scales were coded by the physical distance between the "deep, dreamless sleep" state and the self-report state in centimeters.

Table 3. Result of the presence questions

Questions	Control group M (SD, N)	Experimental group M (SD, N)	Total M (SD, N)
Being there	4.00 (0.89,6)	4.00 (0.63,6)	4.00 (0.74,12)
Real flight	3.83 (1.47,6)	3.67 (1.03,6)	3.75 (1.22,12)
Lab or somewhere	3.67 (0.82,6)	3.50 (0.84,6)	3.58 (0.79,12)
Lab or flight	4.00 (0.89,6)	3.83 (0.75,6)	3.92 (0.79,12)
Sit in lab or flight	3.33 (1.21,6)	3.67 (1.03,6)	3.50 (1.08,12)

The other dependent variable is the objective stress level indicated by the ratio between LF power and HF power of the heart rate variability. User experiments showed that LF/HF ratio and residual heart rate increased/decreased in the stress/relax group during work days (Collins et al., 2005). LF band power (0.04–0.15Hz) and the HF band power (0.15–0.4Hz) are computed based on five minutes of heart rate every hour (start from 7:00pm-7:05pm until next day 5:00–5:05am) with Fast Fourier transformation (Welch's periodogram where the window is 52s with 50% overlap). The LF/HF ratio was calculated by dividing the LF band power by the HF band power.

5.5 Results

5.5.1 *Flight simulation quality*

The presence questionnaire results are reported for both groups together (Table 3). For the first presence question "I had a sense of being in a long haul flight" in the presence questionnaire, the mean is 4.00 (SD = 0.74, N = 12); for the second presence question There were times during the experience when the virtual 'long haul flight' became more real for me compared to the 'real flight', the mean is 3.75 (SD = 1.22, N = 12); for the third presence question "the test bed seems to me to be more like somewhere that I visited", the mean is 3.50 (SD = 0.79, N = 12); for the fourth presence question "I had a stronger sense of being in the virtual reality of the fligh", the mean is 3.92 (SD = 0.79, N = 12); for the fifth presence question "During the experience I had never thought that I was really sitting in the test bed because the long haul flight overwhelmed me", the mean is 3.50 (SD = 1.08, N = 12).

In order to compare flight simulation effects between the control group and experimental group, a one-way ANOVA test was performed (Table 4). None of the five questions shows significant difference. From question (1) to (5), the F and p values are: (1) $F(1,10) = 0.00$, $p = 1.0$; (2) $F(1,10) = 0.05$, $p = 0.83$; (3) $F(1,10) = 0.12$, $p = 0.73$; (4) $F(1,10) = 0.12$, $p = 0.72$; (5) $F(1,10) = 0.26$, $p = 0.62$. These results suggest that the test subjects of the control group and the experimental group share the same "flight experience".

Table 4. ANOVA for presence between the control group and the experimental group

	Groups	Sum of squares	df	Mean Squares	F	p
Being there	Between	0.00	1	0.00	0.00	1.00
	Within	6.00	10	0.60		
	Total	6.00	11			
Real flight	Between	0.08	1	0.08	0.05	0.83
	Within	16.17	10	1.62		
	Total	16.25	11			
Lab or somewhere	Between	0.08	1	0.01	0.12	0.73
	Within	6.83	10	0.68		
	Total	6.92	11			
Lab or flight	Between	0.08	1	0.01	0.12	0.73
	Within	6.83	10	0.68		
	Total	6.92	11			
Sit in lab Or flight	Between	0.33	1	0.33	0.26	0.62
	Within	12.67	10	1.27		
	Total	13.00	11			

5.5.2 *Control variables*

During user experiments, in order to control effects of spurious, intervening, and antecedent variables, measures were taken to try to keep temperature, humidity and noise between the control group and experimental group the same. Measures included using air condition and electric fan to control cabin temperature, the simulation procedure was strictly followed.

The temperature, humidity and noise inside the cabin were recorded each hour from 6:30pm until 4:30am. There were 22 data collected (11 each for the control group and the experimental group). The temperature mean in the control group is 29.35^0C and the standard deviation is 1.76^0C, its mean in the experimental group is 29.45^0C and the standard deviation is 0.82^0C; the humidity mean in the control group is 43% and the standard deviation is 5%, its mean in the experimental group is 42% and the standard deviation is 5%; the noise mean in the control group is 70.18dB and the standard deviation is 0.6dB, its mean in the experimental group is 70.9dB and the standard deviation is 1.04dB (Table 5).

Table 5. Mean, standard deviation and number of temperature, humidity and noise

Control variable	Control group M (SD, N)	Experimental group M (SD, N)
Temperature (^{0}C)	29.35 (1.76, 11)	29.45 (0.82, 11)
Humidity (%)	43 (5, 11)	42 (5, 11)
Noise (dB)	70.18 (0.6, 11)	70.9 (1.04,11)

In order to compare the differences of the temperature, humidity and noise between the control group and experimental group, an independent samples T test was performed (Table 6. For the temperature, the variances between the control group and experimental group are significantly different ($p = 0.007$), however, there is not a significant temperature difference between the control group and experimental group ($t = -0.155$, $df = 20$, $p = 0.879$); for the humidity, the variances between the control group and experimental group are not significantly different ($p = 0.888$), and there is not a significant humidity difference between the control group and experimental group ($t = 0.42$, $df = 20$, $p = 0.967$); for the noise, the

Table 6. Independent samples T test for the temperature, humidity and noise between the control group and the experimental group

		Levene's test for equality of variances		t-test for equality of means		
	Equal variances	**F**	**p**	**t**	**df**	**p (2 tailed)**
Temperature	Assumed	9.132	0.007	–0.155	20	0.878
	Not assumed			–0.155	14.144	0.878
Humidity	Assumed	0.002	0.888	0.42	20	0.967
	Not assumed				20.000	0.967
Noise	Assumed	19.048	0.001	–2.000	20	0.059
	Not assumed			–2.000	16.000	0.063

variances between the control group and experimental group are significantly different ($p=0.001$), however, there is not a significant noise difference between the control group and experimental group ($t=-2.000$, $df=20$, $p=0.063$). Although the noise difference seems to be marginally significant ($p=0.063$), the mean of the noise level is rather close (70.18 vs. 70.9), with very small deviation (.6 and 1.04 respectively). According to Karjalainen (1985), any difference between 1.5dB is hardly perceivable. These results suggest that temperature, humidity and noise were approximately equal between the control group and the experimental group during our experiments.

5.5.3 *Heart rate control*

During the control group experiment, there were 215998 heart rate data collected. The maximum of the heart rate is 149 and the minimum is 41. The mean of the heart rate is 66 and the standard deviation is 11. The median of the heart rate is 65 and the mode is 64. The skewness is 2. The test subjects' heart rate was in the bradycardia state 24.6% of the time. Overall the test subjects' heart rate was in the tachycardia state 7.3% of the time. Figure 12 illustrates the heart rate of the test subjects in the control group.

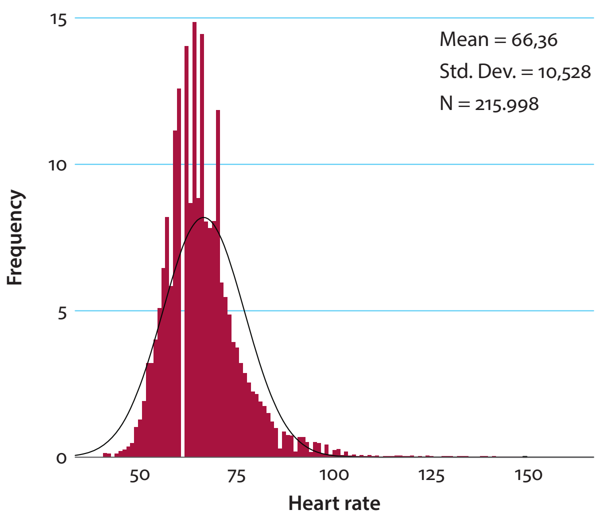

Figure 12. Histogram of heart rate for all test subjects in the control group

During the experimental group experiment, four test subjects listened to ten recommended increasing music playlists. The mean listening time of the increasing music playlists is 22 minutes. The mean of the bradycardia state duration is 6.86 seconds and the standard deviation is 9.02. The number of the bradycardia states is 481. Six test subjects listened to 20 recommended keeping music playlists. The mean listening time of the keeping music playlists is 25 minutes. The mean of the normal state duration is 29.79 seconds and the standard deviation is 53.32. The number of the normal states is 628. Three test subjects listened to three recommended decreasing music playlists. The playing time of the decreasing music playlists is 313, 223 and 1377 seconds respectively. The mean of the tachycardia state duration is 6.53 seconds and the standard deviation is 6.88. The number of the tachycardia states is 49. There were two subjects in the control group listening to music. One listened to music for 35 minutes. Another listened to music for 28 minutes. There were six test subjects sitting in the seat and doing nothing. The time durations were 167, 98, 131, 122,199, and 120 minutes. The mean of the bradycardia state duration for test subjects listening to music and sitting in the seat and doing nothing is 14.78 seconds and the standard deviation is 17.50. The number of the bradycardia states is 957. The mean of the normal state duration for test subjects listening to music and sitting in the seat and doing nothing is 24.66 seconds and the standard deviation is 40.82 seconds. The number of the normal states is 1262. The mean of the tachycardia state duration for test subjects listening to music and sitting in the seat and doing nothing is 13.89 seconds and the standard deviation is 11.87 seconds. The number of the tachycardia states is 186. Table 7 reports the results.

Table 7. Duration in seconds of bradycardia, normal and tachycardia states

Test condition	Heart rate state	M(SD,N)
Control group	Bradycardia	14.78 (17.50,957)
	Normal	24.66 (40.82,1262)
	Tachycardia	13.89 (11.87,186)
	Total	19.89 (32.12,2405)
Experimental group	Bradycardia	6.86 (9.02,481)
	Normal	29.79 (53.32,628)
	Tachycardia	6.53 (6.88,49)
	Total	19.28 (41.32,1158)
Total	Bradycardia	12.13 (15.65,1438)
	Normal	26.36 (45.41,1890)
	Tachycardia	12.35 (11.40,235)
	Total	19.70 (35.37,3563)

A MANOVA examined test condition and heart rate state as fixed factors and the heart rate state duration as the dependent variable. A Univariate analysis was then conducted to check on the heart rate increasing, keeping and decreasing effects. The mean difference of bradycardia, normal and tachycardia state duration between the control group and experimental group is significant ($F(2, 16327.242) = 13.686$, $p = 0.001$). The result suggests that the increasing music reduces the bradycardia state duration from 14.78 seconds in the control group to 6.86 seconds in the experimental group. The keeping music increases the normal state duration from 24.66 seconds in the control group to 29.79 seconds in the experimental group. The decreasing music reduces the tachycardia state duration from 13.89 seconds in the control group to 6.53 seconds in the experimental group. Table 8 reports the differences of the bradycardia, normal and tachycardia state duration in seconds between the control group and experimental group. Figure 13 illustrates the mean differences of the bradycardia, normal and tachycardia state duration in seconds between the control group and the experimental group.

Table 8. Differences of the bradycardia, normal and tachycardia state duration in seconds between the control group and experimental group

Source	Type III Sum of Squares	df	Mean Square	F	p
Corrected model	2121115.551[a]	5	42423.110	35.559	.001
Intercept	297659.230	1	297659.230	249.50	.001
TestCondition	3285.021	1	3285.021	2.754	.097
State	209789.937	2	104894.968	87.92	.001
TestCondition*State	32654.483	2	16327.242	13.686	.001
Error	4243566.263	3557	1193.018		
Total	5837891.000	3563			
Corrected Total	4455681.814	3562			

[a] R Squared = .048 (Adjusted R Squared = .046)

Because there is a relation between the bradycardia and tachycardia state durations with the normal state duration, a reduction of the durations in the bradycardia and tachycardia state means automatically an increase of the normal state duration, we exclude the normal state durations and repeat the univariate analysis. The mean difference of bradycardia and tachycardia state duration between the control group and experimental group is still significant ($F(1,8070.426) = 37.388$, $p = 0.001$). Table 9 reports the differences of the bradycardia and tachycardia state duration in seconds between the control group and experimental group.

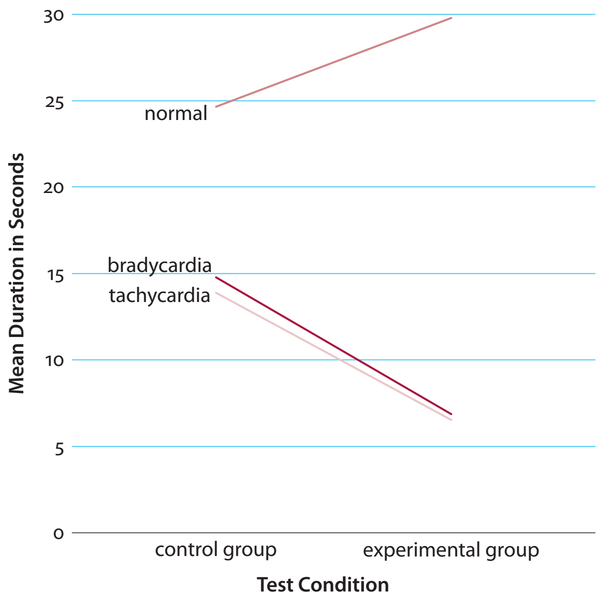

Figure 13. Mean difference of bradycardia, normal and tachycardia duration in seconds between the control group and experimental group

Table 9. Differences of the bradycardia and tachycardia state duration in seconds between the control group and experimental group excluding data for normal state duration in seconds

Source	Type III Sum of Squares	df	Mean Square	F	p
Corrected model	22176.626[b]	3	7392.209	34.264	.001
Intercept	61209.921	1	61209.921	283.570	.001
TestCondition	8070.426	1	8070.426	37.388	.001
State	52.381	1	52.381	.242	.632
TestCondition*State	10.899	1	10.899	.050	.822
Error	360261.841	1669	215.855		
Total	630020.000	1673			
Corrected Total	382438.467	1672			

[b] R Squared = .048 (Adjusted R Squared = .046)

5.5.4 *Stress reduction*

5.5.4.1 Stress scale. For the control group and the experimental group, each of them had 66 (6 subjects *11 times) stress scale data collected. Table 10 reports the mean, standard deviation and number of the stress scale from 7pm until 5am.

In order to examine whether there is a significant difference of the stress scale between the control group and experimental group, a MANOVA examined the stress scale as within-subjects variables and test condition as the between subject factor. Repeated measures analysis was then conducted.

Table 10. Stress scale of the control group and experimental group

Time	Control group M (SD,N)	Experimental group M (SD,N)	Total M (SD,N)
7pm	11.72 (1.53,6)	9.08 (4.27,6)	10.40 (3.35,12)
8pm	10.88 (2.55,6)	10.17 (2.79,6)	10.53 (2.57,12)
9pm	9.88 (2.65,6)	7.40 (4.43,6)	8.64 (3.71,12)
10pm	9.85 (2.69,6)	8.35 (2.87,6)	9.10 (2.77,12)
11pm	11.50 (1.66,6)	7.67 (4.73,6)	9.58 (3.93,12)
12pm	9.82 (2.62,6)	4.08 (4.20,6)	6.95 (4.48,12)
1am	7.88 (3.34,6)	5.90 (4.26,6)	6.89 (3.79,12)
2am	6.97 (5.04,6)	3.68 (4.58,6)	5.33 (4.90,12)
3am	5.10 (2.41,6)	2.08 (1.11,6)	3.59 (2.38,12)
4am	7.63 (2.93,6)	2.67 (1.03,6)	5.15 (3.33,12)
5am	8.25 (3.38,6)	8.22 (3.66,6)	8.23 (3.36,12)

The analysis result indicates that the stress scale difference between the control group and the experimental group is significant ($F(1,7769.404)=264.899$, $p=0.016$). It suggests that listening to the recommended music reduced the test subjects' stress. Table 11 reports the analysis result of the stress scale differences between the control group and experimental group. Figure 14 illustrates the stress scale of the control group and the experimental group from 7pm until 5am.

Table 11. Analysis result of the stress scale differences between the control group and experimental group

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	7769.404	1	7769.404	264.899	.001
TestCondition	248.464	1	248.464	8.471	.016
Error	293.297	10	29.330		

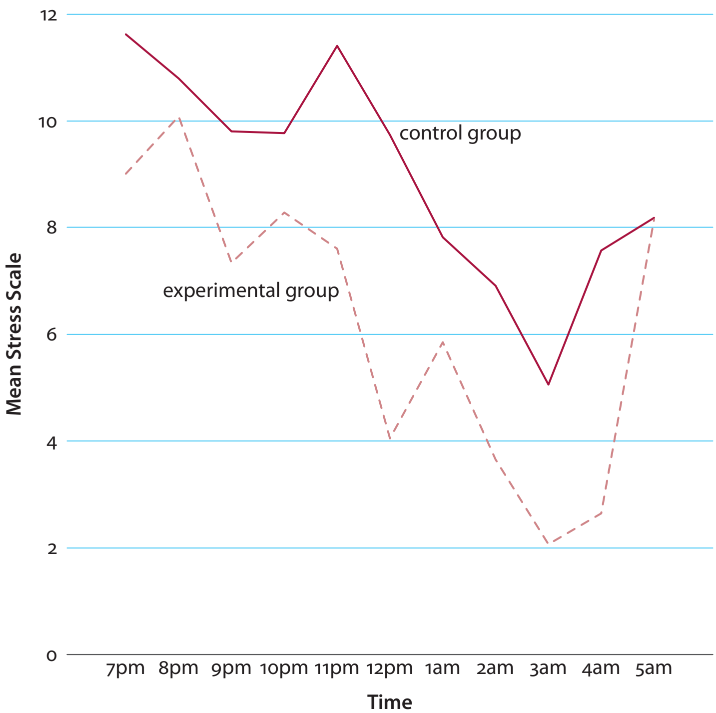

Figure 14. Stress scales of the control group and the experimental group

5.5.4.2 Stress level. For the control group and the experimental group, each of them had 66 (6 subjects *11 times) stress level data computed by the ratio of LF and HF. Table 12 reports the mean, standard deviation and number of stress level from 7pm until 5am.

In order to examine whether there is a significant difference of the stress level between the control group and experimental group, a MANOVA examined the stress level as within-subjects variables and test condition as the between-subject factor. Repeated measures analysis was then conducted. The analysis result indicates that the stress level mean difference between the control group and experimental group is not significant $F(1,14.733)=2.988$, $p=.115$). Table 13 reports the analysis result of the stress level differences between the control group and the experimental group. Figure 15 illustrates the stress level of the control group and the experimental group from 7pm until 5am.

As reported in Table 12 and illustrated in Figure 15 the results showed that except for 2am and 4am the mean LF/HF of the experimental group is lower than that of the control group. According to flight attendant reports and analysis

Table 12. Mean, standard deviation and number of the stress level

Time	Control group M (SD,N)	Experimental group M (SD,N)	Total M (SD,N)
7pm	2.40 (2.88,6)	2.07 (1.57,6)	2.23 (2.22,12)
8pm	2.88 (1.91,6)	2.40 (1.53,6)	2.64 (1.67,12)
9pm	3.45 (1.71,6)	2.75 (2.12,6)	3.10 (1.87,12)
10pm	2.25 (1.81,6)	1.97 (0.94,6)	2.11 (1.39,12)
11pm	2.20 (1.46,6)	1.08 (0.78,6)	1.64 (1.26,12)
12pm	3.27 (2.23,6)	1.83 (1.21,6)	2.55 (1.87,12)
1am	3.00 (1.31,6)	1.60 (1.07,6)	2.30 (1.35,12)
2am	2.65 (2.15,6)	2.67 (1.13,6)	2.66 (1.64,12)
3am	3.90 (3.13,6)	1.55 (1.19,6)	2.73 (2.57,12)
4am	2.10 (1.31,6)	3.00 (1.69,6)	2.55 (1.52,12)
5am	2.62 (1.49,6)	2.45 (1.44,6)	2.53 (1.40,12)

Table 13. Analysis result of the stress level differences between the control group and experimental group

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	797.729	1	797.729	161.805	.001
TestCondition	14.733	1	14.733	2.988	.115
Error	49.302	10	4.930		

of the video recordings, at 2am, there were four test subjects sleeping in the experimental group, while only two test subjects were sleeping in the control group.

Before going deeper into the stress level of sleeping persons, some knowledge of sleep needs to be explored. Sleep is prompted by natural cycles of activity in the brain and consists of two basic states: rapid eye movement (REM) sleep and non-rapid eye movement (NREM) sleep. There are four stages in NREM sleep. During sleep, the body cycles between non-REM sleep and REM sleep. One cycle lasts about 90 to 100 minutes. Non-REM dreams are more likely to consist of brief, fragmentary impressions that are less emotional and less likely to involve visual images than REM sleep dreams. About 20 percent of sleep is REM sleep. Dreams generally occur in the REM stage of sleep. According to Elsenbruch et al. (1999), compared to wake and NREM sleep, REM sleep is associated with decreased HF power, and significantly increased LF power to HF power ratio, which means high stress level.

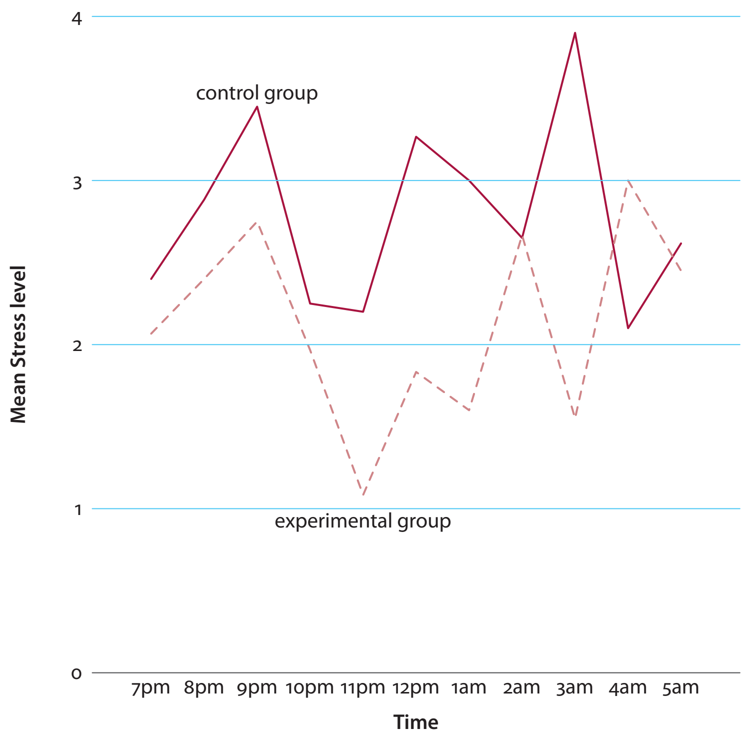

Figure 15. Stress scales of the control group and the experimental group

It is not clear which state (Non-REM and REM) of sleeping test subjects was in at 2:00am. In the control group, one sleeping test subject stress level was 6.4 and the other was 3.5. In the experimental group, the stress level data for sleeping test subjects were 3.7, 3.8, 3.2, and 2.7. For sleeping test subjects, the mean stress level of the experimental group is lower than that of the control group. According to Elsenbruch et al. (1999), the sleeping test subject with 6.4 stress level value was probably in a REM state of sleep, the others were in NREM state of sleep. At 4:00am, there were four test subjects sleeping in the experimental group. There were no test subjects sleeping in the control group. One sleeping test subject's stress level is 6.0 which are much higher than others.

To avoid these uncontrolled sleep effects, we exclude all data between 2am and 4am and repeat the "repeated measures" analysis MANOVA again. The stress level difference between the control group and experimental group becomes significant $F(1,2.531)=5.335$, $p=.044$). The result suggests that the recommended music reduced the test subjects' stress level significantly. Table 14 reports the analysis result of the stress level differences between the control group and experimental group.

Table 14. Analysis result of the stress level differences between the control group and experimental group excluding data for sleeping hours at 2am and 4am

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	70.621	1	70.621	148.861	.001
TestCondition	2.531	1	2.531	5.335	.044
Error	4.744	10	0.474		

6. Limitations

When designing a heart rate controlled music recommendation system to reduce the user's stress, stress measures need to be an accurate indicator of the actual user's stress. To this end, the heart rate variability is a more accurate indicator of stress than the heart rate itself. However, so far we could not find the relations between the heart rate variability and music in literature, therefore heart rate was chosen as the stress indicator. Although we found that by keeping only the passenger's heart rate at normal with music, the user's stress level indicated by heart rate variability can be reduced, the relations between stress, music tempo and heart rate need further investigations.

The system developed in this paper has never been tested with real passengers in real flights. Experiments were only being done in a simulated long haul flight cabin environment, although much attention and effort has been paid to make it as "real" as possible.

7. Conclusions

Based on the achieved results we can conclude that by incorporating the passenger's heart rate signal as an indicator of stress and by feeding this indicator back into a music recommendation process, the stress of the passengers in the economy class cabin of a long haul flight can be significantly reduced.

The presented heart rate controlled music recommendation system contributed to these results by its architectural design. An adaptive framework is designed and implemented, which integrates the concepts of control systems, content-based music recommendation, context-based music recommendation, user profiling, and using music to adjust the user's heart rate. The target of the framework is to keep the long-haul flight passengers heart rate at normal, using heart rate as in indicator of the stress, by recommending personalized increasing/keeping/decreasing music playlists. The framework is implemented with a component-based architecture with five abstraction levels.

These results were observed in two user experiments, simulating a real long haul flight from Europe to Asia. One user experiment was for the control group; the other user experiment was for the experimental group. The difference in the setups between these two experiments was that the experimental group had access to the additional functionality with recommended music playlists. Both the stress scale measured by self reporting and the stress level measured by the heart rate have shown significant reductions.

Although significant results have been found, more research is needed before the technologies developed with this heart rate controlled music recommendation system could be effectively incorporated into real aircraft cabins. Future work should focus on investigating the relations between stress, heart rate, heart rate variability and music, and if possible, conducting user experiments in real flights.

Acknowledgement

This project is sponsored by the European Commission DG H.3 Research, Aeronautics Unit under the 6th Framework Programme SEAT, contract Number: AST5-CT-2006-030958. Emfit Ltd Finland kindly sponsored the project with a generous discount for their Emfit sensors.

References

Abrazhevich, D., Markopoulos, P., & Rauterberg, M. (2009). Designing internet-based payment systems: Guidelines and empirical basis. *Human-Computer Interaction*, *24*(4), 408–443. DOI: 10.1080/07370020903038144

Atluri, M. (2008). *Does music affect blood pressure and heart rate?* Project Number J1103, California State Science Fair.

Baltrunas, L., Kaminskas, M., Ludwig, B., Moling, O., Ricci, F., Aydin, A., Lke, K.-H., & Schwaiger, R. (2011). Incarmusic: Context-aware music recommendations in a car. In *E-Commerce and web technologies*, (pp. 89–100). Springer. DOI: 10.1007/978-3-642-23014-1_8

Bartenwerfer, H. (1969). Einige praktische konsequenzen der aktivierungstheorie. *Zeitschrift fur experimentelle und angewandte Psychologie*, *16*, 195–222.

Bartneck, C., & Hu, J. (2005). Presence in a distributed media environment. In L. Terrenghi, I. Lindt, A. Butz, & M. Kuniavsky (eds.), *User experience design for pervasive computing, pervasive 2005*, München, Munich, Germany: Ludwig-Maximilians-Universitat. Retrived on October 27, 2014. from http://www.fluidum.org/events/experience05/cameraready/bartneck.pdf

Basilico, J., & Hofmann, T. (2004). Unifying collaborative and content-based filtering. In 21st *International conference on Machine learning (ICML)*, 65–72. ACM. DOI: 10.1145/1015330.1015394

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*(4), 445–452. DOI: 10.1136/hrt.2005.064600

Cano, P., Koppenberger, M., & Wack, N. (2005). Content-based music audio recommendation. In 13th *Annual ACM International Conference on Multimedia (MULTIMEDIA '05)*, 211–212. New York, USA: ACM. DOI: 10.1145/1101149.1101181

Çataltepe, Z., & Altinel, B. (2007). Hybrid music recommendation based on different dimensions of audio content and an entropy measure. In 15th *European Signal Processing Conference (EU-SIPCO)*, 936–940.

Chai, W., & Vercoe, B. (2000). Using user models in music information retrieval systems. In 1st *Annual International Symposium on Music Information Retrieval (ISMIR 2000),* Retrieved on October 27, 2014, from http://ciir.cs.umass.edu/music2000/posters/chai.pdf

Collins, S., Karasek, R., & Costas, K. (2005). Job strain and autonomic indices of cardiovascular disease risk. *American Journal of Industrial Medicine*, 48(3): 182–193. DOI: 10.1002/ajim.20204

Dumur, E., Barnard, Y., & Boy, G. (2004). Designing for comfort. *Human Factors in Design*, 111–127.

Elahi, M., Ricci, F., & Rubens, N. (2013). Active learning strategies for rating elicitation in collaborative filtering: A system-wide perspective. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(1), 13. DOI: 10.1145/2542182.2542195

Elsenbruch, S., Harnish, M., & Orr, W. (1999). Heart rate variability during waking and sleep in healthy males and females. *Sleep*, *22*(8): 1067–1071.

Elwess, L., & Vogt, F. (2005). Heart rate and stress in a college setting. *Bioscience*, *31*(4), 20–23.

Golbandi, N., Koren, Y., & Lempel, R. (2011). Adaptive bootstrapping of recommender systems using decision trees. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 595–604. ACM. DOI: 10.1145/1935826.1935910

Haas, M., Rijsdam, J., Thomee, B., & Lew, M. (2004). Relevance feedback: Perceptual learning and retrieval in bio-computing, photos, and video. In *MIR '04: Proceedings of the* 6th *ACM SIGMM International Workshop on Multimedia Information Retrieval*, 151–156, New York, NY, USA, ACM. DOI: 10.1145/1026711.1026737

Hoi, S., Lyu, M., & Jin, R. (2006). A unified log-based relevance feedback scheme for image retrieval. *IEEE Transactions on Knowledge and Data Engineering*, *18*(4), 509–524. DOI: 10.1109/tkde.2006.1599389

Hu, J. (2006). *Design of a distributed architecture for enriching media experience in home theaters*. Phd thesis, Department of Industrial Design, Eindhoven University of Technology.

Hu, J., & Bartneck, C. (2008). Culture matters – a study on presence in an interactive movie. *CyberPsychology and Behavior*, *11*(5), 529–535. DOI: 10.1089/cpb.2007.0093

Iwanaga, M. (1995). Relationship between heart rate and preference for tempo of music. *Perceptual and Motor Skills*, *81*(2), 435–440. DOI: 10.2466/pms.1995.81.2.435

Kaminskas, M., & Ricci, F. (2012). Contextual music information retrieval and recommendation: State of the art and challenges. *Computer Science Review*, *6*(2), 89–119. DOI: 10.1016/j.cosrev.2012.04.002

Karjalainen, M. (1985). A new auditory model for the evaluation of sound quality of audio systems. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'85,* vol. 10, 608–611. IEEE. DOI: 10.1109/icassp.1985.1168376

Knees, P., Pohle, T., Schedl, M., & Widmer, G. (2006). Combining audio-based similarity with web-based data to accelerate automatic music playlist generation. In *Proceedings of the* 8th *ACM International Workshop on Multimedia Information Retrieval*, 147–154. ACM. DOI: 10.1145/1178677.1178699

Knight, W., & Rickard, N. (2001). Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure, and heart rate in healthy males and females. *Journal of Music Therapy*, *38*(4), 254–272. DOI: 10.1093/jmt/38.4.254

Koenemann, J. (1996). Supporting interactive information retrieval through relevance feedback. In *CHI '96: Conference Companion on Human Factors in Computing Systems*, 49–50, New York, NY, USA. ACM. DOI: 10.1145/257089.257131

Kunz, M.G. (2008). *Zombie Notes Bradycardia – Heart Blocks*. Dickson Keanaghan, LLC.

Liu, H. (2010). *Biosignal controlled recommendation in entertainment systems*. Phd thesis, Department of Industrial Design, Eindhoven University of Technology.

Liu, H., Hu, J., & Rauterberg, M. (2008). Airsf: A new entertainment adaptive framework for stress free air travels. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*, 183–186. ACM. DOI: 10.1145/1501750.1501793

Liu, H., Hu, J., & Rauterberg, M. (2009a). Bio-feedback based in-flight music system design to promote heart health. In *International Conference on Machine Learning and Cybernetics (ICMLC)*, 446–450. Baoding, China. Citeseer.

Liu, H., Hu, J., & Rauterberg, M. (2009b). Music playlist recommendation based on user heartbeat and music preference. In *International Conference on Computer Technology and Development (ICCTD'09),* vol. 1, 545–549. IEEE. DOI: 10.1109/icctd.2009.246

Masuhr, J., Klompmaker, F., Reimann, C., & Nebe, K. (2008). Designing context-aware in-car information systems. In *Proceedings of the* 5th *Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services (Mobiquitous '08)*, 1–8. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). DOI: 10.4108/icst.mobiquitous2008.3637

Miluk-Kolasa, B., Matejek, M., & Stupnicki, R. (1996). The effects of music listening on changes in selected physiological parameters in adult pre-surgical patients. *Journal of Music Therapy*, *33*(8), 208–218. DOI: 10.1093/jmt/33.3.208

Mokbel, M., & Levandoski, J. (2009). Toward context and preference-aware location-based services. In *Eighth ACM International Workshop on Data Engineering for Wireless and Mobile Access*, 25–32. ACM. DOI: 10.1145/1594139.1594150

Rauterberg, G. (2006). Hci as an engineering discipline: To be or not to be!? *African Journal of Information and Communication Technology*, *2*(4), 163–184.

Rentfrow, P., & Gosling, S. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*(6), 1236–1256. DOI: 10.1037/0022–3514.84.6.1236

Rubens, N., Kaplan, D., & Sugiyama, M. (2011). Active learning in recommender systems. In *Recommender Systems Handbook*, 735–767. Springer. DOI: 10.1007/978-0-387-85820-3_23

Sarwar, B., Karypis, G., Konstan, J., & Reidl, J. (2001). Item-based collaborative filtering recommendation algorithms. In 10th *International Conference on World Wide Web (WWW '01)*, 285–295, New York. ACM. DOI: 10.1145/371920.372071

Siddiqi, R. (2011). The world of 21st century in-flight entertainment. *Aviation & Tourism (FRIDAY, 19 AUGUST 2011) in The Independent*.

Steelman, V. (1991). Relaxing to the beat: Music therapy in perioperative nursing. *Today's OR Nurse*, *13*(7), 18.

Stratton, V., & Zalanowski, A. (1984). The relationship between music, degree of liking, and self-reported relaxation. *Journal of Music Therapy*, *21*(4), 184–192. DOI: 10.1093/jmt/21.4.184

Su, J., Yeh, H., Yu, P., & Tseng, V. (2010). Music recommendation using content and context information mining. *Intelligent Systems, IEEE*, *25*(1), 16–26. DOI: 10.1109/mis.2010.23

Su, X., & Khoshgoftaar, T. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, *2009*, 1–19 (Article ID 421425). DOI: 10.1155/2009/421425

Suh, Y., Park, Y., Yoon, H., Chang, Y., & Woo, W. (2007). Context-aware mobile ar system for personalization, selective sharing, and interaction of contents in ubiquitous computing environments. In J. Jacko (ed.), *Human-Computer Interaction. Interaction Platforms And Techniques*, Lecture Notes in Computer Science, 966–974. Springer Berlin/Heidelberg. DOI: 10.1007/978-3-540-73107-8_106

Taelman, J., Vandeput, S., Spaepen, A., & Huffel, S. (2009). Influence of mental stress on heart rate and heart rate variability. In J. Sloten, P. Verdonck, M. Nyssen, & J. Haueisen (eds.), 4th *European Conference of the International Federation for Medical and Biological Engineering*, volume 22 of *IFMBE Proceedings*, 1366–1369. Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-89208-3_324

White, J., & Shaw, C. (1991). Music therapy: A means of reducing anxiety in the myocardial infarction patient. *Wisconsin Medical Journal*, *90*(7), 434–437.

Witmer, B., & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240. DOI: 10.1162/105474698565686

Woerndl, W., Schueller, C., & Wojtech, R. (2007). A hybrid recommender system for context-aware recommendations of mobile applications. In *2007 IEEE* 23rd *International Conference on Data Engineering Workshop*, 871–878. Istanbul.

World Health Organization. (2005). *Travel by air: Health considerations, in International travel and health: Situation as on 1 January 2005*. World Health Organization. DOI: 10.1002/0470011815.b2a17156

Yoshii, K., Goto, M., Komatani, K., Ogata, T., & Okuno, H. (2006). Hybrid collaborative and content-based music recommendation using probabilistic model with latent user preferences. In 7th *International Conference on Music Information Retrieval (ISMIR)*, 296–301.

Authors' Addresses

Hao Liu
Xiaoying West 33
Haidian District
Beijing 100085
China

liuh@zhigu.com

Jun Hu
Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5612AZ Eindhoven
The Netherlands
Telephone: +31 (40) 247 8331

Email: j.hu@tue.nl

Matthias Rauterberg
Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2
5612AZ Eindhoven
The Netherlands

G.W.M.Rauterberg@tue.nl

Biographical notes

Dr. Hao Liu is a software engineer with expertise and interest in bio-signal measurements for innovative applications in the areas of Wireless Autonomous Transducer Solutions. He has a Ph.D. in industrial design.

Dr. Jun Hu is a senior member of ACM, and an associate professor at the department of Industrial Design, Eindhoven University of Technology, with a background in computer science and human-computer interaction. His current research interest is in designing for social computing and social interaction.

Prof. Dr. Matthias Rauterberg is a full professor for Interactive System Design, also the head of the Designed Intelligence group, at the department of Industrial Design, Eindhoven University of Technology. His research interest is in human-computer interaction, entertainment computing and cultural computing.