UNIVERSIDADE CATÓLICA PORTUGUESA

SOUND BASED SOCIAL NETWORKS

Dissertation submitted to the Portuguese Catholic University in partial fulfillment of requirements of the Doctoral Degree in Science and Technologies of the Arts – Computer Music

by

João Miguel Magalhães Marcelino Fernandes Cordeiro

ESCOLA DAS ARTES

September 2013

UNIVERSIDADE CATÓLICA PORTUGUESA

SOUND BASED SOCIAL NETWORKS

Dissertation submitted to the Portuguese Catholic University in partial fulfillment of requirements of the Doctoral Degree in Science and Technologies of the Arts – Computer Music.

By João Miguel Magalhães Marcelino Fernandes Cordeiro

Dissertation supervised by Professor Álvaro Mendes Barbosa and co-supervised by Professor Luis Gustavo Pereira Marques Martins.

ESCOLA DAS ARTES

September 2013

This Doctorate Research was partially supported by the Portuguese Science and Technology Foundation (FCT), trough the award of the Doctorate Grant reference SFRH / BD / 47844 / 2008.

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

The author is affiliated with:

CATÓLICA
UNIVERSIDADE CATÓLICA PORTUGUESA | PORTO
Escola das Artes | Som e Imagem

To Cristina

Abstract

The sound environment is an eco of the activity and character of each place, often carrying additional information to that made available to the eyes (both new and redundant). It is, therefore, an intangible and volatile acoustic fingerprint of the place, or simply an acoustic snapshot of a single event. Such rich resource, full of meaning and subtleness, Schaeffer called Soundscape. The exploratory research project presented here addresses the Soundscape in the context of Mobile Online Social Networking, aiming at determining the extent of its applicability regarding the establishment and/or strengthening of new and existing social links. Such research goal demanded an interdisciplinary approach, which we have anchored in three main stems: Soundscapes, Mobile Sound and Social Networking. These three areas pave the scientific ground for this study and are introduced during the first part of the thesis. An extensive survey of the state-of-the-arte projects related with this research is also presented, gathering examples from different but adjacent areas such as mobile sensing, wearable computing, sonification, social media and context-aware computing. This survey validates that our approach is scientifically opportune and unique, at the same time.

Furthermore, in order to assess the role of Soundscapes in the context of Social Networking, an experimental procedure has been implemented based on an Online Social Networking mobile application, enriched with environmental sensing mechanisms, able to capture and analyze the surrounding Soundscape and users' movements. Two main goals guided this prototypal research tool: collecting data regarding users' activity (both sonic and kinetic) and providing users with a real experience using a Sound-Based Social Network, in order to collect informed opinions about this unique type of Social Networking. The application – Hurly-Burly – senses the surrounding Soundscape and analyzes it using machine audition techniques, classifying it according to four categories: speech, music, environmental sounds and silence.

Additionally, it determines the sound pressure level of the sensed Soundscape in $dB(A)_{eq}$. This information is then broadcasted to the entire online social network of the user, allowing each element to visualize and audition a representation of the collected data. An individual record for each user is kept available in a webserver and can be accessed through an online application, displaying the continuous acoustic profile of each user along a timeline graph. The experimental procedure included three different test groups, forming each one a social network with a cluster coefficient equal to one.

After the implementation and result analysis stages we concluded that Soundscapes can have a role in the Online Social Networking paradigm, specially when concerning mobile applications. Has been proven that current off-the-shelf mobile technology is a promising opportunity for accomplishing this kind of tasks (such as continuous monitoring, life logging and environment sensing) but battery limitations and multitasking's constraints are still the bottleneck, hindering the massification of successful applications. Additionally, online privacy is something that users are not enthusiastic in letting go: using captured sound instead of representations of the sound would abstain users from utilizing such applications. We also demonstrated that users who are more aware of the Soundscape concept are also more inclined to assume it as playing an important role in OSN. This means that more pedagogy towards the acoustic phenomenon is needed and this type of research gives a step further in that direction.

Resumo

O ambiente sonoro de um lugar é um eco da sua atividade e carácter, transportando, na maior parte da vezes, informação adicional àquela que é proporcionada à visão (quer seja redundante ou complementar). É, portanto, uma impressão digital acústica - tangível e volátil - do lugar a que pertence, ou simplesmente uma fotografia acústica de um evento pontual. A este opulento recurso, carregado de significados e subtilezas, Schafer chamou de Paisagem-Sonora. O projeto de investigação de carácter exploratório que aqui apresentamos visa o estudo da Paisagem-Sonora no contexto das Redes Sociais Móveis Em-Linha, procurando entender os moldes e limites da sua aplicação, tendo em vista o estabelecimento e/ou reforço de novos ou existente laços sociais, respectivamente. Para satisfazer este objectivo foi necessária uma abordagem multidisciplinar, ancorada em três pilares principais: a Paisagem-Sonora, o Som Móvel e as Redes Sociais. Estas três áreas determinaram a moldura científica de referência em que se enquadrou esta investigação, sendo explanadas na primeira parte da tese. Um extenso levantamento do estado-da-arte referente a projetos relacionados com este estudo é também apresentado, compilando exemplos de áreas distintas mas adjacentes, tais como: *Computação Sensorial Móvel*, *Computação Vestível*, *Sonificação*, *Média Social* e *Computação Contexto-Dependente*. Este levantamento veio confirmar quer a originalidade quer a pertinência científica do projeto apresentado.

Posteriormente, a fim de avaliar o papel da Paisagem-Sonora no contexto das Redes Sociais, foi posto em prática um procedimento experimental baseado numa Rede Social Sonora Em-Linha, desenvolvida de raiz para dispositivos móveis e acrescida de mecanismos sensoriais para estímulos ambientais, capazes de analisar a Paisagem-Sonora envolvente e os movimentos do utilizador. Dois objectivos principais guiaram a produção desta ferramenta de investigação: recolher dados relativos à atividade cinética e sonora dos utilizadores e proporcionar a estes uma experiência real de utilização uma Rede Social Sonora, de modo a recolher opiniões fundamentadas sobre esta tipologia específica de socialização. A aplicação – Hurly-Burly – analisa a Paisagem-Sonora através de algoritmos de *Audição Computacional*, classificando-a de acordo com quatro categorias: diálogo (voz), música, sons ambientais ("ruídos") e silêncio. Adicionalmente, determina o seu nível de pressão sonora em $dB(A)_{eq}$. Esta informação é então distribuída pela rede social dos utilizadores, permitindo a cada elemento visualizar e ouvir uma representação do som analisado. É mantido num servidor Web um registo individual da informação sonora e cinética captada, o qual pode ser acedido através de uma aplicação Web que mostra o perfil sonoro de cada utilizador ao longo do tempo, numa visualização ao estilo *linha-temporal.* O procedimento experimental incluiu três grupos de teste distintos, formando cada um a sua própria rede social com coeficiente de aglomeração igual a um. Após a implementação da experiência e análise de resultados, concluímos que a Paisagem-Sonora pode desempenhar um papel no paradigma das Redes Sociais Em-Linha, em particular no que diz respeito à sua presença nos dispositivos móveis. Ficou provado que os dispositivos móveis comerciais da atualidade apresentam-se com uma oportunidade promissora para desempenhar este tipo de tarefas (tais como: *monitorização contínua, registo quotidiano* e *análise sensorial ambiental*), mas as limitações relacionadas com a autonomia energética e funcionamento em *multitarefa* representam ainda um constrangimento que impede a sua massificação. Além disso, a privacidade no mundo virtual é algo que os utilizadores atuais não estão dispostos a abdicar: partilhar continuamente a Paisagem-Sonora real em detrimento de uma representação de alto nível é algo que refrearia os utilizadores de usar a aplicação. Também demonstrámos que os utilizadores que mais conhecedores do fenómeno da Paisagem-Sonora são também os que consideram esta como importante no contexto das Redes Sociais Em-Linha. Isso significa que uma atitude pedagógica em relação ao fenómeno sonoro é essencial para obter dele o maior ganho possível. Esta investigação propõe-se a dar um passo em frente nessa direção.

Acknowledgments

The process of completing a PhD thesis leaves a mark in most people's life. The same applies to me! The time I started the program coincided with a huge challenge in my professional life: the beginning of my teaching carrier at UCP, which lasted approximately until the moment I finished writing the thesis and decided to move to Macau, to continue teaching in that region. Indeed, this period of time left a mark on me, consequence of the interaction with great people and wonderful places I was lucky to meet and visit. These acknowledgments are dedicated to those people I've met during this period and somehow contributed to the completion of this text. For all of you, my warm gratitude!

I start by acknowledging my adviser Professor Álvaro Barbosa for his enviable “look-aheadness” and thank him for the many doors he has opened along my path, and Professor Luís Gustavo Martins, my co-adviser, for being a beacon both in scientific and ethical fields. Carrying on within UCP’s staff, my recognition goes to all senior researchers and Professors, which have transmitted me and my colleagues, during the curricular year, valuable knowledge in the field of Computer Music, in particular Professor Paulo Ferreira-Lopes, the PhD Program Coordinator. To my class year colleagues (André Rangel, José Luís Ferreira, João Carrilho, Miguel Cardoso, Pedro Patrício, Ricardo Guerreiro, Telmo Marques and Vitor Joaquim) I address my profound admiration (I will always remember our Saturday’s lunches!). The other PhD students enrolled in the Doctoral Program of Science and Technology of The Arts I would like to thank the fruitful exchange of information regarding their individual projects, I have learnt a lot from different and very interesting areas thanks to them. In particular, I would like to show my deepest gratitude to those more “physically” close, sharing (or visiting often) the 0.8 room in UCP, André Baltazar, André Perrotta, Bruno Afonso, Diana Cardoso, Francisco Bernardo, Joana Gomes, José Alberto Gomes, Nicolas Makelberge, Samuel Van Ransbeeck, Vasco Carvalho, Vitor Joaquim and Mailis Rodrigues. I couldn't think of a better team to work with. A special thanks goes to "Perrota" for his help with OpenFrameworks, Vasco for printing the thesis and teaming up during the last 5 years, Baltazar for participating in the Murky Shooting project, Bruno for his help with the classification algorithm, Andrian Santos for participating in the MMucS project, Luís Sarmento for setting up the MySQL database, Jorge Cardoso for his help with the web application and "Nics" for the long chats regarding the slippery combination of art and science. Still within UCP people, I address a word of recognition to Professor Sofia Lourenço for her motivational attitude towards PhD students, Professor Carlos Caires for his good work as coordinator of the Digital Arts Group at CITAR, Professor Henrique Pereira for his example of tenacity and Carla Almeida, Helena Figueiredo e Marta Reis for their companionship!

The international acknowledgments go to Professor Roberto Bresin at the Sound and Music Computing Group – KTH, for hosting my first STSM, which would mark the beginning of my research career, ultimately leading to this this dissertation. A word of gratitude goes directly to the people in CCRMA – Stanford University, in particular Professor Chris Chafe, who has kindly accepted me as a visiting researcher in that institution, Professor Jaroslaw Kapuscinski for his support during my stay and Esteban Maestre and Karim Barkati for the great insights on the topic of my work. To the people in C4DM – Queen Mary University of London I address my recognition, in particular Professor Nick Bryan-Kinns who has hosted me in that institution for three months as a visiting researcher and Ben Bengler, Fiore Martin, Matthias Mauch, Nela Brown, Oussama Metatla and Sebastian Ewert, for their friendly welcoming in the office. Additionally, a special word of recognition goes to Professor Atau Tanaka for the valuable inputs on my research topic and for mediating my visit with C4DM.

A word of gratitude goes to all the participants in the experimental procedure, including people in Macau and Portugal. Additionally, I would like to stress the role of Ana Parada and Katerina Markova in the genesis of the

Hurly-Burly application, which the first outline (including the name) was a result of a group brainstorm during a workshop in Oporto. I also would like to thank Nicholas Carvalho for his dedication in the process of proofreading a significant part of this document.

On the personal side, I would like to thank all my family for their unconditional support, in particular my grandparents Gilda and Júlio and my parents Clara and Domingos for their backing in the most difficult situations. Furthermore, I want to give a word of affection to all my friends who have been constantly there for me, in particular those from Évora (the city where I grew up, sorry, fortunately you are too many to write your names!), the guys from my band Houdini Blues (André, Gonçalo, Hugo e Ricardo), the 49on31 "famili", my "sister" Joana and my old buddy "Cid".

Last but not least, I want to thank my friend and life partner Cristina, for her affection, for contributing to the videos and posters of this project with stunning illustrations and mainly for putting up with all my moods arising from the writing of this thesis (what's next?).

Table of Contents

List of Figures

List of Tables

Acronyms

ACM – Association for Computing Machinery

CASA – Computational Auditory Scene Analysis

CACA – Context-Aware Computing Applications

CCS – Computing Conference System

CITAR – Research Center for Science and Technology of the Arts

CMC – Computer Mediated Communication

CSCW – Computer Supported Cooperative Work

DBMS – Database Management System

FtF – Face-to-Face

GeoSNs - Geo-Social Networks

GUI – Graphic User Interface

HB – Hurly-Burly

HCI – Human-Computer Interaction

ICT – Information and Communication Technologies

MIR – Music Information Retrieval

MMOG – Massive Multiplayer Online Games

MOSN – Mobile Online Social Networks

OSN – Online Social Networks

PARC – Palo Alto Research Center Incorporated

SNS – Social Networking Sites/Services

SM – Social Media

UCP – Universidade Católica Portuguesa

UDID – Unique Device Identifier

USJ – University of Saint Joseph

UGC – User Generated Content

WWW – World Wide Web

"Nothing essential happens in the absence of noise."

Jacques Attali

in Noise: The Political Economy of Music

CHAPTER 1 - Introduction

Humans are naturally inclined to exist in society, to work and live together. Famous philosophers and poets have expressed this idea and made it part of the common sense lexicon. While Aristotle said that the “man is a political animal” (Arist. *Pol.* I.2, 1253ª2-5, trans. Barker) - willing to live in the polys, to find structures and hierarchies, to organize in groups and groups of groups -, John Donne, with a more spiritual oriented discourse, conveyed in his poem “No Man Is An Island”, the importance of every subject for the society.

The topic of this dissertation – Sound Based Social Networks – has much to do with this tendency for humans to live in society but also with the way they actually do it, focusing on their communicational abilities, which underlay and catalyze the social dynamics.

In its most basic scenario, the minimum requirement for an act of communication is comprised of a message being transmitted from an emitter to a receiver through a medium. From this basic act many paradigms can be drawn, depending on the perspective one chooses to. McLuhan has coined the phrase “medium is the message” (McLuhan & Lapham, 1994), asserting that the influence of the medium is such that it represents, itself, a more relevant (and potentially unnoticeable) message than the content it carries. If this assertion may seem too radical, one has at least to admit that the medium strongly affects and shapes the way we perceive messages. It is true that one does not get equally thrilled by watching the same movie on either IMAX or on an iPod, or by receiving sad news via SMS or in person. The idiosyncrasies of the mediums, the way they appeal to our senses and the technological apparatus implied, impact on diverse layers of the message being transmitted, thus the way we perceive it.

Another key element in the communicational chain is language, which in a broad sense can be defined as the “system of communication used by a particular country or community”[1]: a set of practices, codes and syntaxes made common within emitter and receiver, enabling the good understating of the idea being conveyed. Language, for its complex nature, can assume a vast panoply of manifestations, ranging from sign language to oral language, passing through artistic language, which in turn can be divided into several artistic practices like cinema, music, painting or sculpture, each one claiming its own semiotics. Like no other animal on earth, humans have raised the bar of language to a great level - a technology with many layers and purposes that keeps evolving day-by-day. As (Postman, 1993) said, language is a “kind of technique, a machine” running invisible and inconspicuous in background, being not “merely a vehicle of thought but also the driver”.

To conclude with this brief wrap-up on communication, we go back to the beginning – the medium - to introduce Sound, the central point in this text. Sound is typically defined as a physical phenomenon based on the vibration of molecules within a fixed range of frequencies. These vibrations propagate, meaning that they travel in space through different materials, in different physical states. Although air is the most common of the conductors, sound also propagates in solid and liquid materials - like steel and water - showing subtle but relevant peculiarities in each one. The ability of travelling is the key element here since it makes sound suitable for the carrying of messages. Oral language and musical language, for example, exist primarily in the sound domain. One may object that music lives in a score or in peoples mind (remember that Beethoven wrote the 9th in deafness), or that oral messages can be transposed to a written form, both cases with total absence of molecular vibration. However, these examples represent ways for codifying and store latent sonic messages, which ultimately can be considered the point of arrival and departure.

---

[1] "language". Oxford Dictionaries. Oxford University Press. http://oxforddictionaries.com/definition/english/language (accessed June 20, 2013).

The range of frequency that defines sound is not arbitrary and coincides with the audible frequency range of human ears, usually fixed between 20Hz and 20kHz. Therefore, it is acceptable and advantageous for the sake of this study, to assume that sound is more a definition of a stimulus for human perception rather then an exotic physical phenomenon. This way, we are defining sound from an anthropocentric perspective, taking into account its importance for human life.

Striping it down to its inner core, the present dissertation binds together two key concepts - *Socialization* and *Acoustic Communication* - in order to propose an innovative framework for social networking based on the communicational properties of sound, both as a medium and as a message.

1.1 - Related Areas and Scope of this Dissertation

This dissertation was written under the broad umbrella of Computer Music – more generally addressed within the framework of Sound and Music Computing (Serra, Leman, & Widmer, 2007) - a cross-disciplinary field that merges two large areas *per se*: sound and technology. Narrowing the subject down to Sound Based Social Networks, we encounter a sociological dimension related with social networks and furthermore with social media.

Therefore, the core of this dissertation comprises three major scientific disciplines within the broader topic of Computer Music:

- Sound and Music Technology
- Social Networks and Social Media
- Mobile Sound and Context-Aware Computing

1.1.1 - Sound and Music Technology

While the term Sound, as we have seen before, can be defined out of its physical properties (as also from an anthropocentric perspective), Music tends to be a more blurred area and open to some haggling when one tries to crystalize it. That is particularly true after the New Music era (roughly speaking, the music of the second half of the XX sec.), where contribution from Luigi Russolo, Milton Babbitt, Pierre Schaeffer, Stockhausen, Xenákis, Boulez, Cage, Varèse, among others, helped to dissipate any univocal definition of Music (for a better understanding of the history of contemporary music see (Taruskin, 2009) and (Chadabe, 1996) for electronic sound). From all the actors who played an active role in contemporary music history, perhaps the most disruptive and prominent name was John Cage[2], which beyond great unsettling pieces like the 44' 33'' or William's Mix, had accompanied is oeuvre with a consistent theoretical corpus, showing a particular interest in silence (or the absence of it) as musical material (see (Cage, 2011)). Moreover, the term "organized sound" is attributed to Edgar Varèse as a definition of Music as, a fairly broad definition that appears to include (by excess) all the music being done nowadays.

Both, Cage and Varèse's approaches are relevant for this study as they near together the two concepts – sound and music – pointing the focus into sound. Cage asserts that there is no such thing as silence and that became the philosophy beneath his piece 4' 33'', where silence does not occur even when all the instruments of the orchestra are muted. The "music" in this case comprises the sound inherent to the audience and the musicians (breathing, movements) and other environmental sounds. This powerful idea - that sound is a constant presence - underlies our research project and sets the fundamental pillars for using sound as a valid element in the social network paradigm. In turn, Varèse's definition of music as organized sound opens the door to other musical paradigms, away from the traditional trinity: harmony, timbre and rhythm. Two examples would be music concrete, introduced by Pierre Schaeffer, which bases its musical raw material in sound ob-

---

[2] Interesting to note that in a recent article published online by the NPR - *33 Musicians on What John Cage Communicates* - Yoko Ono refers that "History of Western music can be divided into B.C. (Before Cage) and A.C. (After Cage)". http://www.npr.org/2012/08/30/160327305/33-musicians-on-what-john-cage-communicates.

jects without representative musical characteristics (i.e. note/frequency) and soundscape composition, which represents a compositional approach base on environmental sounds, where the "*essence is the artistic, sonic transmission of meanings about place, time, environment and listening perception*". (Westerkamp, 2002)

Regarding the sound/music research field in Computer Music, the core of this dissertation is biased towards the sound rather than music, specifically to the sonic events that occur in a particular place and time. This approach is tied to the definition of Soundscape, which paves the theoretical ground for this research and is described later in sub-chapter 2.3 - Mapping the Soundscape (pp. 23). The role for Soundscapes in this setting is to form part of a context-aware system, able to provide useful cues about the activity of a place in a given time. Information collected this way by a user of such system can be seen as a meta-information element of his/her life experiences.

Nevertheless, a more musical approach is also discussed with regards to Algorithmic Music Composition and Sonification processes, later on sub-chapter 3.3.1 - Sonification (pp. 60). This is important as our framework for a Sound Based Social Network includes an audio feedback module, which takes this compositional method as core.

1.1.2 - Social Networks and Social Media

The sociological aspect of the research is related with social networks, in particular with social media. By studying and extracting the fundamental aspects of what a social network is, as well as the basis of social media, we propose to add a new layer of information to those usually exchanged within user in the cloud (interests, media content, emotional status, location, age, etc.), in order to catalyze further relevant social interactions. The sociological dimension in this research does not occupy the central position; rather it is considered an application target for the research being done in the sound domain. Nevertheless, a good understanding of its principles is essential and is discussed further ahead in Chapter 4 - Being Social in the Information Era (pp. 69).

1.1.3 - Mobile Sound, Context A, Ubi Comp, Mobile Comp.

Mobile Computing is a field of Human-Computer Interaction characterized by providing computing processes in portable devices (Imielinski, 1996). Such technology became possible thanks to developments in hardware, which allowed more processing power in smaller devices; in parallel with the production of improved long lasting batteries and wireless network capabilities. The expansion of this area demanded new software and interaction paradigms and opened the doors to new opportunities in HCI. Smartphones are a good example of such evolution, with an incredible market penetration over the last decade. A comparative study revealed that in the United States of America, from 2007 to 2012, smartphone penetration increased from approximately 9 million users to 110 million Americans[3].

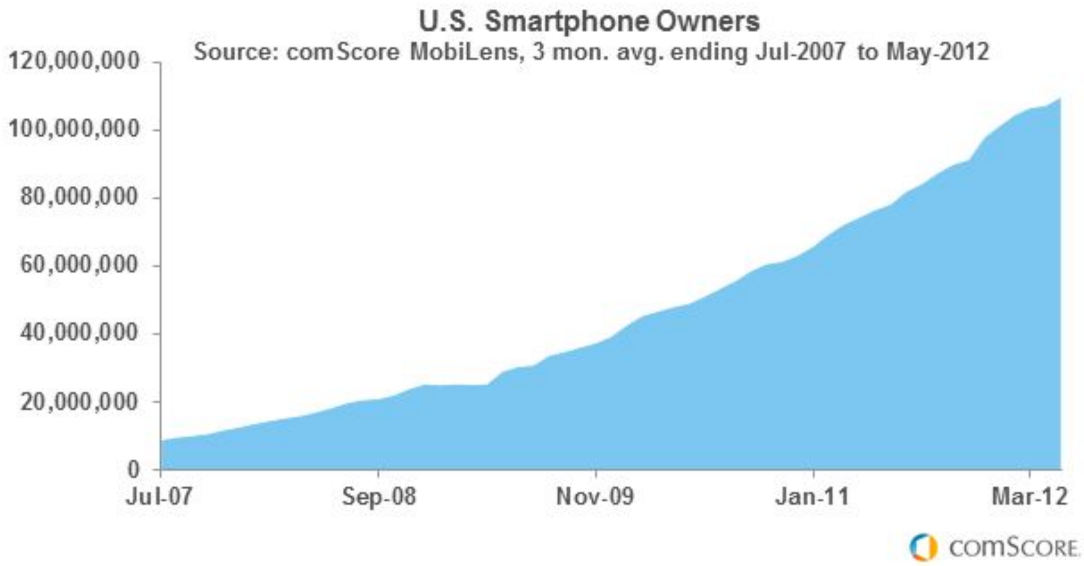


Figure 1 - U.S. Smartphone Owners from 2007 to 2012 (source_ ComScore)

[3] According to a study conducted by ComScore in 2012 http://www.comscoredatamine.com/2012/07/getting-smart-u-s-smartphone-population-reaches-110-million-consumers/ [accessed June 14, 2013].

The possibilities opened by Mobile Computing commonly overlap and merge with Ubiquitous Computing technologies, which are characterized by computational processing integrated (in the most invisible way possible) in everyday objects, with human-computer interaction distant from inline commands or menu driven GUI (Weiser, 1993). Smartphones are one example where this overlap may occur, though not necessarily, since most of its computational tasks are menu driven, accessed according to more traditional desktop paradigm. Some ubiquitous computing behavior on a smartphone would be, for example, an application running on background that sends a message to the user every time he/she passes by a determined location. Such operations, also baptized calm computing (Weiser & Brown, 1996a), everyware computing (Greenfield, 2006), pervasive computing, sentient computing (Addlesee et al., 2001) or disappearing/invisible computing (Norman, 1999) are important for our study as they commonly represent the base for Context-Aware Computing Applications (CACA), which characterize the modus operandi of our system, described in detail in Chapter 8 -Hurly-Burly: Prototype System for Experimental Research (pp.151)

Context-Aware Computing Applications, close related with Context Sensing (Ferscha, Vogl, & Beer, 2002) and Context Enabled Applications (Salber, Dey, & Abowd, 1999), represent an area of HCI characterized by systems that “examine and react to an individual’s changing context” (Bill Schilit, Adams, & Want, 1994). The same authors also underline that “This form of computing is broader than mobile computing because it concerns mobile people, not just mobile computers”(Bill Schilit et al., 1994). The distinction makes clear that CACA focuses on the user experience design rather then the mere hardware portability, although only a tenuous line keeps both apart. In the same manner as Ubiquitous Computing, also CACA can be part of non mobile devices - like public displays for example - however, CACA tends to integrate mobile computing systems, as they allow computational processing in different contexts, making them suitable for this HCI paradigm.

The last topic being introduced in this brief overview of technological areas adjacent to our research is Mobile Sound, which is in fact a designation not commonly found in literature. An example of such a paradigm would be Scribe4Me, a mobile sound translator for the deaf which analyses samples of environmental audio and sends a text description for the user (Matthews, Carter, Pai, Fong, & Mankoff, 2006). Although not common in literature, Mobile Sound can be used to address mobile electronic devices able to produce, record and/or process sound (Behrendt, 2010). As expected, most of the audio tasks happening in mobile computing devices are music related: play, record, analyze and retrieve, store, mix, etc. In fact, Mobile Music is a much more common name in modern literature and relates to many different projects, ranging from music production applications like ZooZbeat (Weinberg, Godfrey, & Beck, 2010), musical instruments like the Ocarina (Wang, 2009) or intelligent music players like Zagora (Camurri et al., 2010).

Before we move on, it is relevant to stress the distinction made so far between Mobile Computing (which focus on the portability aspect), Ubiquitous Computing (computing integrated into everyday objects), Context Aware Computing Applications (focused on the software ability to interpret and respond to the context), and Mobile Sound (sound related tasks happening on a mobile computing device). These topics will be discussed in more depth on Chapter 3 -Mobile Technology, Soundscapes and Context (pp. 39).

1.2 - Motivations - research question - objectives – statement

In the media domain, sound tends to be regarded as the poor relative, the least considered on the media production pipeline. As sound designers we have learned this lesson the hard way. For example, during the production process of a movie, sound is usually left for the last stage of production. Although not a problem *per se*, the issues tend to arise when there is also no involvement of the sound crew during the concept and pre-production stages. By the end of the production pipeline, sound designers are mostly called upon to solve problems, rather than to give creative input, and to include sound metaphors into the narrative. Although this scenario has gradually changed towards a more participative role of the sound designer in the whole production process, many people still think that sound is not as fundamental as other elements for the quality of the message being transmitted, such as photography, acting, costume design, etc. Those who think this way are usually unaware of how a movie is made and what is involved in the sound track making of a movie. Moreover, entering in the psychoacoustics domain, there are also quite relevant aspects of how sound can affect us and influence the way we perceive everyday events, which are beyond the knowledge of non-experts.

The underlying aims of this research can be regarded as a quest to reveal sound's importance in everyday life, a study to bring people and sound together. Such general aim is also present in other projects running parallel to this research. For example, in Murky Shooting (Cordeiro, 2011) we have produced a sound-based game where the player has to hit the target relying on audible cues only, shifting the paradigm from visual-centered towards audio-centered games.

Despite this primary goal, all our research grounds in scientific premises and attempt to answer more specific problems. In the present research, the main goal is put as: define and access the potential of environmental sound as a catalyzer of interpersonal relationships. With the research question underlying this aim being: Can environmental sound play a role in social media as a catalyst for interpersonal interaction?

The objective of the research is then to ascertain the correlation between environmental sound (specifically in soundscapes) and social interrelations. The unusual and unpredictable nature of the topic brings this project into the realm of exploratory scientific research.

After a systematic analysis of the problem, several sub-questions have arisen and are listed below, according to the scientific areas of our theoretical framework.

1) *Soundscapes and Acoustic communication*

- Is there a match between environmental sound and people's activity?
- How can we keep track of our personal soundscape?
- How much is the average person aware of its sound environment?
- Do people valorize the potential of sound for HCI?
- What is environmental sound ontological potential?
- (How) Can environmental sound be used as a meaningful feature for social networking?

2) *Social Networks and Social Media*

- What lies in the base of a social network?
- What are the main characteristics of a social network operating in the cyberspace (online social network)?
- How do people integrate sound in their social media interactions?

3) *Mobile Computing*

- Is it possible do design and implement audio monitoring tools in everyday mobile computing devices?
- What are the intrinsic limitations of this computing paradigm?

4) *Auditory Display (sonification)*

- Does it make sense to apply sonification to sound-based datasets?
- What mapping/approach to follow?

5) *Data visualization*

- What approach to follow, regarding data visualization of sound events?

The list above was the starting point to trace the methodological plan for our research, which is summarized in sub-chapter 1.3 - Methodology Overview (pp. 11).

1.3 - Methodology Overview

In this sub-chapter we present an overview of our research design, listing and describing the most relevant steps that led to the conclusion of this project.

1) Formulation of the research problem

Regarding the motivations described in sub-chapter 1.2 - Motivations - research question - objectives – statement (pp. 8), we have chosen the topic Sound Based Social Networks. This topic was first hypothesized during a workshop on Design for Mobile Devices, promoted by Future Places Festival in the year 2010, in a workgroup composed by Ana Parada, João Cordeiro and Katerina Markova. Throughout the first year of research the topic experienced its maturation process, being iteratively discussed with colleagues, advisor and finally submitted as central topic to the scientific council.

2) Extensive literature review

The first two years were dedicated to literature review and state-of-the-art research, which encompassed the main areas of Social Networks and Social Media, Acoustic Communication a Soundscapes, Mobile Computing and Context-Aware Applications.

3) Preliminary Research Experiments

In order to ground our main research in scientific evidences, we have developed three preliminary experiments that assessed the potential of sound in several domains and showed how often people overlook the sonic phenomenon. Moreover, these experiments provided us with some technological knowledge in Sonic Interaction Design, Mobile Software Development and Mobile Sound.

4) Objectives formulation

After a good understanding of the theoretical background, we outlined the conceptual framework and defined the research methodology for data collection according to the main goal and sub-objectives.

5) Experimental research project (proof-of-concept, sample and data collection)

The chosen approach was based on empirical research and consisted in the development of a proof-of-concept mobile application to assess the validity of the initial hypothesis – *that environmental sound is able to catalyze social relations*. The application was then distributed over a group of individuals who formed the research sample, representative of a larger population of social media users. The research data was collected from the application records (history of environmental sound descriptions of all the users, along the timespan of the experience) combined with individual closed-end surveys and focus group interviews.

6) Processing and analyses

The data collected by the methods described above – a mix of qualitative and quantitative data – was analyzed using a computer: edited (detect and errors removed), tabled (distributed in tables) and statistically analyzed (correlations, averages, etc.).

7) Discussion of the results

The current thesis reports the process and represents the outcome of the whole research, including a critical analysis of the major findings.

1.4 - Main challenges

It is common to say that the best way to deal with problems is to regard them as opportunities and that might well be the case in our project. Along the research process we faced challenges that iteratively made us evaluate and rethink the research design process. While most of the challenges were expected (thus more easily overcome), others appeared suddenly and demanded an extra amount of attention. In either case, we can divide the problems into two categories: 1) those related with technological issues and 2) the ones regarding research method issues.

On the technological side we had to face the overwhelming evolution of mobile technology, which makes every move quickly outdated. Besides that, mobile computing has always struggled against energy issues. Adapting a monitoring tool to fit in a personal mobile device dramatically decreases the battery charge cycle. These and other limitation are discussed ahead.

Considering the research methodology, we had to deal with privacy issues regarding the use of our proof-of-concept application: the users had to expose themselves in a way they weren't familiar with, by sharing their environmental sound information. An extra effort was put into explaining to the participants what the study was about and how the collected information would be managed. The way younger generations which grown immersed in digital technology – usually named *digital natives* after (Prensky, 2001) – related with privacy is more open than their previous counterparts, mostly due to the possibilities brought by new technology (mobile in particular), which is increasingly being tailored to social media, allowing the broadcasting of private information in large scale. However, they understand now that sharing everything is not a risk-free activity. An additional issue, this time related with the theoretical background, was the sparse systematical literature on the topic *sound based social networks*, which does not represent a research field *per se*. To finish, we would also point out that the effort to bring sound into the spotlight is, somehow, an unnatural procedure on an image-centered society and evoked some perplexity on the participants, which afterwards gained a new understating and further acceptance of the medium.

Nevertheless, all the mentioned difficulties hold a latent potential, as they deal with present technologies and regard the current (and fashionable), social media phenomena. The limitations might actually become opportunities for this study to prove its scientific validity and set the basis for future projects.

1.5 - Applications of this study

The study of sound in the establishment of social links it is not a new endeavor. Telephones and radio broadcasting have been doing that for more than a century. However, what we propose is a different approach to sound communicational properties, regarding the *causal listening* rather than *semantic listening* (Chion, 1994) as discussed in Chapter 2 -Sound as a Medium, Sound as a Message (pp. 17). Such proposal unveils some innovative applications for the use of sound in everyday ICT. Being proved that environmental sound can convey useful information about context and peoples activity (with low demanding computational processing or expertise from the user's side), there are plenty of everyday situations that could benefit from the application of our research outcomes. Next we provide a list of some hypothetic applications[4].

- **The Intercom System Baby Monitor** – Such devices are a paradigmatic technology, since they already make use of environmental sound in order to establish an elementary social bond between parents and children babies. Typically, these devices are not computer based and do not interpret the sound events they capture (that analysis is left to the caretaker). A computational analysis of the sound could be used to detect different events (someone entering the room, the wakening, slight movements) and send alert messages accordingly (sms, tv pop-up messages, alert sounds).
- **The Hub Monitor** – A similar system to the Intercom System Baby Monitor could be used in office buildings, to assess the usage of communal social areas like the hub or canteen. Using a system that interprets the sound

[4] From this list we exclude our proof-of-concept application explained in detail in Chapter 8 -Hurly-Burly: Prototype System for Experimental Research (pp.136).

source instead of broadcasting the actual sound can have advantages in protecting users privacy, yet giving valuable information on the place's occupancy.

- **The Club Monitor** – Entering the MIR world, nightclub costumers could benefit from a real-time feed of music metadata (title of the song, artist, genre) playing at the moment in each club of the city. This information could be effortlessly collected and broadcasted collaboratively by the costumers of each place, using their smartphones. If gathered from different clubs in the area, this could work as a recommendation service based on the music being played.

1.6 - Structure of the thesis

In this sub-chapter we present the structure of the thesis, along with a brief description of each chapter, providing a map for the reader.

The thesis is divided in two major conceptual sections: contextualization and experimental process.

1) Contextualization

In this section we present the results of systematic literature review on the thesis topic, divided in three chapters:

- *Chapter 2 -* Sound as a Medium, Sound as a Message covers the literature review and contextualization of the research in its relation with sound.
- Chapter 3 - Mobile Technology, Soundscapes and Context introduces the topic of mobile devices as casual sensors for monitoring environment and human activity.
- Chapter 4 - Being Social in the Information Era covers the systematic study of social media applications.
- Chapter 5 - State-Of-The-Art On Sound-Based Social Interaction Systems presents a survey on the state-of-the-art technologies regarded as a references to our study.

2) Experimental Process

In this section the detailed description of the experimental process and development of the proof-of-concept/research tool is presented, along with an analysis and discussion on the collected data and further findings.

- *Chapter 6* - Preliminary Research On Soundscapes And Sonic Interaction is a chapter dedicated to present two research projects developed in parallel, by the author, during the course of this doctoral study. The projects are not directly related with the main topic of the dissertation but provided important insights on important areas such as Mobile Computing and Soundscapes;

- Chapter 7 - Research Methodology describes the methodology followed in the course of this interdisciplinary research. Several approaches, from divers scientific fields, contributed with methods and techniques for our study.

- Chapter 8 - Hurly-Burly: Prototype System for Experimental Research is dedicated to present all the details regarding the conceptualization, development and implementation of the proposed research tool / proof-of-concept.

- Chapter 9 - System Evaluation and Result Analysis is dedicated to evaluate the results of the questionnaires made to the subjects of the experience.

- Chapter 10 - Conclusions. Chapter dedicated to present the main contributions of this research project and future work to be done in this field.

CHAPTER 2 - Sound as a Medium, Sound as a Message

The current chapter is dedicated to expatiate on the general topic of sound and soundscapes in particular, which have been widely discussed in literature, tackled from different perspectives and aimed at different goals. Our approach in this chapter aims to provide a good overall understanding of these scientific areas, favoring the aspects that related, more or less directly, to our research agenda. Thus, the title of the chapter encloses an abstraction regarding the role of sound in our study and should be understood in the following way: sound is usually used as medium to convey semantic messages (speech, alarms, music, etc.) but sonic events that were not originally conceived as messages (the sound of a car, the sound of wind, etc.) hold a communicative potential which we propose to explore.

2.1 - Sound: physics and meanings

Foremost, sound is regarded as a physical phenomenon, which is characterized by the vibration of molecules in a given medium in a way that is perceived by the human auditory system. Some of the early systematizations on the sound phenomenon were related to the musical practice and date from the 6th century BC, when Greek philosopher Pythagoras undertook a study on musical intervals and their mathematical relations. Aristotle also dedicated thoughts on the topic in his writings compiled under the name Poetics (Aristotle, 1987), where he talks about the sound differences between consonants in speech, grouping them according to their sonic characteristics, and in the work De Audibilibus, now usually ascribed to Strato (Aristotle, 1881), in which a formulation of a wave theory is mooted. Later, during the 1st century BC, sound was also debated from the architectural perspective, on the work of Vitruvius – On Architecture (Vitruvius, 1931) – were concerns about room acoustics are expressed and the wave theory is suggested (Kilgour, 1963).

However, it was only in the 19th century that major systematizations in the field happened, mostly by the hand of English scientist and Nobel Prize in Physics Lord Rayleigh with his two volumes book The Theory of Sound (Rayleigh, 1877, 1878) and with German scientist Hermann von Helmholtz (Helmholtz, 1913) on the field of psychoacoustics.

This superficial historical review shows that the field of sound is twofold: on one side lie the sound sources, on the other how it is perceived; or as (Olson, 1967) put it, two major scientific fields emerge in sound, the physical acoustics and physiological acoustics. The first one deals with the physical properties of the vibration while the second deals with the physiological operation of the auditory system, responsible for transducing physical movement into electrical stimuli for the brain. Nevertheless, the aural experience would not be complete without the understanding of how stimuli are perceived and organized by the brain. This complex phenomenon is mainly studied under psychoacoustics, yet extending to other scientific areas such as anthropology, musicology, psychology and cognition (Erlmann, 2004). Today it is clear for the scientific community that the meaning we attribute to sound events differ from listener to listener, as a consequence of the cognitive dimension implicit on the act of listing. For example, a musical piece will resonate differently on distinctive listeners according to their musical background and affective relation with that particular piece or musical genre, in the same way that the sound of a forest will give much more information to someone used to living in the countryside than to someone from the city, generally speaking (McAdams & Bigand, 1993).

2.2 - Sound awareness

Sound is therefore, a complex phenomenon and it represents a daunting task when one tries to cover it in all its dimensions. But despite this complexity, the auditory experience happens effortlessly and most of the time unnoticed. Perhaps because of this, sound and our auditory system tend to be less considerate when compared to image and vision. While sounds can be considered indexical from a semiotic perspective, as their meanings are quickly understood by most humans because they are largely given by their obvious physical origin (Saldanha, 2009), the average person, normally operating on a casual listening mode (Chion, 1994), has an overall low awareness of his or her sonic milieu. It often takes a disruptive sound to perceive and become conscious of changes in a soundscape. These disruptive sounds frequently owe their conspicuousness to the lack of reference to any visual cues which can be true both with sounds operating clearly out of context and/or because a sound source is out of one's sight. Other forms of disruptive sounds include those that listeners dislike but are forced to live with, while not being completely unaware of its source (ex. a neighbor's dog barking outside your window at 6 o'clock in the morning). This lack of awareness towards sound is also noticeable when people are confronted with the production process of a sound for a movie. From our experience delivering sound design courses we conclude that people in general do not recognize neither the techniques involved in the production pipeline nor the extension of the impact that sound has in the dramatic load of a movie. Besides, when someone rarely mentions the extraordinary quality of a movie soundtrack is usually referring to the music of the film and not to the whole sound track, which is comprised of a much richer soundscape, with subtle nuances arising from the fine mixing of human voice, music, sound effects and ambient sounds.

Sound as a sensorial stimulus has peculiarities that make it unique and highly apposite for accomplishing daily life tasks. To begin with, audition is the vigilant sense, it never shuts off even when asleep. It was due to this feature that humans (and of course other animals) were able to survive in the wild, among all sorts of predators who would attack at night or outside of the viewing angle. Sound is three-dimensional, our auditory system is accurate enough to identify the source locus wherever it presents itself to the listener, with a moderate increase in precision when the listener is faced directly toward the source. The process is not analogous to localizing a visual source on the fovea of the retina (Middlebrooks & Green, 1991).

> *"Hearing, together with its active complement, listening, is a means by which we sense the events of life, aurally visualize spatial geometry, propagate cultural symbols, stimulate emotions, communicate aural information, experience the movement of time, build social relationships, and retain a memory of experiences."* (Blesser & Salter, 2007)

These aspects in the nature of sound had impact not only in the way humans survived in the wild but also on how recent technologies involved. Two brief examples of this fact would be the redefinition of radiobroadcast after the arriving of television and the new trends in sonic interaction design.

For more than three decades, radiobroadcasting was the preeminent medium in households in western civilization. Families would gather around the radio device to listen to the several radio programs whether they were informative, entertainment or educational. Radio, by its technical characteristics – one antenna emitting spherically-shaped electromagnetic waves into the air (hence, the name radio/radius) - helped creating a sense of community, as those near the antenna would receive the signal while those further away would not. This resembles the church bell effect, which used to delimit the parish territory in countryside villages and even in more urban territories, creating what Schafer called and Acoustic Space that potentially embraces an Acoustic Community (Schafer, 1994) (this topic is of most importance to our thesis, hence we will come back to it later in sub-chapter 2.4.1 - The acoustic community (pp. 31)). Most would agree that the magic of radio was the magical world of the ear stated by McLuhan when referring to the early oral tradition of the tale-tellers (McLuhan, 1962). Mainly, because radio is a medium that demands an active roll from the listener, who has to paint the narrative with his or her own mental images. With the advent of television in the mid-fifties, radio would rapidly lose its place in the limelight, as the spectators where starving for a more "multi-modal" experience in the media. Meanwhile, radio would not disappear for good, as it still represents a ubiquitous presence nowadays. What happened was that radio had to find in the unique characteristics of sound a means for its survival. That twist consisted in moving from the household to the street and mainly into the car. A technological innovation by Bell Laboratories scientists in 1948 resulted in the invention of the transistor. This novelty offered radio manufacturers the chance to produce miniature portable receivers, being the first electric mobile sound device. The proliferation of radio devices inside cars makes total sense as when driving all the interactions occurring between the driver and the environment (except the road) should be, as much as possible, eyes-free. This way, radio revealed one of its superior virtues: being able to coexist with other human activities without representing an obstacle.

The other relevant example taking advantage of the highly regarded sound characteristics is the recent Sonic Interaction Design research field. With pervasive computing taking over our daily life, adding computational power and connectivity in situations that would be impossible just a few years ago, it is expected that new ways of interaction with computers develop beyond the traditional mouse and keyboard. Here sound can play a significant role in multi-modal interactive systems, as it is suitable for interactions without demanding tangible or visual contact with the user. Sonic Interaction Design is represented in three major fields: Auditory Display (giving audible representation to information, events and processes), Interactive Sonification (sonification of real-time processes based in sound models) and Soundscape Design (design the sound for a place, either real or virtual). Regarding Sonic Interaction Design, the Road Map for Sound and Music Computing says:

> *"Sound-based interactive systems can be considered from several points of view and several perspectives: content creators, producers, providers and consumers of various kinds, all in a variety of contexts. Sound is becoming more and more important in interaction design, in multimodal interactive systems, in novel multi- media technologies which allow broad, scalable and customized delivery and consumption of active content. In these scenarios, some relevant trends are emerging that are*

likely to have a deep impact on sound related scientific and technological research in the coming years. Thanks to research in Auditory Display, Interactive Sonification and Soundscape Design, sound is becoming an increasingly important part of Interaction Design and Human-Computer Interaction." (Serra et al., 2007)

Despite the potential that sound holds in different scenarios of our life, some authors state that we live in imposing image-preponderant society, where old listening habits associated with rural lifestyle and pre-industrial era are disappearing.

In 1962, Marshal McLuhan writes an enlightened text about the prevalence of the visual realm over the aural, as a result of the invention of typography. For him it was clear that "*the interiorization of the technology of the phonetic alphabet would translate man from the magical world of the ear to the neutral visual world"* (McLuhan, 1962). This statement (somehow embedded in prophetic vision) would be amplified by developments in the printing technologies, which made images ubiquitous in our society.

Jaques Attali in his book - Noise: The political economy of music – also shares this concern and goes even further, saying that:

"For twenty-five centuries, Western knowledge has tried to look upon the world. It has failed to understand that the world is not for the beholding. It is for hearing. It is not legible, but audible." (Attali, 1985)

For Attali, noise, and music in particular (which he considers to be an organization of noise), are the most accurate gauge and premonition factor for socioeconomic change.

"We can, for example, toy with the idea that it is not by chance that the half-tone found acceptance during the Renaissance, at

precisely the same time the merchant class was expanding; that it is not by coincidence that Russolo wrote his Arte Dei Rumori ("The Art of Noise") in 1913; that noise entered music and industry entered painting just before the outbursts and wars of the twentieth century, before the rise of social noise." (Attali, 1985).

Both Attali and McLuhan share a concern about the lake of sound awareness in current society. While McLuhan is concerned in describing the source of the problem Attali is more attentive as to how sound can help us to understand society's status and directions. Yet, it would be Murray Schafer, to pay a great attention on the history of sounds and its relation with society. Like Attali he agrees that

(…) music is an indicator of the age, revealing, for those who know how to read its symptomatic messages, a means of fixing social and even political events". (Schafer, 1994)

2.3 - Mapping the Soundscape

The Soundscape term was first coined by Murray Schafer to describe the sound of a particular location in a particular time, analogously to what happens in the visual realm with the term landscape (Schafer, 1977).

Schafer was the promoter of a new scientific subject called Acoustic Ecology, which debates the relation between sound and environment. From his publications, The Soundscape: Our Sonic Environment and the Tuning of the World (first published in 1977 under the name The Tuning of the World) (Schafer, 1994) is probably the most enlightening as it compiles and systematizes material from previous research in the filed, accomplished in the World Soundscape Project and published as The New Soundscape and The Book of Noise. Drawing from his background in music composition, Murray Schafer assumes the global soundscape as a musical piece, where all humans are regarded as musicians and potential composers. He aspires to a positive balance between human-produced soundscapes and the natural soundscapes of earth (Schafer, 1994). This balance is, by necessity, anthropocentric as it regards man as the disturbance and victim at the same time, and demands from individuals an increase on their level of awareness of the acoustic environment at any given time, that is to say, their Sonological Competence (Laske, 1974).

In order to establish a theoretical background for this topic, Schafer deconstructed the sound events existing in our daily life (present and past) by analyzing its sources, purposes, consequences and morphologies. Also, as mentioned before, he overlaid his musical background to this study, aiming to find a musical intonation for the global soundscape. From this research he drew a set of key concepts that help assess the current cacophonies on big urban centers but also the intricate tapestry of rural environments; along with a set of pedagogical technics like the Ear Cleaning exercises.

As a way to characterize a sonic environment he established three great categories for sounds according to their relation with the environment:

> *Keynote Sounds are those sounds that are heard by a particular society continuously or frequently enough to form a background against which other sounds are perceived.*
>
> *Sound Mark is derived from landmark to refer to a community sound, which is unique and possesses qualities that make it specially regarded or noticed by the people in that community.*
>
> *Sound Signal is any sound to which the attention is particularly directed. In soundscape studies sound signals are contrasted by Keynotes Sounds, in much the same way as figure and ground are contrasted in visual perception* (Schafer, 1994).

While these categories are important for study and annotation purposes from an ethnographical and ecological point of view, other approaches may also be effective in describing and assessing a soundscape in the context of sound design for cinema or design of auditory displays.

Michel Chion - by this time a professor at University of Paris II: Sorbonne Nouvelle - is a composer of experimental music and a theoretician on the audio-visual relationships, responsible for some of the most regarded publications in this field. For analyzing the soundtrack of a movie – i.e. the soundscape of the movie – he defines three major groups:

<u>Music</u> - all the music present in the movie, diegetic and non-diegetic, typically used to set the tension for a scene;

<u>Ambient Sound</u> – all sounds that by its low loudness remain in the background, and which describe the scene context;

<u>Dialog</u> – speech present in the movie that is meant to transmit a semantic message (excluding voice in music and *walla*) in the foreground;

<u>Sound Effects</u> – sounds that are added over the remaining sounds, used to highlight natural events or invoke unnatural characteristics (Chion, 1994).

Although such methodology has been draw from the observation of the production pipeline of sound films (not the other way around), this systematization provides an interesting framework for the soundscape analysis, as it is more focused on the source classification of the sound events (casual listening) and their physical properties (reduced listening) rather than on the semantic meaning on a given context (Chion, 1994)

Depending on the purpose of the project – if it's the sound design production for a movie, designing an artistic installation or an urban soundscape – the categories might have so be set differently. For example, in a research project aiming for the observation of the workspace in order to suggest new auditory interfaces, authors have characterized the soundscape along three dimensions:

Sound type - speech, music, non-speech/everyday, non-speech/abstract;

Acoustical information level - background, foreground and contextual;

Information category - visible/hidden/imagined entities/events, passing of time, position in space, patterns in entities/events and emotions (Macaulay & Crerar, 1998).

In our study, we will assume two major categories for the soundscape classification/analysis: the sound type and sound intensity. These categories and how they will be assessed are discussed later in Chapter 8 - Hurly-Burly: Prototype System for Experimental Research (pp.151).

2.3.1 - Soundscape and Environmental Noise Control

With the industrial revolution taking place during the 17th century, the city soundscape dramatically changed. New mechanical sounds were injected into the acoustic environment, increasing the panoply up to an overpopulation of sounds, which would end in a lo-fi soundscape. For Schafer the lo-fi soundscape surges as an opposition to the hi-fi concept, which defines a system with a low signal-to-noise ratio. A lo-fi soundscape is that in which sounds carrying meaning (signals) can't be distinguish from the blur amalgam of background sounds (noise). The result is a much-constrained acoustic horizon (confined to a few meters radius), where people can't listen to sound events occurring far from their location nor subtle and less conspicuous sounds close to them. This phenomenon is usually found in big cities and is considered potentially harmful (De Vos & Van Beek, 2011), able to induce sleep disorder, provoke annoyance, stress and even be the cause of minor psychiatric disorders (Stansfeld & Clark, 2011). Advances in society from demographic, geographic and technological points of view, are not always accompanied by an environmental consciousness, *especially in emergent under-developed countries.*

It has been proven that disturbances in the soundscape are prone to provoke physiological and psychological diseases in humans and animals, by prolonged exposure to it (Chuengsatiansup, 1999). Studies assessing the quality of the soundscape of a given place are important to ensure that the soundscape is in consonance with the place, which embodied it. In urban environments, mainly in big large-sized cities, it is common to observe that the sonic expectations for a given place are defrauded by overwhelming soundscape of automobile flux (Szeremeta & Zannin, 2009). Humans are usually the source and victims of the pollution problem, being harmed by the lack of environment-friendly urban planning and unawareness of their own ecological footprint. A study conducted at Muir Woods National Monument, California, supports this idea: visitors were surveyed regarding which sounds they considered more annoying, and the results concluded that these were the utterances from visitors themselves (Pilcher, Newman, & Manning, 2009).

The scientific community has long warned for the problem of the quality of soundscapes and tried to find ways to mitigate noise problems (Birgitta Berglund & Nilsson, 2006). A closer look into noise abatement laws is a good indicator on how sound has been dealt with along time in different states, unveiling the acoustic phobias of each era and each society. In order to deal with this problematic it is important to tackle it since the early beginning, that is, from the design stage of urban planning. Some complex studies suggest modeling subjective evaluation using artificial neural networks, creating mechanisms to assess the soundscape quality for different scenarios in different contexts (Yu & Kang, 2009).

Our urban spaces are now naturally boisterous, with road traffic noise identified as the worst cause of sound pollution in the city. In some years, with the eventual dissemination of electric cars, the soundscape of the city would change again into a quieter place. While this represents a gain in society's life quality, other problems arise from this new car design: people, even not aware, are used to listening to cars and to avoid collisions by sensing them this way. This road safety issue was already identified and some governments have legislated accordingly, as in the case of the United States of America with the S. 841 - Pedestrian Safety Enhancement Act of 2010:

> *S. 841 would require the Department of Transportation to study and establish a motor vehicle safety standard that provides for a means of alerting blind and other pedestrians of motor vehicle operation for otherwise silent vehicles, such as hybrids.*[5]

Scientific studies have identified the minimal sound level for safety warnings, in order to avoid unnecessary noise pollution (Kerber & Fastl, 2008), and some car brands such as Toyota, Ford or Hyundai have started to include warning sounds in their hybrid and electric cars.

Admitting a problem exists is the first step towards a solution; the next would be studying its nature and finding the right tools to deal with it. For sound pollution it is definitely imperative to increase awareness for the sound phenomenon and a soundscape approach (Kang, 2011) is a promising way to deal with this harm instead of simplistic approaches based exclusively in measuring acoustic pressure, like the Community Noise program (B Berglund, Lindvall, & Schwela, 1999).

2.4 - Acoustic Communication

In this section the focus is pointed towards the communicational properties of sound, both as a medium for transmitting messages as well as sound events as messages themselves. In order to achieve this goal, attention will be given to the Acoustic Communication theory defined by (Truax, 2001) which approaches the soundscape from what it is able to communicate rather than the energy exchanges involved in it. In line with this perspective, other studies regarding the language of environmental sounds (Ballas & Howard, 1987), the sound of social life (Mehl & Pennebaker, 2003), the auditory events

[5] Retrieved online on October 9, 2012 http://www.gop.gov/bill/111/2/s841

perception (W. Gaver, 1993) and the relation between sound and space will be brought into discussion.

Space and Time are two dimensions that are inevitably a part of our life. Everything happens in a given place and in a particular time, and sound holds an interesting relation with both variables since it emanates from physical sources that occupy space and, more importantly, it needs time to exist as a physical phenomenon. That is why music belongs to the category of temporal arts, like dance, cinema or theater, unlike painting or photography. Sound needs time to be, to live and to be processed either by a computer or by the human auditory system. Though studies have proven that less than a second is enough for humans to recognize musical genres (Gjerdingen & Perrott, 2008), the truth is that time is always present in the equation and the longer one listens to it, the more accurate the evaluation. This 'slowness' of sound makes it more suitable for global appreciation of context, while image is better in attending to the details. The reverberation fingerprint of a place takes time to reveal to the listener, in the same way that the recognition of an event by its sonic outcome (the sound event) takes its time. A study conducted by (Warren & Verbrugge, 1984) showed that for the proper distinguishing between a bottle breaking or bouncing "*(…) removal of initial noise from breaking tokens does not reduce perceptual performance. This finding for the natural cases indicates that the burst is not necessary for categorization of the two events*."

Comparative studies between auditory and visual perception proved that the object recognition task is faster using vision than using audition, and a congruent multi-model cue system provides the fastest way for the recognition task. Additionally, the same study reports that biological relevant auditory distractors would confuse the listener and demand more time for the recognition task (Suied & Viaud-Delmon, 2009). From this simple but unerring example, it is fairly reasonable to corroborate that soundscapes have a real impact on the way we perceive the world.

The acoustic communication deals with this problematic as it departs from the semantic properties of sounds, " *(…) It doesn't deal with sound in isolation from the cognitive process that understand it.*" (Truax, 2001). The soundscape in this sense is our understanding (the way we feel and think) of the sounds we are listening to, while the sonic environment is the 'deaf' description of those sounds.

This interdependency of visual and auditory perceptual models is also accounted for in (Chion, 1994) film theory, regarding what he calls the audio-visual contract, that is, the way vision influences the way we listen and vice-versa. Other studies on this interdependency have indicated that people have a tendency to establish such links when sound and image are congruent (Woszczyk, Bech, & Hansen, 1995), although (and not questioning the obvious influence), the observance of a linear correspondence between influence and congruence would not necessarily be true for all combinations of auditory/visual scenes (Viollon, Lavandier, & Drake, 2002) (Anderson, Mulligan, Goodman, & Regen, 1983).

A misplaced soundscape is as annoying as an undesirable speech. If one is visiting a relative seriously ill in the hospital it is not appropriate to talk to them about recent deaths from failed cures. Soundscapes are, therefore, impregnated with meaning(s), and is this meaning that primarily account for the evaluation we do about them.

For many years, the scientific community has paid more attention to the energetic aspect of the soundscape rather than its communicational properties. The increase of noise pollution and the annoyance and potential health problems arising from careless urban planning prompted this route. Although crucial, reducing noise levels is not always feasible and cost-effective and will not necessarily improve quality of life. This approach is what (Schafer, 1994) calls a negative methodology towards acoustic design, as it departs from forbiddance instead of addition in conformity with context and people; because it conceives sound as waste instead of a resource. However, some studies (mainly in recent years) in noise abatement and auditory perception seek to relate other variables like "*regional differences in noise reactions, non-negligible impacts from contextual modifiers, and results that differ qualitatively and quantitatively from those predicted by standard static exposure–annoyance curves*" (Klæboe, 2011). For example, when comparing annoyance motivated by railway noise with traffic and aircraft noise at the same sound levels, the former seems to be more easily tolerable (Fields & Walker, 1982a), and tends to diminish among long-life residents of railway neighboring (Fields & Walker, 1982b). The perception of an environment as stressful also changes accordingly to the degree of control that individuals have upon the pollution source (Guski, 2001) as well as with the geographic situation of subjects, depending if they are used to live in more or less busy environments (Kang, 2011).

2.4.1 - The acoustic community

Sound is not necessarily bad just because it is loud. Church bells are considered loud sounds and nonetheless are usually welcomed by populations, mainly when their meaning is/was essential for the life of the community. Sound is above all a mediator between the listener and the environment, creating relationships, fostering understandings about what the community is, who is involved, how many, how much, how long. Like (Attali, 1985) said, "nothing essential happens in the absence of noise" and "After all, the soundscape is not an alien force but a reflection of ourselves" (Truax, 2001).

Such reflection of ourselves is also the reflection of the space we inhabit, we cross, we shape. The memory we hold from places we visited is also a memory of the sounds we heard, the people we spoke with, the music we listened to. The most significant places that sound can describe are those that no longer exist (like the old house we grew up, or the bar we used to go to), the places that exist only as part of our memories and are evoked by (and live in) the sounds that were part of them.

> *"The sense of place of a location is the collection of experiences and associations the location provokes in its users. As phenomenology comprehends well, in order to experience a place fully, a range of bodily senses come into play, among them hearing."* (Saldanha, 2009)

When a soundscape identity is coherent we may say we are in the presence of an acoustic community. Such a soundscape should be pervasive and play an important role in the life of the community, conveying signals and cues about its rhythms and relevant events.

> *Therefore, the boundary of the community is arbitrary and may be as small as a room of people, a home or building, or as large as an urban community, a broadcast area, or any other system of electroacoustic communication.* (Truax, 2001)

The concept of community is, by definition, a group of living organisms sharing common interests or resources, interacting with each other and usually inhabiting the same location. The acoustic community shares the same sound both as resource and as place. It is the sound that feeds the acoustic community and creates a sense of belonging to a restricted and delineated conceptual space. A radio show creates a sense of community as it talks to a particular target that shares common cultural traces. The sense of physical territory was more evident when no electronic technology was involved in the transmission of sound signals. With Internet streaming and podcasts, the territorial aspect is no longer determinant and the dynamics of the systems behavior change, as the usual space/time limitations are no longer observed. Nevertheless radio shows still create a (displaced) community of listeners by sharing and promoting particular aspects that resonate within the group of listeners (like the musical genre, political interest, religious beliefs, etc.). It is not surprising that community and communication both share the same radical "commun-". Communication is defined as making something common within a group, which, in turn, potentially gives rise to (and underlay) a community.

Computers not only allow for acoustic communities based on TV or radio broadcast. An example of a shared acoustic space in the musical domain can be found in the project Public Sound Object by (Barbosa, 2005) where a virtual environment is set up for remote musical performance. In this case musicians share the same auditory feedback (the musical result) but perform in different locations of the planet.

The impact of new technologies in the soundscape is sometimes devastating, but one has to attempt on the opportunities they can provide in terms of new communicational paradigms or extensions of the existing ones. However, it is important to have in mind that these new acoustic communities are accessed only by means of technology and those who don't have technology may not aim to be part of them.

2.4.2 - The language of the soundscape

Our proposal, although seeking a communitarian approach based on sound, reveals slight differences from the examples discussed above, as it does not share sound among the community but an abstraction of the individual soundscapes of each peer of the network. An analogy for this would be the difference between "abstract music" and *musique concrète*, a term coined by Schaeffer to define his own music, made of assemblage of several recorded sounds:

> *"when I proposed the term 'musique concrète,' I intended … to point out an opposition with the way musical work usually goes. Instead of notating musical ideas on paper with the symbols of solfege and entrusting their realization to well-known instruments, the question was to collect concrete sounds, wherever they came from, and to abstract the musical values they were potentially containing."* (Reydellet, 1996)

Similarly to the opposition between "abstract music" and *musique concrète,* our system captures concrete sounds and classifies them according to conceptual categories (music, speech and environmental noise), creating an abstraction from a concrete example. This mechanism may also be found in the linguistic process defined by (Saussure, 2011), where words/sounds (signifier) are assigned to concepts (signified) - objects, actions, emotions, qualities or entities – thus forming a sign. Our conceptual framework conceives sounds as being the "words" of a purely descriptive language, detached from any previous meaning imposed by a code like oral language, Morse code, music, etc. In other words, a language emerging from a *causal listening*, as defined by (Chion, 1994), or *everyday listening* coined by (W. Gaver, 1993), as a listening mode where source recognition prevails rather then the description of physical/acoustic properties of sound (named *reduced listening* by (Chion, 1994) and *musical listening* by (W. Gaver, 1993)) or the semantic meaning of the pattern (*semantic listening* (Chion, 1994)). For example, the sound events comprised of a female woman shouting "help" and "hello" with an identical intonation, would both hold the same meaning: a woman shouting. This is the language of the soundscape; sound events (signs) have a meaning resulting from its source, and are understood within a certain context.

This is somehow coherent with results attained from (Ballas & Howard, 1987) on their study on perception of environmental sounds and how they can be regarded as language. This study covers only environmental sounds, which they define as sounds that are produced by real events and that have meaning by virtue of casual events. On the relation between language and environmental sounds they say:

> *"(…) language refers to meaning by virtue of social convention. Environmental sound refers to events by virtue of physical law. The connection between an environmental sound and its cause is fixed by the physical laws of acoustics."*

However, in a later study (Ballas & Mullins, 1991), one of the authors contradicts his first judgment and states that:

"Studies have shown that generalizations of context effects found in language research to non speech sound must be made with caution because exact parallels do not exist. In fact, the differences between language and everyday sounds are so great that an argument could be made that comparisons are fundamentally invalid."

This denial of a language of environmental sound is based on several premises (for example, the lack of hierarchical organization found in language perception - phonemes, morphemes, words, phrases, sentences) but focus mainly on the impact of context in the identification of environmental sounds "*which had significant negative effects and only minor positive effects*", unlike in speech.

These findings, although revealing, don't contradict our approach of assuming an underlying language mechanism on the soundscape, in the way that such language acts on a high descriptive level and with very limited scope. Its communicational approach is applied to communication between the environment and humans (or computers) and scarcely between human-human (or computer-computer). Nevertheless, (W. Gaver, 1993) organized everyday sound events under an hierarchical structure in order to expand "*traditional accounts of the primitive physical features of sound"* using an ecological acoustic approach "*which describes the acoustic properties of sounds that convey information about the things we hear"*. Moreover, the studies presented before concern the narrow and delimited human perception field and seem to frivolously mix language and linguistics. A study on the effects of linguistic mediation on the identification of environmental sounds (Dick, Bussiere, & Saygin, 2002) introduces the idea that "*there is little, if any, linguistic mediation in most subjects' processing of this type [environmental sounds] of auditory stimuli."*. The sound of a shot may not be misunderstood or easily confused with other sound whether in the presence of contextual sounds or not, but the meaning we attribute to it (and here we are moving towards a semantic listening) is definitely biased by context: in a western movie it means one of the cowboys won the duel, in an athletics competition we know a race is starting.

2.4.3 - Subliminal power of sound: pros and cons

The communication of emotion is another powerful feature of sound that lies in the foundations of musical expression. But not exclusively. Emotions can be sonically expressed (or present) in non-musical sounds like the knock on a door (Vitale & Bresin, 2008) and especially in the way we speak. The role that paralanguage plays is essential in the way the message is conveyed “In a conversation, non-verbal communication carries an important information like intention of the speaker. In addition to the message conveyed through text, the manner in which the words are spoken conveys essential non-linguistic information. The same textual message would be conveyed with different semantics (meaning) by incorporating appropriate emotions.” (Koolagudi & Rao, 2012). Sound is known by the Auditory Display community as a modest indicator of absolute values (unless semantic codes like verbal speech or Morse codes are used), but an excellent vehicle to express relative values and unveiling tendencies (for example in large datasets). This is actually a key distinguish between sonification and warnings or exact judgments of data values. These features seem to go well with evaluation of emotions since happiness, passion or sadness are not objective measurable magnitudes. Listening acuity may demand practice and previous knowledge of the sound source(s), but once acquired, it represents a unique way to gather information that otherwise would demand expensive resources or remain obscure. It is fairly easy to recognize the emotional status of a close friend just by ear, or to identify whose relative is entering in the house by the way the key and the footsteps sound. In the same way that a doctor listens to the breath of his patients or the mechanic listens to the roar of an engine, represent ancient diagnosis methods that are quite powerful when combined with a multimodal analysis.

The power and ubiquity of sound may be unfamiliar to most people but it was not unnoticed by those in the media realm, who have understood its use for propaganda and commercial purposes. The "subliminal inculcation of values" stated by (Truax, 2001) occurs in a daily base through radio and television. Messages flood our acoustic space in loop, whether one is in a shopping mall, work place, home or transportation. The act of "turning on the TV to have a sense of companionship" is mainly provided by sound and is common to a large number of people. Apparently harmless messages enter our house, carefully crafted by skilled sound designers, using "catchy" sound effects, warm voices and are supported by thorough market and perceptual studies. As (Schwartz, 1974) wrote on is seminal book about the impact of media on people – The Responsive Chord:

> *"Today, there is a nearly constant flow of information at all times. Indeed, one has to expend considerable effort hypothesizing a situation in our culture in with communication does not regularly occur. We take in electronically mediated auditory and visual information as part of our life process. It is part of our immediate physical surrounding, and we sit in it, absorbing information constantly. The vital question is: What are the characteristics of the process whereby we organize, store, and act upon the patterned information that is constantly flowing into our brain? Further, given these processes, how do we tune communication to achieve the desired effect for someone creating a message?"*

This passage is enlightening as it expresses three important ideas: 1) We are surrounded by information, 2) How do we perceive it from a cognitive perspective? and 3) How can the media producer optimize his or her communication accordingly?

For our research and previous discussion, point 2) holds the most interest as it regards the effect of sound in our daily life. The presence of continuous streams of information in the background may seem innocuous but it is absorbed and processed by our brain by means of a “distracted listening“ (Truax, 2001) mode, where sound is there to suppress needs of companionship and joy while conveying subliminal messages.

The same happens with other mechanisms like “sound perfumes” and “sound walls” (like the noise of the air conditioners) used to constrain our acoustic horizon, or Moozak (also called “elevator music”) - generic music tailored to induce specific emotions regarding a specific purpose (promote consumption, productivity, stress, etc.).

CHAPTER 3 - Mobile Technology, Soundscapes and Context

In the last fifty years, with the profusion of computer processing, technological apparatus have evolved in a way that would sound science fiction a few years ago, serving humans on a vast panoply of tasks, within a diverse range of areas. Undoubtedly, one of the areas that most grasped this technological development was the broad area of Communication. This fact does not sound surprising as we were all used to major advances in this field even before computers had arrived.

The first audio wireless broadcasting is usually attributed to Reginald Fessenden when in the Christmas Eve 1906 he transmitted a message of him playing O Holy Night on the violin and reading a passage from the bible. Sixteen years after this experiment the United States made radio licenses available to broadcasters, and only two years later there were one thousand and four hundred radio stations in the country (Ruben, 2010), a number that kept growing every year, even after the arrival of television and internet.

A similar phenomenon was also observed in television broadcasting. On November 26th 1936 the first public transmission of a high-definition television program took place at Radiolympia in London. In the years that followed this event, particularly after WWII, television gained an unprecedented market penetration becoming the most influential mass media.

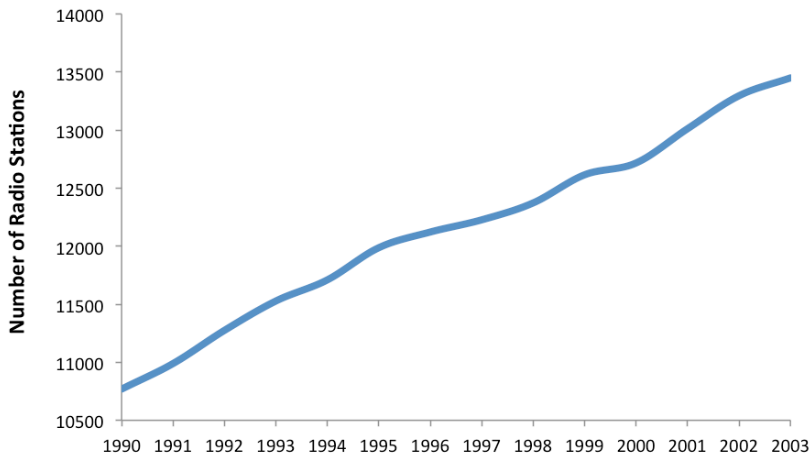


Figure 2 - Number Of Licensed Broadcast Radio Stations in U.S. - 1990 to 2003 (source: Federal Communications Commission, www.fcc.gov)

This craze happens in the communication realm because people are keen to communicate. This circumstance naturally induces the increment of information being exchanged to follow the increase of technologies that support these exchanges, as they alter the information-per-unit-of-time ratio. While a send and reply process of a letter would take the time of writing plus the time for the post transmission (which ultimately would take much longer than the writing process), with e-mail technology this last portion of the equation is almost virtually removed. The result: more communicational events occur and the flux of information that floods our life tends to grow. “Time is erased in the new communication system when past, present, and future can be programmed to interact with each other in the same message.” (Castells, 2007).

Nonetheless, our personal threshold for the allocation of information seems to be constantly being pushed forward. The risk of an Information Overload (a.k.a. infobesity) is real and might eventually turn against itself one day, but for now multidisciplinary IT teams are refocusing the problem in order to keep it under control, taking human cognition as the limiting factor and working “toward the implicit goal of assisting everyday life and not overwhelming it.” (Abowd, Mynatt, & Rodden, 2002)

Despite all the advances achieved so far, some of the technology we have today and which we justly consider top-of-the-line was raised many years before as a scientific aim or pure fictional abstraction exercise. One example that can be found in (Verne, 1889) was an audio-video conference system regarded as a household appliance.

> *"The telephote! Here is another of the great triumphs of science in our time. The transmission of speech is an old story; the transmission of images by means of sensitive mirrors connected by wires is a thing but of yesterday. A valuable invention indeed, and Mr. Smith this morning was not niggard of blessings for the inventor, when by its aid he was able distinctly to see his wife notwithstanding the distance that separated him from her."*

Whereas some sci-fi predictions from last century didn't find any suitability in today's world, some did become parte of our everyday life as the one just described. Another visionary man, this time a scientist named Mark Wiser, also incurred in some of the most interesting futuristic prediction, chiefly in his article The Computer of the 21st Century published in 1991, where he describes the trajectory towards the ubiquity of computers. More than twenty years after, some of his prophecies form an integral part of our life, while others are still far way or will never engage.

> *"The telltale by the door that Sal programmed her first day on the job is blinking: fresh coffee. She heads for the coffee machine. Coming back to her office, Sal picks up a tab and "waves" it to her friend Joe in the design group, with whom she has a joint assignment. They are sharing a virtual office for a few weeks. The sharing can take many forms -- in this case, the two have given each other access to their location detectors and to each other's screen contents and location. Sal chooses to keep miniature versions of all Joe's tabs and pads in view and three-dimensionally correct in a little suite of tabs in the back corner of*

her desk. She can't see what anything says, but she feels more in touch with his work when noticing the displays change out of the corner of her eye, and she can easily enlarge anything if necessary." (Weiser, 1991)

The main distinction between Weiser and Verne was that the former did dedicate his life to accomplish what he envisioned. The ubiquity of computation is a cornerstone on our project, mainly in the form of mobile devices, and is described in the next subchapters.

3.1 - Ubiquitous Computing

We can assume that, nowadays, computers are already quite an ubiquitous device, being present in many of our daily moments whether we are working, relaxing or socializing. However, the term ubiquitous computing (ubicom) – coined during the late eighties in PARC (Weiser, Gold, & Brown, 1999) - doesn't refer primarily to the existence of many computers spread all over the place. It rather concerns the invisible and pervasive presence of computers in our daily life, merged within the everyday objects we use, in the clothes we wear, in the places we inhabit. At the risk of sounding too theological, we may say it's an inconspicuous omnipresence.

Throughout the two last decades, this perspective on computing has gained several nicknames such as pervasive, wearable, augmented, invisible, disappearing, calm, everyware, everyday computing, and so forth, depending on the research focus or the angle from which it is tackled. But fundamentally, in its basis, all these approaches share the common vision of (Weiser, 1991):

We are therefore trying to conceive a new way of thinking about computers, one that takes into account the human world and allows the computers themselves to vanish into the background.

The background is usually regarded as something we don't dedicate our main attention to; it is the setting where something more important occurs.

Computers, by fading into the background, step away from their explicit involvement on a specific task. The way how this vanishing would happen was also anticipated by (Weiser, 1993) stating that “In the long run, the personal computer and the workstation will become practically obsolete because computing access will be everywhere: in the walls, on your wrist, and in “scrap computers” (like scrap paper) lying about to be used as needed.”. Almost twenty years after this declaration, computers did become more ubiquitous, forming part of our households with intelligent houses technologies, promoting our mobile phones to smartphones and integrating our cars assisting us in navigation tasks, parking and secure driving. Nonetheless, many of the interactions with computers still happen within the mouse, keyboard, and screen paradigm. Even on a mobile device like a smartphone, with a multi-touch screen and no physical keyboard, most of the HCI interactions rely on a keyboard GUI, meeting and only slightly upgrading the former paradigm. Computers, therefore, didn’t completely vanish into the background as Weiser predicted. To accomplish an effective ubiquity, computing and HCI should address, according to (Abowd et al., 2002), three main features: *context-sensing*, *capture and access* and *constant connectivity*. We find this list relevant, as it is the backbone of our project, which entails these dimensions plus others discussed next.

3.1.1 - Mobile to Mobility

The physical size of computers dramatically changed since its inception. While at first computers would require a medium size room to accomplish computational tasks that sound ridiculous to today’s standards, current state-of-the-art micro computers can easily fit in the palm of our hand, holding much more processing power than their ancestors. The progressive reduction in size observed in computer hardware throughout time turned real the plans for a portable and personal computer/notebook designed by (Kay, 1972) – the Dynabook:

The size should be no larger than a notebook; weigh less than 4 lbs.; the visual display should be able to present at least 4000 printing; quality characters with contrast ratios approaching that of a book; dynamic graphics of reasonable quality should be possible; there should be removable local file storage of at least one million characters (about 500 ordinary book pages) traded off against several hours of audio (voice/music) files.

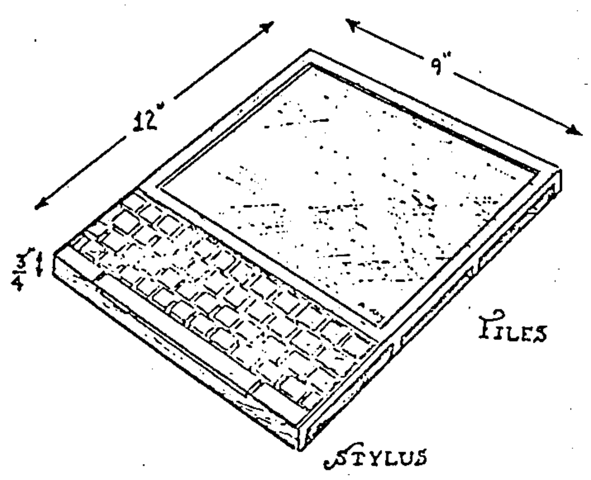


Figure 3 - DynaBook by Alan Kay, 1972

The features described in this paragraph can nowadays be easily found in any portable computer, table or even smartphone. The laptop paradigm, which exploded during the nineties, was based on the idea of using former desktop computer in any place. This would mean to be able to use the same computer, with the same information and setup in most places, which represented at that time (and still today) a great advantage. Nevertheless, ubiquitous computing requires a different type of mobility to address its aims. The approach suggested by (Bill Schilit, Adams, & Want, 1994) defined a Mobile Distributed Computing System as “a collection of mobile and stationary computing devices that are communicating and cooperating on the user's behalf”. The idea of mobility is focused on the user rather than the device “because it concerns mobile people not just mobile computers. These systems aim to provide ubiquitous access to information, communication, and computing" (Bill Schilit et al., 1994). With later advances in mobile technology the devices became so small and powerful that (Schmidt, Beigl, & Gellersen, 1999) used the term Ultra-Mobile Devices to define "computing devices that are operational and operated while on the move, and characterized by a shift from general-purpose computing to task-specific support". Users move from one side to another and actually use their devices while moving, whether to listen to music, navigate, take notes, pictures, socialize or communicate. During these tasks, users change their physical location, consequently changing their social and environmental context. If a smartphone is good to take pictures of the places one visits, the desktop would be more suitable to organize them in folders and the TV set to view them in family, always accessing the same database, the same pictures.

The distributed nature of this network of devices makes it suitable for ubiquitous computing since each one holds a different set of inputs and outputs that make them individually tailored to specific tasks and contexts (e.g. the TV Set has a bigger screen, the smartphone is lighter and has a built-in camera, the PC has a mouse and proper keyboard). However, dealing with mobile devices involves constraints that have to be taken in consideration when designing new systems, these are: the unpredictable variation in network quality, limitations on local resources imposed by weight and size constraints, concern for battery power consumption, limited input and output.

3.2 - Input: context sensing to context-awareness applications

Context, as described in the previous chapter, has an important impact on the way we perceive events and consequently on the way we react to them, having in mind the given context. The first problem one faces when dealing with context is to find an all-purpose definition that fits different scenarios. From all the designations found in literature, we will use the one suggested by (Dey, 2001), by its broader range that encompasses most of the situations we can think of.

> *Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*

This definition goes further and expand other suggested by (BN Schilit & Theimer, 1994) which restrain context to the location and closeness to object and people.

The gathering and recognition of a context by a device is called context-sensing and can be achieved implicitly by using sensors and tracking implicit user interaction or explicitly by requiring the user to specify it. One example would be the location of a user in a map occurring automatically through GPS versus the manual introduction by its user through the introduction of an address.

The panoply of sensors used to gather information about the surrounding environment is diverse and includes GPS antennas (for location), light and vision (for environment luminosity, people/objects activity and recognition), microphones (ambient sound, music and speech), accelerometers and gyroscopes (movement, device orientation), Wi-Fi / Bluetooth antennas (Wi-Fi networks, BT devices), atmospheric (humidity, temperature and air pressure) and bio-sensors (heartbeat, galvanic skin response, etc.). The use of implicit input by tracking user logical interaction with the computer is sometimes regarded as a software sensor approach (Pathan & Reiff-Marganiec, 2009).

The process of context-sensing happens in two different stages: feature gathering and context matching. In the first stage, data arriving from the sensors is collected as raw records, in the second stage this information is processed and matched to a specific context.

For more accurate context identification, sensor fusion is usually applied during the first stage, which consists in the combination of the data arriving from several sensors simultaneously. The second stage describes a machine perception task and requires previous knowledge and memory. The machine must be told which are the different possible contexts and what the feature combinations and thresholds for each one are. After that, every new sensor input is analyzed and the result is outputted using machine learning and data mining processes. This process goes from trivial k-NN algorithm to advanced methods including polling and inferences from other devices in proximity, as described by (Miluzzo et al., 2010). Anyway, the classification is based on “the assumption that similar situations (considered as one context) are represented by similar stimuli. Therefore, sensors may be used to determine contexts based on the assumption that in similar contexts the sensory input of the characterizing features is similar." (Schmidt, 2012).

After this wrap-up on the topic of context-sensing and implicit interaction a question emerges: what is the real advantage of following such approaches? The answer is typically “to increase a system’s understanding of its surrounding environment in order to take the right adaptive behavior” (Schmidt, 2012). This goes in line with the definition of *context-aware computing* coined by (BN Schilit & Theimer, 1994) which describes applications that “adapt according to their location of use, the collection of nearby people and objects, as well as the changes to those objects over time.” and relates to the idea that “our natural interactions with the physical environment provide sufficient input to a variety of attendant services, without any further user intervention” (Abowd et al., 2002). In fact, some of our non-computerized appliances are context-aware, for example, a heater that triggers when certain temperature is reached, a light that comes on when someone enters the room, the automatic fire sprinkler that triggers in the presence of a fire or the radio in the car that automatically shifts channel when traffic news is being broadcasted on a different channel. All these examples respond to an exchange in the environment that is sensed by different kinds of mechanisms. Computer applications that are considered context-aware react to some exchange in the surrounding environment sensed by the means described before. Some basic context-aware adaptive behavior in a computer device would be the change from portrait to landscape view of a GUI on a tablet when this is turned or the dim of screen luminosity with the increase of background light. The reactions and adaptations to context can occur in different ways and assume several forms. (Schmidt, 2012) discriminates between the following types of context-awareness:

- Context-adaptive systems - proactive applications, function triggers, adaptive applications, intelligent displays;
- Adaptive and context-aware user interfaces – user interfaces that adapt to the idiosyncrasies of each utilization (e.g. the GUI rotation mention above);
- Managing interruptions based on situations and sharing context and context communication - context as a source to schedule interruption and communication (e.g. mobile phone doesn't ring when user is in a meeting or provides this information to the caller to deicide either to continue with the call or postpone it);
- Generated data for metadata - data that is collected automatically and added to documents produced by user (ex. notes);
- Implicitly user-generated content - data that is collected and stored automatically based on user actions and context (e.g. keep a record of user localization and output the average velocity on a trip). This should support later retrieval;
- Context-aware resource management – use information from context and user interaction to manage resources (e.g. switch from paid network connection to free connections when available or change *energy saver* settings when connected to power cord).

In order to provide a reliable context-aware service it is important that systems have access to constant connectivity, as connectivity is the key to share current contexts and write and access data. Understanding and evaluating a context remains a hard task for machines mainly for the subjectivity underlying the process. Even though, for a neat interaction between humans and ubiquitous devices, a good sense of context and human activity will represent a valuable asset.

The application we describe in Chapter 8 - Hurly-Burly: Prototype System for Experimental Research (pp.151) is primarily considered a context-sensing application but also context-aware since it implements two of the features described above, namely: context-aware user interface and implicit user-generated content.

3.2.1 - Groupware towards Group-Awareness

Computer Supported Cooperative Work (CSCW) is a field of HCI devoted to understanding group work processes and developing tools to enhance collaborative efforts (Greif, 1988). Its importance is so remarkable that it was a focus of the former Special Interest Group on Groupware (SIG-GROUP) from the Association for Computing Machinery Organization (ACM). The core applications of such a group includes services as: collaborative text editing processors, intelligent conference systems and music collaboration, such as the Public Sound Objects (Barbosa, 2005), where users create music together by sharing the same visual and sonic space.

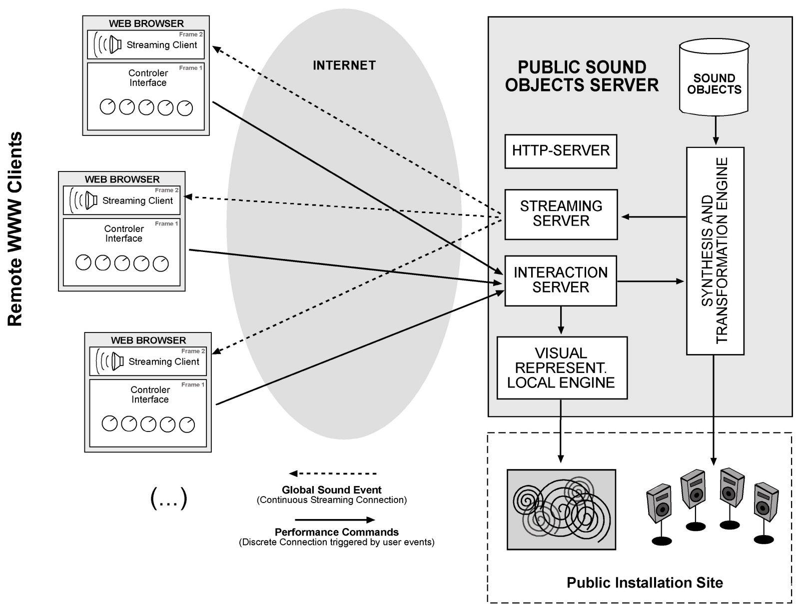


Figure 4 - Public Sound Objects setup (Barbosa, 2006)

The reason why we included this chapter is because our application shares some features with traditional CSCW systems and may even be considered to be one of those. The software application is designed for network use, where all users of the same group share the same visual space, thus complying with two of the requisites of a CSCW application. The third requirement is the engaging of all users in the same task. While our application doesn't try to address a specific task (like editing a text document) we must agree on the existence of an underlying goal: socialization and group-awareness, a concept closely related to co-presence (Koskinen & Battarbee, 2006), which is based on the knowledge of other members activities and status. This feature gives the user a sense of belonging to a group, while it allows him/her to keep track of his/her friend's sonic activities. All these ideas are inline with the vision of Peter Hoschka who fifteen years ago urged for a change of focus in CSCW towards a social-web:

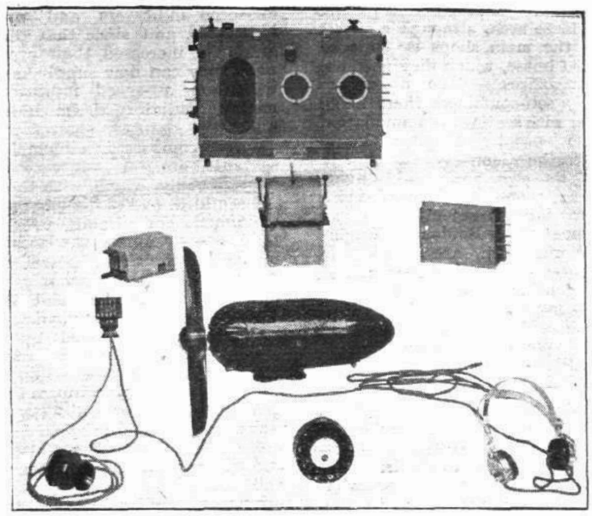
> *"[…] it may be said that the development of basic tools for co-operation support is by now well established in industry, and is less suited as a topic for a research organization. We have therefore shifted the focus of our work: from basic cooperation tools to the computer as a social medium."* (Hoschka, 1998)

3.2.2 - Mobile Sound: sound as context

We have already listed some of the different ways by which implicit input can happen. For the sake of this research project, we narrowed the focus of attention into sound as an input of environmental features, and how it can be used to describe activities and promote context-aware applications.

The presence of audio in mobile devices usually regards receiving/transmitting audio or listening to recorded sound. Both technologies have more than ninety years of existence and have not always shared the same housing as nowadays. The first audio wireless devices being commercialized were telephones for aircrafts, ships and trains. One example is the Aircraft Wireless Telephone Type AD2, by Marconi Company, produced in England in

1920 (Figure 5). This type of instruments could travel in space but were far from being practical mobile devices.



Marconi's Aircraft Wireless Telephone, Type AD2. This set weighs only 60 lbs., yet has a telephone range under normal conditions of about 150 miles.

Figure 5 - One of the first wireless telephones, built in 1920 (Spooner, 1922)

First recording and playback commercial device date from late 1880s. The phonograph was invented by Thomas Edison and made use of wax cylinders to store the sound recordings. This device was no conceived as mobile audio equipment but its compact dimension would allow moving from one place to another, recoding and playing back audio in different scenarios.

Further developments in both areas led to some of the most revolutionary and common equipment used today in modern societies. Regarding wireless sound communication, mobile phones (based on cellular networks) became the standard, with a market penetration in Europe surpassing the mark of one mobile phone per person, already in the year 2006 (Comission of the European Communities, 2009) (Figure 6). This is a remarkable mark if we consider that Motorola produced the first commercial mobile phone that was not attached to a vehicle of transportation only in 1984[6] (Figure 7).

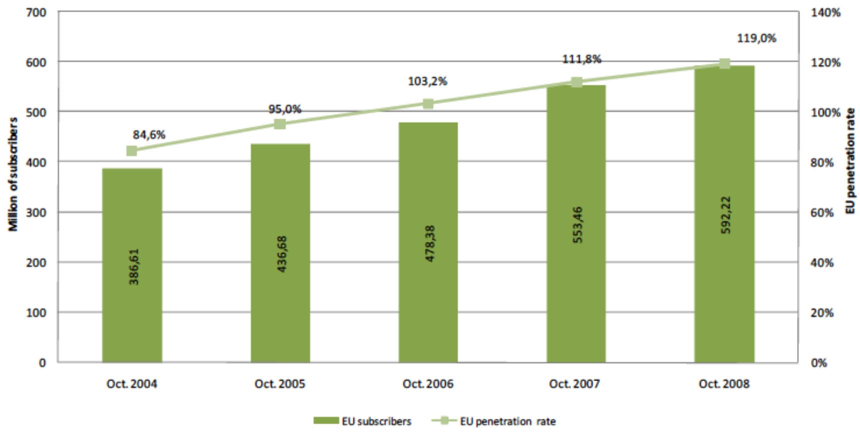


Figure 6 - Mobile Subscribers Penetration in Europe (based on active mobile subscribers) - 2009

Figure 7 - Motorola DynaTAC was the first commercial mobile phone using cellular networks available on the market in 1984.

On the other hand, the portability of music players was always more easily achieved as the first recorders and reproducers (the phonograph) were already small-sized devices. Later, vinyl players saw the portability as a commodity and some brands release proper battery operated vinyl players, aiming for mobility and selling the idea of having music everywhere. However, the individual experience of listening music (or sound) would arrive only after the

[6]The DynaTAC phone: http://www.motorola.com/us/consumers/About_Motorola-History-Timeline/About_Motorola-History-Timeline,en_US,pg.html#1980

invention of transistors-based radios in 1954 (Lane & Lane, 1994) (Figure 8). Small pocket sized devices could then be carried and used on the move, virtually anywhere. But radios were not music players, in the sense they didn't allow their users to load their own music. Instead, they were mere receivers of a synchronous remote music broadcast system, allowing limited control to the user but the on/off and frequency switch (i.e. tune into a different radio station).

Figure 8 - Regency TR-1, the first commercial transistor radio, available in 1954.

Only in the 1979 the concept of listening to your own music on individual mobile devices turned into realty, with the advent of Walkman, produced by Sony[7] (Figure 9). The Walkman is a pocket size cassette player that can be use on the move. Typically Walkman devices didn't have speakers so the listening ended up becoming a personal activity, by the exclusive use of headphones, transforming the way people consumed music towards an individual experience and promoting what would become known as the Walkman Effect (Hosokawa, 1984).

[7] TPS-L2 was the first Walkman by Sony: http://www.walkmancentral.com/products/tps-l2

Figure 9 - TPS-L2 was the first Walkman available on the market, in 1979, produced by Sony.

As we mentioned in the Introduction chapter, names and concepts concerning portable audio are not always coherent within scientific literature and the audio industry. For this reason, we have decided to adopt the term Mobile Sound to define audio computation being held on a mobile computing device.

In our project we use the sound of the environment surrounding the user as an implicit input for the system, in order to foster social interactions among his/her network of friends and acquaintances. The use of sound as an input for context-sensing and context-awareness applications is found in several research projects and commercial applications (for examples, see Section 3.2 -). Typically, sound is used within a sensor fusion system aiming to provide extra information to increase the confidence level of context identification, in other cases sound is used as the primary source of information and on other occasions it is also used as an explicit input.

A list of the most relevant projects using sound as an input on mobile computing devices, with emphasis on networked systems aiming for social relations, group-awareness and communication are presented in detail in

Chapter 5 - State-Of-The-Art On Sound-Based Social Interaction Systems (pp. 99).

3.2.3 - Computational Auditory Scene Analysis

For a human it is trivial to distinguish different sounds in a soundscape, identify their sources, pointing out their general sonic properties and their positions in the physical space. This is even possible in a busy acoustic environment when a myriad of sounds compete for conspicuity, what (Cherry, 1953) has named as the *cocktail party effect*. This human ability has been studied under the field of Auditory Scene Analysis and consists on discerning the organization of sound scenes according to their inferred sources. (Bergman, 1990) gave an important contribution to this area with his seminal book Auditory Scene Analysis, drawing an analogy between the perception of auditory scenes and visual scenes.

However, for a machine, this task is all but easy and results are far from being excellent. The research that deals with this matter is called Machine Audition (also called Computer Audition or Machine Listening) and the emphasis on auditory scenes is called Computational Auditory Scene Analysis (CASA) (Brown & Cooke, 1994). In this subchapter we briefly characterize the field and present the methodological overview for a typical CASA task, as it represents a key element on our system.

The main approach for CASA is to model the human auditory system. Therefore, the starting point is usually to mimic the behavior of its individual biological parts, such as the ear (outer, middle and inner) and later the brain. This method is valid and may be useful to shed some light over the functioning of human auditory system.

Moreover, the chosen approach for a CASA system is usually twofold; it first depends on the resources of the system (modeling human perception is computationally demanding) and second on the aims of the project. We may generally assume that all humans are born with fairly equal auditory systems, but the fact is that with aging people shape their auditory system either by ruining it or by educating it in a certain direction. Musicians will develop better musical hearing (e.g. with more accurate tone discrimination) whilst a mechanic will tend to have higher accuracy in engine-like sounds.

When a CASA system is developed, it typically aims at a specific task, which demands a specific type of hearing; analogously to the musician and the mechanic. Some of the most common applications for CASA can be divided in the next four categories:

- Segmentation is the ability to determine the onsets and offsets of logical sounds. Some examples are: time stamping different radio contents or different speakers on a TV debate to assess who talked longer (Tzanetakis & Cook, 1999);
- Source segregation is the ability to isolate a specific sound source within a mix of sounds on the same track. One example is to isolate a melody from one instrument (Martins, 2008);
- Classification consists in determining the source of the different sounds in a mix. It can also be extended to musical genre recognition (Tzanetakis & Cook, 2002);
- Speech recognition is the ability to transcribe speech into text, one application is automatic subtitling for films (Gauvain & Lamel, 2000).

A classic approach for a CASA system entails two main steps:

- Digital Signal Processing, which may include audio feature extraction (use of audio descriptors like MFCC, Centroid, etc. to extract characteristics of the signal), filtering and other manipulations.
- Machine Learning, algorithms to give logical meaning to the extracted features (classification, segmentation, retrieval, etc.).

A subset of the CASA field is the Computational Auditory Scene Recognition (CASR), where the "*focus is in recognizing the context, or environment, instead of analyzing and interpreting discrete sound events*" (Peltonen, Tuomi, Klapuri, Huopaniemi, & Sorsa, 2002). For example, the Bag-of-Frames approach "*proposes to directly recognize soundscapes as a whole, without the prior identification of constituent sound sources. In these works, soundscapes are modeled as the long-term accumulative distribution of frame-based spectral features*." (Aucouturier, Defreville, & Pachet, 2007)

The system we propose aims at a general listening classification task, by distinguish between three different soundscapes mainly composed of: environmental sounds, musical sounds and speech. Thus can be considered a CASR task. Details of its implementation are presented at point 8.2 - The Mobile Application (pp. 153).

3.3 - Output: towards Ubiquitous Displays

Internet, and mainly the World Wide Web at the beginning of the nineties, made the world smaller and sowed the seed for what (Castells, 2007) would be called the Information Era. The hypertext was the new train. People were a click away from new and undiscovered virtual sites and were able to exchange messages and documents in real-time. This had an impact on the way people interact and inhabit the world, mainly in developed countries where this technology was/is available on a democratic basis.

One of the side effects for this new condition is the torrent of information that floods our everyday life. The issue is no longer how to gain access to information, but rather how to filter the information, distinguishing the essential from the accessory.

Although Web is available free and sets no boundaries to its own organization (i.e. everyone can set up a website without further respect for implementation rules, only content), the emergent structure turned out not to be a random network. Instead, hubs and clusters form its underlying structure.

Web portals help to keep related information tied together under the same website and search engines allow to retrieve content that shares the same search criteria by crawling the entire network. As users of the Web, it is easy to agree that deprived of such instruments, surfing on the Web would be a discouraging undertaking.

Nonetheless, the endeavor to keep information overload under control is not unique to Web organization. Let us climb the abstraction ladder and make an analogy between a website (e.g. a web portal) and a physical place (e.g. a living room). In the living room we have things that are background and other that are foreground, for instance, we are not constantly paying attention to the telephone, we do it only when it rings. Similarly, if we live in the city, there may be some street noise filling the acoustic space of our living room, to which we do not pay constant attention. When designing web interfaces, designers and information architects also have to tend to the underlying hierarchy of information, in order to distinguish the foreground from background and aid the user in the mediation process.

Ubiquitous interfaces aim to take this endeavor a step further and make HCI as flawless as possible, although (Sá, 2010), in a remarkable systematization on the subject, concluded that "*the interface, as mediation process, should oscillate between transparency (associated to immediacy) and opacity (associated with hypermediacy), existing commutatively in both modes*.".

In the previous section we discussed the input process of ubiquitous computing, with implicit and explicit context-sensing. In the output stage, ubiquitous computing also aspires to an invisible interface, a non-imposing way to convey information.

> *"[…] we want displays that are less demanding of our attention. We'll achieve Weiserian invisibility by designing output that provides for peripheral awareness of information out of the foreground of our conscious attention. Researchers have explored*

the trend toward peripheral output for a particular class of displays—ambient displays." (Abowd et al., 2002)

This regards the control of information on the interface side. An approach grounded in human cognition range, aimed to provide information from multiple sources that rest at the periphery of our senses and afford qualitative, ambient forms of communication.

An historic example of a calm computing display is the artistic installation from Natalie Jeremijenko, called Dangling String. This piece consisted of a thin piece of wire attached to an electric motor near the ceiling of an office that rotated according to the actual network activity. The information was not presented in a graph nor displayed as a number; rather the velocity of rotation gave it qualitatively. "*First, by placing things in the periphery we are able to attune to many more things than we could if everything had to be at the center. […] the periphery is informing without overburdening*." (Weiser & Brown, 1996b). The physicality of the object also helped moving the display into the periphery of the brain and the sound coming from it could be heard in many offices, without being obstructive.

Wind Chimes, a traditional decorative artifact, represent another paradigmatic example of calm technology. Even though the original goal is religious, Wind Chimes are, in fact, a visual and sonic gauge for wind activity. It would be very difficult, maybe even impossible, to precisely measure wind velocity and direction with one of these devices. However, by observing and listening to it, it becomes easy to deduce a rough, generic wind activity, which in most cases is more than enough. In addition, it holds an aesthetic dimension, which is important in output displays, either visual (Tractinsky & Eytam, 2012) or auditory (Barrass & Vickers, 2011).

Both projects, presented as examples of calm displays, are relevant for our study as they encompass processes for information visualization and sonification with an emphases on aesthetics. These two research areas are briefly introduced in the following sections.

3.3.1 - Sonification

In this section a description of the Auditory Displays (AD) field is presented, with emphasis on continuous Sonification processes. This overview has the purpose of contextualizing the reader on this topic, as it represents the theoretical and practical basis for a significant layer of our research project.

An Auditory Display concerns the communication of information through auditory output. It represents a vast area of research, which encompasses all aspects of an auditory-based HCI system, such as: the computational setup, sound equipment, modes of interaction with the system and all technical solutions involved in obtaining sound from data. In turn, Sonification is a recent subtype (yet a core component) of AD that *"[…] seeks to translate relationships in data or information into sound(s) that exploits the auditory perceptual abilities of human beings such that the data relationships are comprehensible*." (Hermann, Hunt, & Neuhoff, 2011b)

As a research field, AD and Sonification can be considered as still maturing. Contributions to this field arise mainly from the ICAD community (International Conference on Auditory Displays), which had its first gathering only twenty years to this part in Santa Fe, New Mexico in 1992. A recent publication called The Sonification Handbook (Hermann, Hunt, & Neuhoff, 2011a) provided a great contribution to the field, gathering insights from more than thirty prolific researchers. This can be considered a landmark as it takes a step further in systematizing and indexing a field that per se is averse to absolute truths and unequivocal jargon, stemming from its multidisciplinary nature, which encompasses disciplines as Computer Science, Music, Sound Design, Psychology, Sound Engineering, Physics, Psychoacoustics, Audiology, Product Design, among others.

Auditory Displays have been present in many of our everyday devices, ranging from the car to the alarm clock, as well as computers, medical equipment or household appliances. Mobile devices, which are akin to our project, are also on the list and represent a great window of opportunity to AD to shine, as the small screens and great computational power afford and demand auditory interactions. The strength of Sonification derives from the unique properties of sound discussed in Chapter 2 - Sound as a Medium, Sound as a Message, in particular the perceptual and cognitive features of the human listening system, which exceeds most modern computers in pattern recognition tasks, identifying sound sources, spoken words, and melodies, even under noisy settings.

The applications for AD are many and have generally been defined by the community in terms of four broad categories (as appears in (Hermann et al., 2011b)):

- Alarms, alerts and warnings – sounds used to indicate the occurrence of a particular event, which may or not demand user action. Alarm clock, microwaves countdown or fire alarm are examples of this type of AD;

- Status, process and monitoring messages – sounds used to provide a continuous monitoring over an ongoing status of a system. Most of the sounds might be alarm-like sounds, triggered by reaching a defined threshold, while others would need an uninterrupted presence in order to express subtle variations in data flow. These include sound such as medical life status equipment or telephone hold time;

- Data exploration – this category is frequently referred to as Sonification, as it concerns the exploration of large datasets by means of sound. It conveys a more holistic representation of the data rather than individual details. Examples of sonifications under this category are auditory graphs and model-based sonifications.

- Art, entertainment, sport and exercise – use of sound to accomplish or enhance entertaining tasks, like, audiogames and aid visually-impaired people in group activities. Artistic creation has also been using sonification techniques for musical composition, as data represents meaningful extra-musical material for composers.

An important aspect of any AD system is how users can interact and control de auditory output, which ranges from *non-interactive* – “the traditional ‘concert model’ of sound presentation, which treats the listener as a passive recipient of the audio presentation” (Walker & Kramer, 1996); to *user-driven display*, where the user has control on the flux of interaction which occurs within the system constraints. For example, Model-based Sonification is more suitable for user-driven interactions, as it does not simply map data-to-audio but instead creates a model whose sonic response is based on data and triggered from user interactions (Hermann & Ritter, 1999). On the other hand, the most basic mapping between data and sound is called *Audification*, where waveform-like datasets are transposed in pitch and time-stretched, in order to be acoustically perceived (and hopefully meaningful) by the human listening system.

The design path for the construction of an AD is an interactive process that usually follows the steps described next:

- Task Assessment – What is the task (monitoring, process-awareness, data exploration, point estimation and point comparison, trend identification, data structure identification, exploratory inspection)? Why is the AD needed and where is it going to take place (good definition of the context), who is going to use the AD system (age, social background, musical ability, listening/visual impairments) and when is it going to be used?

- Data – Define the dataset involved: the variation range, real-time vs. fixed, qualitative (verbal) vs. quantitative (numeric);

- Mapping – How is data going to be processed (filtered, shifted, rounded, grouped, translated, scaled, polarity, etc.), what sounds or sound parameters are associated with data (variation, threshold, event, metaphors, semiotics, etc.), is it continuous or event triggered, *audification* or music-like, multiple or single data streams, etc.

- Interaction – How users interact with the AD: model-based, non-interactive, contextual cues or non-contextual cues,
- Sound rendering – How the sound is produced and delivered to the user (it is mainly a technological issue): synthesis vs. sample; local vs. server; speakers vs. headphones; mono vs. multichannel, etc.
- Prototyping – Create a functional instance of the system.
- Perception and Evaluation – The response of the user to the system: does it work? Does it enhance interaction? How closely does the perceived message pair with the one initially projected? Does it require training or any particular perceptual ability (e.g. absolute pitch)?

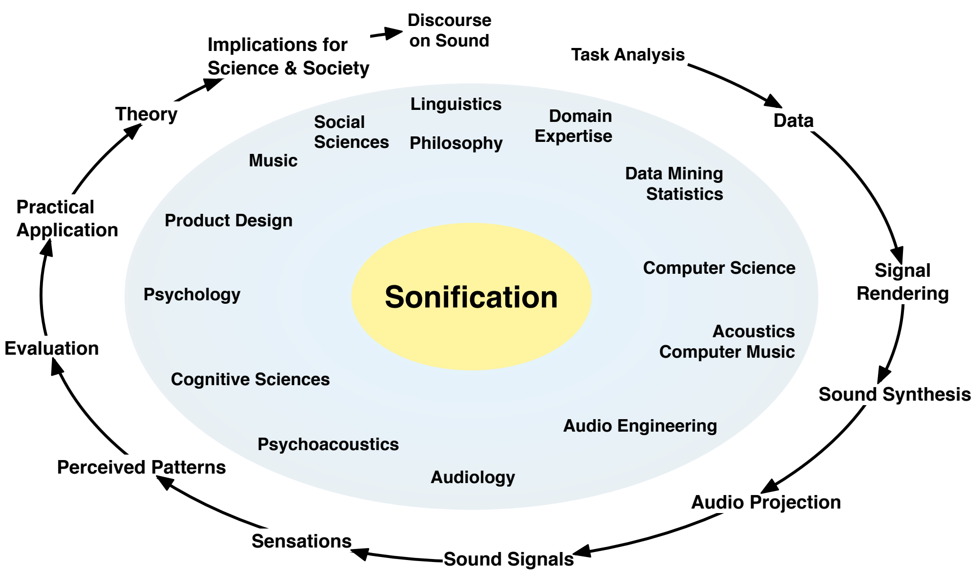


Figure 10 - The sonification production cycle (source: (Hermann et al., 2011a))

As ADs are based on sonic output, an aesthetic concern may be regarded as one valid feature for engaging the user. Nonetheless, it is important to attend to what (Edworthy, 1998) has commented about the use of AD during monitoring tasks :

> *"[…] it can sometimes be the case that a particular type of auditory interface is selected for use—in preference either to other possible interfaces or to no interface at all—because it is aesthetically pleasing and/or technically available, and not because it has necessarily been shown to improve performance, or reduce stress levels, or demonstrate some other benefit."*

In our opinion, the balance between aesthetics and more "descriptive" approaches depends on the nature and goal of the task evolving the AD. Like the wind chimes – that populate a soundscape with pleasant musical sounds while conveying rough information on wind activity – some AD systems may aim for a more aesthetical experience based on data rather than a high-resolution display. Moreover, a musical approach to AD "*forces us to move from being hearers to being listeners. What then becomes important for sonification designers is not how 'musical' their work sounds, but how easy they have made it for the audience to listen to it, and by listen we mean 'attend carefully'.*" (Vickers & Hogg, 2006). Mainly in monitoring tasks, "*the aesthetics play an important role in reducing fatigue and annoyance*." (Vickers & Hogg, 2006).

Drifting to the visual display realm, where graphic artists and designers have a significant part on the design of visual displays, (Tractinsky & Eytam, 2012) have recently stated that aesthetically meaningful ubiquitous displays enhance the HCI process, as (1) beauty is associated with quality products, (2) aesthetics induces pleasure, satisfies human needs and impacts on the first impression one has of the product and (3) represents a positive differentiating factor between interactive products. Although not properly assessed, these same principles may apply to the AD domain. This methodology was followed by (Kirsty & Samuel, 2009) who came out with a sonification toolkit which proposes a user-centered customization of the aesthetic representation and scope of the data; as well as a prioritization of "*aesthetic variables and controls in the interface, derived from musical practice, aesthetics in information design and responses to experimental user evaluations to inform the* *design of the sounds and display*". The forgoing shed some light on the importance of an aesthetical concern on the design of AD and Sonifications, though not without proper user testing (for a broader discussion on Computing Aesthetics see (Fishwick, 2008)).

Our research project is deeply weaved within the soundscape field; therefore, our sonification applies the principles of Soundscape Sonification, a subtype of sonification that shares the principles enunciated by (Schafer, 1994) and is described in Chapter 2 - Sound as a Medium, Sound as a Message.

Typically, a Soundscape Sonification is used for Process Monitoring Sonifications and is composed by overlapping different layers of sound, creating a continuous sound stream, which holds information about data being monitored. A good soundscape should be balanced from an Acoustic Ecology perspective, this is, the sounds should be coherent both with the nature of the task being monitored and the sonic universe they belong to.

A Process Monitoring Sonification is usually meant to be used as a background activity while the user is focused on other tasks, or to enhance a visual display monitoring. When used as a secondary task, attention can be *peripheral* or *serendipitous-peripheral* (Vickers, 2011), which corresponds to the distinction between information that is vital for the task and information that is useful but not essential. These definitions are both aligned with the concept of calm computing outlined by (Weiser & Brown, 1996a) and described before.

Soundscape Sonification usually use concrete sounds, which hold a close relation to the source of the data being sonified. However, this is not always possible, as some data does not have a sound associated to it. In these cases, either metaphors or a musical approach can be used.

Some research showed that using natural/concrete or even metaphoric sounds (instead of musical sounds) would reduce the annoyance and may lead to a faster understating of the message being transmitted (W. Gaver, Smith, & O'Shea, 1991), while others concluded that musical sonifications like WebMelody (Barra et al., 2001), would provide a pleasant experience, specially when the user is allowed to customize the sonification by choosing his/her own music. Furthermore, the concept of music is *per se* of difficult definition, and the boundaries between traditional melodic music and noise-like music are blurred. In any case, whether for music, concrete sounds or something in between, the idea of customization is powerful and might be the key for pleasant sonifications and computing interaction in general.

The problem of annoyance and sonic intrusion regarding Soundscape Sonification was addressed by (Kilander & Lönnqvist, 2001, 2002) which introduced the concept of WISP weakly intrusive ambient soundscape, described as:

> *"[…] an ambient soundscape in which events and states in the computing and physical environment are reflected in the form of non-intrusive audio cues. A WISP is meant to service a physical space, like a personal office or a common room. The listening experience is intended to convey intuition rather than interruption; each signal should be sufficiently non-intrusive to be accepted without disturbing the focus of the task at hand, while distinctive enough to be separable from the other cues."*

Also (Eggen et al., 2008) designed soundscapes for physical shared places and found that "*a soundscape can not only serve as an information carrier, but that the sounds can also have a decorative value […] Its informational and aesthetic qualities got their meaning in this workspace and were appreciated by its inhabitants.*"

Sound in general and Soundscapes in particular are not the best ways to convey absolute values (excluding speech, of course), however they represent a valid way to communicate variations and tendencies. Peep (Gilfix & Couch, 2000) is an application that used the principle of soundscape to monitor the network state. Several natural sounds (birds, wind, rain…) and their intensities were mapped to some events and variations on the network.

Nevertheless, (remember Duke Ellington "*If it sounds good, is good*") it is not always advantageous to set fixed rules and dogmatic "dos and don'ts". The variables present in a sonification process (or should we say in any auditory processes) are so countless and diverse that a certain degree of intuition should be held.

CHAPTER 4 - Being Social in the Information Era

> *"We live in a society exquisitely dependent on science and technology, in which hardly anyone knows anything about science and technology." (Sagan, 1990)*

These eloquent words from Carl Sagan are part of an article called Why We Need To Understand Science, where the author addresses the issue of scientific illiteracy found in the US during the late eighties. Thought things may have improved since then (with internet providing tools to empower the consumers, ultimately transforming them in *prosumers*), the fact is that people are usually not aware of the technological breakthroughs behind the gadgets they consume and, more important, the underlying and long-term impact of their use.

In this chapter we addresses the impact of digital social media on society, with particular emphasis on how it shapes social networks (i.e. how impacts the way people interact with each other). In order to achieve this goal, the first section is dedicated to the concept of society and community; the second will present an overview of Network Theory and the third chapter – Online Social Networks - will combine the two previous topics and add *digital technology* to the equation.

4.1 - Society and Community

Live organisms of different species and realms tend to live together, either as colonies, bands, shoals, herds, and other kind of communities. This natural phenomenon finds its main *raison d'etre* on the reproduction of the species, which generally occurs between two elements of opposite sex within the same specie. Though autogamy is present in the natural world, it is mainly circumscribed to most flowering plants, numerous protozoans, and many invertebrates. Cross-fertilization, being the dominant reproduction system among living organisms, has shaped the structure of societies by determining how the simplest (and probably the most important) form of social network is formed: the husband-wife dyad. A dyad is the unity of social networks: a group formed by two individuals with a link in-between. Any large-scale social organization can be decomposed into its multiple dyads, which may differ depending on the perspective from where the observer is analyzing the community. Regarding familiar connections – or *links,* as discussed on the next point – the fundamental dyad can be considered to be the couple, but other dyads can be extended to mother-son, father-son, brother-sister and all other possible combinations within two elements of a household.

Monogamy paves the away for societies based on husband-wife dyads, which are generally found in abundant cultures of the world. When this unity is expanded in order to incorporate one more element, it is called a triad, which corresponds to a group of three people connected and sharing some common elements. In a romantic relation, the inclusion of a third element – regarding monogamy families – is often a disruptive factor, impacting the stability of the network by diminishing its connection strength, thus promoting the disaggregation of the structure. Such phenomenon it is not only observed in couples or other familiar relations. In general, triads are known to be less stable then dyads because, and despite the closeness of the group, there are moments where two elements tend to unite and regard the third as an invader. Ultimately, this behavior can be extrapolated from individuals to estates, countries or other groups. George Simmel, in his seminal book Soziologie: Untersuchungen über die formen der vergesellschaftung (Simmel, 1908) was one of the first authors to establish the theoretical field for the problematic of triads as an expansion of dyads.

> *"No matter how close a triad may be, there is always the occasion on which two of the three members regard the third as an intruder. The reason may be the mere fact that he shares in certain moods which can unfold in all their intensity and tenderness only when two can meet without distraction: the sensitive union*

> *of two is always irritated by the spectator. It may also be noted how extraordinarily difficult and rare it is for three people to attain a really uniform mood when visiting a museum, for instance, or looking at a landscape and how much more easily such a mood emerges between two."* (Wolf, 2011)[8].

More recently, other authors have dedicate themselves to the study of interactions within small groups, which lays on the base of microsociology, a micro-level approach to sociological analysis, regarding everyday personal interactions (see (Caplow, 1968) for a deeper understanding of dyads and triads). Such topic is of most importance to this research as new social media technologies tend to be focused on individuals, promoting their public presence on a wider network, by exposing a personal profile, an avatar, images of their life, cultural and social affiliations, social connections (romantic, familiar, professional) and other elements that contribute to the construction of a digital identity, which might serve to provide a truthful image of the actual person, or, by the contrary, builds on the assumption that digital identity should move away from the actual person, towards the creation of a illusory *digital persona,* which (Clarke, 1994) defines as *a model of an individual's public personality based on data and maintained by transactions, and intended for use as a proxy for the individual.*

The definitions of society and community partially overlap, and borders in their definition are somehow difficult to trace. In a broad sense, society is used to define any group of people in which there are organized relations and mutual exchange of services (Auroux & Weil, 1991). A community also refers to a group of people living together and/or having or sharing something in common (Simpson & Weiner, 1989). In both definitions, what seems to be shifting or getting outdated in present era is the notion of place, which generally underlays the definition of community, society or neighborhood. Place is

---

[8] This bibliographic reference corresponds to a translation and compilation of several texts of Simmel, including the previously mentioned book *Soziologie: Untersuchungen über die formen der vergesellschaftung.*

regarded as a physical location where people live, work or have other regular activity. A place is a geographic location impregnated with meaning for the people who generally occupy it. Thus, it is more than a mere physical location or an architectural space; the notion of place incorporates a sense of belonging in its inhabitants, kept alive thru mechanisms of individual and social memories.

> *"Place incarnates the experiences and aspirations of a people. Place is not only a fact to be explained in the broader frame of space but it is also a reality to be clarified and understood from the perspectives of people who have given it meaning."* (Tuan, 1979)

According to this definition, which clearly adds a human cognitive dimension to a cold appreciation of a geographic location, a place usually sets the stage for a *social space* to happen. Countries, Regions, Cities, Parishes and Homes are all social spaces where people, through dynamics of neighborhood and proximity, interact and organized themselves in groups, which ultimately are called communities.

The reason why the notion of place is changing in the definition of communities and societies is essentially a reflection of the advances in media technologies, particularly those adapted for self-directed use by a given community, somewhere referred as *community media* (Cross, 2007). Former community media technology was usually one-to-all and had a geographic epicenter. Local radios and newspapers, for example, were (and still are) important to promote and define cultural identity of communities. The fact that radiobroadcasting technology is based on transmission through antennas, defined a broadcast converge perimeter that ultimately shaped the community's social space. Also with newspapers, the transmission of the information was physically constrained by transport pathways. However, in Castell's *Information Era* people no longer need to root their community in terms of geographical vicinity. The social space is more prone to accommodate distant connections, since the communication is now more facilitated then ever. Instead of physical proximity, the new communities tend to aggregate around common interests, practices and values, supported by online digital technologies, which allow the creation of groups, forums, chat rooms and other methodologies for gathering and exchange of information.

> *A community of interest is a gathering of people assembled around a topic of common interest. Its members take part in the community to exchange information, to obtain answers to personal questions or problems, to improve their understanding of a subject, to share common passions or to play.* (Henri & Pudelko, 2003)

Co-location (here defined as being physically in the same place, within the reach of human sensory system) is regarded to be less crucial than before for the maintenance of a community and new tools appear everyday to foster new connections and promote community identity without demanding the physical gathering of their members. Nonetheless, this form of communities focused on shared interests is far from being a novelty or *a sign of the times*. Occupational communities such as the academic community are a good example of how groups of people gather around the same concerns, and they exist for centuries. However, it is undeniable that new online technologies have shaped how this community interacts and share information; a good example is the production of online academic conferences: meetings of scientists and academics located around the globe using online technologies, to discuss a specific topic, a phenomenon that would sound science fiction three decades ago. Such communities that migrated into the virtual world grounded on a common former practice were defined by (Henri & Pudelko, 2003) as "*community of practice, i.e. practice the same trade or share the same working conditions. The collective of practice emerges from collective activity. It does not constitute an aim in itself but is the result of the involvement of individuals in the actions of professional practice."*

Nowadays, some of the virtual communities existing in cyber space are an expansion of actual communities that were created outside the virtual domain. One example of this phenomenon would be a discussion forum in an e-learning platform (Learning Content Management System), composed of the whole students attending the actual class. These kind of virtual communities have the purpose of supporting and expanding the communicational resources at the community disposal. Also, they work as repository for past interactions and useful information for the group. Thus, such virtual communities can be classified as *community of practice* since the relation with the professional practice is obvious. On the other hand, others less official and structured virtual communities can emerge spontaneously within the class. Examples of this would be a Facebook group created for the class by the self-initiative of one of students, or just a tight *cluster* that emerge gradually in time, as more students start to establish common links around the classmates (clusters will be discussed later in this chapter). This virtual presence is the unfolding of an actual community, and the primary reason why it exists might not be related solely with professional aims but also with other social aspects of school life. In this scenario, the concept of *community practice* might not be enough to define such communities.

The idea of having virtual communities strictly based on technological mediated interactions is something that not every one finds possible. (Weinreich, 1997) stated that the definition of community cannot encompass groups existing solely on the virtual domain.

> *"What I want to ask is this: How far can mediated contacts constitute community? I believe they cannot. You may get to know other people through CMC [Computer-Mediated Communication], the Net will provide the means to maintain contact and interconnections between people and organizations. But they won't constitute communities because CMC cannot substitute for the sensual experience of meeting one another face-to-face. Trust, cooperation, friendship and community are based on con-*

tacts in the sensual world. You communicate through networks but you don't live in them."

Sixteen years after this article was written, the virtual world became more and more prominent in advanced societies and the word *community* was definitely expanded to encompass groups of people who never established face-to-face contact. The reason why this happens might well be a consequence of the evolution of interfaces, which have blurred the borders of co-presence. In the last decade, the increase of the bandwidth in Internet connections and the ubiquity of online devices made possible for the providers to announce high quality audio and video transmitted effortlessly and seamlessly every where, every time. The inclusion of these visual and auditory stimuli in CMC narrowed the sensorial gap felt between co-located and remote communication. When using a computer to communicate, people are no longer constrained to a chat window, which allows nothing but text and emoticons. Video cameras and microphones replace people's eyes and ears, while wider/3D screens, surround sound and augmented-reality glasses deliver these audio-visual stimuli in a more faithful and immersive manner. Tangible user interfaces have also been study in the field of CMC and *"results suggest that haptic feedback significantly improves the task performance and contributes to the feeling of "sense of togetherness" in SVEs [Shared Virtual Environments]."* (Basdogan, Ho, Srinivasan, & Slater, 2000); commercially speaking, tactile interfaces have found some popularity in the realm of cybersex, where tactile and haptic stimuli are important to achieve sexual arousal[9].

Shared Virtual Environments represent a particular method of CMC, where users share and inhabit a virtual space, usually embodied by an avatar. The main goal of SVEs is to achieve a state of co-presence or "*virtual togetherness",* which (Durlach & Slater, 2000) defined as "*the sense of people being together in a shared space*". Generally speaking, SVEs can be divided into two main groups: one where the users share their audio and video, the other

[9] See Mojowijo for an example (http://www.mojowijo.com)

where the users exist as an avatar. Regarding the use of sound in fostering co-presence in SVEs, this is usually relegated to the use of voice. However, some incursions have successfully been implemented in CSCW regarding collaborative music production (Barbosa, 2005).

Our research takes these premises in consideration and proposes the creation of a SVE where users exist as simplified avatar, which embodies the sound environment surrounding the actual user. This experiment does not intend to restrict communication in any means, only to improve the experience of co-presence by adding a layer of information regarding actual physical qualities of users locations.

4.2 - Introduction to Network Theory

Network Theory (or Network Science) is an area of applied mathematics related to the study of discrete objects and the way they interact and form networks. It is rooted in the Graph Theory, which regards graphs as mathematical structures used to model and represent relations between objects.

In the Graph Theory, graphs are sketched using vertices (also called nodes, site or actors) to represent objects or entities, and edges (also called links, bonds or ties) to represent the connections between them. By using this strategy, it becomes easier to pin down certain kind of problems and systematize solutions. Leonhard Euler raised this theory in 1736, when he gave the mathematical solution to the Konigsberg Bridges problem. Konigsberg is a city not far from St. Petersburg, which is crossed by the river Pregel. In a period of economic prosperity, seven bridges were built to serve the city, linking different areas and giving rise to a popular brain-teaser among the inhabitants of the city: is it possible to cross all the seven bridges without crossing the same bridge twice? The layout of the city at that time had five bridges linking the natural island of Kneiphof and two more bridges linking other parts of the city (Figure 11).

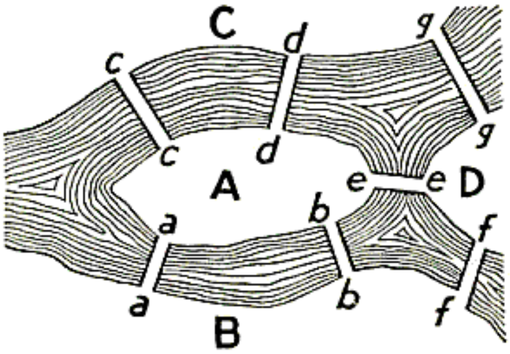


Figure 11 – The Konigsberg Bridges Problem ((Kraitchik, 1953) quoted in (Weisstein, n.d.)).

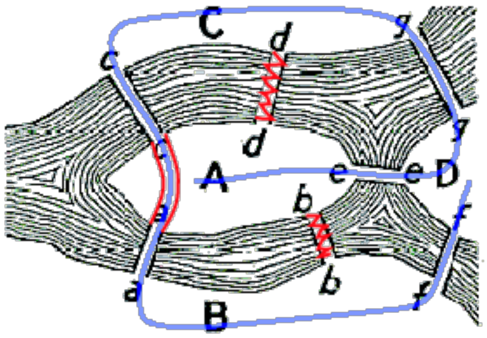


Figure 12 - The Konigsberg Bridges Solution ((Kraitchik, 1953) quoted in (Weisstein, n.d.)).

In order to solve the problem, Euler made a graph representing each area of the city as a node with letters and the bridges as lines, linking the nodes (Figure 12). This systematization would allow him to find the answer to the problem and establish a rule about networks: nodes with an odd number of links must always be the starting point or the endpoint of a pathway. By looking into the graph and applying this law the answer is negative, but by adding one more bridge the answer turns to positive. The fundamental aspect to retain about this example is that networks have properties and by changing the layout of networks we change the properties of these networks, which has an impact on their behavior. Graph Theory became the main mathematical language to study the morphology and behavior of networks.

Nowadays, Graph Theory, and more particularly, Network Theory, are used to study the connections and relations in different scientific fields like sociology and social anthropology (Hage & Harary, 1983), chemistry (Hansen &

Jurs, 1988), neuroscience (Sporns, 2003), computers and communications (Faloutsos, Faloutsos, & Faloutsos, 1999) and so on. They are utilized to find all sorts of properties and solve problems rooted in the deep mechanics of the underlying networks. Some applications are related with the study of propagation of informatics virus and the structure of the World Wide Web (Han & Tan, 2010), spread of infectious diseases (Hethcote, 2000) and scientific collaborations (Barabási et al., 2002).

The social network analysis accomplished in sociological and anthropological studies aims to understand and systematize behavioral patterns within groups involved in some kind of interaction. In order to accomplish this task, social scientists adopted and recognized the mathematics behind the Graph Theory as a useful resource towards a more quantitative methodological approach.

> *"Over the past decade, there has been an explosion of interest in network research across the physical and social sciences. For social scientists, the theory of networks has been a gold mine, yielding explanations for social phenomena in a wide variety of disciplines from psychology to economics."* (Borgatti, Mehra, Brass, & Labianca, 2009)

As mentioned before, a network is a set of vertices connected by edges. In social sciences, the vertices usually represent people and are designated by actor and edges represent a certain kind of interaction happening within people, called ties. For example, in the Facebook application users represent the actors and ties are the friendship relations. In such a case, when someone is friends with another it means that the relation is reciprocal or symmetric: both are friends of each other (undirected edges). But if we are talking of a geographical network of roads and places, where the former are considered to be the edges and the last the vertices, some edges might be one-way only (directed edges), since not all roads have two-directions. Moreover, depending, for example, on the average of traffic, edges might present different weights and be classified differently according to the road type (for example, distinguishing between highways and roads). It is also possible that one edge has the same vertex as starting and end points, take for example a short road used to reverse direction. On the other hand, the same network might allow different types of nodes. Keeping with the example of the road, we could distinguish between two different places within the same network: gas stations and hamburger restaurants. These two places would have different representations on the graph, for example, by using a color code for the vertices.

Each vertex can have several connections, either direct or indirect, classified as *in* or *out* according to its direction (*out edges* pointing away from the vertex and vice-versa). The number of connections of each vertex is called the degree of connectivity, which can be classified as *in-degree* or *out-degree* depending on the direction of the edges. A vertex with a degree clearly above the average of the network is considered to be a *hub*. The existence of hubs and the way they connect among themselves in a network will lead to distinct networks topologies, assuming singular properties and behaviors (e.g. Figure 13).

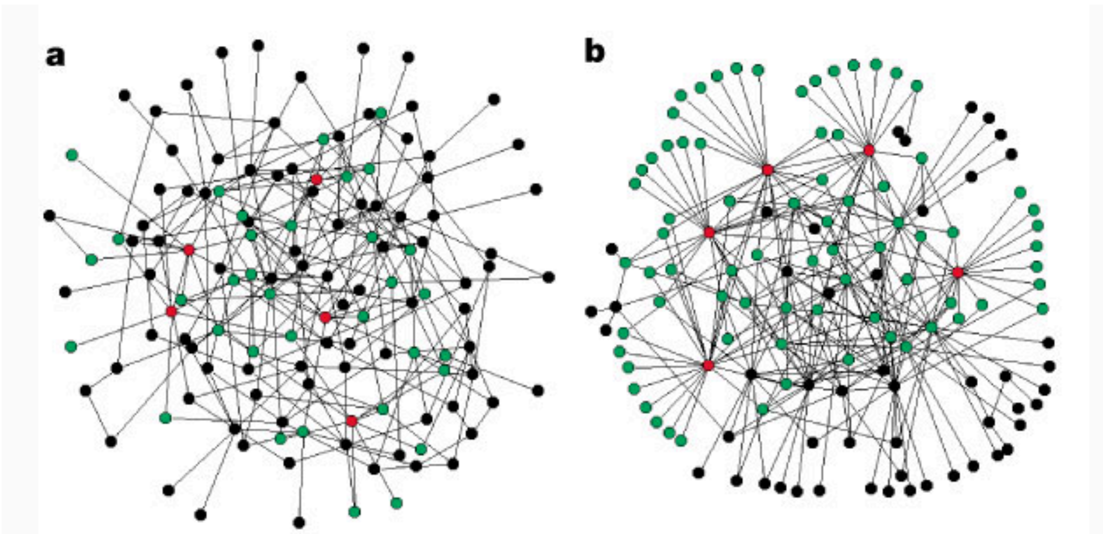


Figure 13 – Example of two network topologies. "***a,*** *The exponential network is homogeneous: most nodes have approximately the same number of links.* ***b,*** *The scale-free network is inhomogeneous: the majority of the nodes have one or two links but a few nodes have a large number of links, guaranteeing that the system is fully connected"* ((Albert, Jeong, & Barabási, 2000)).

The *component* to which a vertex belongs is comprised of all the vertices that can be linked to it throughout a path of edges, regardless of its number. In a directed graph (where edges possess directionality) there is an *in-component* and an *out-component*. In computer networks, the number of edges from one vertex to another is called the *hop count* (i.e. the number of steps between intermediate devices such as routers), while the shortest path (number of edges) between any two vertices of a network is called the *geodesic path.* Finding the longest geodesic path within a network gives the *diameter* of the network while the *average path length* is calculated by averaging the entire geodesic path within a network. This property or measurement represents the average number of steps one has to go in order to link any two members of a network. It is an important value but used isolated provides little knowledge about the network topology.

The *size* of a network corresponds to the total number of nodes; wile the *density* is defined as the ratio between the existing edges and the maximum possible edges. This measurement is closely related to the average of all nodes' degrees mention before and both measures are important to assess the *connectedness* of a network, which relates to its topology. This is also associated with the concept of *clustering coefficient*, which measures the density of a node's immediate neighbors. In other words, the clustering coefficient tells you how many of your friends are also friends with each other, it is the ratio between the existing links and the maximum number of those links. The clustering coefficient of a whole network is the average of the cluster coefficients of all individual nodes.

4.2.1 - Properties of networks

We know by now that networks are nodes that can connect to each other through edges. Moreover, we know it is possible to measure, classify and extract meaningful information from the way these nodes and edges agglomerate. Now it is time to focus on how real-life networks work and what their topologies are like.

Paul Erdös and Alfréd Rénys in 1959 were the first to present a model for the formation of networks, known as *Erdős–Rényi model*. Their model explained that a network of interconnected nodes emerges by connecting an edge between each pair of nodes with equal probability (Figure 14). For them, networks had a random distribution with small variations on the degree of each node. Their contribution was important, since they concluded that with high-density networks, messages could travel through the entire network, demanding only one link for each node. But the application of the *random graphs* proposed by Erdős–Rényi model is of little application in our complex world. As became clear in later years, networks rarely have a random distribution.

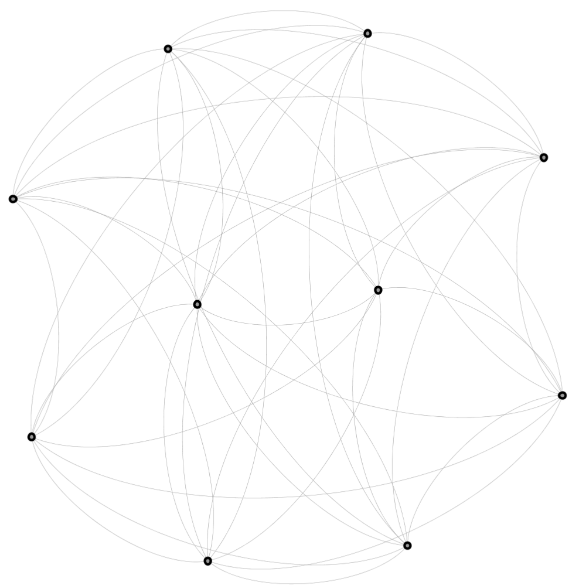


Figure 14 - Random Network

The Small-World Experiment (Milgram, 1967) conducted by the American sociologist Stanley Milgram (best known for the polemical Milgram Experiment regarding obedience to authority), paved the way for the six-degrees of separation theory, which states that all people in the world are just six "handshakes" away from anybody else. Considering inhabitants of the planet as nodes of a social network, it would take only six hops to connect any two people, regarding their race, country or age. Participants in the study were asked to send a letter to a particular person, without addressing him directly. Instead, they had to send the letter to someone they knew on a first-basis which would repeat the procedure without breaking the rule. The results show that "*chains varied from two to ten intermediate acquaintances, with the median at five. […] is, in certain ways impressive, considering the distances traversed*." (Milgram, 1967)

More than thirty years later, (Watts & Strogatz, 1998) devoted themselves to the study of small-world networks and discovered that what makes the six-degrees of separation possible[10] is that we live in small-world networks which are "*[…] partly disordered but not random, are highly clustered but have small characteristic path lengths, like random graphs.*" (Watts & Strogatz, 1998) (Figure 15). In a small-world typology, nodes are never many hops away from other nodes because few non-adjacent edges are enough to abruptly decrease the geodesic distance. However, these conclusions lead to a model for the formation of networks that can explain some behaviors but didn't provide an answer for how those small-worlds networks emerge.

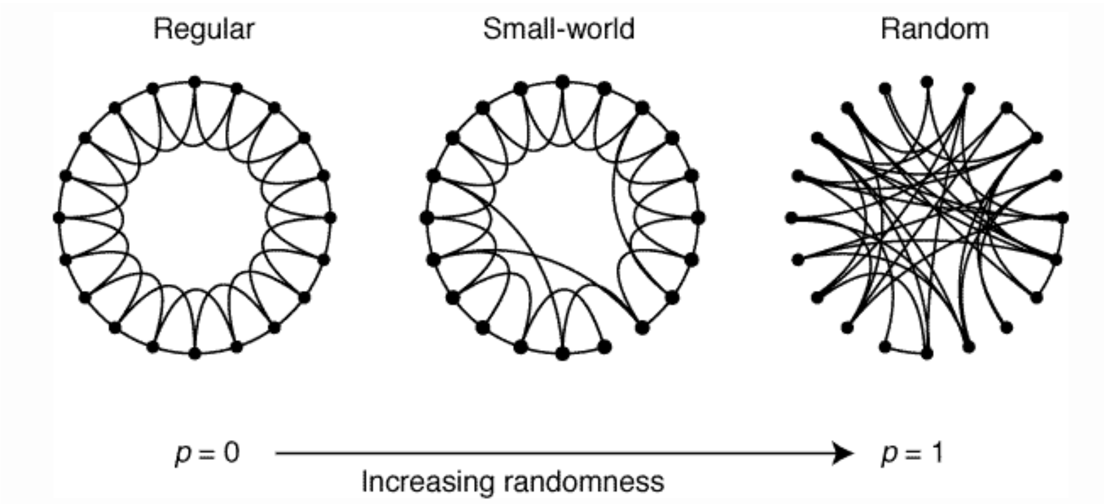


Figure 15 - The Small World Effect (middle) and Random Network (right) (Watts & Strogatz, 1998).

Using the same Watts and Strogatz data, Albert-László Barabási found the answer for the formation of networks that match the topologies of most re-

[10] Some authors contest this number. Actual Facebook average geodesic distance "[…] is 4.74, corresponding to 3.74 intermediaries or degrees of separation" 9/28/13 2:47 AM.

al-life networks, like the Internet, WWW or social networks. This model is called the Barabási-Albert model and provides an explanation for the creation of hubs, by means of *preferential attachment.* In this model, an edge has a higher probability to connect with nodes with higher degree, which tend to make “rich nodes richer” (Figure 13, b). These networks follow a *power-law* degree distribution, which results in a less steep decay than random networks, thus such networks are called *scale-free* since no “typical node” exists. They usually have very short average path length and the clustering coefficient tends toward zero.

Most common networks show a distribution proposed by Barabási, with few highly connected nodes (hubs) and a long tail of poorly connect nodes. But the way hubs connect among them is also an important and distinctive property of networks. In order to better understand this mechanics let us introduce the concept of *assortative mixing* (or *homophily*), which is “[…] *the tendency for vertices in networks to be connected to other vertices that are like (or unlike) them in some way. We consider mixing according to discrete characteristics such as language or race in social networks and scalar characteristics such as age*.” (Newman, 2003). The degree of individual nodes of the network can also be used to analyze its distribution pattern. When applying this criterion to social networks one observes that *“[…] social networks are often assortatively mixed, but that technological and biological networks tend to be disassortative.”* (Newman, 2002). This means that popular people are friends with popular people.

4.3 - Online Social Networks

After attaining on the mechanics of networks, in particular the scale-free networks that seem to pair and explain the typology and behavior of the world-wide-web and social networks, it is time to narrow into the topic of online social networks (OSN) and social media, which represent a practical application of these theoretical concepts to real-life situations. What are OSN and what is the fundamental roll of social media in our society? Why have these phenomena become so preeminent in so many areas of our lives, ranging from education (Maranto & Barton, 2010) to business (Kaplan & Haenlein, 2010), entertainment (Cheung, Chiu, & Lee, 2011) or activism (Ghannam, 2011), just to mention a few?

The first step is to distinguish between the different concepts surrounding this topic. The stem concept is 'Social Network', which, as we have discussed thoroughly in the point before, can be described as a network of people. That is to say: a group of people that have some kind of link between them. The nature of this link is not fundamental, as it can take on a myriad of typologies: physical/geographical (places visited), emotional (friendship, romantic), professional (co-authoring in scientific publications), entertainment (MMOG), etc. In any case, for the sake of this study, we will address Social Networks mainly as a network of people tied by friendship relations, which can vary in depth and nature (from best friends to acquaintances; from family to work partners).

Adding the prefix 'Online' to Social Networks, we are addressing Social Networks which have a particular mode of coalescing their links, by utilizing Information and Communication(s) Technologies (ICT) to accomplish and maintain the interactions[11]. OSNs are typically built on top of web-based services, allowing users to sign in and create personal profiles as part of a digital identity, which is complemented by their "*patterns of social interaction in (…) [the online] community life*" (Zhang, Jiang, & Carroll, 2010)[12]. In order to establish a network, users are normally able to add "friends" (other people) with whom they share a connection and exchange different layers of information,

---

[11] In 9/28/13 2:47 AM ICT is described "as an extended synonym for information technology (IT) to emphasis the role of unified (tele)communications, the integration of telecommunications with computers, as along with the necessary software, middleware, storage and audio-visual systems that enable users to create, access, store, transmit and manipulate information". Address the document for a discussion on the concepts of ICT and related terms.

[12] Digital identity in social networks can also be addressed from the Social Identity Theory perspective, which "assumes that on parte of the self-concept is defined by our belonging to social groups." 9/28/13 2:47 AM

not only confined to their profile (name, gender, age, place of birth) but also regular publications regarding their current life.

Wile OSNs is the term used to describe this concept, Social Network Sites (SNS) is the name given to the different websites that provide this service, which are sometimes defined as Social Web Sites (W. Kim, Jeong, & Lee, 2010) and show a close relation to Social Computing, defined as *"(…) any type of computing application in which software serves as an intermediary or a focus for a social relation."* (Schuler, 1994)

These websites (like Facebook or Twitter) are also generally addressed as Social Networking Services, although this term encompasses more than only web technologies.

4.3.1 - Why are we going online?

As we have seen before, people can engage in OSN for a great number of reasons and goals, but what seems to be the pattern is that most of the networks established in cyberspace are virtual versions of social networks existing offline. A study concluded that *"(…) emerging adults' offline and online worlds are connected and they use online communication for offline issues, and to connect with people in their offline lives."* (Subrahmanyam & Reich, 2008). This conclusion points the direction of a parallel online world that does not substitute but rather expands the offline world. This is also supported by studies using geo-social metrics which show that *"there is a vast portion of users with short-distance links and that clusters of friends are often geographically close."* (Scellato, Mascolo, Musolesi, & Latora, 2010). However, one it is not a perfect mirror of the other, the overlap is imperfect and the most active interactions in offline worlds do not always find a match in the online world. Nevertheless, some websites with social network characteristics are indeed specialized in gathering strangers that are eager to meet new people and create new links. That is often the case of dating websites were people go to find a romantic relationship. Such websites do promote social relations, by linking people with individual online profiles and matching expectations, but they usually lack the network component, where each user can be linked to other users, and browse these other users connections. Such websites stay middle way between social network site and a recommendation site, since the goal is well defined and can be usually defined as: finding the best match. On the other hand, "*what makes social network sites unique is not that they allow individuals to meet strangers, but rather that they enable users to articulate and make visible their social networks.*" (Boyd & Ellison, 2007)

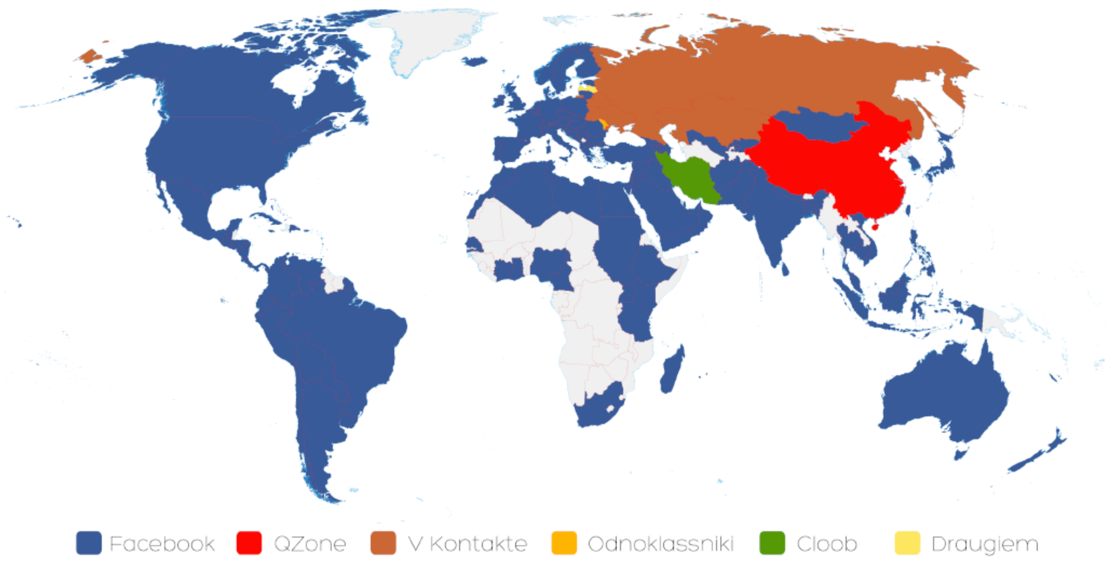


Figure 16 - World Map of Social Networks - June 2013. Facebook has now 1.155 billion monthly active users. Africa and Lat. America are the largest continents with 346 million users, Asia 339 million, Europe 272 million, US & Canada 198 million users (Cosenza, 2013).

There are a growing number of SNSs available and they have become one of the most influential technologies of our era (Figure 16). A curious indicator of how this technology has spread is the number of US patents granted, which include 'Social Network' as part of its name. The data shows an increase from 2 patents in 2006 up to 77 in 2012 (Figure 17).

It is quite reasonable for us to think of our time as the one when online social networks are more predominant than ever. However, tracing down the origins of online social networks it is not an easy task. The roots are definitely related with the invention of the Internet and first Computer Mediated Communication (CMC) systems. Both technologies paved the way for the emergence of online communities, that would succeed later on when computers became more powerful and the bandwidth increased.

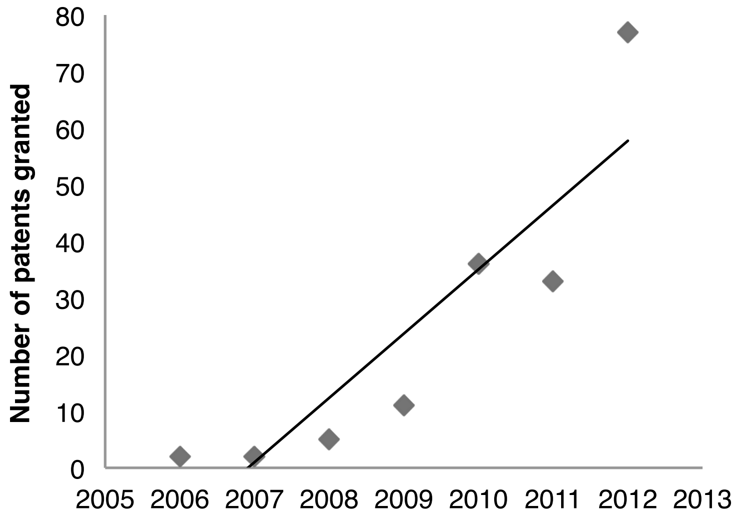


Figure 17 - Number of patents granted in the USA containing the word “social network” in the name.

An example of a former CMC technology was the Telenet (Mathison, 2012) which in 1978 allowed US costumers to dial into a Computing Conference System (CCS) with local calls in any of the 90 major cities. By doing so, users were informed which of their connections was online and ready to engage in a chat. Besides, users could read old messages and engage in “discussions” that correspond to actual group forums. (Hiltz & Turoff, 1993) By that time, the first steps of online social networking had been taken, though the focus was originally pointed towards CSCW. Yet, in France, in 1983, another pre-World Wide Web technology was about to take off: the Minitel. This videotext online service allowed users to buy transportation tickets, keep track of their phone contacts and engage in social interactions. Meanwhile, an unexpected twist for the company was the massive emergence of adult chatting (*messageries roses)* which overloaded the entire French national network during their first flood of acceptance. About this event, (Rheingold, 1993) would write what could justify todays craze for online social networking:

“(…) big institutions often think of CMC as a kind of database, a way of broadcasting information on screens to large populations

who spend their time interacting with information, but populations of citizens almost always use CMC to communicate with each other in new ways unforeseen by the system's original designers. People everywhere seem more interested in communicating with each other than with databases.“

With this in mind, and by the same time, when the explosion of the WWW and broadband took place, (Hiltz & Turoff, 1993) foresaw what is now, more than ever, a picture of our urban social reality.

"(…) we propose that the widespread of availability of human communication via computer will mean the ultimate replacement of urban networks as a basic for of social organization in postindustrial society by national and eventually international networks.” (Hiltz & Turoff, 1993)

After the first CMC services that would show some of the features found in actual social networks, by the mid nineties, the first online social network sites started to emerge (Figure 18).

“(…) the first recognizable social network site launched in 1997. SixDegrees.com allowed users to create profiles, list their Friends and, beginning in 1998, surf the Friends lists. Each of these features existed in some form before SixDegrees, of course. Profiles existed on most major dating sites and many community sites. AIM and ICQ buddy lists supported lists of Friends, although those Friends were not visible to others. Classmates.com allowed people to affiliate with their high school or college and surf the network for others who were also affiliated, but users could not create profiles or list Friends until years later. SixDegrees was the first.” (Boyd & Ellison, 2007)

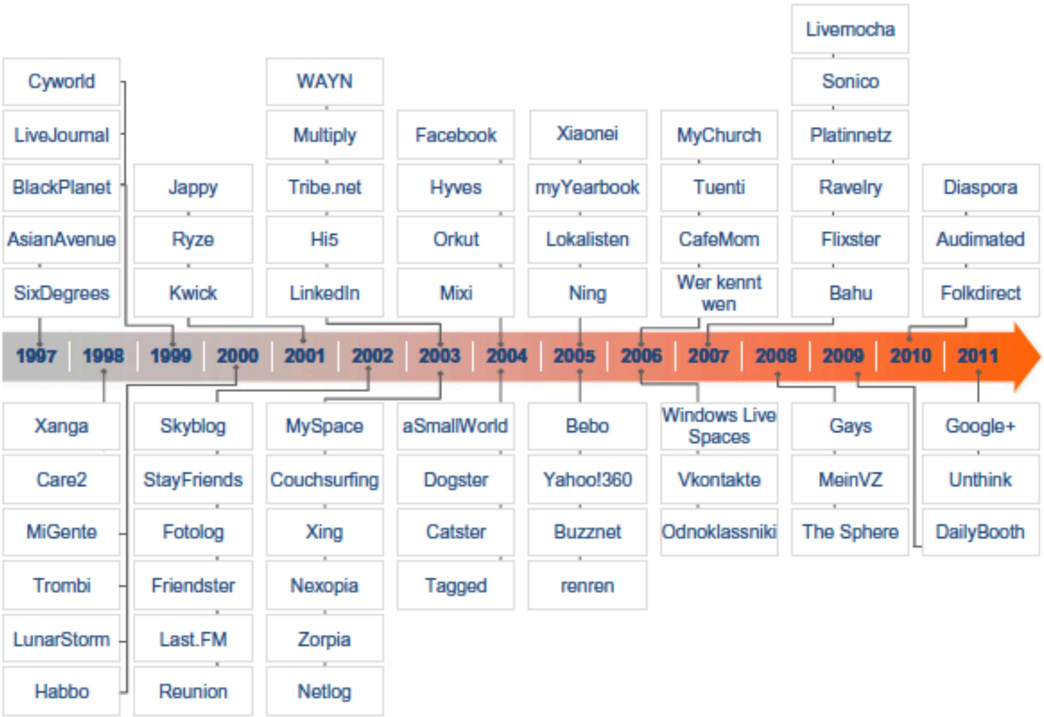


Figure 18 – A timeline of the foundation of selected online social networks from 1997 to 2011 ((Heidemann, Klier, & Probst, 2012))

4.3.2 - Social Presence and Media Richness

The actual landscape of SNS is dense and its distribution follows geographical patterns, supporting the idea that ONS work fundamentally as a layer on top of offline social networks. Modern interactions in the SNS are still following the same pattern of sharing interests and activities. However, besides the original textual messages present in the early form of ONS and still widely used nowadays (ex: status updates on Facebook), other types of data can be exchanged between peers of the network. These include but are not limited to photos, videos, animations, emoticons, sounds (voice messages, music, field-recordings), judgments (likes, '+', poles, ratings, etc.) and geo-localization.

According to the social presence theory developed by (Short et al., 1976), where social presence is defined as "*the degree of salience of the other person in the interaction and the consequent salience of interpersonal relationships*", different communication media have different degrees of social presence. We can think of it as the awareness one has of the other during a communication process. Face-to-face (FtF) interactions represent the highest state of social presence, since all the cues conveyed by both interlocutors can be sensed by each other (voice, facial expressions, body language). The effectiveness of a medium can thus be classified by the degree of social presence it affords, which in turn is impacted by the intimacy and immediacy of the medium. Intimacy is used to describe the level of mediation occurring between the emitter and receiver. In FtF communication one finds a great level of intimacy compared, for example, to a voice over IP conversation. A midway scenario would be a visiting area in a prison where people are separated by a wall of glass. In this case, there is no visual mediation but the acoustic communication is totally mediated by the use of a telephone. Immediacy is related with the synchronousness of the communication and also affects the degree of social presence of a medium, for example, an e-mail message is considered to have a low immediacy level when compared to an instant messaging chat.

In the media richness theory, richness is defined as "*the potential information carrying capacity of data. If the communication of an item of data, such as a wink, provides substantial new understanding, it would be considered rich. If the datum provides little understanding, it would be low in richness.*" (Daft & Lengel, 1983). A richer medium is considered more effective as it accounts for the decrease of uncertainty by the resolution of ambiguity. Adding a layer of geo-location data on top of a text-message with a friend's address, will aid the understanding of the information, as it can be visualized directly on a map and the best route can be presented.

Both social presence and media richness theories are important as they can be applied to virtual environments and online social networking. On our study we propose a new layer of information being exchanged between peers of a social network, which promotes the increase of richness and social presence.

4.3.3 - Mobile in online social networking

It is clear by now that advances in technology have a determinant impact in the way societies are molded and evolve. Regarding communication technologies (Rheingold, 2001) pushes the bar stating that "*Social revolutions follow communications revolutions.*". One of these revolutions – widely discussed in this text so far – is the mobile technology. We have seen how mobile devices have altered the way people communicate, consume music and interact with the environment. Mobile technology allows an always-connected experience, bringing computers and communications into the people's spaces, instead of the opposite. Therefore, as expected, mobile computing and mobile communications influenced how virtual communities gathered, and impacted on the dynamics of OSN. (Rheingold, 2001)

Furthermore, some researchers argue that mobile devices used in communication, because they shrink our network by keeping people closer, have become affective technologies.

> *"That is, objects which mediate the expression, display, experience and communication of feelings and emotions. Users enjoy an affective relationship with their phones and feel attached to them. This is partly due to the intrinsic affective character of human communication, and also because mobile phones are close to the body. They are an extension of the human body at the same time that they extend and augment its abilities. Emotional attachment is enacted in the personalization of handheld devices and services. Mobile phones are not only an extension of the owner's presence, but they also allow the virtual presence of those linked to us by phone communication. Thus, they become an important element in the building and maintaining of groups and communities."* (Lasen, 2004)

Online location-based games[13] are a good example of how mobile technology may enhance the online social networking experience. In this case, interaction within nodes of the network has a specific goal determined by the game design, which is location-dependent and thus reliant on of mobile technology. Moreover, traditional social network sites started to include location features that were target to the mobile technology, transforming into geo-social networks. For example, one of those features is the geo-tagging of pictures, which adds to the picture metadata the location where it was taken. Other services include the sharing of users actual location within their friends in the network, which may lead to casual encounters when both are at the same geographic area (Foursquare, Twitter, BrightKite, Facebook Check-in). Besides location-based features, mobility also enhances the offline interaction. For example, in the classroom it is now usual that students use their mobile devices (and a SNS) to share documents and materials related with the topic of the project being taught and communicate one each other when in different rooms.

4.3.4 - Social Media

The concept of Social Media (SM) is trendy but not always clear what it designates. We have decided to include it here since ONS can be considered part of the Social Media phenomenon, which is broader and may provide the fundamentals to ground the former concept. We may define Social Media as "*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content.*" (Kaplan & Haenlein, 2010).

Web 2.0 is the participatory web where users are at the same time consumers and producers. It is the Web of the crowd, of the ratings and comments on products and content of others. One may say it is the democra-

---

[13] *"(...) mobile games in the more specific sense that exploit spatial location as part of the game design"*9/28/13 2:47 AM 9/28/13 2:47 AM. See, for example, the game Ingress by Google (http://www.ingress.com).

tization of the online world, where social intelligence takes over individual intelligence. An example of Web 2.0 in opposition to Web 1.0 is the collaborative and open encyclopedia Wikipedia against the Encyclopedia Britannica. Web 2.0 is a concept for the WWW, supported by special technological solutions, such as Adobe Flash, RSS and AJAX, which allowed the participation and interaction with users. Since 2009 the term Web 2.0 started to decay in popularity and SM took off, supported by the increase popularity of SNS, this is particularly noticeable when comparing the Google search trend for both words (Figure 19).

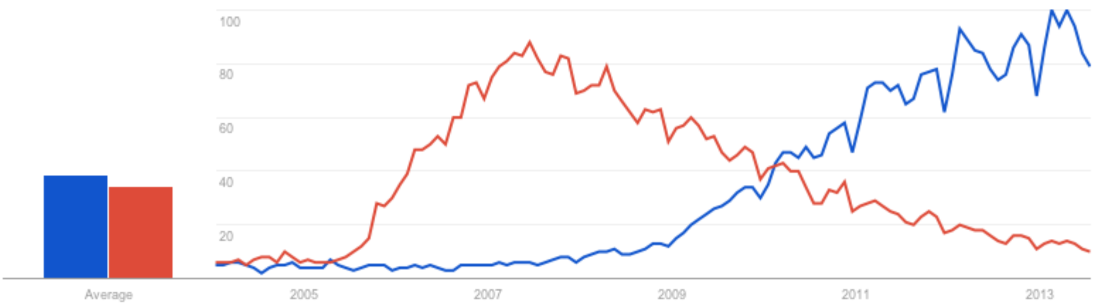


Figure 19 - Google search trends for words Web 2.0 (red) and Social Media (blue).

On the other hand, User Generated Content (UGC) is the outcome of the technological and conceptual foundation of the Web 2.0. It is the designation of users as content producers, instead of the typical broadcast and publishing agencies. Websites like YouTube, Vimeo, Wikipedia or Blogspot work as hosting services for housing all sort of rich media content.

In the same way, SNSs like Facebook and Twitter are also considered SM applications since they allow the upload and distribution of media by their users. Worth noticing is the fact that most popular services of SM are free of charge for basic users and have their business models based on premium accounts and/or advertising.

4.3.5 - Privacy issues

The protection of personal data is one of the holly grails of the network society. In an era where data is a commodity and everything seems to be able to be digitized (thus turning into data), privacy gains a whole new meaning. The Social Web of OSN and SM are constantly defying the boundaries of privacy, to a point where it is plausible to question if the distinction between private space and public space will still be valid in the future. That is the question raised by (Manach, 2009) in the article *La vie privée, un problème de vieux cons?*[14] where the author debates how the new generations of digital natives will understand privacy under the light of frivolous broadcasting on SM and OSN. Professor Ravi Sandhu compares the *data revolution* felt by the first generations of Web 2.0 users with the sexual revolution experienced by young couples during the 60s and 70s *"For a while, there were very few inhibitions and people did extremely risky things. That's where we are now with sharing information. (…) Over time, people learned, 'Hey, this is not risk-free.' At first the freedom was attractive, but then everybody realized the risks."* (quoted in (Peppers & Rogers, 2009)). Personal data can assume various forms and natures, ranging from the banc account number, to the private sexual pictures exchanged between lovers over the web. In any case, the discussion in this scope it's not so much on 'how safe is my information?' but rather 'how private is my information and how willing am I to make it public?'. That accounts for *self-preservation* and *self-discloser* in ONS, the former states that for any social interaction one always expect to controls others impressions about itself, and the later is the conscious or unconscious exposure of personal data. For example, communities of naturists[15] [16] do exist offline and are willing to expose their naked bodies online without great concern for privacy, they show a particular high degree of self discloser, when compared with other people.

---

[14] Private life, a fool old men's issue? (translation by the author). Due to the popularity of this article, the same was extended and publish in book one year later. Manach, J. M. (2010). La vie privée, un problème de vieux cons? (Vol. 10). Fyp.

[15] http://naturistcommunity.com

[16] http://perspective-numerique.net/wakka.php?wiki=ClubDesNaturistesNumeriques

This is an extreme example that most people would consider confined to a particular group of people, but the truth is that such behavior may be more common than one expects. According to a studied conducted at Harvard University, one youngster in every five and one young adult in every three already shared online personal photographs or videos naked or semi-naked (Palfrey, Sacco, & Boyd, 2008). This is particularly worrying if we think that those kind of audiovisual contents usually represent the currency for blackmail in cyber bulling and online harassment.

A study conducted in 2008 showed that users in general are not aware, precisely, how much of their personal information is available for the others online. One of the reasons may be the fact that many social apps (add-ons) for SNS like Facebook, request access to personal information in order to install and run. The authors defend that *"OSNs must clearly indicate the bare minimum of private information needed for a particular set of interactions. If an external application requires access to list of friends and nothing else, then the default should be that bare minimum."* (Krishnamurthy & Wills, 2008).

The use of location-based services[17] is one of the technologies that most frightens people regarding their online exposure. When combined to OSN it gets the name of Geo-Social Networks (GeoSN)[18] and two main issues emerge: *location privacy* and *absence privacy*. *"The former concerns the availability of information about the presence of users in specific locations at given times, while the later concerns the availability of information about the absence of an individual from specific locations during given periods of time."* (Freni, Vicente, & Mascetti, 2010). Besides location and privacy, there is also *co-location* and *identity* privacy issues, which result from the intersection of several bits of data gathered from different users, allowing the inference of a person's encounters and identification (Ruiz Vicente, Freni, Bettini, & Jensen, 2011). This is possible because data nowadays is highly interconnected and

[17] Mobile Location-based services can be achieved with several types of techologies, such as GPS, Cell ID, aGPS, Broeadband Satellite Network, etc. 9/28/13 2:47 AM.

[18] Also named Geo-Aware Social Networks.

our data is aggregated with others data, allowing prediction algorithms to run over large datasets, decreasing the error deviation. *"How machines see us depends on how our data connects to others. The tastes and interests of people who don't yet exist within systems can be easily predicted based on the patterns of others. And, when machines have access to a person's social network, the predictions are even stronger."* (Boyd, 2012)

According to a report from 2010, done by Forester Research, *"just 4 percent of Americans have tried location-based services, and 1 percent use them weekly"* (quoted in (Miller & Wortham, 2010)). This technology, though raising privacy concerns, is of great interest for marketers, which soon realized its potential to deliver an immersive multichannel experience for in-store shoppers. As before, the success of this technology will probably be driven by marketing demands. Shopkick is an example of a geo-location application (with some OSN characteristics), dedicated to offer bonuses to costumers who walk into stores. On the company's website[19] one can read:

> *"You just do what you love doing anyway—walking into your favorite stores like Target, Macy's, Best Buy, Crate&Barrel, Old Navy, Exxon and Mobil convenience stores, and more—and you automatically rack up über-versatile points called "kicks." No applications to fill out, and no purchase required. That's right—you just walk in with your smartphone and instantly earn kicks."*

Wile the market pushes the boundaries of self-discloser; new generations of digital natives seem more opened to accept these technologies and use them in their own social advantage.

Another privacy issue concerning ONSs is related to job hunting. It is getting increasingly more commons that employers check applicant's profile on OSN before hiring. On the other hand, several cases of firings have been repeatedly reported in the media, after employers checking employee's pro-

[19] http://www.shopkick.com/about

files and find non-acceptable or deviant behaviors according to companies principles (W. P. Smith & Kidder, 2010).

From this brief discussion on privacy issues regarding OSN, we conclude that the problem will not have a straight technological solution based on data control[20]. Developers determined to produce SNS have to have in mind that wile the boundaries are being pushed, there are still blocking thresholds that may refrain people to use a particular technology due do privacy issues. In any case, once again, technology is shaping the way we understand the world and re-think established concepts. Solutions for privacy problems will probably redound in a redefinition of privacy itself.

[20] As a consequent of recent scandal evolving NSA spying practices, Vint Cerf, one of the fathers of the Internet, asserted that there are no technological cure for privacy issues.
http://gigaom.com/2013/07/09/internet-inventor-no-technological-cure-for-privacy-ills/

CHAPTER 5 - State-Of-The-Art On Sound-Based Social Interaction Systems

In this point we present a survey of projects that show some scientific or commercial relation with the study of online social networks and are representative of the state-of-the-art R&D being produced in this domain. The criteria to include them in this list were: the overall quality of the projects (both in scientific and commercial terms) and the contiguity with some dimension of our project, in particular the mobility and the relation with the sound environment. A large number of projects that we consider also relevant for our research were left out of this list since they touch only very specific aspects of our project. For example, machine audition algorithms specialized in environmental sounds (e.g. (D. Smith, Ma, & Ryan, 2005)) or social networks based on uploaded sounds (e.g. Freesoud.org (Akkermans et al., 2011)).

5.1 - Survey of projects related with the field of OSN

❖ CenceMe

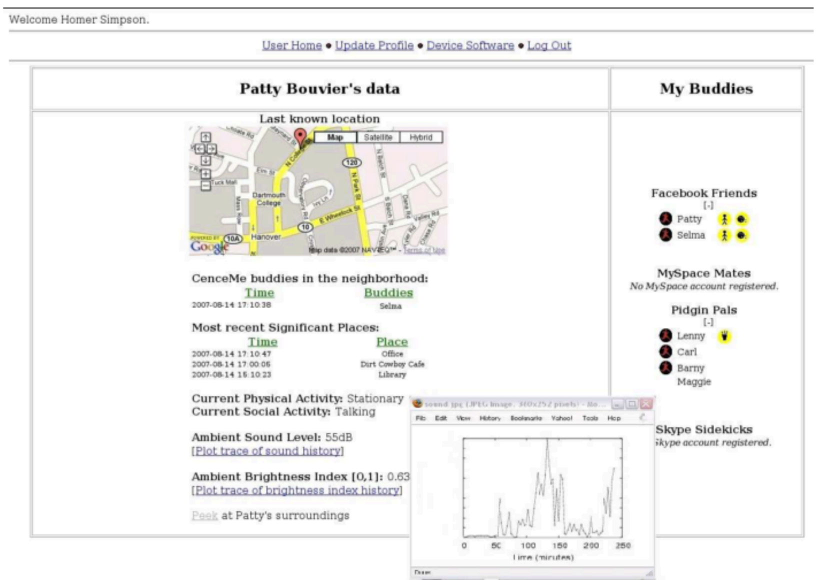


Figure 20 - CenceMe Web Portal

The CenceMe project (Miluzzo, Lane, Eisenman, & Campbell, 2007) is, so far and to our knowledge, one of the most relevant examples produced in the context of our research field. The authors have designed, implemented and tested key features regarding the three main areas of our study.

The project consists of a system that injects users sensing presence into the social networking interaction. The goals of the system are: "*(i) to provide information to individuals about their life patterns; and (ii) to provide more texture to interpersonal communication (both direct and indirect) using information derived from hardware and software sensors on user devices*". Though the main goal of the research does not address the study of soundscapes and its latent potential as an acoustic communication vehicle, the authors do use sound as one of the many sensed stimuli, which include: movement, color, temperature and location. All this data is gathered and properly mined in order to infer about user current status regarding: activity (standing/walking/running), mobility (stationary/walking/driving), indoor/outdoor and conversation/non-conversation.

The sound plays an accessory roll, especially dedicated to infer about the oral communication status of the user or groups of users. The application architecture is based on client-server paradigm and users can access the information regarding their friends sensing presence both through a website or common OSN plug-ins (Facebook, Gaim/Pidgin). Besides physical sensing (through built-in cameras, microphones, external accelerometers) the application also tracks virtual events such as profile changes in social networks or played songs on the device.

❖ Soundwalks

Figure 21 - Soundwalks map view

The Soundwalks project (Fink, Mechtley, Wichern, Liu, & Thornburg, 2010) is a social network website[21] based on a system for automated re-sonification of geographic sound activity. The goal is to virtually synthesize the soundscape of a place using an ontological framework based on acoustic, semantic, and social information. The data uploaded by users concerns geo-tagged audio files and manual tagging (concepts). The framework relies on the prior knowledge of acoustic and semantic ontologies combined with community-defined social links between sounds and concepts. To re-sonify specified locations through the playback of sounds in a database, the ontological framework is used to create a graph-based generative soundscape model. This project is particularly interesting for our research for three reasons: 1) it has an online community side, 2) regards soundscapes and its computational analysis through machine listening tasks and 3) translates sound into sound.

[21] www.soundwalks.org (not available by the time this thesis was submitted)

❖ Assessing contextual mood in public transport

This project (Maur, Costa, Pitt, Galvão, & e Cunha, 2013) is a pilot study which uses a mobile application to measure the user experience satisfaction on public transportation (or quality of experience QoE), through mobile sensing of the environment (sound and movement) and personal data gathered with a short feedback form. This project is interesting for our research since it uses a mobile sensing system based on the same contextual cues of Hurly-Burly.

❖ Ear-Phone: An end-to-end participatory urban noise mapping system

In this system developed by (Rana, Chou, Kanhere, Bulusu, & Hu, 2010) sound maps are created through participatory crowdsourced sensor data collection using mobile phones;

❖ Widenoise[22]

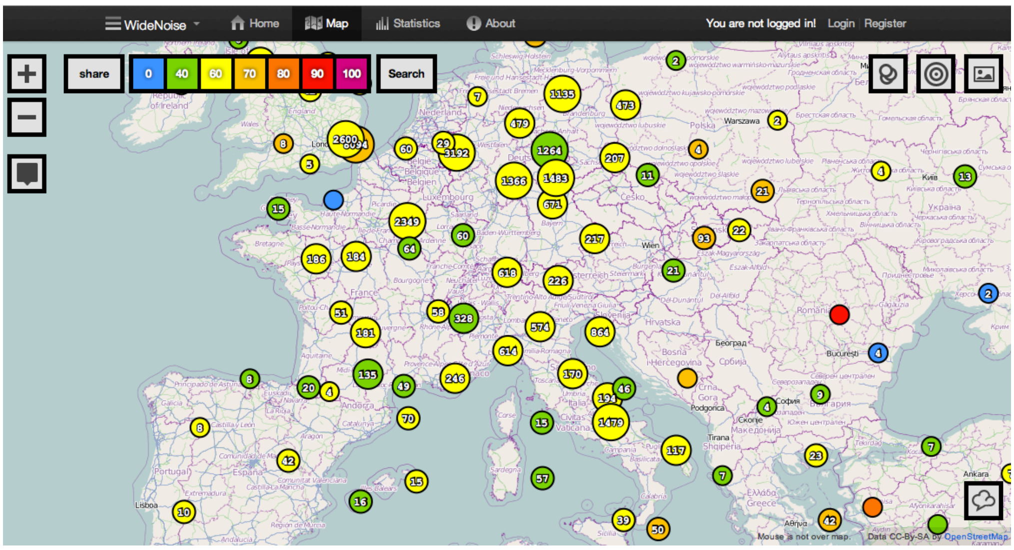


Figure 22 - Widenoise Sound Map

This project consists of an application for iOS and Android operative systems that transforms the devices in calibrated sound meters, measuring

---

[22] http://www.widetag.com/widenoise/ (accessed 12 June, 2013)

and creating a sound intensity map that can be accessed by users, in order to learn about their surrounding soundscape. Widenoise is jointly developed by the University of Wuerzburg and the University of Kassel in the context of the EU project EveryAware[23]. This application is also part of the UBICON platform, "*which provides a framework for the creation and hosting of ubiquitous and social applications for diverse tasks and projects, featuring the collection and analysis of both physical and social activities of users for enabling interconnected applications in ubiquitous and social contexts*" (Atzmueller et al., 2012).

❖ NoiseSPY

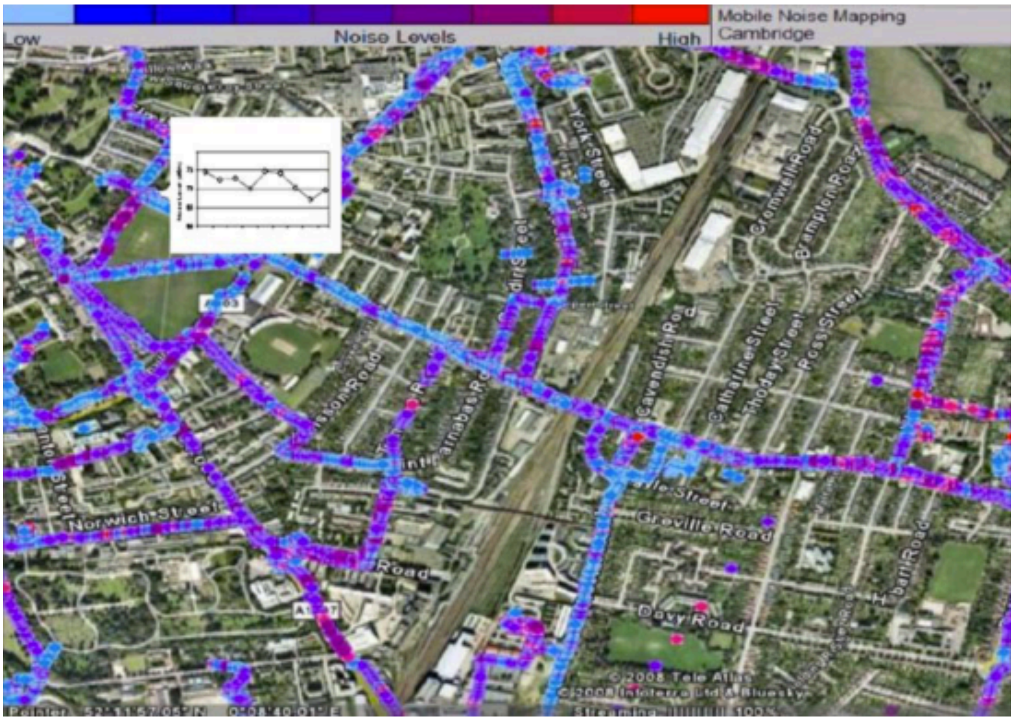


Figure 23 NoiseSPY Noise Mapping. Colored trace represent the loudness in dBA (Kanjo, 2009).

This project by (Kanjo, 2009) is a real-time mobile phone application for participatory urban noise monitoring and mapping. It transforms conventional mobile devices into low-cost data logger for monitoring environmental noise.

---

[23] http://cs.everyaware.eu/event/widenoise (accessed 12 June, 2013)

The system allows the visualization of collective data (sound intensity and GPS), allowing users to know in real-time which are the most quiet/noisy areas.

❖ The Quiet Walk[24]

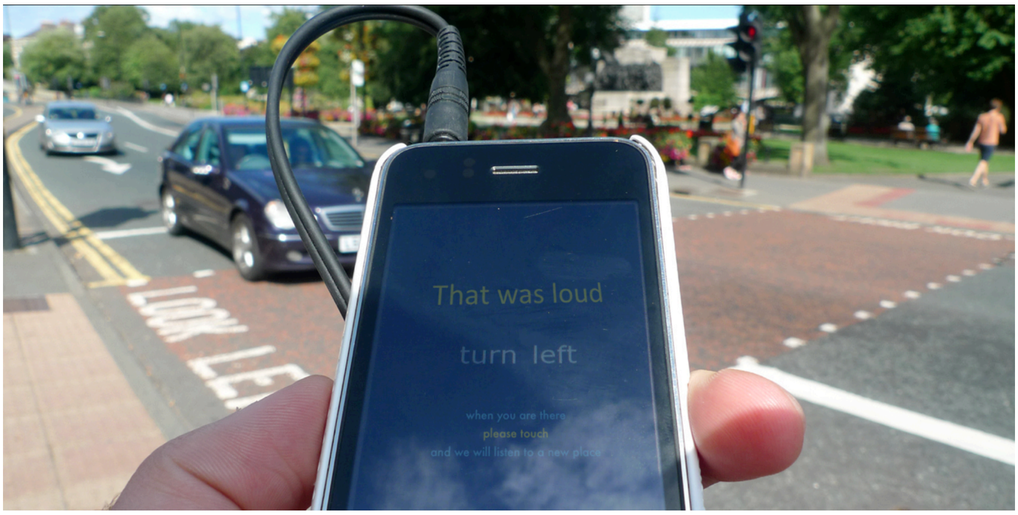


Figure 24 - The Quiet Walk GUI.

This project by Alessandro Altavilla, is a mobile application that allows the creation and update of a sound intensity map, collaboratively. The users can then drift around the urban space following the directions displayed by the application, which are calculated considering the level thresholds of the acoustic environment.

❖ Nericell

The Nericell project (Mohan, Padmanabhan, & Ramjee, 2008), from Microsoft Research India, is a system for rich monitoring of road and traffic conditions using mobile smartphones, where the recordings accomplished with the built-in microphone constitute the input of a honk-detection algorithm,

---

[24] http://dm.ncl.ac.uk/thequietwalk/The_Quiet_Walk/Alessandro_Altavilla.html (accessed, 12 June, 2013). This project is a work in progress and therefore does not feature in any scientific publication, yet; we are including it here considering the relevance of the institution (Culture Lab) and the author's portfolio/curriculum.

used to estimate the current traffic conditions of an area. Besides the sound data, other sensors are also used in order to draw a more precise picture of the road and traffic conditions, such as GPS and accelerometer (for bump detection). This project demonstrates how the sound source analysis can be used to describe a everyday situation/activity.

❖ SurroundSense

Figure 25 - SurroundSense, ambience fingerprinting system (Azizyan, Constandache, & Roy Choudhury, 2009).

SurroundSense (Azizyan et al., 2009) is a mobile phone based system that attempts logical localization using ambient light, color, sound, and user activity. By combining these contextual cues (including environmental sound) the system produces an optical-acoustic fingerprinting of the different places, which can be called to identify user location within the GPS margin of error.

❖ SoundSense

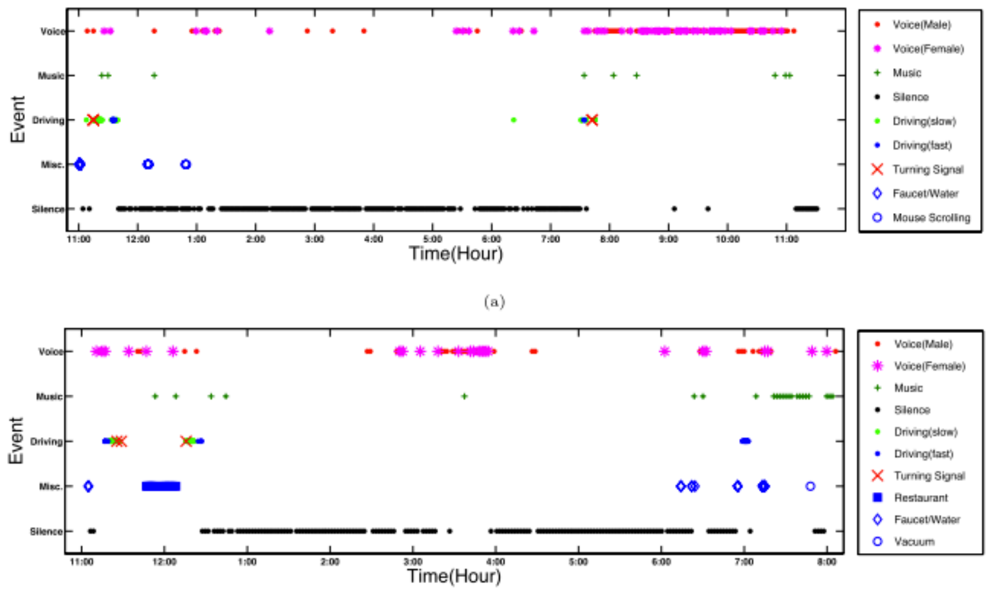


Figure 26 - SoundSense activity log visualization on two different days (H. Lu, Pan, Lane, Choudhury, & Campbell, 2009).

SoundSense (H. Lu et al., 2009) is a scalable framework for modeling sound events on mobile phones, implemented on the Apple iPhone and specifically designed to work on resource limited phones. SoundSense uses a combination of supervised and unsupervised learning techniques to classify both general sound types (e.g., music, voice) and discover novel sound events specific to individual users.

❖ Energy-efficient Ambient Sound Sensing and Classification Using Smart Phones

This project by (Chen, Du, Liu, Liu, & Mai, 2011) proposes an optimized algorithm for detecting indoor vs. outdoor locations based on the sound captured by the built-in microphone of smartphones. The project aims at finding the best tradeoff between classification accuracy and energy consumption through numerous experiments.

❖ VibN

Figure 27 GUI of VibN in Live (left) and Historical (right) views (Miluzzo et al., 2011).

The VibN project (Miluzzo et al., 2011) is a continuous sensing application for smartphones running on iOS and Android devices. Regarding the audio features, this application allows users to actively record comments and broadcast them, however, because the system is mainly based on opportunistic sensing (i.e. without the active participation of the user), it also records audio in background mode. The audio recorded this way is analyzed and all the parts containing speech are deleted in order to protect users privacy. The outcome of the participatory sensing is outputted in real-time as hotspots in a map, which the authors call Live Points of Interest (LPOI). Moreover, the application also affords an “historical mode”, where a record of the user sensing is kept.

❖ EmotionSense

EmotionSense (Rachuri et al., 2010) is a mobile sensing platform for social psychology studies based on mobile phones which, among other contextual cues, analyzes environmental sound captured by the built-in microphone. The system has the ability of sensing individual emotions as well as activities, verbal and proximity interactions among members of social groups by processing the outputs from the sensors of off-the-shelf smartphones.

❖ Darwin Phones

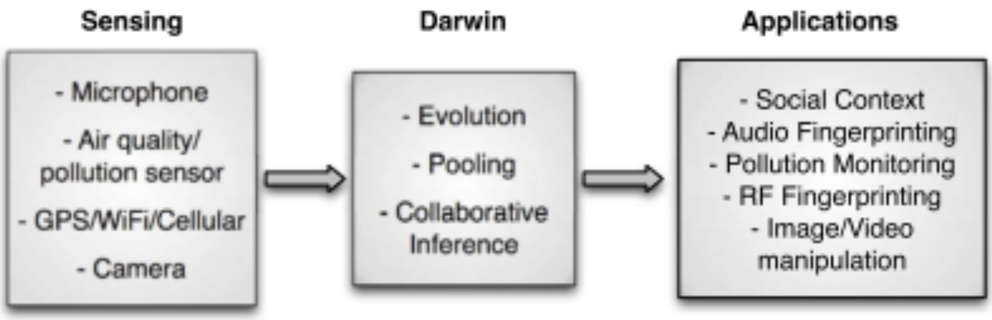


Figure 28 - Darwin system outline and applications (Miluzzo et al., 2010)

Darwin Phones (Miluzzo et al., 2010) is a framework for mobile phone sensing that combines cooperative sensing and classification techniques to reason about human behavior. The sensing options are wide and should be understood as an abstraction rather than specific procedures; nonetheless, authors evaluate the system using a voice recognition task to which eight different devices contribute to infer a result.

❖ Zagora

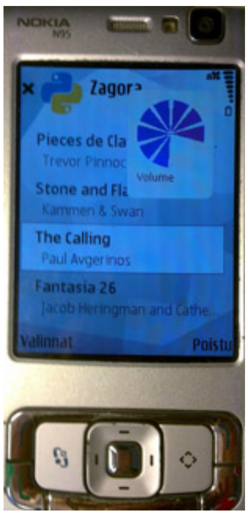


Figure 29 - Zagora GUI - Context-Aware Music Player Mobile Application.

Zagora (Oksanen, Kleimola, & Valimaki, 2009) is a context-aware mobile music player application that senses the environmental sound and selects a playlist accordingly. It performs advanced audio processing to differentiate between situations like street, restaurant, car, office, and meeting, and uses the situation information to filter down an online music catalog. This project is interesting for our research since both input and output are happen in a mobile device and are based on sound.

- **NoiseTube**[25]

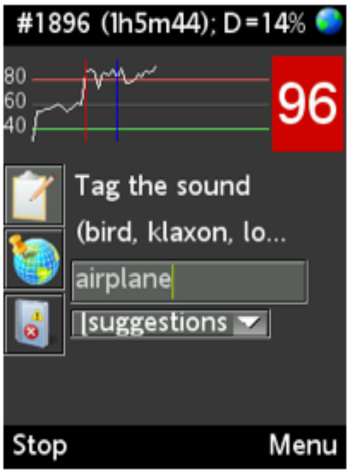


Figure 30 - NoiseTube Mobile GUI.

NoiseTube (Maisonneuve & Stevens, 2009) is a project which aims at turning GPS-equipped mobile phones into noise sensors, enabling people to measure their individual exposer to noise level in their everyday environment. NoiseTube has a communitarian dimension: users are invited to sign in and create a profile, which is complemented by annotations and a history of their sound level measurements. The main outcome is a participatory sound intensity map.

[25] http://www.noisetube.net/ (accessed 12 June 2013)

❖ Smule's Ocarina

Figure 31 - Smule's Ocarina World Listener view

Smule's Ocarina (Wang, 2009), is a mobile application which emulates the experience of playing an ocarina by blowing genteelly into the devices' microphone and tapping on the screen. Besides being a mobile sound application, the interest of this application for our research arises from its social dimension. Ocarina features a *World Listener* mode, which allows users to see themselves (and the other Ocarina players) represented on a world map view, and listening/sharing snippets of performance data, which are rendered into sound in their devices. This acoustic community is also available through a web portal, where users can login and share Ocarina scores.

❖ SoundFishing

SoundFishing (Midolo, 2008), is a prototypal application that senses the sonic environment and, according to pre-established rules, extracts particular samples of it. Regarding the relevance of this project for our research, we highlight its underlying goal, which aims at creating awareness for the richness of the daily-soundscape by framing parts of it that usually go unnoticed. Although the concept was designed for mobile devices, the implementation and evaluation were accomplished using a laptop computer.

❖ Scribe4Me

Figure 32 - Scribe4Me GUI - Example of an audio-to-text transcription.

Scribe4Me (Matthews, Carter, Pai, Fong, & Mankoff, 2006), is a mobile sound transcription tool for the deaf and hard-of-hearing, which is continually recording a 30s buffer of sound environment, analyzing it upon request and delivering a description via SMS to the user. This application is particularly interesting for our study, since it also focus on unconstrained sound sources rather than speech or music recognition. It was implemented on a PDA, though a human transcriber accomplished the classification manually.

❖ Recognizing workshop activity using body worn microphones and accelerometers

This project by (Lukowicz et al., 2004) is a research project where a wearable system is designed and implemented aiming to identify situations and activities based on the sound environment and movement of the users. This project though not implemented on mobile devices, uses the same inputs (contextual cues) as Hurly-Burly.

❖ Sonic City

This project by (Gaye, Mazé, & Holmquist, 2003) is also a wearable system that analyses several contextual cues in order to produce a sonic artistic output laying in the area of sound design, soundscape composition and electronic music. This project is interesting for our study since it focus on mobility and shares the idea of using sonic contextual cues (among others) to determine a sonic output.

CHAPTER 6 - Preliminary Research On Soundscapes And Sonic Interaction Design

In this chapter we present preliminary research related with the study of soudscapes and sonic interaction design. Two projects are described, representing peripheral research accomplished by the author and his team in the course of the PhD program. We have decided to include them as they provided us with clues and knowledge on particular phases and specific scientific areas related with the main topic. Moreover, and probably the most significant aspect of these projects, is that they all share the underlying goal of our main research: to promote sound awareness among people.

The two projects presented in detail are:

- *Murky Shooting – audiogame for mobile devices* (Cordeiro, Baltazar, & Barbosa, 2012; Cordeiro, 2011)
- *Massive multichannel setup for soundscape assessment* (Cordeiro, Barbosa, & Santos, 2013)

Both projects were subjected to academic scrutiny when presented in peer-reviewed conferences.

6.1 - Murky Shooting – audiogame for mobile devices

In this research project we developed an effective approach for game design based on auditory, non-speech, feedback. Next we'll describe the design and implementation of a mobile application – Murky Shooting - that makes use of sound to engage users in a shooting game experience. Further on, we discuss the pros and cons of the mobile version compared with the former desktop version and conclude with a presentation of preliminary results, which prove that auditory cues solely are sufficient to provide an effective way of interaction, allowing the user to improve his/her performance and keep immersion in the game narrative.

6.1.1 - Introduction

Videogames have conquered a respectful place in the media domain, engaging people from all around the world, within all age groups and social statuses. Recent reports reveal that 72% of American families play videogames and the average age of players is 37 year old, revealing that not only young people make use of videogames (Entertainment Software Association, 2011). The history of videogames show us that since the very beginning a huge effort has been put into turning the game experience as portable and ubiquitous as possible, thus allowing users to play everywhere. If it is true that some games are based on collaborative dynamics, inviting people to play in-group, one may also say that the game experience always entails a personal dimension, in which technology places the player against his/her own limits.

The portability of the gaming experience is reaching new levels of penetration due both to the evolution of handheld consoles and the dissemination of mobile games, which are videogames played on portable devices like mobile phones, smartphones, portable media displays, etc. Typically, mobile games face more technical constraints than the previous since the target devices are not primarily dedicated to accomplishing such demanding tasks (e.g. mobile phones). However, the ubiquity of such devices has made them suitable to be played by a growing number of people and, consequently, games have been developed to serve that demand. These constraints usually lead to simplifications of some aspects of the game like graphics, interactions, sound and/or narrative complexity.

The game we present takes this into consideration and pursues new ways of interaction between the game and the player, exploring the sound dimension not only as an add-on to the game but as the main vehicle for conveying information about the gameplay, crossing the thin boarder between videogames and audiogames (Friberg & Friberg, 2004) (Gärdenfors, 2003).

There are already videogames on the market that put the focus on the sound dimension, mainly by exploring music as the driving force for the game experience. However, the graphic dimension in this type of games still plays a significant role. In our project the human computer interaction happens in two different ways: 1) the user communicates with the computer using gestures, 2) and the computer responds with non-speech, non-musical sound. Typically, audiogames are oriented towards blind people. The use of sound as the primary medium in computer-to-human communication makes audiogames suitable for those with visual impairments. However, preliminary tests with sighted people (discussed ahead) showed us that the game is welcome and interesting also for this group. Eighty percent (80%) of the players, when confronted with the open question “Comments on the game?”, showed great appreciation and declared no significant faults. Some hypotheses for this fact are: 1) the game still holds a challenging dimension; 2) users appreciate novel ways of interaction. Furthermore, as a sense, sound has particular features that make it suitable to engage players in gamming activities. While using headphones, the player gets immersed in a personal sonic space, which disconnects him/her from the outside world, converging his/her attention to the game. Additionally, sound is very competent when transmitting relative (instead of absolute) cues. For example, between two notes it is easier to distinguish which one has the higher pitch and which has the lower, rather than identifying their individual frequencies (Brewster, 2002). Our game makes use of this perceptual evidence, allowing players to improve their skills over time.

6.1.2 - Murky Shooting - Game Design

Murky Shooting falls into the shooter game category. Like others, the goal of the game is to hit a target as many times as possible over a time span. A noisy crow perched on a high voltage cable represents the target. Its position on the cable changes over time and is made known to the player through the crow’s caw (sample sound), which changes its panoramic according to the crow’s horizontal position on the cable.

Different game levels are set by varying the amount of time the crow remains still in one place. Decreasing the time increases the difficulty. On the entry level the crow remains still during five seconds before updating its position, decreasing one second per each level.

The player has access to unlimited ammunition, yet it has to recharge his/her weapon every eight shots. This control feature was included to discourage players from indiscriminate shooting and is implemented differently in the desktop version and mobile version.

In order to aid the player during the hunting task there are auxiliary sound guides that indicate the crosshairs position. These guides are blip-like sounds (synthesized sound) that change pitch and frequency of occurrence consistently with the distance to the target. Moving away from the crow's position lowers the pitch and decreases the number of blips. When the crosshairs are on sight with the crow, a burst of noise is triggered. The player then has the opportunity to hit the crow and increment his/her score. A background sound composed of a quiet city ambient was added to the game in order to set the context for the action. A celebration sound plays every time the player hits the crow and a rifle sound is heard at every shot.

The game presents two distinct modes: 1) on "daylight mode" the player can see the crow and the crosshairs moving on the screen. This mode is used for practicing, since it allows the user to understand the match between sound and game action. 2) On "night mode" the player doesn't have any visual feedback besides the score and countdown.

6.1.3 - Desktop Vs. Mobile Versions

The game was formerly developed for desktop/laptop computers (Cordeiro, 2011). In the previous version, the player interacts with the game using a mouse or a joystick, although other interfaces such as the ©Nintendo Wii remote controller were also implemented with success.

On the desktop/laptop version the “load weapon” action is performed by pushing and holding the “c” key on the keyboard for three seconds.

The horizontal position of the crow is transmitted to the player by using binaural HRTF(Algazi, Duda, Thompson, & Avendano, 2001), instead of simple panning. The system provides three different head models in order to match the player-auditioning model as well as possible. The use of such a technique allows a more natural feeling of space but demands more processing power due to the underlying audio convolution process, and needs more hard drive to host the files for the HRTF database (4 MB). All audio in the former version (SFX and HRTF) is in AIFF format, 16 bit, 44.100Hz, occupying more than 40Mb of audio files.

The game was initially programmed using ©Cycling74 Max/MSP 5 visual programming language. The graphical user interface consists of a square illustration with the name of the game, control settings, the image of the bird and the crosshairs in daylight mode (Figure 33). This version was tested within a group of seven players which improved their score by 47,83% after playing four times, being consistent with the initial assumption that the game can be learned and mastered (Figure 34).

Figure 33 - GUI for “practice mode” (desktop version). The crow and the crosshairs are visible, along with the score, time left and other settings.

The mobile version of the game preserves the same fundamental aspects of the former desktop version, showing a great concern on the efficiency side. It was important to keep the same flow during gameplay, without any extreme downsizing on graphics, sound interaction or game narrative.

In order to achieve that goal, we chose the ©Apple iPhone/iPod devices to run the game. These devices are popular among our target users and demonstrate great processing performance. The game was programmed in C++ using a software framework called openFrameworks[26].

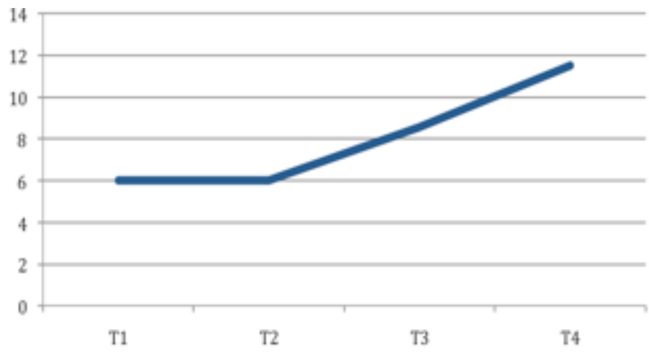


Figure 34 - Evolution of the score along four consecutive times played (x-axis: gameplays; y-axis: score).

The GUI (Figure 35) was resized to fit on the iPhone/iPod resolution (960 x 640) and most of the game settings became obsolete and were removed (head type, sound levels, audio DSP and controls). Both modes - daylight and night - were kept. The auxiliary guide sounds consist of FM synthesis and are processed in real-time on the device. The sample sounds (crow, shoot, celebrate) were exported as .wav mono files instead of stereo to save space. The background sounds, because of their longer duration, were exported as MP3 320kbs.

[26] http://www.openframeworks.cc

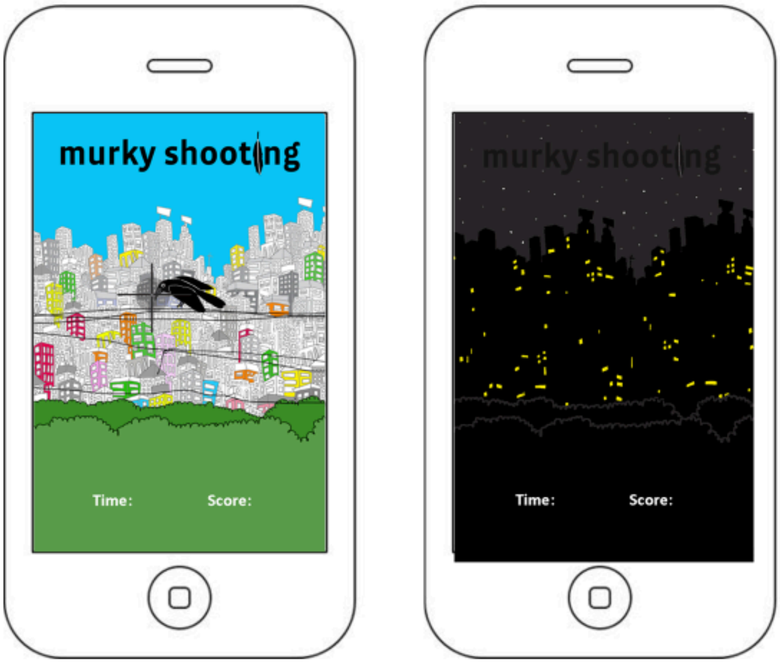


Figure 35 – GUI for mobile version. Left – practice mode; right – Game Mode

The most significant downsize was the change from binaural sound (HRFT) to simple stereo panning. This decision was made to keep the application less demanding in terms of processing power and smaller in size (avoiding the HRTF sound database).

The user input control also changed considerably. Instead of using a mouse or a joystick, the player uses the game device as a controller itself through its on-board three-axis accelerometer and touch screen. The cross-hairs position changes by tilting the iPhone/iPod on the y-axis, while tapping the screen to fire the shot. Loading the weapon requires the user to tilt the device 90º in the x-axis and return to the original position (if accomplished successfully a loading sound is triggered).

6.1.4 - System Evaluation And Assessment

For the system evaluation two tests were conducted. One to assess the binaural versus stereo sound environments, and another to compare the different gamming controls (laptop's trackpad vs. iPod/iPhone devices). The sample was made up of 13 people (12 male and 1 female), ranging from 17 to 38 years old, all with moderate-to-high game experience.

In the first test the subjects played both versions (binaural and stereo) using the trackpad, without knowing if there was an alteration. Later, they had to answer if they noticed any kind of difference. The results showed that 54% of the tested sample did notice a difference in sound and interestingly 71% of these actually preferred the stereo version, stating that finding the bird seamed easier.

For the second test, the subjects were asked again to play the game, this time using the iPod as controller. Afterwards they answered which controller they preferred and why – 25% chose the trackpad and 75% the iPod. The main reasons stated in favor of the trackpad were that is "more precise and easier to control". In favor of the iPod the main reasons were that "is more intuitive", "more interactive" and "more fun to play" with the downside being a "little more difficult".

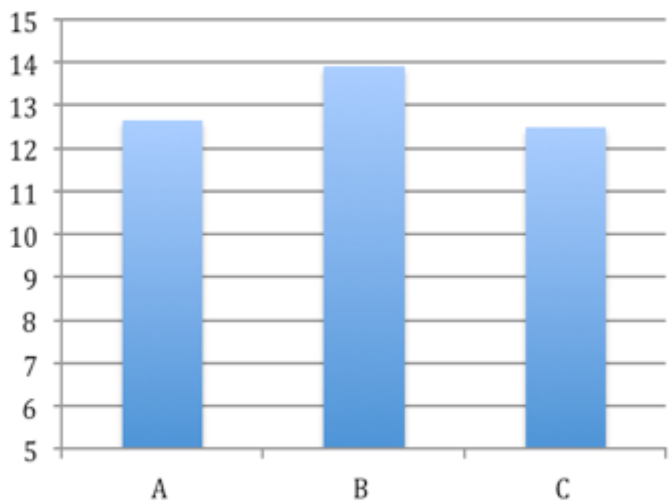


Figure 36 - Comparison of the score's average (y-axis) on different setups (x-axis): A – binaural/trackpad; B – stereo/trackpad; C – stereo/iPod.

To finish the evaluation one compared the average scores on the different setups: A - Binaural/trackpad; B - Stereo/trackpad; C - Binaural/iPod (Figure 36). These preliminary results show us that changing the audio from binaural to simple stereo and using different game controllers does not impact significantly on the game score.

6.1.5 - Conclusions

The use of sound as a complementary output interface for software applications is widely accepted among videogames companies and the scientific community (Collins, 2008). Murky Shooting uses sound as the main output interface, proving that sound by itself is enough to guarantee a neat interaction between game and player. The mobile version of the game uses the same game algorithm in a device with more technical constraints (reduced processing power, lower physical memory and screen dimension). In this new framework, the game proved to be equally efficient and engaging: 1) the gestural interface provides a tangible and intuitive way for the player to interact with the game; 2) the use of auditory feedback, although an abridged version, keeps the interaction fluent and coherent. Sound has features that help to overcome some of the visual constraints found in mobile devices, mainly the small screen dimensions which are a relevant downside when playing videogames, both at the input and output stages. The main contributions from this project to our main research were the validation of sound as a vehicle for communication between user and the environment (in this case, the game environment) and the possibility to test and gain technical proficiency regarding the use of openFrameworks in iOS and audio programming in mobile devices.

6.2 - A Soundscape Assessment Tool Based on a Massive Multichannel Loudspeaker Setup

This research is twofold: 1) it describes a comparative study between two listening tests methodologies and 2) presents a new soundscape assessment tool based on a multichannel loudspeaker setup and sonic interaction. The goal of this research is to enhance soundscape assessment methodologies by improving the quality and validity of laboratorial testing. Our proposal builds upon the concepts of “ecological validity” of auditory stimuli and “representative design” of the listening experience, both introduced by (Brunswik, 1956) in the context of psychological experiments on perception. Results suggest that representative design (mimicking real-world settings in laboratory) in soundscapes listening tests is better achieved through surround systems than binaural systems (using headphones). Moreover, users agree that the interactive system for real-time soundscape design helps raising awareness for the soundscape and auditory phenomena.

6.2.1 - Introduction

Soundscape is a term coined by Murray Schaffer as a synonym of “sonic environment” (Schafer, 1994). However, this term is also used to express a new positive attitude towards the sonic phenomenon, assuming sound as a resource rather than a waste. This perspective implies the enrolment of the listener on the assessment of a sonic environment, since the focus is no longer on the sound itself but rather on how people perceive it. Therefore, it requires the contribution of a large number of disciplines to fully comprehend the phenomenon, including – but not limited to - Psychology, Philosophy, Anthropology, Cognitive Science, Sociology, Acoustics, Composition, Sound Design, Physics and Biology.

Regarding noise pollution issues, some specialists (Kang, 2011; Lercher & Schulte-Fortkamp, 2003; Schulte-Fortkamp, Brooks, & Bray, 2007) have followed Schaffer’s footsteps and proposed solutions based on a soundscape approach, as an alternative to *deaf* solutions such as Community Noise (Donovan, 1975). They state that sonic welfare is not achieved exclusively by keeping noise levels down but by assessing how different communities understand and relate to their sonic environment, by identifying meanings, patterns and dynamics. In order to follow a soundscape approach for the mitigation of noise annoyance problems and promote the development of friendly sonic environments, measuring sound pressure levels is only one side of the process. The other side is accomplished through listening tests, using people to assess different dimensions of a soundscape.

This paper addresses the methodology for laboratorial listening tests and suggests the representative design approach and search for ecological validity as key elements for its success.

❖ Listening Tests

Soundscape studies make use of listening tests as a common research tool to assess and evaluate several dimensions of soundscapes, such as quality (Guillén & Barrio, 2007), annoyance (Schulte-Fortkamp, 2002) or loudness (Suhanek & Djurek, 2012). Listening tests can be accomplished in different ways and formats, depending on the circumstances and aims of the experience. They can be done in situ – in the actual scenario, in presence of the original soundscape being assessed - or in laboratory, where the sonic environment is recreated artificially.

Listening tests *in situ* are usually preferable but not always possible, since some studies consider non-real places or focus on particular dimensions that demand artificial manipulation (e.g. *auralization*). Typical approaches include sound walks (Birgitta Berglund & Nilsson, 2006), where listeners walk through a place describing/evaluating its soundscape or narrative interviews to people who are/were at the place (Szeremeta & Zannin, 2009).

Listening tests in laboratory are usually easier to control but they lack the sense of realism observed in *in situ* tests, both because of the artificial scenario and artificial acoustics provided by the playback systems. Typical approaches for this type of tests include headphone listening (stereo and binaural recordings or synthesis), listening to real sources brought into the lab (Steffens, 2003) and listening through loudspeakers (usually with a limited channel count) (Guastavino, Katz, Polack, & Levitin, 2005).

Listening tests are normally complemented with a questionnaire or other survey method.

❖ Representative Design and Ecological Validity

The lack or realism observed in perception tests occurring in a laboratory may compromise the legitimacy of the experiments, hindering scientific truth. This idea is expressed in (Brunswik, 1956) where the author presents the concept "representative design" as preparation for experimental conditions according to behavioral setting to which the results are intended to apply. That is, to approximate the laboratory situation to the real environment as far as possible, without oversampling improbable variables. This approach defies classical approaches where interference from the environment should always be avoided in order to optimally assess the proper variable. According to (Hammond, 1993) this position was firmly defended, inter alios, by Helmholtz or Wundt.

In turn, "ecological validity" is a concept also introduced in (Brunswik, 1956) to define the correlation between a perceptual cue and the distal variable to which it is related. For example, it defines how much an auditory stimulus can assist in identifying the actual source (or other variable: position, loudness). In (Gibson, 1979) Gibson presents an ecological approach to visual perception and defends that "The laboratory must be like life". However, this concept is often misused in scientific literature as a synonym for "representative design" (see (Araujo, Davids, & Passos, 2007) and (Hammond, 1998) for a discussion).

We see as necessary the use of representative design experiments to assess soundscapes in laboratory, since the definition of soundscape is highly dependent on the idea of environment and context of action. Therefore, it is important not only to match the acoustic environment of the original location by recreating other stimuli considered relevant for the variable being studied. For example, studies have demonstrated the influence of visual stimuli on listening (and vice-versa), stating that what we see influences what we listen to and how we listen ((Viollon et al., 2002)).

In the following experiments, in order to diminish visual bias and accomplish a representative design for the listening tests we have recreated not only similar acoustic stimuli but also coherent visual cues of the soundscape's original place.

6.2.2 - Experimental Procedure

In order to provide an answer for the stated problem, an experiment was conducted involving listening tests with 25 subjects. The experiment was divided in two parts: *Part I* - a comparative study between two sound reproduction systems; *Part II* - user test for an interactive soundscape design tool.

❖ Overview

The soundscapes used in the listening tests were recorded at a university cafeteria, on a Friday, between 12h and 13h. The cafeteria was showing light activity (some meals were starting to be served), with an occupancy rate around 25%. The place is approximately 110 m^2; ceiling height of 4m, stone walls and tiled floors. It contains 15 tables and 40 chairs. It is a highly reverberant space and the noise of chairs dragging on the floor plus chatting makes it a boisterous place at peak hours.

The experiment took place in a former warehouse recently converted into a MoCap studio, equipped with a multichannel surround system of 16 discrete speakers (Genelec 6010A). The room area is approximately 75m^2, with stone walls, concrete floors and minor acoustic treatment (black flannel curtain converging the surrounding walls). Though not measured, subjective analyses indicate that the reverberation time is shorter than the cafeteria.

In order to follow a representative design approach, by giving a sense of the place represented in the listening test, some of the cafeteria conditions were replicated at the experiment set. Subjects were sitting at a cafeteria table with some dishes and a soft drink can on top. A large canvas was suspended in front of the subjects, to receive a video projection (Figure 37 - Left). The video consisted of footage from the cafeteria, recorded from the point-of-view of a person sitting at one of the tables. This way, subjects participating in the experiment where looking at a real-size video image of the cafeteria, matching the perspective of someone actually sitting there.

Previous studies (Guastavino et al., 2005) reported that seeing the speakers would influence how people perceive the sound. In order to avoid visual bias the room was darkened, featuring only one spotlight hanging on the celling, pointing down at the subject's table (Figure 38).

❖ Subjects

Twenty-five sets of listening testes and questionnaires were conducted. Subjects were college students or university staff, aged from 20 to 37, who agreed to participate without payment. All participants were familiar with the cafeteria.

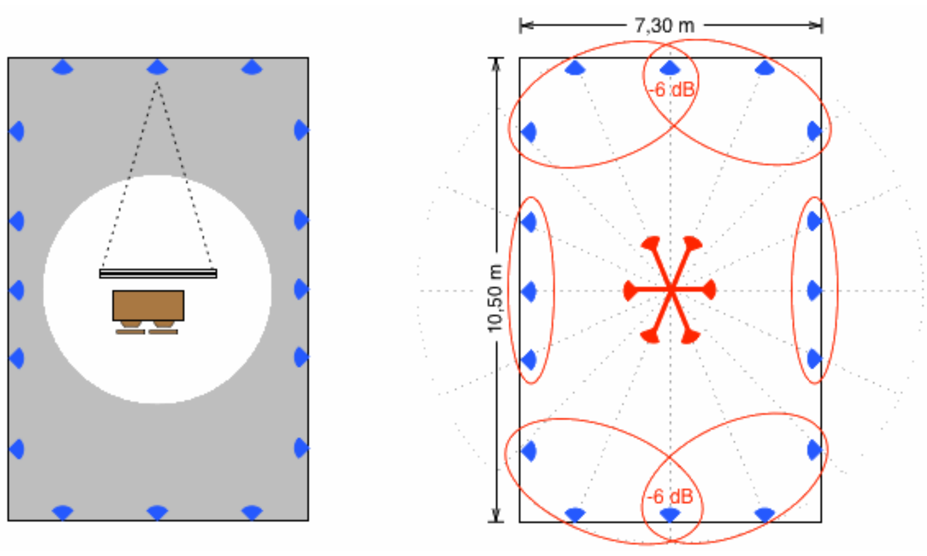


Figure 37. Experiment room plan. Left: table, canvas and spotlight. Right: microphone array arranged according the speakers' distribution (red cross in the center), speakers (blue) and signal distribution (red ellipses, one for each channel/microphone).

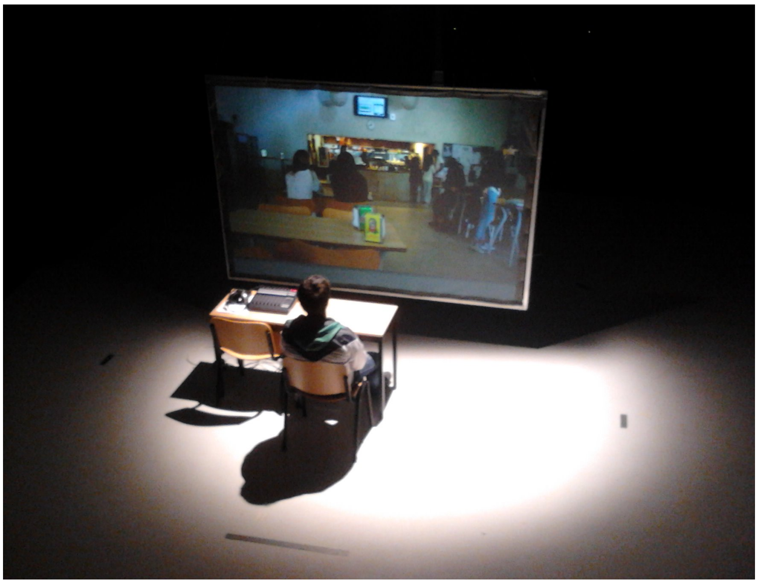


Figure 38. Experimental setup.

- **Part I – Binaural vs. Surround**

  ◊ **Method**

In Part I we compared two sound recording/reproduction systems: binaural and surround. Both audio recordings were accomplished at the same time and location using two different methods (described next). Each sample was two minutes long. Subjects listened one after the other with approximately 20 seconds interval. The two samples were played in random order to avoid a sequence bias. While the sound was playing, subjects were asked to engage in leisure activities, such as reading a magazine or playing a tablet game (neither activities demanding the use of audition). The goal was to distract the subjects, avoiding attentive listening and promoting background listening. After the listening test, subjects were asked to answer a structured questionnaire orally, which aimed at characterizing their listening experience, including their preference regarding the two systems.

Sound pressure level was regularly measured during the recordings (SPL – C-Weighted).

◊ **Binaural setup**

The binaural recording was made using two omnidirectional blocked ear canal microphones. We preferred this technique based on (Møller, Hammershøi, Johnson, & Sørensen, 1999) research and empirical tests comparing it with a DIY dummy head available at the moment of the test. A research assistant was sitting still at a table wearing the microphones, a video camera was recording footage from his point-of-view (subjective camera) and a sound recorder device was capturing the sound from the microphones.

During the test, participants listened to the binaural recording using a pair of closed-back circumaural headphones (Sennheiser HD320). Subjective measurements were made to match the reproduction pressure level with the sound pressure level measured during the recording session.

◊ **Multichannel setup**

The multichannel recording was made using an array of six cardioid, small capsule, and condenser microphones. The arrangement of the microphones was based on the IRT Cross (Theile, 2000) (also known as “atmos cross” for being suited for recording ambient sound). We have decided to expand this technique by adding an extra pair of microphones, since our reproduction system had sixteen discrete speakers instead of the usual five or seven found in commercial surround systems. Moreover, the angles of the microphones were setup having in mind the placement of the speakers in the reproduction room (Figure 37).

The reproduction of the audio files (six channels) was done using a setup with sixteen speakers (described before). The system was previously calibrated and followed the signal distribution mention in (Figure 37 - Right), using custom-made software[27], which handled the distribution of de signal, including phase delays. The volume of the speakers placed in front of the lis-

[27] Software was programed in Max/MSP 5 and adapted the spatial audio external Ambimonitor by Philippe Kocher and Jan Schacher (www.icst.net).

tener (behind the canvas) was 3dB louder, to compensate for the sound blockage caused by the video canvas. Since all the speakers were playing slightly different signals, a rich sound field was achieved. No subwoofer was used.

◊ Results

When subjects were asked to identify the listening experience that more closely matches the real situation, from an auditory perspective, 48% chose the surround system, 40% the binaural system and 12% felt no difference. When subjects were asked the same question but focusing on the global experience, the answers were: 60% for the surround system, 36% for the binaural system and 4% did not feel any difference.

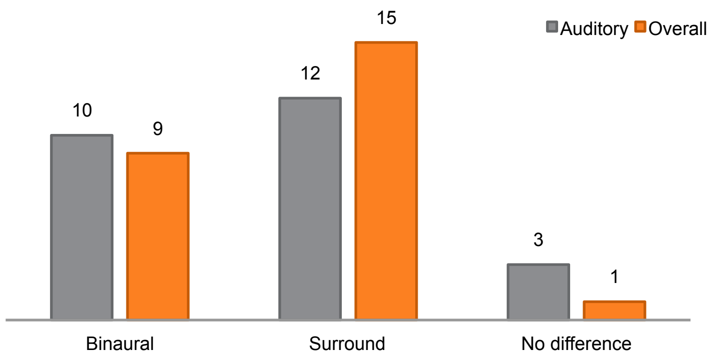


Figure 39 Subjects preferences regarding the auditory and overall experiences, with binaural and surround systems (number of responses).

The following question was about the reasons of their choice regarding the overall experience (it was a closed question allowing multiple choice). The reason that stood out was the "sensation of space", while all the other reasons were represented roughly equally and far below the most popular. In order to better understand this data we analyzed the answers separately for the subjects who selected the binaural system and the surround system. Those who chose the surround system clearly did it due to "the sensation of space provided" and "free from headphones use". Those who chose the binaural system were not so assertive choosing their motivations, pointing out several other reasons related to the auditory experience, such as “better source localization”, “more immersive”, “higher sound fidelity”, “higher sound quality” and “more natural” (Table 1 - Reasons why subjects chose binaural/surround as the most global real experience (number of answers).).

Reasons	Both resp.	Binaural	Surround
Headphones free	6	0	6
Sense of space	18	7	10
Fidelity	9	6	3
Immersiveness	8	5	3
Naturalness	5	3	2
Source localization	9	5	3
Sound quality	6	4	2

Table 1 - Reasons why subjects chose binaural/surround as the most global real experience (number of answers).

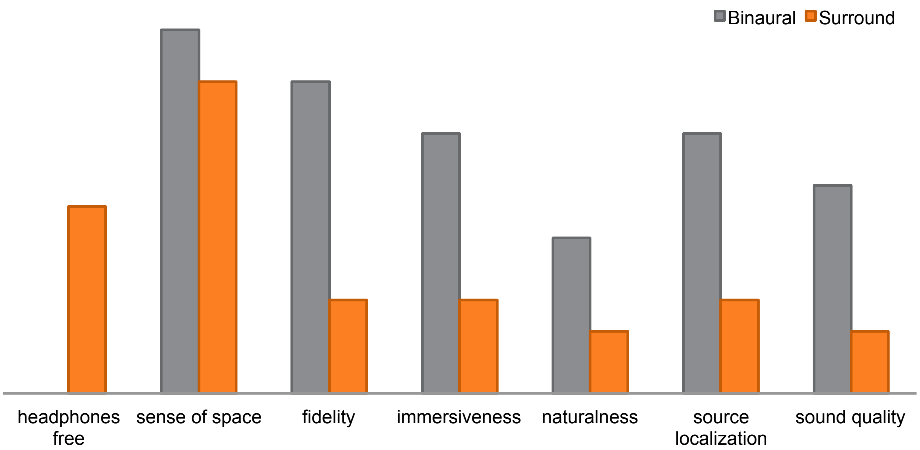


Figure 40 - Reasons why subjects chose binaural/surround as the most global real experience (normalized values, just for visual comparison).

❖ Part II – Soundscape Designing Tool

◊ Method

Typically, soundscapes and noise control studies are based on listening tests followed by structured questionnaires or free interviews, which are then analyzed through linguistic exploration of verbal data. In experiment’s

Part II we have ran preliminary tests to assess a methodology based on active design of soundscapes and semi-structured questionnaires. After responding to Part I questionnaire, users were invited to create their favorite soundscape for the cafeteria. To accomplish the task, custom-made software was designed to accommodate four audio tracks, each one bearing four 6-channel surround samples of the same sound category (dialogs, radio/tv, nature, cafeteria). The system allowed the user to select one sound for each track/category and change its volume (a typical audio mixing process). No GUI was used; subjects interacted with the system using a tangible MIDI controller[28]. The sounds were reproduced using the surround system described before.

◊ **Results**

After creating their favorite soundscape, participants were asked to rate from 1 to 5 (nothing to very), how much this type of exercise contributed to raise their awareness for the sound environment (Figure 41). Measurements of SPL for each submitted soundscape showed that the average level from all participants is 72,9dB (C-Weight, slow response), roughly 2dB lower than the levels measured in the cafeteria during the experiment.

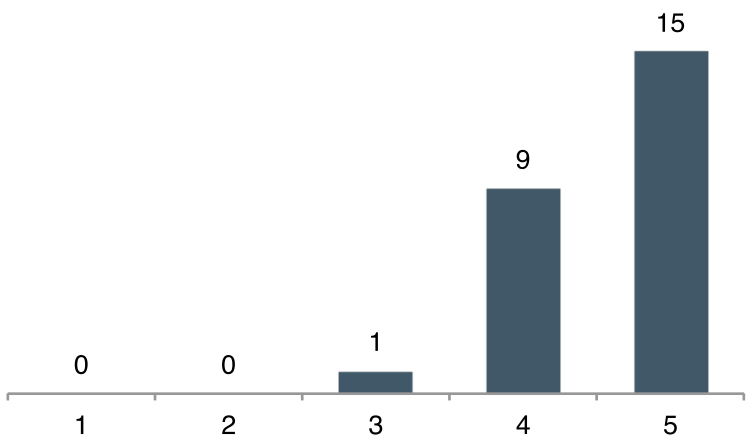


Figure 41 - Evaluation of the exercise regarding its ability to promote environmental sound awareness.

[28] Behringer BCF2000

The last question asked for a comment on the experiment. In general, participants enjoyed doing the experiment since it provided them with the opportunity to listen to the same sound in different formats. Furthermore, the system for soundscape design allowed them to think and “craft” the sound for a place, which was something that most of the subjects had never thought about. Some subjects also pointed out the effectiveness of the real-size video canvas in providing a sense of place.

6.2.3 - Conclusions

Part I: A comparative study between binaural listening and a massive multichannel surround listening was done, using the soundscape of a cafeteria as testing material. The results showed that participants are divided as to which system is able to deliver the most realistic auditory experience but elect the surround system as the one able to provide the most global realistic experience. Paradoxically, both participants pointed “the sense of space” as the main reason for their choice, which lead us to conclude that this concept is problematic, since it does not mean the same for every person, and it should therefore be avoided or previously explained to subjects in future tests. Further analysis revealed that participants who choose the binaural system, do it for its good auditory properties (fidelity, quality, localization) while those who choose the surround system, privilege the fact that they do not have to wear headphones, over the argument of sound quality. Therefore, from the overall results, we conclude that in listening tests on soundscapes, it is preferable to prioritize the use of surround systems over headphone listening. It provides better sense of reality, which moves towards the representative design approach.

Part II: An interactive system for soundscape design was tested with the massive multichannel speaker setup. Users agreed on the potential of the system to promote sound awareness among people. We believe that this methodology based on sonic interaction provides a better way to communicate about sound with non-experts, since it eliminates the linguistic intermediation, which is error prone (as demonstrated before during this study). Therefore, it should be considered in soundscape studies as a complement to traditional listening tests and questionnaires.

6.2.4 - Further Research

In the long run we envision an upgrade to the system, consisting of the accommodation of video projection on three surrounding walls and the addition of six speakers mounted on the ceiling pointing down, creating a dome-like setup. Consequently, future research will explore different multichannel recording/reproduction techniques, in order to increase the ecological validity improve representative design methods in soundscape assessment.

CHAPTER 7 - Research Methodology

Research methodology is the set of procedures one has to systematically complete in order to find the solution for a scientific problem. These are standardized techniques known and shared by the scientific community, ensuring the reproducibility of results when all the original conditions are replicated. Furthermore, they allow for the identification and quantification of the systematic errors existing in each procedure (for example, when measuring something). Different sciences have necessarily different methodologies and techniques: one cannot expect to apply a questionnaire to a colony of bacteria nor kill someone's relatives to study how the subject deals with the loss. A consolidated methodology and a stable terminology[29] are necessary conditions for the production of cumulative knowledge.

In this chapter we start by describing the overall research approach used in throughout our study and present, in finer detail, the methodology for an experimental mobile application conceived as a research tool.

7.1 - Towards an interdisciplinary research methodology

The choice of a research method always depends on the phenomenon being studied and the scientific field in which it is pursued. More traditional sciences – like physics or astronomy - have long established their methods, while younger and emergent sciences – like HCI or network science (presented in point 4.2 - (pp. 76)) - are still consolidating their fundamental procedures. Our research project lies in a multidisciplinary field that aggregates different areas of knowledge such as – but not limited to - HCI, Sociology and Psychology. Therefore, the designated methodology should comply with the

[29] About terminology, we would like to highlight a passage from 9/28/13 2:47 AM "(…) scientists do not use the common language in the common way; they assign special definitions to specific terms. And when that occurs, and when those special definitions acquire stable and significant meanings to many workers in the discipline over a period of time, then indifference to established usage together with arbitrary redefinition become obstacles to progress."

best practices observed in each of these sciences, merging and adapting to the project's demands.

A research tool is a methodological resource often used to assess an undetermined phenomenon; it represents a way to collect data regarding the variables participating of that phenomenon, which can later be analyzed in order to draw conclusions. From the social sciences perspective, there are three main groups of research tools: experiments, surveys and ethnography. All tools are meant to be applied in groups of people (samples), who should be representative of a broader population.

An *experiment* is usually a laboratorial test where most of the variables are kept under control, such as the listening tests described in 6.2 - A Soundscape Assessment Tool Based on a Massive Multichannel Loudspeaker Setup (pp.121). An *experiment* is able to provide great internal validity since all the systematic errors are taken into consideration wile designing the study. However, the possibility of inferring universal laws are constrained to the laboratorial settings, and real-world applications are sometimes hindered by the lack of real and more complex settings during the study. Our preliminary research mentioned above aimed at improving laboratorial experiments to match real-world settings. The outcome of an *experiment* is usually based on quantitative data.

Surveys are one of the most common tools used in social sciences research. They can range from closed questionnaires to open interviews. In the former, respondents are asked to answer a fixed number of questions, each one with a fixed number of answers; in the later, interviews may have a structural script, which adapts to the answers being given. Closed questionnaires produce numeric quantitative data and are more easily compared with each other, while interviews require a linguistic analysis and tend to produce more qualitative data. Nowadays, using web technologies, surveys can have an instantly wide spread across a population at a very low cost.

Ethnography is the scientific description of the customs and habits of a population. It is, therefore, a methodology based on the observation of individuals in their natural habitat. The scientist has no control over independent variables, thus it is difficult to assess a case-effect dynamic. On the other hand, these studies are more prone to having external validity, since the experimental setting corresponds to a real-world scenario[30].

Due to conceptual and technological aspects of our research it is important to consider HCI as being a stem knowledge area. HCI is still a young scientific discipline; established methodologies started during the 70's in the Silicone Valley but new assessment methods are still being introduced. HCI has emerged from the combination of core computer science with cognitive psychology, which *"encompasses many sub disciplines with different research questions and methods"* (Boring, 2002), such as sociology, anthropology or communication. HCI is, therefore, related to human factors[31]. In this field, most of the research methods rely on empirical studies, since the phenomena assessed are always somehow related with the use of technology.

Some research approaches in HCI include experimental psychology, task analysis, conversation analysis, activity theory, design and ethnography (see (Monk & Gilbert, 1995) for an overview), which are implemented in the form of case studies, field studies, action research, laboratory experiments, survey research, applied research, basic research and normative writings (Kjeldskov & Graham, 2003). From the combination of different approaches and the emphasis on *using technology*, some new methodologies have emerged, expanding more traditional techniques. Some examples include cultural probes (B. Gaver, Dunne, & Pacenti, 1999), eye tracking (Poole & Ball, 2006) and electroencephalography monitoring (Lee & Tan, 2006).

---

[30] Some authors denominate ecological validity as this match between the experimental setting and a real-world scenario. We prefer not to use this term because of its historical ambiguity, as explained before in 6.2.1 - Introduction (pp.109).

[31] In fact, the annual conference that the ACM dedicated to this topic is called CHI (Conference on Human Factors).

Mobile HCI is a particular case of HCI dedicated to the interaction with mobile devices. Due to its specificity and relative novelty within the field, research methodologies are more scattered and *“research methods examining phenomena in context such as case studies are not widely used”* (Kjeldskov & Graham, 2003).

Our approach puts an emphasis on mobile HCI and combines several methodologies in an interdisciplinary fashion, where borders of distinct knowledge are defied, aiming at the integration of different methods into a science of its own.

7.2 - Methodological framework

After the literature review and definition of research question stages, we implemented some research tools, in order to collect data that could support or refute our hypothesis.

In Chapter 6 - Preliminary Research On Soundscapes And Sonic Interaction Design (pp.113) we have already presented some methodologies for assessing interaction with mobile devices vs. desktop; and listening tests between different reproduction systems. Both experiments were focused on the interaction with sound and assisted the gathering of both quantitative data (through score counting, interaction log, SPL measurements and closed questionnaires) and qualitative data (obtained in semi-structured questionnaires). Next we will focus on the methodology implied on the main study of this thesis.

In our main research, the methodology was based on the development of a mobile application, which users utilized in their daily activity. This application was our main research tool, providing us with quantitative data from user logs; and setting the context for surveying the users, after those having used the application (granting us with qualitative data). The methodology suffers a great conceptual influence from *cultural probes*[32] design methodology, developed by (B. Gaver et al., 1999), which tackles the problems from an "artist-designer" tradition rather than purely "engineering-based" approaches. The authors define it in the following manner:

> *"Unlike much research, we don't emphasize precise analyses or carefully controlled methodologies; instead, we concentrate on aesthetic control, the cultural implications of our designs, and ways to open new spaces for design. (…) Unlike most design, we don't focus on commercial products, but on new understandings of technology. This allows us—even requires us—to be speculative in our designs, as trying to extend the boundaries of current technologies demands that we explore functions, experiences, and cultural placements quite outside the norm. (…) The artist–designer approach is openly subjective, only partly guided by any "objective" problem statement. Thus we were after "inspirational data" with the probes, to stimulate our imaginations rather than define a set of problems. (…) Our designs could offer them [the users] opportunities to appreciate their environments—social, urban, and natural—in new and intriguing ways."*

As in Cultural Probes methodology, we are also not searching for a solution to a precise problem, but instead give people the opportunity to experience their environment and technology in novel ways and assess their feedback, in the hope of getting valuable inputs for future models an theories, that can then be materialized into everyday objects or practices. Furthermore, we highlight and relate to the importance given to the "*tone and aesthetics of the probe materials"* (which in our case is the mobile application, project's website

[32] Cultural Probe is the name given by the authors to a set of crafted materials (not typical research materials such questionnaires, but postcards, maps, stickers, disposable photographic cameras and sketch books), which are distributed among groups of participants in the experiment, in order to document and register their ideas about a place or a phenomenon. The probes are later returned to the researchers to be analysed.

and questionnaires) in order to reduce the distance between general public and the vernacular scientific world.

Notwithstanding, and to better postulate a definition for our experimental research regarding the software application, we find it useful to apply the first stage of (Basili, Selby, & Hutchens, 1986) framework for experimentation in software engineering, which gives a structure for "*the experimental process and provides a classification scheme for understanding and evaluating experimental studies*" (Figure 42). It is divided into five categories:

I. Definition					
Motivation	Object	Purpose	Perspective	Domain	Scope
Understand	Product	Characterize	Developer	Programmer	Single project
Assess	Process	Evaluate	Modifier	Program/project	Multi-project
Manage	Model	Predict	Maintainer		Replicated project
Engineer	Metric	Motivate	Project manager		Blocked subject-project
Learn	Theory		Corporate manager		
Improve			Customer		
Validate			User		
Assure			Researcher		

Figure 42 – Definition stage of (Basili et al., 1986) framework for experimentation in software engineering.

Motivation | Our main research motivation is to better <u>understand</u> the role of soundscapes in social networking. This is a very wide and broad goal, which does not coalesce into an equation or a percentage. The pursue is not to conceive a state-of-the-art mobile application that flawlessly infiltrates into people's device and reach the 100 Top Apps rank.

Object | The object of the study can be described as a <u>model</u> of interaction, based on sound, regarding the enhancement of social networking and use of mobile devices.

Purpose | The purpose of the study is to gather a set of scattered knowledge from divergent areas and <u>motivate</u> in an exploratory fashion to pave the way for forthcoming theories and frameworks on the topic of acoustic communication and sonic interaction.

Perspective | The perspective from which we approach the study is clearly from a researcher point of view, we aim at a research tool and not at a final product, where reliability and optimization are mandatory.

Domain/Scope | The domain and scope are restricted to the research project/software application.

On the possible need to describe the methodology, we would say that it combines some characteristics from field studies, case studies and surveys.

7.3 - Methodological procedure

We will now describe the methodological procedures regarding the experimental part of our study, based on the design, development and implementation of an OSN application that can sense the surrounding acoustic environment and broadcasts this information to the *friends* in the network (described in detail in Chapter 8 - Hurly-Burly: Prototype System for Experimental Research, pp. 151).

7.3.1 - Subjects Selection

The first step was the selection of a sample of subjects to participate in the experiment. In order to ensure external validity, i.e. the ability to reach generalizable results, the sample had to be statistically representative of the population to whom the study concerns. Ideally, the results of a study should be as general as possible, especially when aiming at high-level conclusions (high abstraction level, not focusing on a specific problem). In practice, however, research always requires a certain amount of compromise, and the constraints of our study dictated the size contraction of the population to whom the study concerns. The requisites for people to be selected as subjects of the sample were:

- *Familiar with mobile online social network (MOSNs)*. We focused on people to whom running OSN on the mobile devices was already an established habit. Having to learn and adapt to this practice could negatively influence the results. In 2012 the age demographics for people using MOSNs revealed the 25-34-years-old group as the most active and MOSN are becoming the preferable way to access OSN[33].

- *Owner of an iOS device, supporting iOS <4.3*. This excludes the first two iPhone and iPod models (iPhone 1, iPhone 3G, iPod Touch 1, iPod Touch 2). In the 1st quarter of 2013, 39% of smartphone owners in the U.S had an iPhone[34], which is smaller than the penetration market of Android devices (multi brands). Nevertheless, this device is among the most expensive ones and not always owned by potential users.

- *Belonging to a highly clustered group*. The experiment was implemented in groups of participants. In each test group, all participants had to know each other beforehand, meaning that the online social network mirrored a potential offline social network. Our aim was to assess how members of an established social network would react to a new layer of information being exchanged within the network. In our opinion, using sound to convey information about strangers would not provide meaningful information nor foster significant social relations, thus we have decided to not assess that particular dimension. All test groups knew everyone in their network, from close friends to acquaintances, showing a clustering coefficient = 1. Moreover, asking people to share personal and real-time information with strangers would mean a privacy probation that could refrain people to engage. Assuming this methodology made the formation of test groups difficult but provided an optimal research ground.

---

[33] In USA. See 9/28/13 2:47 AM for more statistics .

[34] According to http://www.nielsen.com/us/en/top10s.html (accessed on July 28, 2013)

- *Voluntary participants.* No payment was ever involved. Only a motivation prize of €50 was to the participant who used the app for the longest period in a group of undergraduate students.

- *Willing to share private data with friends*: As we mentioned before, sharing private data corresponds to privacy probation. Not all the participants that were initially invited agreed, arguing privacy issues. A great deal of effort was put into the explanation that no audio was ever shared during the application usage, but even so some of the subjects quitted after understating the overall concept. That attitude reveals suspicion towards the use of sound in context-sensing applications.

- *Willing to test unknown and experimental software on their devices:* Before installing the app, users were informed about the installation issues, which included sharing their UDIDs and installing some new certificates. Sharing UDIDs is a common practice among Apple beta-testers but is always intimidating for non-experts. In fact, sharing UDIDs was not risk-free[35] until Apple decided to deprecate the use of UDIDs by the release of iOS5 early in 2011, and started to reject one-year later App store submissions for applications, which query the iOS UDID.

7.3.2 - Test Groups

Taking into account the constraints mention in the previous point, three tests groups were formed, counting a total of 23 subjects, which ran the application and filled in the survey. Theses groups were defined according their offline proximity, usually rooted in a common occupation (school class, work, common project, see Figure 43). Some users were present in more than one group. The author of the thesis was part of all networks but did not fill in a questionnaire.

---

[35] See 9/28/13 2:47 AM for a discussion on the use of UDID by commercial applications.

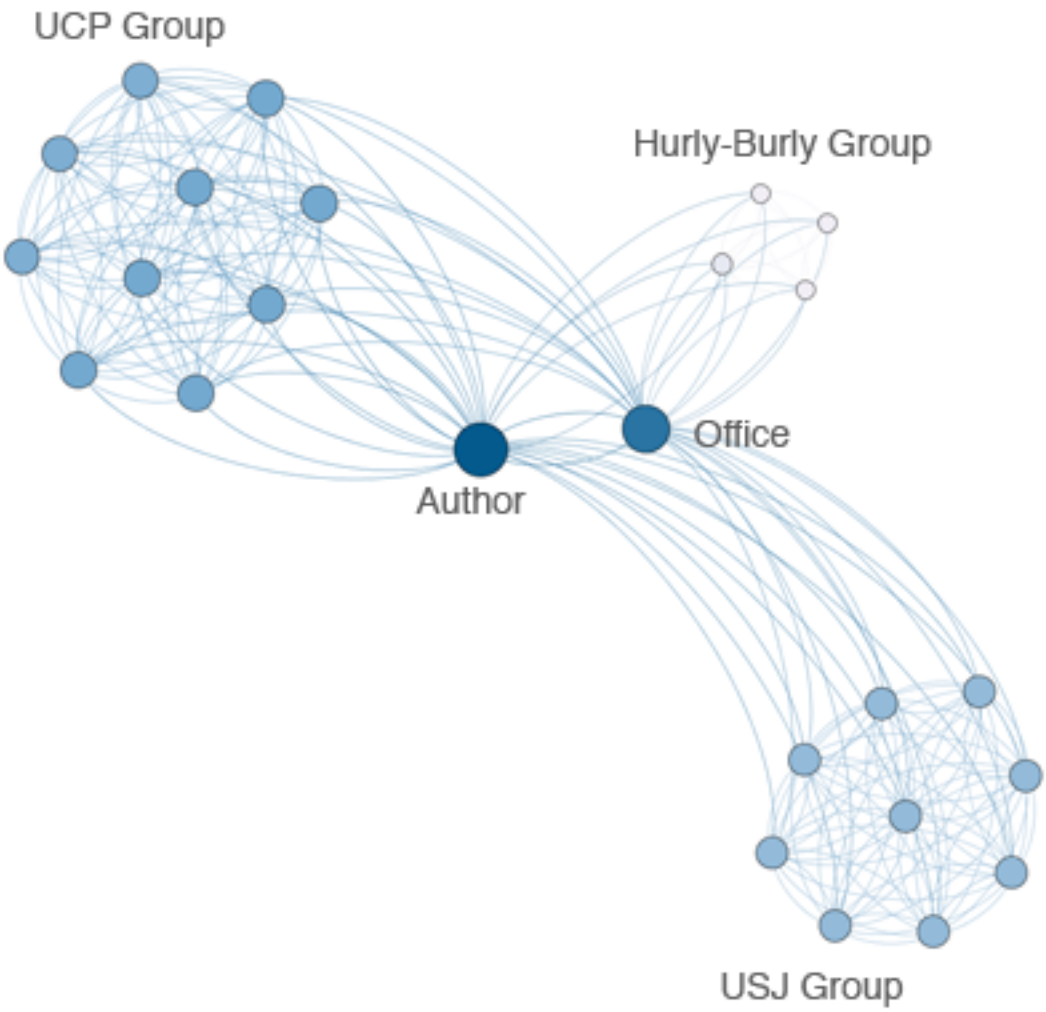


Figure 43 – Network of the subjects of the experience.

Generally speaking, subjects can be described as *young adults, both male and female, engaged in social networking, belonging to the middle class and having some relation to the media world*. One third of the subjects live in Asia and the rest in Europe; one subject was living in Canada during the experiment. The groups are not representative of the whole worldwide population but they significantly represent the user profile mention above. A description of the individual groups is presented bellow; real names will not be displayed for privacy issues.

❖ **Group 1 – Communication and Media Students / USJ**

This was the first group to test the application. The group was formed by 9 students from Communication and Media Licentiate in University of Saint

Joseph, Macau. They are aged between 19 and 23, 3 male and 6 female subjects, all students and most working part-time. Initially, this group was larger, but two students declined to participate, alleging privacy concerns; they said: *"People cannot ear our conversations, that is very dangerous!"*. There is a large market penetration of iOS devices in this class and students are very keen on online social networking, with Facebook dominating the online social network landscape. However, its knowledge about mobile technology is mainly from a user perspective.

ID	Name	Age	Device	Genre
25	User 25	20	iPhone	Male
27	User 27	21	iPhone	Female
31	User 31	21	iPhone	Male
32	User 32	21	iPod	Female
33	User 33	19	iPod	Female
34	User 34	20	iPhone	Female
37	User 37	23	iPhone	Male
42	User 42	19	iPhone	Female
44	User 44	21	iPhone	Female

Table 2 – Test Group 1, USJ Students

- **Group 2 – Staff from Sound and Image dep. - UCP**

Teachers and researchers on the Sound and Image Department and CITAR – Research Center for Science and Technology of the Arts, in the Portuguese Catholic University, in Porto, formed this group with 11 members. It is aged between 27 and 43, and has a gender distribution of 7 males and 4 females subjects. There is a considerable market penetration of iOS devices, though some of them belong to the university and are used for development. It is a group that uses online social networks in a moderate way and most of the subjects were already connected through Facebook using closed/interest groups and LinkedIn.

ID	Name	Age	Device	Genre
6	User 6	33	iPad	Male
7	User 7	33	iPod	Male
8	User 8	29	iPod	Male
11	User 11	27	iPhone	Female
12	User 12	29	iPhone	Female
18	User 18	43	iPhone	Male
22	User 22	39	iPad	Female
54	User 54	30	iPad	Female
65	User 65	35	iPhone	Male
66	User 66	42	iPhone	Male
67	User 67	34	iPad	Male

Table 3 – Test Group 2 – UCP Staff.

❖ Group 3 – Hurly-Burly Team

This group was made up of people that worked on the conceptualization and graphic interface of the application. It is a group specialized in interaction design, therefore expert in the field of HCI but not necessarily in the sound domain. This was the smallest group of the whole experience, with three members only, aged from 24 to 32, one male and 2 female. They were all regular users of online social networks.

ID	Name	Age	Device	Genre
46	User 46	27	iPhone	Female
53	User 53	24	iPod	Female
60	User 60	32	iPhone	Male

Table 4 – Test Group 3 – Hurly Burly team.

❖ Monitoring Devices

The research team used two devices as monitoring equipment. One iPod was carried around and an iPad remained still at the group's research office. All the other groups were connected to these devices, seeing it as a friend in their devices.

7.3.3 - Software deployment

❖ Submitting devices for beta testing

After selecting the subjects' sample and defining the groups, participants were invited to join this experiment and provide their UDIDs. This request was sent by e-mail along with a video introducing the project, a link for a webpage with more information (including scientific publications) and a video tutorial on how to find the UDID using iTunes. This task was not as trivial as expected; some users requested our accompaniment to perform this action. More than 90% replied positively to this request.

Following the UDID collection, we introduced users UDID in our developers account in the Apple Developers Database. The program allows up to 100 devices for testing. Afterwards, a *provisioning profile for distribution* was released for each group, granting to each device the permission to run the designated application.

❖ Upload users on database

The next step was to create a web database for the application, in order to manage the users and collected data. This database served three purposes: 1) Register all users individually, providing them with a user ID, a username and a password. 2) Create a network by establishing the link between different users (who is friends with who) and 3) Prepare a container for the data collected by the application during its usage (this database is described in detail in 8.3 - System Database, pp.167).

❖ Compiling the application

In order to be distributed among the subjects of the test, the application was compiled independently for every user, creating an IPA file for ad-hoc distribution. Each *built* included an individual *user ID, username* and *password,* hardcoded in application code. This daunting task avoided the use of a *login* module in the application, which could abstain subjects from using it (for example, if they forgot the login credits). Additionally, combining each *built* with a specific UDID, resulted in individual versions that could only run on a specific device, associating a user to a device, avoiding the risk of getting a mixture of incompatible data.

❖ **Distributing the App**

The last step was the distribution of the application through the test subjects. The Apple ad-hoc distribution allows the distribution of experimental applications, up to 100 users, without submitting them to the App Store (the only drawback is the need for the users UDID, which can be avoided with an enterprise account). In order o accomplish this task, we have used the free version of TestFlight[36] service, which proved to be extremely reliable and convenient, allowing the distribution by e-mail and installation bypassing iTunes. All users, except one, were able to install the application.

7.3.4 - Experimental period

❖ **Running the application**

The tests occurred separately in each group and followed the same order of the group's description (see above). The time frame for running the test in each group was approximately three days. During this period, users were asked to run the application as long as possible, including during the night. An indicative period of 24 hours of use was set as suitable. At this step, some users did not continue in the experiment.

❖ **Distributing the questionnaire**

After running the test, a questionnaire was delivered to the subjects, in order to assess their experience using the application and collect information

[36] http://testflightapp.com

about the users and their relation with the topic of mobile sound, online social networking, self-disclosure and privacy. The questionnaire was produced in the digital format, created and distributed by e-mail using the free service Google Drive (Form document).

Accompanying the survey was a link for a webpage containing an interactive graph, with a history plot of users sonic and movement activity. The graph afforded zooming in and out, and timeline navigation. This webpage was only sent to the users showing a significant amount of data and was necessary for answering one question of the survey. A detailed presentation of the questionnaire takes place in 9.1 - Questionnaire Design (pp. 179). After filling in the questionnaire the experimental procedure was over and users could uninstall the application from their devices.

CHAPTER 8 - Hurly-Burly: Prototype System for Experimental Research

In this chapter we present in detail the design and development of a mobile application used as a prototype system for experimental research on the topic of sound-base social networks.

8.1 - System Overview

The application is named Hurly-burly, a synonym for “boisterous activity”, and fits in the category of MOSNs. Users have the opportunity to create links with other users, sharing and receiving information about themselves and the environment, on a mobile device. In this particular case, the information being shared regards the soundscapes of the places visited by each user during his/her daily life, complemented with a rough description of his/her movements. The sound is sensed in a “calm computing” style; the application is working inconspicuously in background, sensing the sonic environment and feeding the social network, wile keeping a personal record of the data. Each user acts as a terminal of a sensor network linked through the use of sound.

The system is comprised of a mobile application, a web application and a webserver combined in a typical client-server configuration (Figure 44). The mobile application is responsible for input and output tasks, covering the visual and auditory display of real-time information as well as the soundscape sensing. The web application displays the sonic profile (corresponding to long-term information) and the webserver saves and retrieves information based on forms/queries (represents a functional piece of the system, potentially invisible for the user).

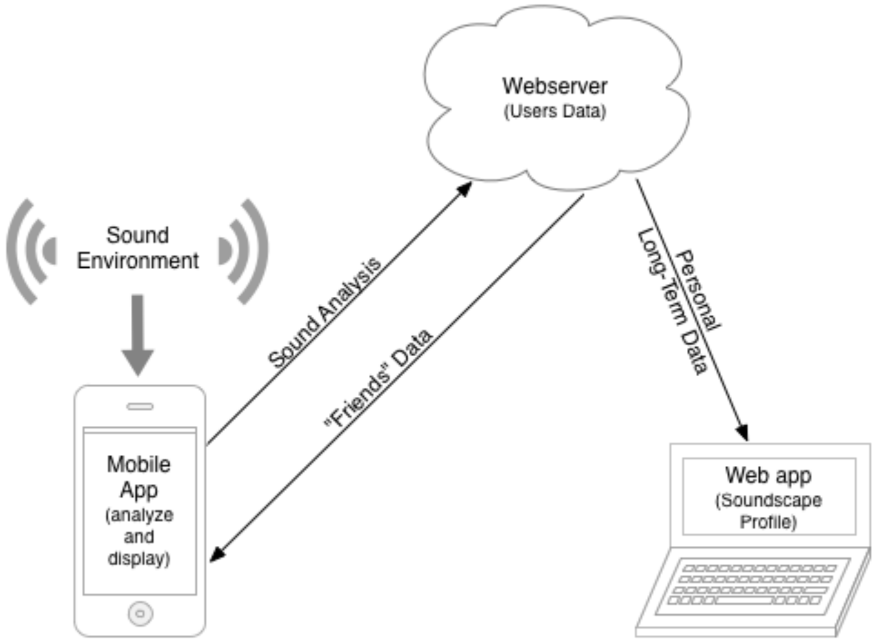


Figure 44 - The three main elements of Hurly-Burly System: mobile application, webserver, and web application.

8.1.1 - Conceptual framework: towards a sound based social network

Soundscape studies have shown that Environmental Sound is a rich resource for understanding the social context of a place, its dynamics, problems and virtues, assuming sound as a resource rather than waste. The Soundscape approach to solve noise annoyance problems has proven to be a valid solution, since it approaches sound from an holistic perspective, which takes meaning and context in consideration rather than focusing solely on sound level measurements, as usually found in Community Noise strategies. Nonetheless, this approach is still confined to geographic places, whether interior or exterior, private or public, natural or artificial. In our system we propose a redefinition of the concept *place* by shifting the point-of-audition from a static geographical *place* to a dynamic *place*: the user.

During their daily life, people travel through a sea of sounds, eventually without even changing their geographic location. Passively or actively, Soundscapes vary along the hours, days, months and years, characterizing in each moment the sonic context of a place. By extending the analysis time span, sonic patterns about the place are unveiled, contributing to the characterization of its sonic profile. This data can then be extrapolated to other layers of significance, mainly when correlated with data gathered through multi-modal analysis (geolocation, time, weather conditions, etc.).

Shifting the *place* from static to dynamic, we are characterizing not a geographical location but a node of a social network, which drifts in space, time and network position. With the proposed system, the user is able to keep track of his/her personal sonic profile. If more data is collected, more detailed and accurate is the profile. This information when evaluated in context can be of great relevance for daily social interactions, since the short-term and long-term analyses show different (but complementary) aspects of the social behaviour of the nodes of a network.

8.2 - The Mobile Application

Within the Hurly-Burly system, the mobile application is responsible for analysing the sound and movement, uploading the information on a server and displaying auditory/visual information to the user. It was built as a prototype version, yet a great effort was put on its aesthetic, as a key element of our methodological approach (figure - logo).

The application was developed for the iOS platform, being compatible, at the completion of this thesis, with iOS +4.3 systems.

We have chosen the iOS platform for four main reasons:

1) <u>Processing Power</u>: The nature of the software application demanded a powerful system in order to perform advanced tasks such as sound classification. iOS devices (iPhone, iPad and iPod) are among the most powerful mobile computing devices in the market.

2) <u>Programming language:</u> Official programing language for iOS is the C-based language Objective C. However, a popular C++ framework among media artist called OpenFrameworks allows

programmers to write code in C++, which is later compiled for the iOS platform. Besides being an efficient programming language, it was familiar to us and has a great community supporting the development of Add-ons that fulfilled most or our needs.

3) Penetration market: iOS devices are highly appreciated among our target audience (young adults, higher education, engaged in MOSNs), therefore a common device among our sample.

4) Multitasking: Latest modes of iOS devices allow multitasking; including sound capture wile the application is not running in foreground.

The mobile application is comprised of five modules: 1) a soundscape-sensing module, 2) a movement detection module; 3) a client-server communication module; 4) a visualization module and 5) a sonification module. These will be described in the next points.

8.2.1 - Mobile Application: How it works

In this point we will attain on how the application works, describing the application modes and user status, followed by a practical example of utilization.

❖ Application mode and user status

When users have the application running on their devices, it can either be in background or foreground. Working in foreground means that the application GUI is displayed on the screen and is the main application being managed by the system. Working in *background* means that the application is represented on the screen only by a red bar on top, continuing the sound and movement analyses but not sending this data to the server and *friends,* consequently. If the application is in foreground and the device screen locks, the application passes immediately to background, returning to foreground when unlocking the screen.

User status varies from online and offline, depending on the application mode and other factors. When the application is in foreground mode and a connection with the server is established, the user status is online. When the application is working in background, user is always offline and when the application is in foreground but a connection with the serve is not established (either by the lack of an Internet connection or server problems), the user status is offline.

Every time the user status is offline, the information regarding the sound and movement is not sent to the server in real-time. Instead, the information is recorded on the device and uploaded to the server when the user status changes to online. When the user is online, the information is constantly being uploaded on the server and made available for all user *friends.* Whenever a status change occurs, this bite of information is updated on the server.

❖ A practical example of utilization

User *Antonio* has downloaded and installed Hurly-Burly application successfully through the TestFlight distribution system. The application has his login credentials (username, ID number and password) hardcoded, so there is no need for manual login. His social network was pre-established on the server database according to the experimental methodology, so he does not have to invite friends or be invited.

He is at home – a quiet place in a calm neighbourhood – sitting down on a confortable armchair, reading a book wile chill-out music is playing on background. *Antonio* has a 24/7 Internet plan on his iPhone so he is always connected. The webserver hosting the database is working well.

1) He holds his iPhone with one hand and starts the Hurly-Burly application.
2) His friends are immediately shown on screen represented by waveforms with a nametag (GUI is thoroughly explained later).
3) A message is sent to the server saying that Antonio is online.

4) Five seconds later the application captures three seconds of an audio samples and proceeds with the analysis, returning “music - 40dB”.
5) It also analyses his movement returning “steady”, since Antonio was sitting down and had held firmly the device during the analysis.
6) A waveform representing him is displayed accordingly; positioned on top of the list.
7) Because Antonio is connected to the Internet, the server is up and the application is in foreground, the data collected is sent to the server and written on the database. His entire friends see this data updated on their devices.
8) Antonio, puts his phone on a table.
9) A few minutes after the screen auto-lock activates, the application starts working in background.
10) Antonio status changes to offline. A message is sent to the server, updating his status on the database.
11) Every three minutes the application analyses a sound sample and the device’s movement.
12) This information is recorded on the device.
13) After twenty minutes Antonio gets back to his phone and unlocks the screen.
14) The application shows up immediately on the screen (foreground).
15) Antonio’s status changes to online and a message is sent to the server updating this information.
16) All the data recorded until then is uploaded on the server.
17) On the screen Antonio can see his friends current soundscapes. Curiously, John, Rick and Sebastian show all a loud environmental soundscape. Antonio sends John a text message “what are you up for the afternoon?”, John replies “Hi! I’m with Rick and Sebs at the pub, stop by if you want”.

8.2.2 - Mobile Application: Soundscape Sensing Module

This module is responsible for analysing the sound that gets into the device through its microphone, performing typical sound analysis and machine audition tasks, using established techniques and software algorithms. It is a two-fold process comprised of: 1) a sound level analysis and 2) a sound classification task limited to three general categories: music, speech and environmental sound.

Since a mobile device has limited memory resources and processing power (when compared with high-end desktop units), the analysis task required a compromise between accuracy/precision and device's resources. Moreover, the goal of this implementation was to achieve a satisfactory analysis and classification tasks that could provide a coarse approximation to the real acoustic environment. Producing a state-of-the-arte machine audition algorithm was out of the scope of this project and beyond our field of expertise.

The analysis is always performed on ±2,79 seconds mono audio samples (44.1kHz, 16bit), recorded every 3 minutes (or whenever the user double tap on the device' screen), using the built-in microphone of the mobile device. The soundscape-sensing module was built using visual programming language Pure Data[37] and embedded in the main C++ code using a software wrapper called LibPD[38].

❖ Sound Level Measurement

The goal of this feature is to be more precise than accurate, that is, to achieve better relative measurements than absolute measurements. Users do not have access to the measured value on the application, only a graphical *icon-like* visualization (described later). The sound level measurement of the captured sound is accomplished using *equivalent continuous sound pressure level* (Leq,T) which sums up the total energy over a given time period (T), providing a level equivalent to the average sound energy during that period. In order to provide more significant results in the context of human audition, these measurements are A-weighted, which intend to approximate the frequency response of our hearing system. Although being tailored for more continuous sounds and derived from the 40dB Equal Loudness Contour Curve, preliminary tests proved that such metric was more convincing than a linear Leq, which didn't respond well for quieter sounds. Since all measurements are

---

[37] Originally developed by Miller Puckette, http://puredata.info

[38] Official webpage: http://libpd.cc.

made over a period of 2,79s, this is notated as LAeq, (2,79s). Bellow a threshold of 20dB SPL, the sounds were labelled as silence. We used for this task an algorithm extracted from PureMeasurment (a collection of patches for Pure Data intended to perform acoustical measurements)[39]

❖ Sound Classification Module

This module is responsible for the classification of incoming sound, in three major classes: musical sounds, speech sounds and environmental sounds. This three major groups where chosen regarding the differences on the production processes (sound sources / activity) and underlying emotional potential.

- Speech is a human activity and is usually associated with the exchanged of information. When speech label is activated, it means that the background noise is low. User is potentially listening or talking in a quiet environment (at school, at home watching TV/radio, at the office, having a conversation, etc.)

- Music is also a result from human activity and is usually related with leisure activities. Generally, music is found on the acoustic environment when festive activities occur (parties, celebrations), in artistic events (concerts), at private places (car, home), in the workplace (musicians) and in commercial places (shopping centres, stores, etc.). Music classification is based mainly on the pitch content of the audio material. When the music label is activated it means that there is a low level of background noise and music with a defined pitch content is playing.

- Environmental Sounds can be described as audible acoustic events caused by motions in the ordinary human environment, have real events as their sources (and thus are meaningful), are usually more

[39] Developed at Institute of Hearing Technology and Audiology – Jade University (http://tinyurl.com/onthmz3)

"complex" than laboratory sinusoids and are not part of a communication system (VanDerveer, 1979). In the scope of our application, environmental sounds are all the sounds not classified as music or speech or, in some cases, when more then one category is present. This is, therefore, the most usual category in a normal setting.

As mention before in 3.2.3 - Computational Auditory Scene Analysis (pp.55), the sound classification process is usually composed by a feature extraction task followed by a machine learning classification algorithm. More than ten audio features were tested, including: spectral brightness, spectral flatness, spectral roll-off, spectral flux, spectral centroid, zero crossing rate, mel-frequency cepstral coefficients, bark-frequency cepstral coefficients, spectral magnitude, spectral skewness and spectral kurtosis. The chosen classifier was the k-nearest neighbour. (k-NN) machine-learning algorithm; selected by its easy implementation, low demand of processing power and satisfactory results. Other more complex algorithms (as Neural Networks, Support Vector Machine or Hidden Markov Models) were not tested but are regarded as an option for these type of task (Álvarez, Llerena, & Alexandre, 2011; Couvreur, Gaunard, Mubikangiey, & Fontaine, 1998; Guo & Li, 2003).

The implementation of this module was also accomplished using PD, with a set of patches and objects called TimberID, created by William Brent. For this particular project, we have based our methodology on (L. Lu, Jiang, & Zhang, 2001) followed by trial and error tests. This approach consists of two main sequential classification tasks: first sounds are classified as speech and non-speech; secondly, the non-speech sounds are classified as music or environmental sounds. The two steps are detailed next.

1) Discrimination between speech and non-speech sounds

F<u>eature extraction</u>: For this task we used spectral irregularity, sigmund~ (pitch) and sigmund~ (envelope). The ±2,79 seconds sample (122880 audio samples at 44.1kHz) was divided in three smaller files with 40960 samples each, analysed with a Hann window size of 4096 samples and a hop size of the same value. The analysis of each sample of ~1s length, produced a feature vector of 30 values. A short description of each one follows:

- *Spectral Irregularity* inspects a spectrum and assesses how much each frequency bin compares to the immediate neighbours. Spiky spectra will have higher Spectral Irregularity then smooth spectra (Jensen, 1999).

- *Sigmund~* is a Pure Data object created by Miller Puckette which analyses an incoming sound into sinusoidal components. The output is two-fold: pitch estimation and envelope tracking, describing roughly the mean amplitude of the sample.

Classification task: The first step of the classification task was the training phase, which consists of storing the feature vectors and class labels of the training samples, producing a multi dimensional vector space. The features were normalized from 0-1. A set of 649 audio samples was analyzed and clustered in two groups: Speech (0-444) and Non-Speech (445-648).

- The second phase is the classification of the incoming sample; the feature vector extracted from the analysis of the sound sample is compared against the space vector of the test samples. Three training samples nearest to the query point are considered (k=3) and a Manhattan[40] distance was chosen based on the results of comparative tests. The tID object outputs a confidence value for each classification, which we use as follows: if confidence level is <=2 the sample is rejected and a new one is captured, else the resulting classification is assumed.

[40] Also named Taxicab geometry.

2) Separation of non-speech sounds in musical sounds and environmental sounds.

Feature extraction: For this task we used sigmund~ (pitch), spectral spread, spectral kurtosis and spectral flatness audio features. The ±2,79 seconds sample (122880 audio samples at 44.1kHz) was divided in three smaller files with 40960 samples each, analyzed with a Hann window size of 4096 samples and a hop size of the same value. The analysis of each sample of ~1s length, produced a feature vector of 30 values. A short description of each one follows:

- *Spectral Spread* is a measure of the concentration of a spectrum's energy around its spectral centroid as defined in MPEG-7 standard (H. G. Kim, Moreau, & Sikora, 2005). A single sine wave shows a bandwidth of zero and white noise is close to infinite bandwidth. It is extracted by taking the root-mean-square (RMS) deviation of the spectrum from its centroid in each frame. This same measure is named Bandwidth elsewhere (Mitrovic, Zeppelzauer, & Breiteneder, 2010).

- *Spectral Kurtosis* measures the flatness of a distribution around its mean value, indicating the peakedness and flatness of a distribution (Peeters, 2004). In other words, it is the smoothness of the spectrum compared to a Gaussian distribution (4^{th} moment) (Pachet & Roy, 2009). The Spectral Kurtosis value for a sinusoid will be much higher than for white noise.

- *Spectral Flatness* is the ratio of the geometric mean of magnitude spectrum to the arithmetic mean of magnitude spectrum (Peeters, 2004), it measures the similarity of the spectrum to a white noise (Pachet & Roy, 2009). The spectrum of white noise should have a high flatness value, close to 1.0.

Classification task: The first step of the classification task was the training phase. This classifier was trained with 445 sound samples, clustered in two groups: Environmental Sounds (0-223) and Music (224 - 445). The features were normalized from 0-1. If the confidence level reported by the classifier is =>1 then the classification is accepted, else a decision logic takes place, based on the pitchness of the sound reported by sigmund~ object. If the quantity of successfully recognized pitches is <= 8 then the sound is classified as environmental sound, else, if the mean of the recognized pitches is >= 59, the sound is classified as environmental, else it is classified as music (Figure 45).

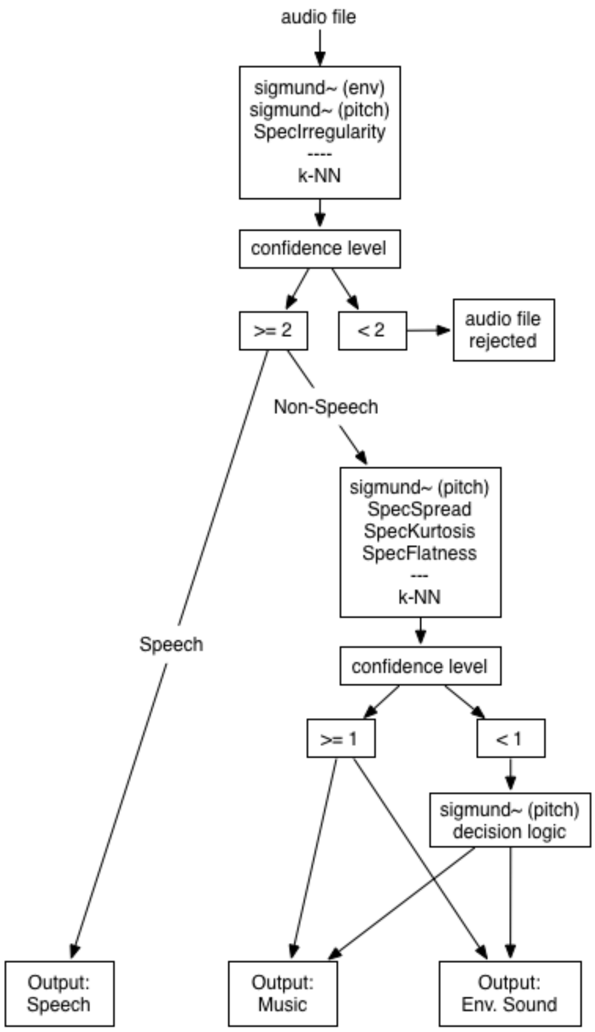


Figure 45 - Sound classification flow.

The classification task was successfully integrated in the iOS application. Performance testes are presented in 8.5.2 - Classification algorithm (pp.174).

8.2.3 - Mobile Application: client-server communication

The application is responsible for managing the bidirectional data flow between the device and the database hosted in the server. Details of the implementation are presented next.

- **From application to server**

Changes on user status and the result of sound and movement analysis are uploaded every three minutes onto a database hosted in a server. This is done through the HTTP protocol, using a FORM of type POST. The data obtained from the sound and movement analysis in PD is arranged in five strings and sent via five individual FORM fields, which incorporate the user ID, username and a timestamp of when the sound was captured (Figure 47). A new form with the classification data is created every time a sound classification event occurs. Before sending the data, the application checks if its connection to the server is available (i.e. if there is an Internet connection available and if the server is on), by sending a request to the server. Whenever the user status changes (from online to offline and vice-versa) a message is sent to the server updating this information (Figure 46).

```
// starts a form
ofxHttpForm form;

// set the action url
form.action = "http://labs.artes.ucp.pt/.../uploadNewTag.php";

// sets the method: POST or GET
form.method = OFX_HTTP_POST;

// creates a field for the formdraw
form.addFormField("user_id", t_ID);
form.addFormField("date_time", t_td);
form.addFormField("tag", t_t);
form.addFormField("volume", t_l);
form.addFormField("username", t_n);

// submits the form
httpUtils.addForm(form); // submits the form
```

Figure 47 – Code excerpt used for uploading data on the server regarding classification and movement.

```
// starts a form
ofxHttpForm form;

// set the action url
form.action = "http://labs.artes.ucp.pt/.../updateStatus.php";

// sets the method: POST or GET
form.method = OFX_HTTP_POST;

// creates a field for the form
form.addFormField("user_id", ofToString(u_id));
form.addFormField("status", ofToString(myUserCnt->getStatus()));

// submits the form
httpUtils.addForm(form); // submits the form
```

Figure 46 – Code excerpt used for updating user status.

❖ From server to application

Every thirty seconds the application reads information from the database, related with *friends'* status, sound and movement. It starts by checking if a valid connection to the server exists, using a HTTP request and response methods; then proceeds with the different requests using HTTP FORM of type

POST. All the communication is made with a web service in PHP (described later).

8.2.4 - Mobile Application: visualization module

This module is responsible for the GUI of the application, which is largely based on the display of the information regarding the user and his friends' activity (Figure 48). Each user (and his/her *friends*) is represented by an animated waveform, which has a specific shape and amplitude according to the type of sound, amplitude of the sound, movement of the device and user status. We used the analogy of waveforms so that the information was conveyed to the user in a fast, and meaningful way.

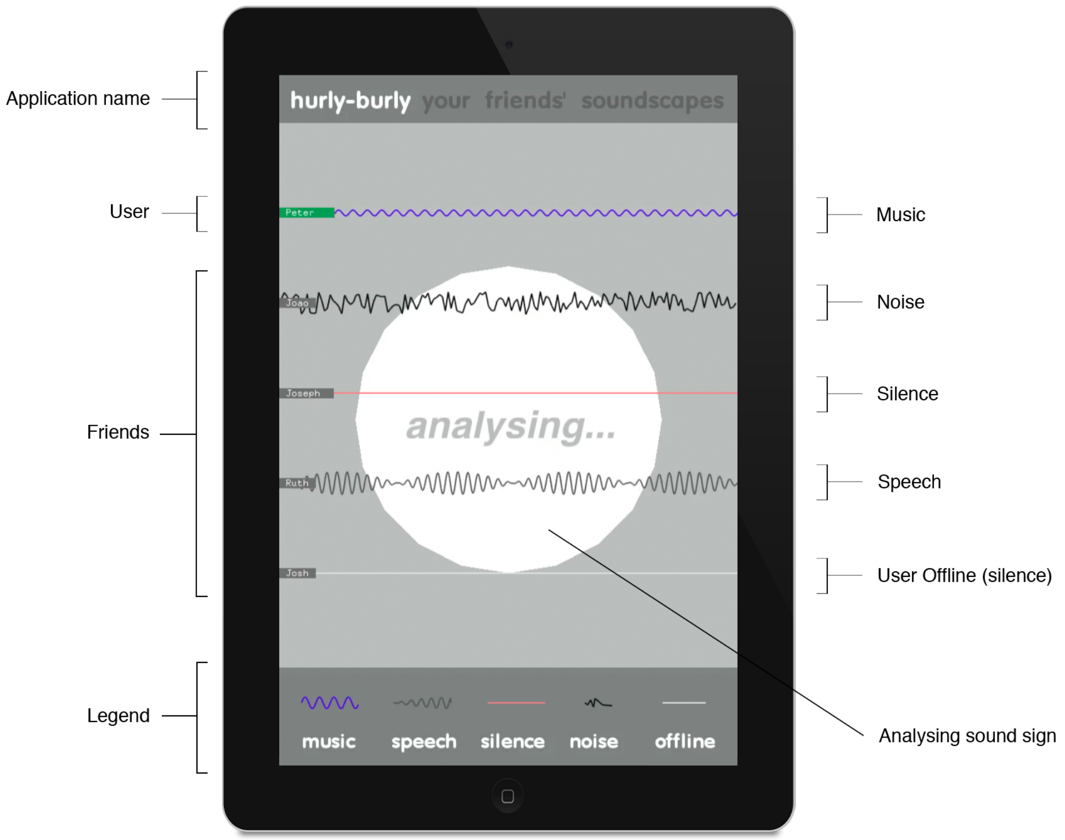


Figure 48 - Hurly-Burly's GUI explained.

The name of the user is displayed inside a grey box on the left end of the waveform, while his/her *friends'* names are inside grey boxes. Waveform's amplitude changes according to the intensity of the analysed sound: the louder the soundscape, the higher the wave's amplitude. When a waveform regards an offline user, its colour is always white, which means that the information is not up-to-date. When the user is offline, all waveforms display in white.

- Grey Sine Wave – Music
- Black Noise Wave – Environmental Sound
- Blue Sine Wave with AM – Speech
- Pink flat line – Silence

The GUI affords scrolling up and down through the list of users, passing one finger vertically across the screen. Other graphical elements of the GUI are: the bottom bar with a legend (explaining the meaning of the four types of waves and offline status), the top bar (with the name of the application and the slogan "your friends' soundscapes") and a white circle marking the analysis period.

8.2.5 - Mobile Application: sonification module

The goal of this module is to provide an auditory display for the information being conveyed to the user. By sonifying this information we are translating sound into sound and providing an audible feedback about the actual sonic status of user's social network.

An individual sound is assigned to each *friend*, depending on his/her sound type and status. The output volume of each *friend's* sound is proportional to the amplitude of his/her soundscape. A low-pass filter affects the sound of offline friends. Passing two fingers on the screen, from top to bottom, triggers the mix of all sounds. The sound stops immediately if the fingers remain on the screen after the movement. If the fingers are removed right after the movement, the sound remains for a few more seconds, depending on the acceleration of the movement. This sonification process is called a model-based sonification (Hermann, 2011), since the produced sound is a result of the system reaction to a physical interaction and is based on the properties of sonified data.

This module was not implemented on the version of the application that was used as a research tool; therefore we could not assess how users reacted to this feature. Nevertheless, a prototype of this system was built in laboratory using two distinct mappings: one with concrete sounds matching the three sound categories (dialogs for speech, traffic noise for environmental sounds and melodies from acoustic instruments for music), generating a more "real" soundscape; the other using electronic musical/noisy sounds, producing a more artificial soundscape. The outcomes of this prototype were very satisfactory. The sonification provided a good alternative to visual display, in particular when assessing the overall behaviour of the *friends'* activity.

8.3 - System Database

The client-server model is a computational structure for distributed applications, which distributes tasks between the provider of a service (server) and the requester (client). The Hurly-Burly system is based on the *client-server* model, where the application running on the device is the *client* and the database containing information regarding the users is the *server*. A database is a crucial element on the majority of web services, such as community websites, blogs and online social works. On *client-server* modules, the database is the core of the system, to which the clients access to read or write data. Hurly-Burly users – in order to properly operate with the system - have to be previously inserted on a database stored on a server. This database, besides the list of users, is responsible for saving all the data generated by those during the regular use of the application. A technical description of the implementation is given below.

8.3.1 - Technology used on the implementation

The database was created using a database management system (DBMS) called MySql, a free and open source system, owned by Oracle Corporation and regarded as one of the most popular open source DBMS used worldwide[41]. The computer server is located in the UCP facilities, in Porto, and is configured with the web server software Apache HTTP Server[42], developed by the Apache Software Foundation and also a free and open source service.

The web service used to access the database was PHP 5.2.4[43], a server-side scripting language, used as an interface between the database and the application request-response messages. The database administration was done using phpMyAdmin[44], a free and opens source tool written in PHP that provides a convenient GUI for managing MySQL databases.

❖ **Database layout**

In order to manage the users of the system and the data collected by these, three tables were created: users, users-friends and data. Bellow is a short description of each table and in Table 5 are presented the fields and field types for each table.

- Users: table containing the list and credentials of all the users of the system, including their current status.
- User-friends: table containing the relations/links between all users of the system (who is *friend* with who).

---

[41] http://db-engines.com/en/ranking/relational+dbms (retrived 4/8/2013)

[42]Software characteristics: Apache/2.2.6 (Win32) DAV/2 mod_ssl/2.2.6 OpenSSL/0.9.8e mod_autoindex_color PHP/5.2.4

[43] PHP stands for PHP: Hipertext Processor and is supported by The PHP Group. http://pt1.php.net (retrieved 4/5/2013)

[44] Developed by The phpMyAdmin Project. http://www.phpmyadmin.net/home_page/index.php (retrieved)

- Data: table containing the entries provided by each client, resulting from the analysis of the sound and movement.

Table	Fields	Type	Description
users	id	int(11)	A manual assigned ID for each user
	name	varchar(50)	Name used for authentication
	password	varchar(60)	Generic password, hardcoded on the application
	status	int(11)	Online or offline
users-friends	id	int(11)	Auto incremented, not really used
	user	int(11)	User ID (from user table)
	friend	int(11)	User ID (from user table)
data	id	int(11)	Auto incremented, not really used
	date_time	datetime	Timestamp for each new entry aaaammddhhmmss
	tag	int(11)	Type of sound
	volume	int(11)	Volume of the sound analysed
	id_user	int(11)	User ID (from user table)
	username	varchar(20)	Name (from user table)

Table 5 – Tables layout for Hurly-Burly database.

❖ Server-side API

The server-side API written in PHP was comprised of the following functions:

Client: Hurly-Burly Application

- userCertification() – queries the database for matching credentials
- uploadNewTag() – client application adds new entry on data table
- updateStatus() – updates user status on every change

- returnDataFromFriends() – queries the database for last entry of user's friends

Client: Web application for soundscape history visualization

- returnDataFromMyself() – queries the database for all the user data (for plotting purposes)

8.4 - The web application – visualising soundscape history

A web application was created to plot a timeline graph, showing the evolution of user's soundscapes and movement (Figure 49). This application is accessed on a web browser, displaying an interactive graph that users can scroll and zoom. With this resource, users can analyze their soundscapes' temporal evolution, identifying sonic patterns that otherwise would remain veiled. Each point represents an analysis event; the vertical axis is the amplitude of the sound captured. Each color represents a type of sound. The blue lien on top represents movement, when the value equals 100db means that the sample was captured while the device was moving and 80dB means that the device was steady. We used this information as part of the questionnaire presented to subjects of the study at the final stage of the experimental procedure.

The timeline graph was created with Google Visualization API45, a JavaScript API for web developers used to turn structured data into graphs. The interface between the MySQL database and the Google Visualization API was comprised of PHP and GV Streamer46, a server side API, turns data into a compliant data source for Google visualizations.

[45] https://developers.google.com/chart/ (retrieved 4/8/2013)

[46] http://www.gvstreamer.com (retrieved 4/8/2013)

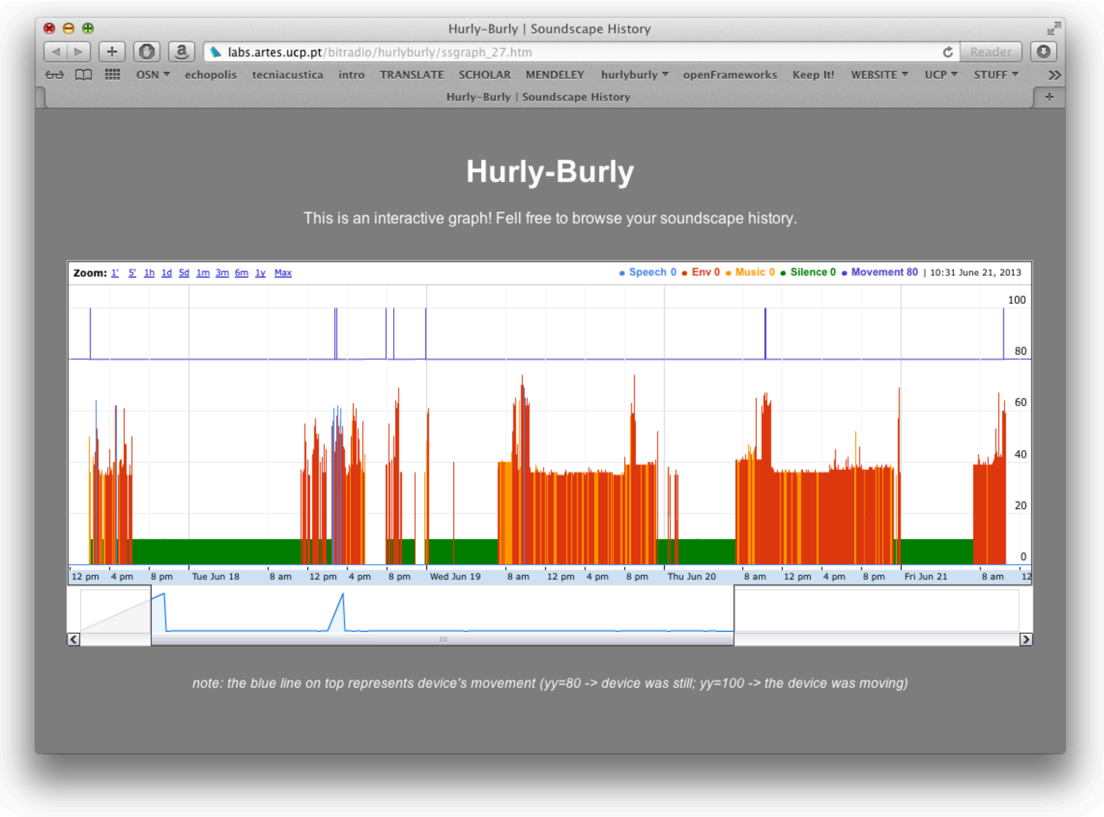


Figure 49 - Hurly-Burly: GUI for the web application.

8.5 - Evaluation of the system implementation

In this point we present some key tests about system's performance. The outcome of the tests dictated much of the implementation decisions undertaken during the development process. In general, the tests assessed the computational side of the system, such as battery life, sound classification tasks and SPL measurements.

8.5.1 - Sound Pressure Level Test

In order to assess the measurement performance and proper calibrate the application, several measurement tests were conducted, comparing the output SPL (A-Fast) from the application with a professional SPL meter. For this test we used an iPod 4th generation. At least, ten measures were made for each test signal, with intervals of 5dB (Figure 51). The results show that for pink noise and environmental sounds the variations are similar, the application tends to add 5,4dB (average) consistently along the different amplitude values, up to 80dB. From this point onwards, the application becomes less sensible, revealing a drop on the values, gradually reporting amplitudes below the SPL measurements (Figure 51, graph 6).

The test with the 100Hz sinusoidal signal shows that the application is much less sensible than the SPL Meter, due to the existence of a low pass filter applied to the input signal on the device, used to mitigate wind noise and other unwanted low frequency sounds (Figure 51, graph 2). The test with the 1000 Hz sinusoidal signal shows that the application is accurate between 40 and 80 dB, presenting an average of 0.2 dB variance. From that point onward, the application becomes suddenly much more sensible than the SPL Meter, contradicting the tendency of low frequency signals.

We have supported the calibration on the results obtain from complex sounds, since the soundscapes are usually composed of a rich cacophony of complex sounds. We have attenuated in 5dB the values obtained by the application between 40dB and 80dB and linearly decreased the signal compensation up to -6dB at 100dB (Figure 50).

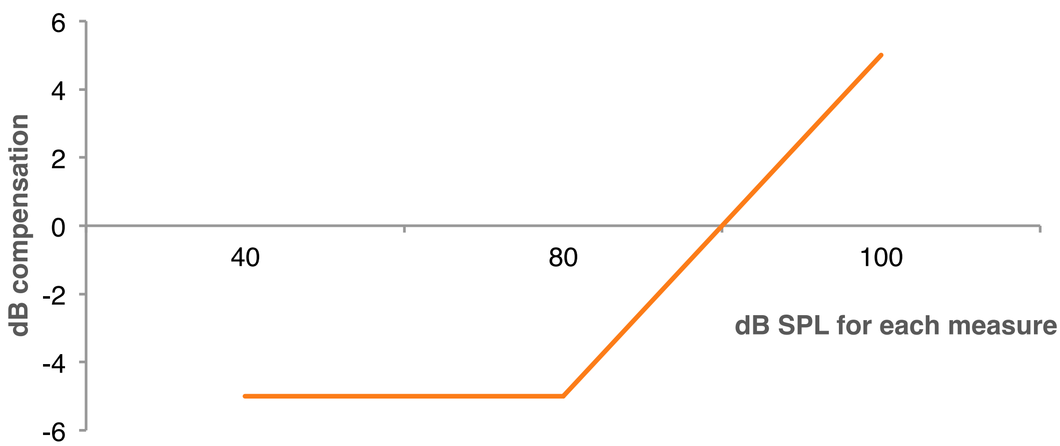


Figure 50 - SPL compensation curve for the application

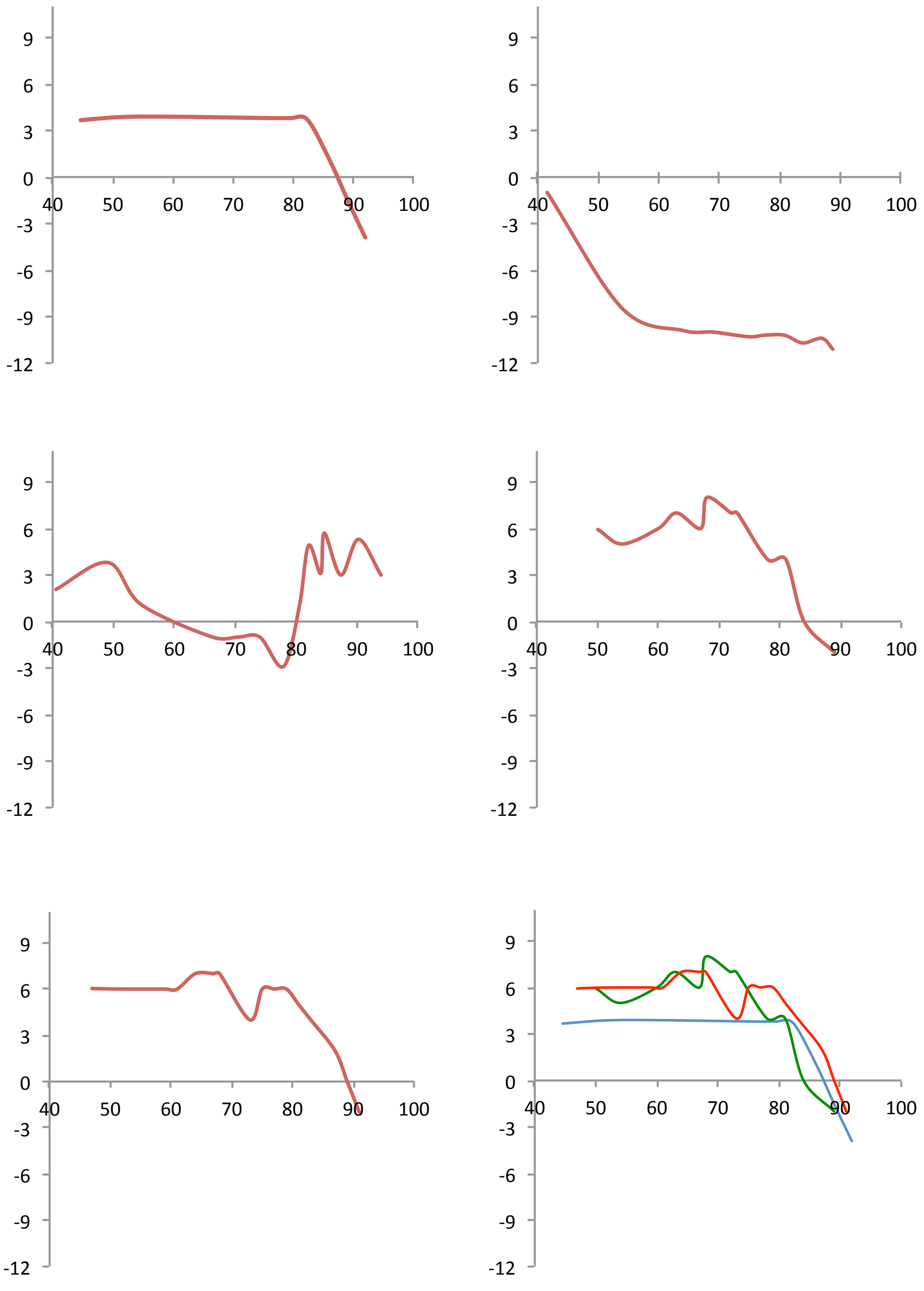


Figure 51 - SPL (A-Fast) difference between the application measurements and a SPL professional meter. Graphs 1 to 5 (left to right, top to bottom) represent the following test signals: 1) pink noise, 2) 100Hz, 3) 1000Hz, 4) pub/bar vivid ambience, and 5) lawn mower. Graph 6 represents an overlap of graphs 4 to 6 (blue: pink noise; green: pub/bar; red: mower). The horizontal axis corresponds to the dB SPL range of the test (40 to 90 dB); vertical axis corresponds to the variance (or error) of the application measurement (dB).

After calibrating the application based on the test results obtained with the iPod Touch, we used this device to compare with other iOS devices, and proceeded with a rough empirical compensations. For example, the iPad microphone is significantelly less sensible so adjustments had to be made made.

8.5.2 - Classification algorithm

Tests were conducted in laboratory to assess the performance of the classification algorithm. A database of 106 sounds (not present in the training examples) was used, containing an equal number of samples of the three categories (music, speech, environmental sounds). Part of the sounds was recorded with an iPod, while other part was obtained from freesound[47] database. The sounds that were not recorded with the iPod were equalized with a hi-pass filter (80Hz, −20 dB/octave.), in order to approximate the sample sounds to the coloration created by iPod. In addition, we favoured sounds with "natural" character, recorded in live situations, instead of purely electronic sounds and "dry" sounds recorded in studio. The tests were accomplished using a laptop unity running a modified version of the PD patch embedded in the application, only improved with a log feature.

After intensive testing, changing systematically the feature design and classifier options (distance, K index, normalization), we achieved a suitable performance for the system. From the 106 test sounds, 10 were rejected (9%) and 96 were accepted (91%), from which 92 (87%) were classified correctly and 4 (4%) were misclassified.

[47] http://freesound.org

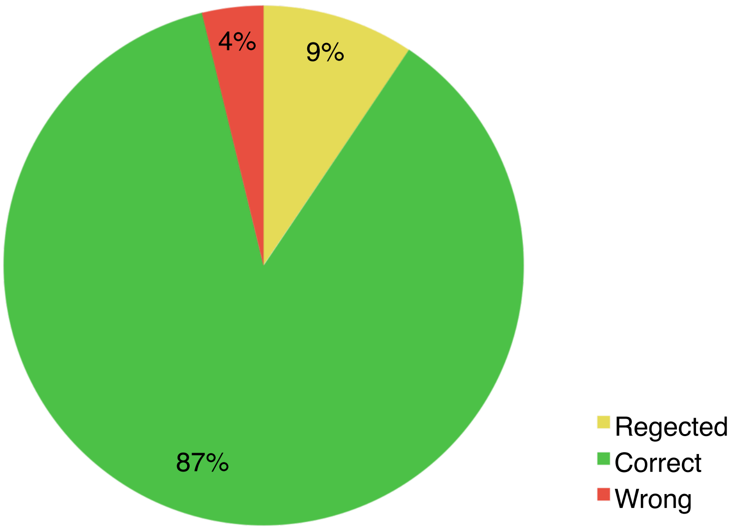


Figure 52 – Results for the performance of the sound classification algorithm.

8.5.3 - Discussion

❖ Classification algorithm

The results of the tests in laboratorial conditions proved that the algorithm performed very well, especially considering the simplified approach followed in the implementation of a machine listen system. In real life situations, when running on an actual iOS device, the algorithm maintained a good performance when classifying speech sounds, but in face of steady environmental sounds (the so called flat line sounds, such as air cons or computer fans) we noticed that sometimes were classified as music. The same way, some noisy music (with a higher spectral flatness) was sometimes classified as environmental sound. We believe that this happens because in our approach we have used, mainly, a tonal-base criterion to distinguish between musical and non-musical sounds, instead of attaining on the rhythmic characteristics of the sound, for example. In addiction, the training stage of the K-NN classifier would have benefited from a larger database of training examples, exclusively comprised of sounds captured with an iOS device in real-life environments.

Achieving a benchmark in this area was completely out of the scope of the investigation. In general, the performance of the algorithm was quite reasonable and served the purpose of a system prototype for experimental research. Some limitation of the algorithm did not hinder the scientific experiment.

◊ **SPL measurement**

The main limitation of the SPL measurement module is its low external validity. For logistic reasons, we were not able to test individually all 12 models of iOS devices that could run the application. The calibration was done with one device (iPod Touch 5th generation) and adjusted for the iPad 2 and iPhone 5. We know from experience that the microphone coloration and sensibility changes significantly among the several models.

In any case, the application was not developed as a measuring device (actually, it did not output a value for the user, only on the web application was possible to read numerical results), and the most important goal was achieved: the device was able to track relative changes in the sound pressure level. Again, we do not consider this minor constraint as an impediment for the scientific experiment validity.

◊ **Other features**

The major issue related with the use of the application in a real scenario (i.e. for a commercial application) is probably the amount of energy it consumes, which already adds on top of highly demanding devices. The fact that the application runs continually during the day and night, capturing and analysing sound and constantly communicating with the database turns this application almost unusable from the user experience perspective. In general terms, one can say that the energy consumption doubled, when compared with the normal usage of the devices. For our study, that did not constitute a real issue because we asked the participants to carry a power cable with them to avoid battery drain.

Other constraint also related with the device was the impracticality of multitasking audio. iOS devices are able to compute audio in background while handling other tasks, since these task are not audio related. When users of an iPhone received a call, the application was shut off. After the call, users had to re-start the application.

8.6 - Chapter conclusion

In this chapter we have presented in detail the design and implementation stages of the prototype system – Hurly-Burly - used as a tool in the experimental research. It was not a trivial process as it involved the participation of different areas of knowledge, ranging from machine listening to mobile computing, database and network engineering, sonification and information visualization. The final product was the outcome of an interactive process, interleaving design and implementation tasks. At every stage, implementation constraints and opportunities defied the original design, forcing us to adapt and think further. For example, the movement-tracking module was only added after the first user tests, when we realized that high levels of environmental noise corresponded not to the actual soundscape but to the noise of scratching and rattle inside people's pockets and handbags. Adding information regarding the device's movement helps understanding sound in context.

Incorrect sound classification that might eventually occur has been partially overcome by adding a manual classification option, where the users can trigger a new classification event, by double tapping on the screen. This way, whenever users spot a misclassification, they have the opportunity to request the system a new one.

Considering all the qualities and constraints of the system, we believe that the design and implementation was successfully achieved in the context of the research aims. The other options for continuous monitoring of context and social interactions would include the use of devices with which users were not familiarized (see for example the sociometric badges (Olguín &

Pentland, 2008)). Instead, ”mobile phones have become affective technologies. That is, objects which mediate the expression, display, experience and communication of feelings and emotions. Users enjoy an affective relationship with their phones and feel attached to them.” (Lasen, 2004). The affectionateness and familiarity with mobile devices was determinant for engaging users in this experiment, besides it solved logistic problems involved in the production of custom-made devices with the processing power and input/output features of modern mobile devices.

CHAPTER 9 - System Evaluation and Result Analysis

In this chapter we present the evaluation of the system from the perspective of its use among the participants of the experience. Technical analysis and performance testing were addressed in 8.5 - Evaluation of the system implementation (pp.171), therefore they are excluded from this chapter.

The analysis presented here is grounded in two different sources of data: the questionnaires and the soundscape history of the users. Both are related since the soundscape history is used as an integral part of the questionnaire. However, their essence is different and so the analytic approach will be.

We start by deconstructing and explaining the different sections of the questionnaire, clarifying what information we aimed to acquire and why. This section of the chapter could also fit in the methodology chapter, but presenting it at this stage – after exposing the user to the application features - can provide the reader with a better understanding of both, the questionnaire and the results.

9.1 - Questionnaire Design[48]

A questionnaire is a research tools widely used in several scientific areas, in particular those related with human factors. Designing a questionnaire is a process that should not be overlooked, since poorly formulated questions lead to useless data. In order to verify the validity of the questionnaire it is important producing it iteratively, always testing the questionnaire with small groups before distributing it among the final audience. We followed this approach since the first time, testing the questionnaire with individual people before distributing it among the first test group. However, when analyzing the results, we agreed that all the questions were pertinent but more were needed.

[48] The full version of the questionnaire can be found in Appendix C.

Therefore, the first round of questionnaires had less four questions than the second and third rounds (questions 1.6; 4.8; 4.9; 4.10).

The majority of the questions offer a four-point Likert scale, "forcing" the respondents towards a positive or negative answer by removing the neutral central option (although we usually offer a "n.a." option).

9.1.1 - Section Zero – Respondent Identification

The goal of this section is to collect basic identification elements from users, such as: name, age, nationality and gender (open question). Although the questionnaire is anonymous (i.e. the identification of respondents it is not made public), it is important to collect users basic identification to proceed with age studies and other demographic analysis, and be able to group them according to the corresponding test groups.

9.1.2 - Section One – Mobile Devices

The goal of this chapter is to characterize the participants' relation with mobile devices, in particularly how they connect to the Internet and deal with mobile sound. It was important to access their habits in this matter, for the screening of potential systematic bias. For example, mobile phones and music players – both could run the app – have different features and usages regarding mobile sound and Internet connection, which necessarily impacts the study. The first four questions in this section address, specifically, how users have used their devices during the test period.

Questions in this section include:

- Which device did you use to run the application?
- During the experiment, was the device always with you?
- During the experiment, how often was your device connected to the Internet?

- During the experiment, how did you connect your device to the Internet?
- When do you generally use sound in your device?
- How often did you use de app during the test period?

9.1.3 - Section two - Social Media

The goal of this section is to characterize the user's relation with social media and online social networks. Since Hurly-Burly is an OSN, it was important to trace the profile of users regarding their ONS habits. Different profiles might contribute to different appreciations and usages of the application. For example, the devices that people use to access ONS range from mobile to desktop, which definitely impacts the performance of an application as Hurly-Burly.

Questions in this section include:

- What is your favourite social media application?
- Why do you use social media applications?
- What is your MAIN goal when using social networking applications?
- Regarding social networking, what information do you find more useful to share and receive?
- How often do you use social media applications?
- How do you prefer to access social media applications?

9.1.4 - Section Three - Soundscapes

This section is comprised of only one question, which seeks to assess the knowledge of users regarding the term *soundscape*. This was important only to understand if the user is aware of the concept of "sound of a place", if user is aware of the sonic phenomenon that surrounds him/her everyday. The term *soundscape* is very specific (unlike landscape, for example) and those who know what it means can have a better chance to understand the application goals.

Question in this section:

- Do you know what a soundscape is?

9.1.5 - Section Four - Hurly Burly Application

In this section we address the opinion of the participants about Hurly-Burly application and the quality of their *user experience*, such as usability, easiness, GUI and performance. Furthermore, we present hypothetical scenarios for the evolution of the application, which challenge the concept of privacy and address the noise pollution issue.

Questions in this section include:

- Did you understand the main goal of the application?
- Was it easy to use?
- Did the application worked without problems?
- From your experience using the application, what were the main problems you detected?
- Do you think that knowing your friends soundscapes may enhance social interactions
- Imagine that instead of sharing sound information, you were sharing the actual sound. What do you think of that?
- Noise problems affect the lives of numerous people in urban and industrial centres. Do you think that applications like Hurly-Burly can be used as tools to monitor and raise awareness for this problem?
- What would be the first reason to quit the app?
- What do you think about the graphic representation of the soud-scapes (waveforms)?

9.1.6 - Section five – The Soundscape History

This is also a one-question section, dedicated to the soundscape history plotted on the web application. Users are asked to check the plot of their soundscape history and report if they find any relation with their real life. In other words, if they can identify in the patterns of the graph their activity during the test period.

Question in this section:

- By looking into your soundscape history graph, can you identify a match between the graphic representation of the sound and your activity during the experiment?

9.2 - Data Evaluation – Questionnaire

In this point we present the results of the data collected through the questionnaires. The information is separated in three logical parts that do not necessarily correspond to the original sections of the questionnaire. The parts are:

1) *Characterization of the sample* regarding its demographics, relation with social media, mobile sound and soundscapes;

2) *The user experience* of the application.

3) *User's opinion about soundscapes and OSN*.

The three test groups are different in size; therefore, we will use normalized graphic representations whenever we compare one against others. Additionally, not all the questions are subject to group comparison and some are not discussed in this section. In both cases, omissions occur because no added value would have been brought to the study.

9.2.1 - Characterization of the sample

❖ **Gender**

The sample is comprised of 48% male respondents against 52% female. While there is a balance in the overall sample, the group comparison show that USJ and HB groups have a higher rate of female participants than UCP.

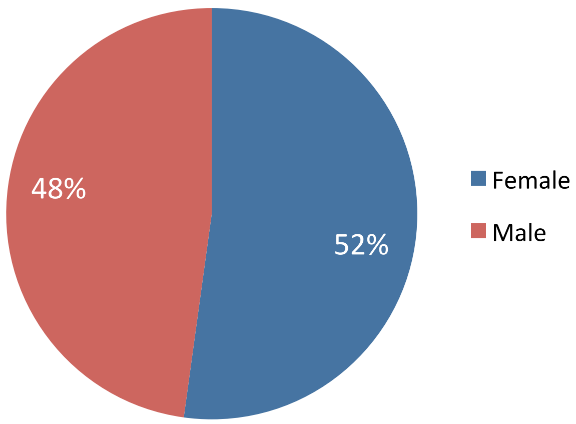


Figure 53 - Gender distribution of the sample

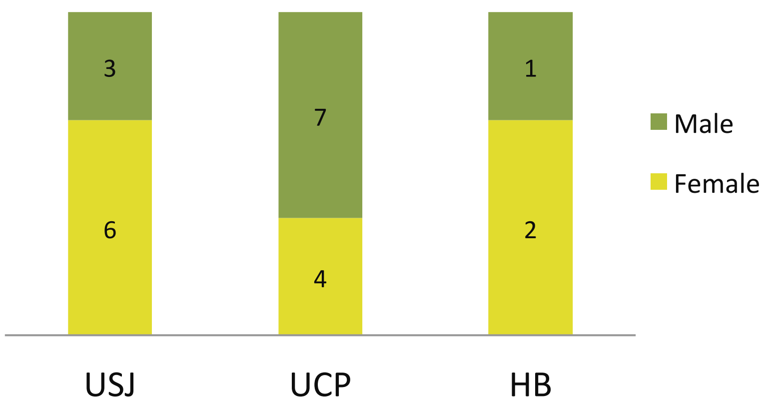


Figure 54 - Gender distribution in each group (normalized graph)

❖ Age

Regarding the age of the participants, the groups are also distinct, in particular USJ and UCP. The former is comprised of undergraduate students with an average age of 20 years old. The last is comprised of teachers and researchers, showing an average age of 34. There is a gap of fourteen years between these two groups (Figure 55). HB group is situated approximately in this gap, presenting an average age of 27 years old.

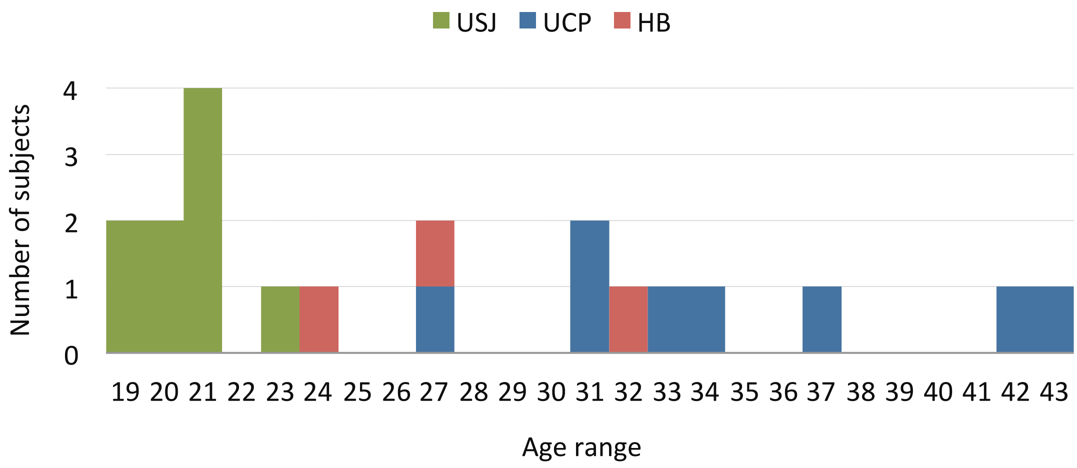


Figure 55 - Age distribution of the subjects from different groups

❖ Nationality

The nationality of the subjects is very diverse due to the fact that the experience took place in two different continents. Moreover, universities tend to be a place of international confluence. Also in this variable the groups differ: HB and UCP tend to mainly constituted by Portuguese people, while the USJ group is mainly comprised of Asian nationalities.

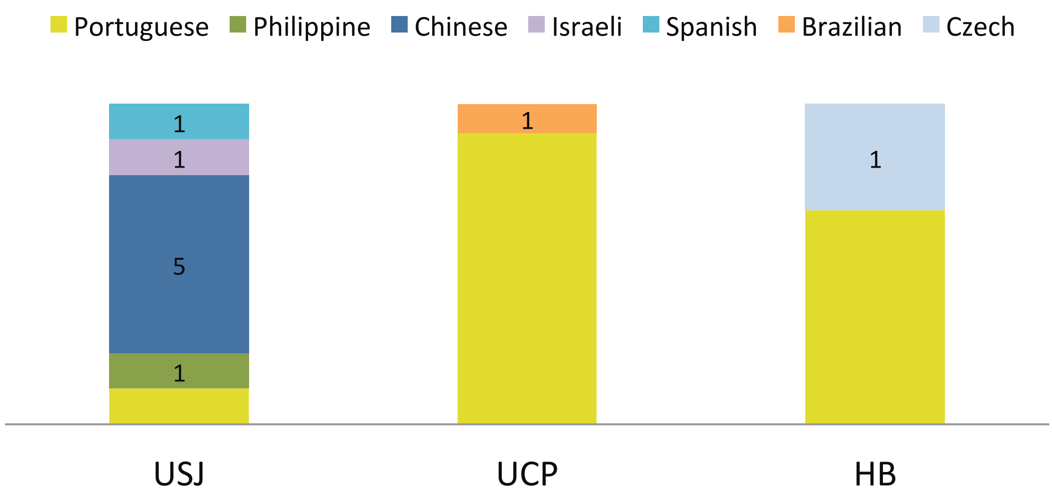


Figure 56 – Subjects nationality distribution (per group - number of answers / normalized graph)

❖ Social Media and Mobile Devices Habits

Question 2.1 asked participants what was their favorite SM application. Facebook was, by far, the most popular SM among participants of every groups (Figure 57).

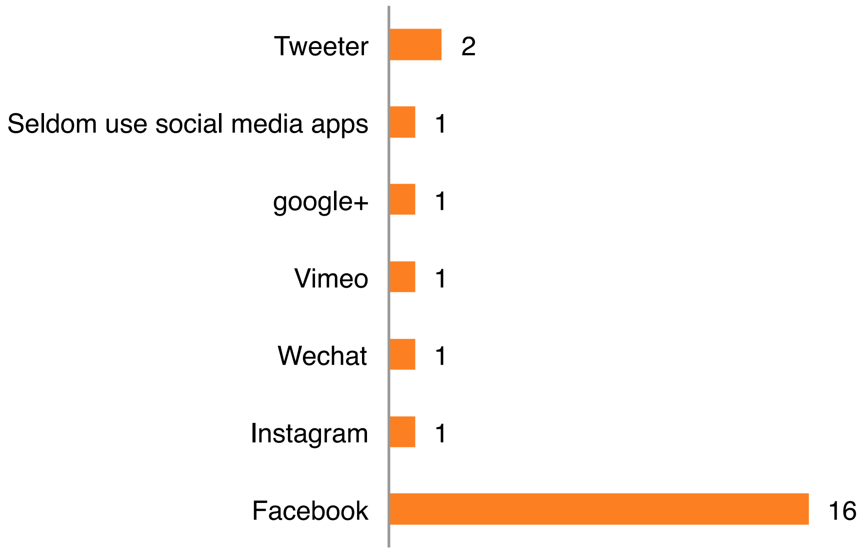


Figure 57 - Social Media Landscape among subjects (number of answers)

When enquired about the regularity in which they use SM (question 2.5), users demonstrate a high commitment with this kind of services, regardless of the test group they belong to (Figure 58), with 78% saying that they use them everyday.

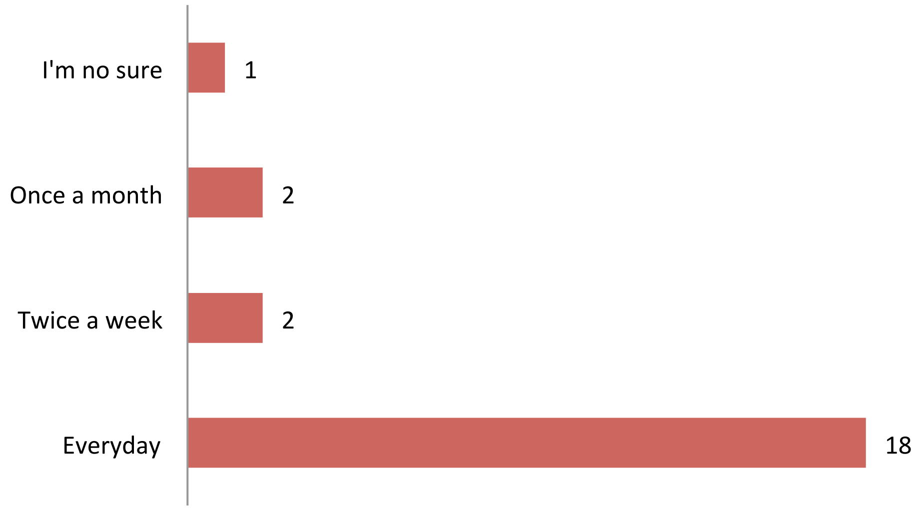


Figure 58 - Characterization of the sample according to the usage frequency of OSN (number of answers)

The reasons why people use SM tend naturally to the “social” dimension of these services, characterized by providing an online link with their peers through text messages, audio, *likes*, voice chats, etc. (Figure 59). However, this focus is not completely homogeneous through all test groups. For example, the younger group of USJ clearly prioritizes *leisure* as the second reason for using SM, while the older groups are more focused on the selective information they can collect via this service (Figure 60).

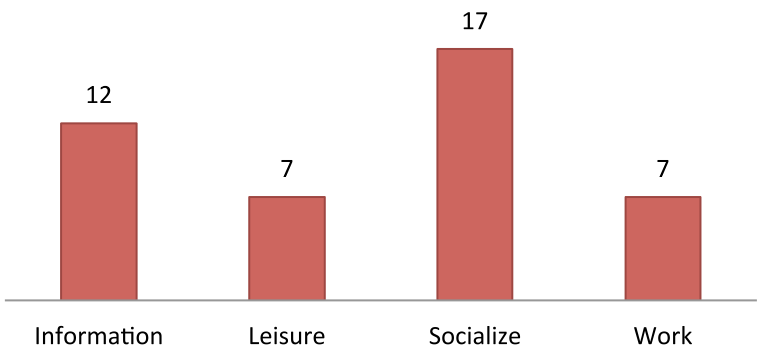


Figure 59 - Main goals of using SM (number of answers)

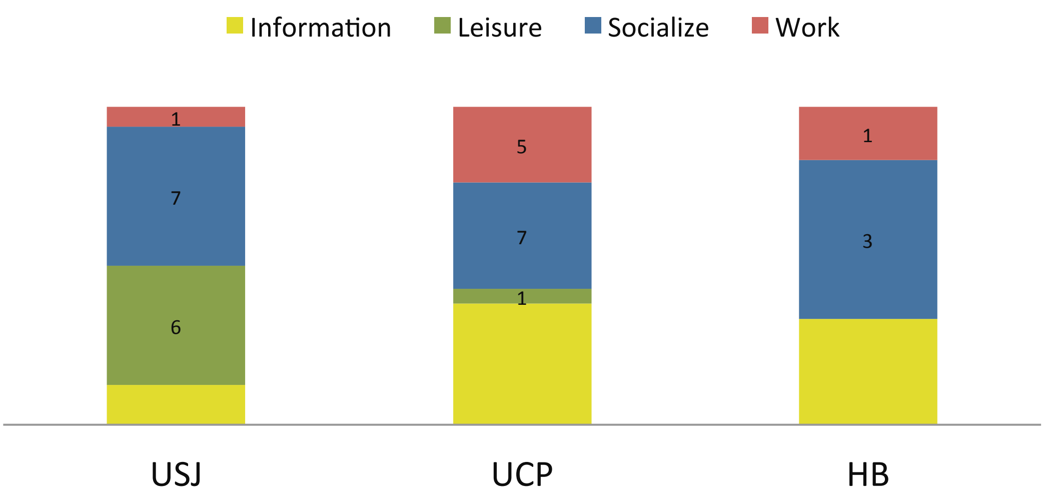


Figure 60 - Main goals of using SM (per groups / number of answers / normalized graph)

Focusing on the social dimension of SM, we asked participants what they were looking for: meeting new people or mirroring their offline social network (question 2.3). In general, the answers were unanimous: participants use OSNs mainly as an extension in the virtual domain of their offline social network (Figure 61).

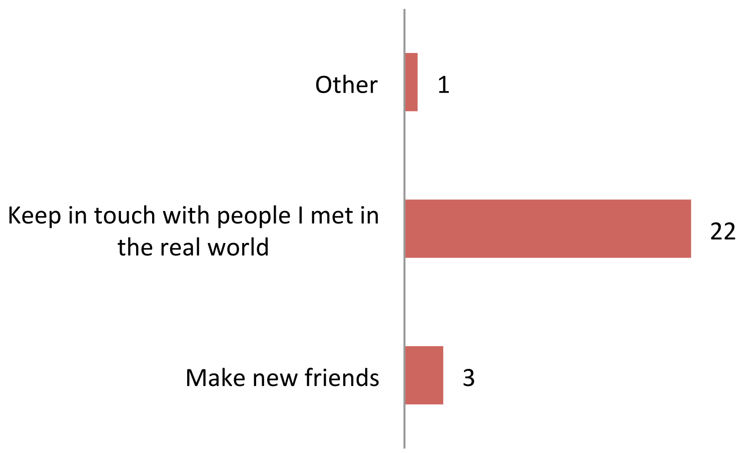


Figure 61 - Social motivations to use OSN

When asked about the content they like to share and receive on SM (question 2.4), answers tend to be homogeneous among the groups, electing *photos* as the most popular media, followed by text and *likes* (Figure 62).

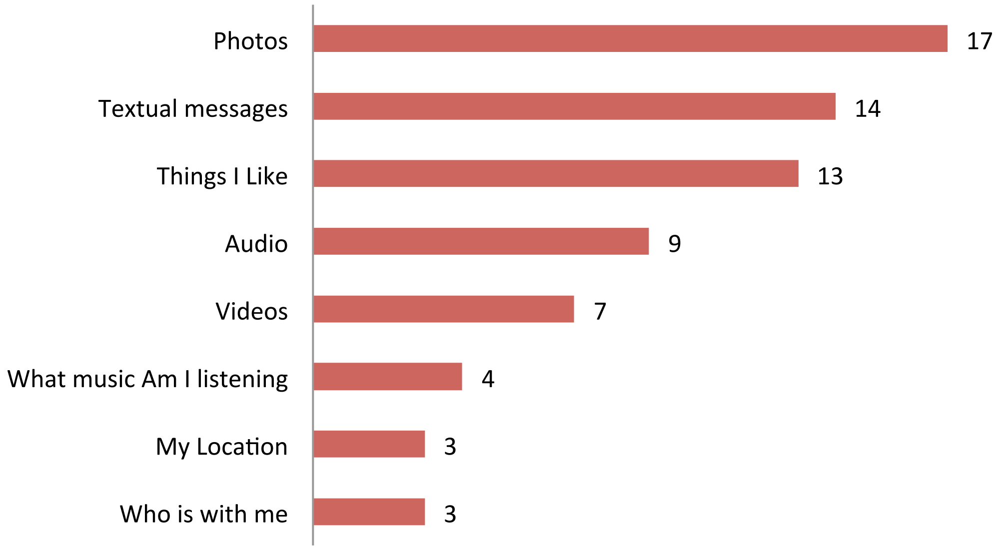


Figure 62 – User's sharing preferences in SM (number of answers)

Regarding the devices used to access SM, users do prefer laptop computers to mobile devices (question 2.6). However, if considering tablets as mobile devices, the statistics tend to even. This distribution is relatively homogeneous among the three test groups.

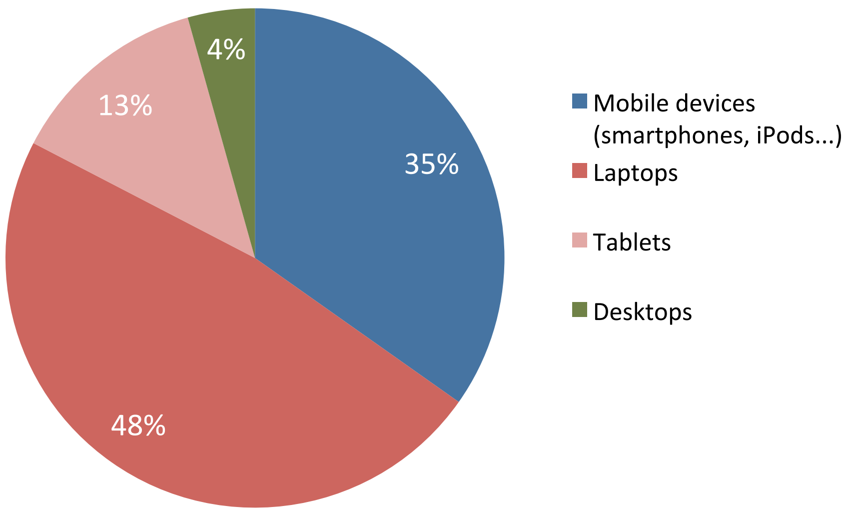


Figure 63 - Kind of devices used to access SM

When enquired about the use of sound in mobile devices (question 1.5), users nominate *listening to music* as the principal activity. This appears ahead of *calls* in the rank (Figure 64). However, when differentiating iPhone users from the other users, the former do elect *calls* as the main sound-related activity.

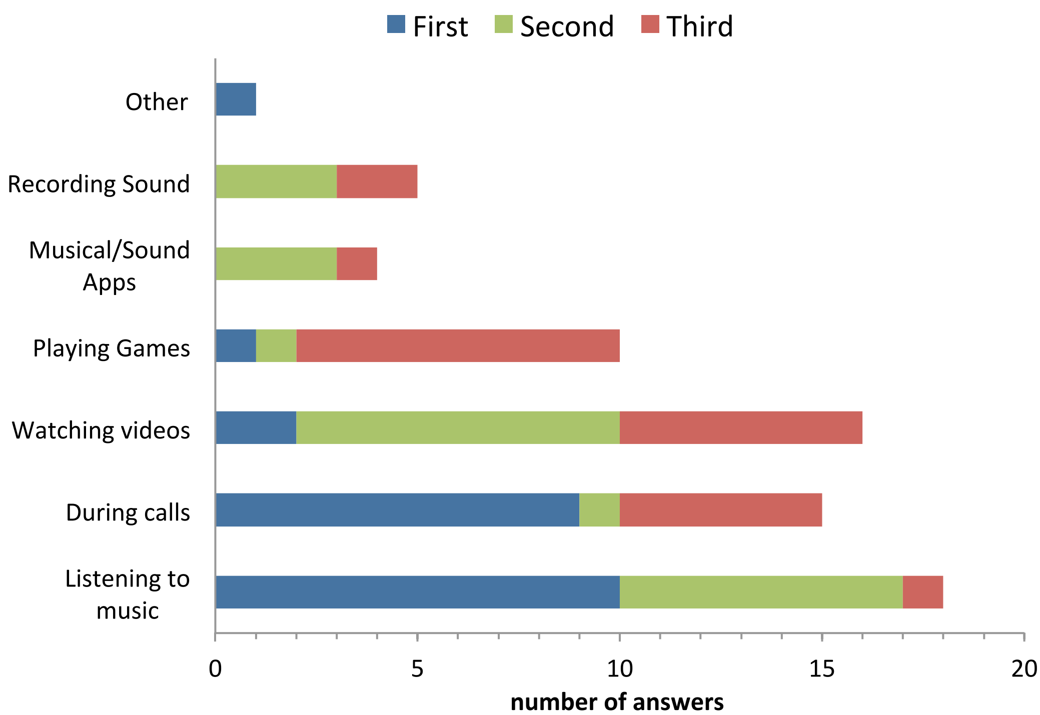


Figure 64 - Usage of sound in mobile devices (ranking)

9.2.2 - User experience

In this point we analyze the results from the questions regarding the overall *user experience* on Hurly-Burly application. We start with the questions focusing on the device used to run the application and its Internet connectivity and we finish with questions regarding the use of the application. A discussion of the results is presented in point 9.4 - Discussion (pp. 204).

❖ Device and Internet Connection

The iPhone was the main device used among participants, 65% against 17% for the iPad and 18% for the iPod (question 1.1).

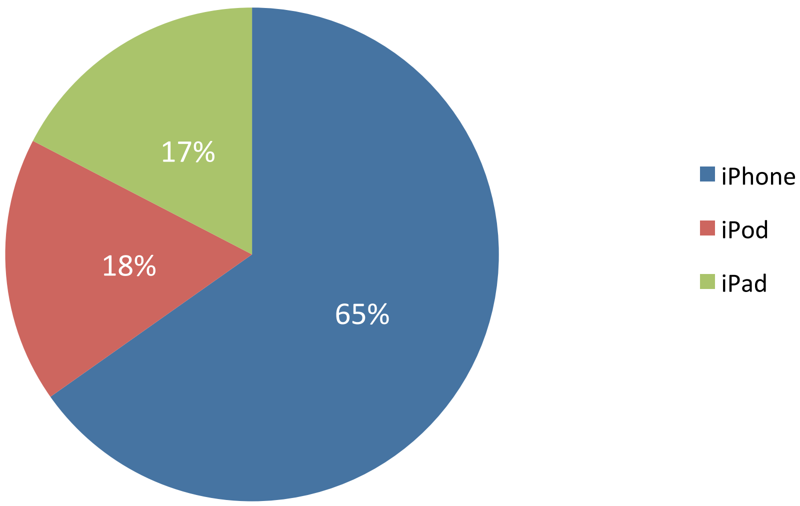


Figure 65 - Device distribution (during the experience)

During the experiment, 48% of the users carried the devices always with them and 52% did it most of the time (question 1.2).

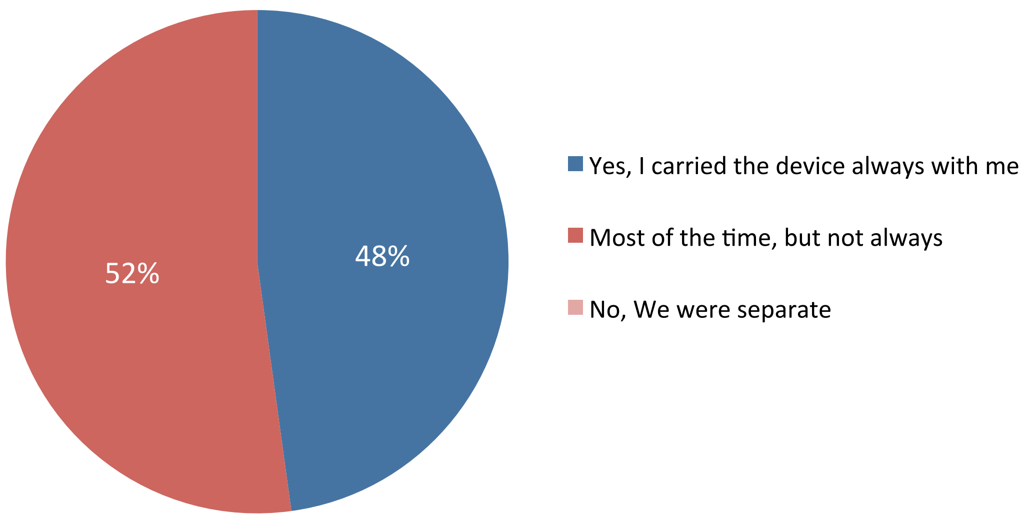


Figure 66 - Proximity between user and device during the experiment

Regarding the connection of the device to the Internet, 78% of the users had the device always or almost always connected, while 18% were connected half the day and 4% only occasionally connected to the Internet (question 1.3 - Figure 67).

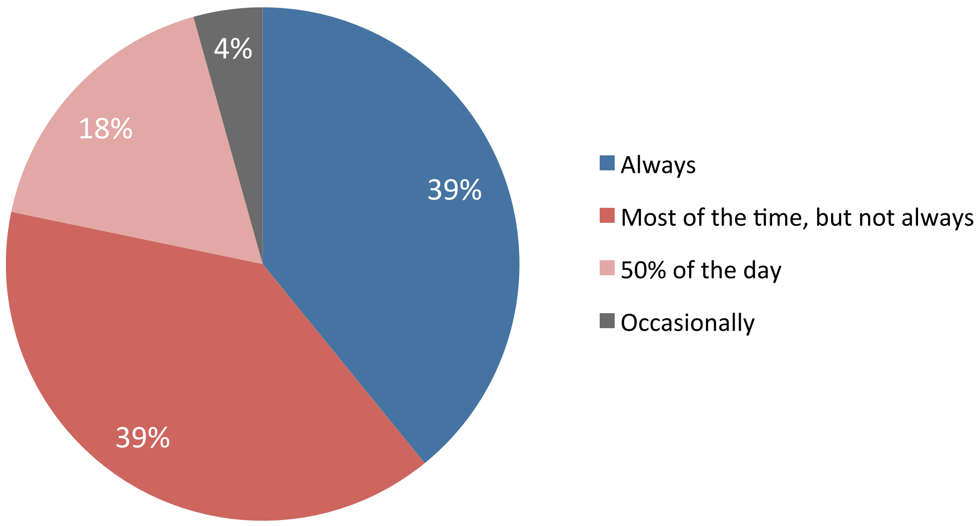


Figure 67 – Characterization of the device's connection to the Internet during the experiment

❖ User experience

This section is related to the users' appreciation of their experience while interacting with the application, generally called *user experience*. The first question assessing this aspect is regarding the users' general understanding of the application goals (question 4.1). From a scale of 1 to 4, where 1 is "understood completely" and 4 is "I did not understood at all", 52% of the enquiries evaluated their knowledge as "1" (excellent) and 39% as "2". Only 9% replied "2" and no one reported total ignorance about the application goals (Figure 68).

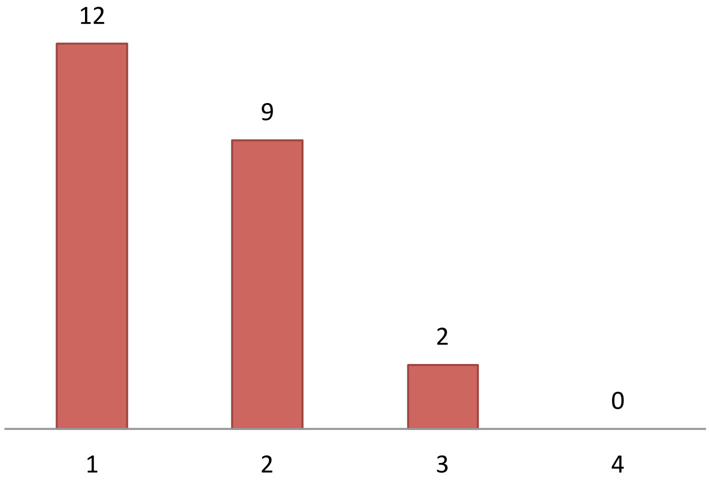


Figure 68 - Understanding of the application goals (1 – very, 4 – nothing / number of answers)

Regarding the easiness of working with the application, in a scale of 1 to 4, the results were also positive, with 61% of users evaluating as “1”, 17% as “2” and 22% as “3” (question 4.2 – Figure 69).

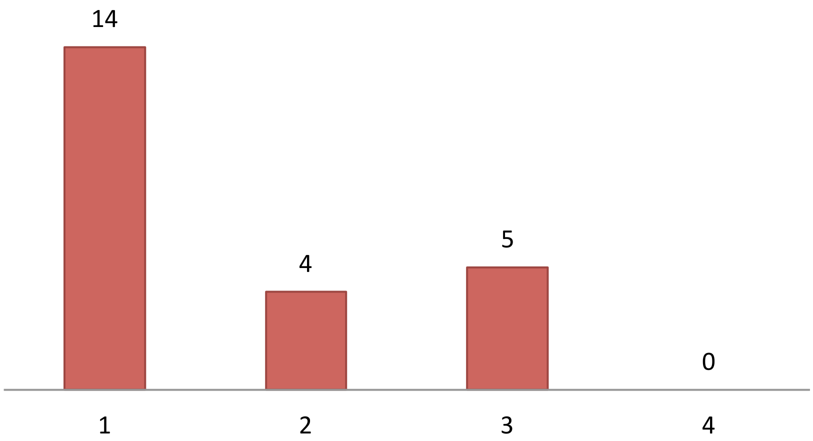


Figure 69 - Evolution of the application easiness (1- easy, 4 hard / number of answers)

Concerning the application performance (question 4.3), the majority of the users evaluated it as “working well” (48%) or “working with minor problems” (34%), while 9% reported “major problems” and 9% did not produce any evaluation (Figure 70). However, separating the evaluation results by the users of different devices, the results show that iPhones users reported significantly more problems than other users (Figure 71).

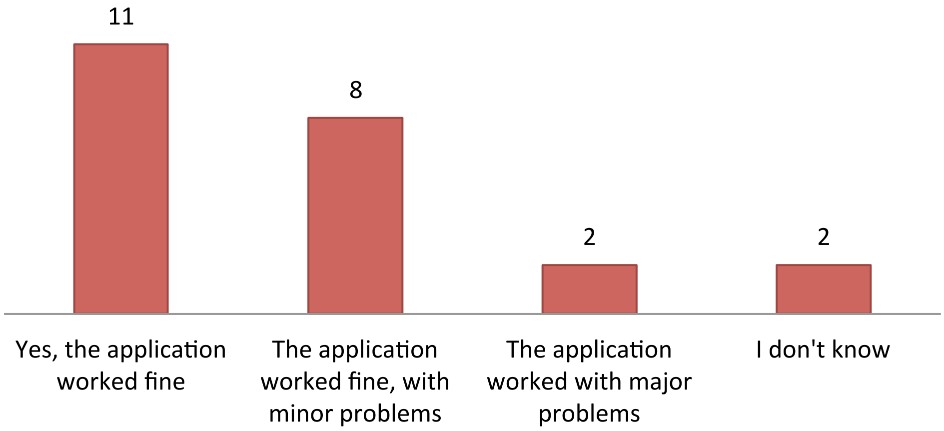


Figure 70 - Evaluation of the application performance (number of answers)

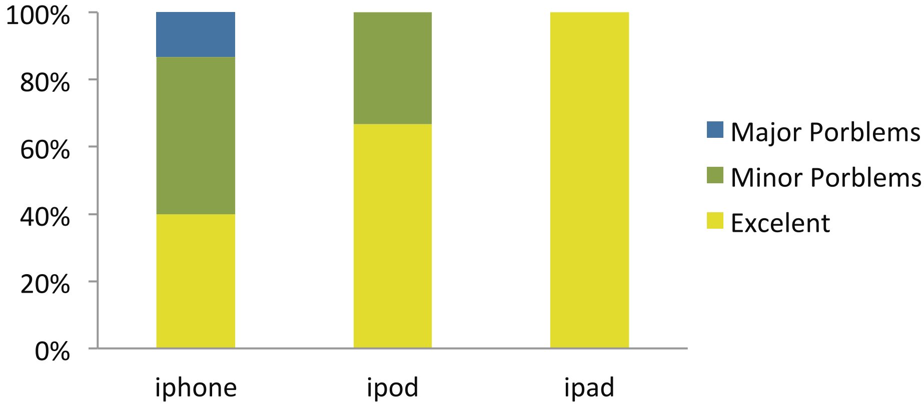


Figure 71 - Evaluation of the application performance (distribution according the devices types / normalized graph)

When questioned about the first reason that would take them to quit the application (question 4.8), 72% of the users replied “low battery” issues, 21% evoked “privacy issues” and 7% mention “all the reasons” as valid to quit the application (this question was posed only to the group UCP and HB – Figure 72).

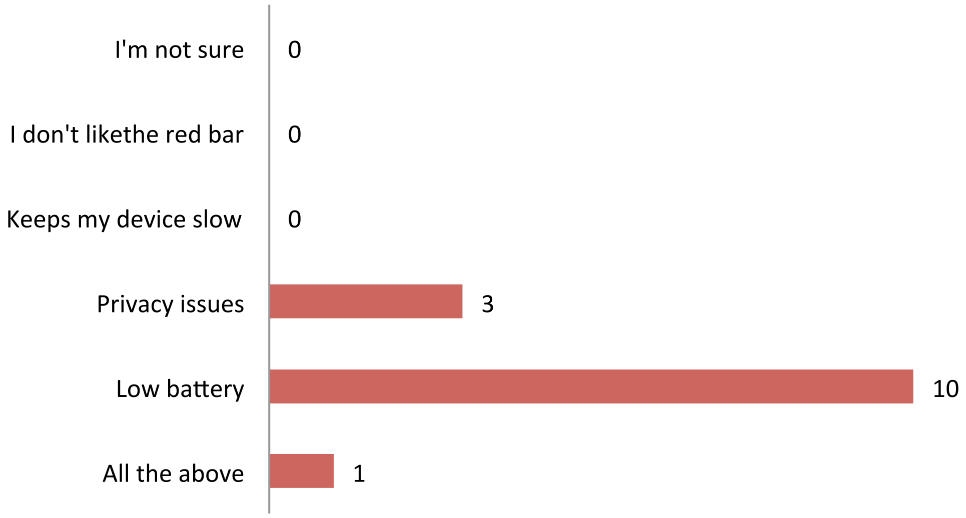


Figure 72 - First reason to shut down the application (number of answers)

When enquired about the GUI, in particular the representation of the sound, 69% of the users replied that they like the representation and understand the match, 15% understood but did not like, 8% did not understand and 8% was not sure (this question was posed only to the group UCP and HB – Figure 73).

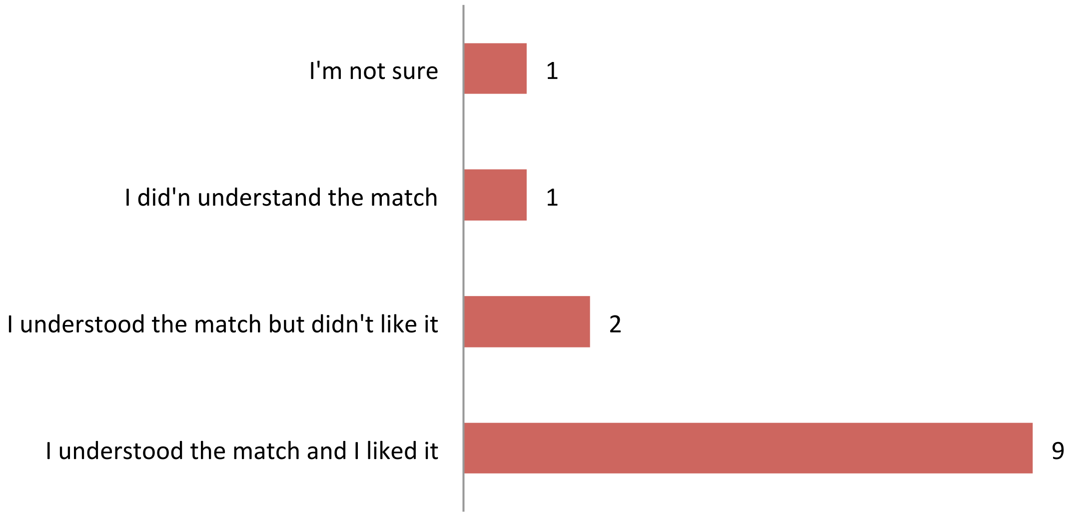


Figure 73 - Evaluation of the visual output regarding users' sonic activity (number of answers)

❖ Users' opinion on soundscapes and SM

In this point we present the results from questions addressing the topic of soundscapes in relation with mobile media and social media. We started by asking people if they were familiar with the concept of soundscape (question 3.1). This was important since knowing this concept suggests that the person have had contact with theories on environmental sound and therefore is aware of its existence in a more formalized way. Additionally, most of the subjects were students, researchers or professors of media studies and we knew that some, at least, were familiar with the topic. The results of the first question matched our expectations, 70% said they knew the concept of soundscape, 26% *more or less* and only 4% ignored the concept (Figure 74).

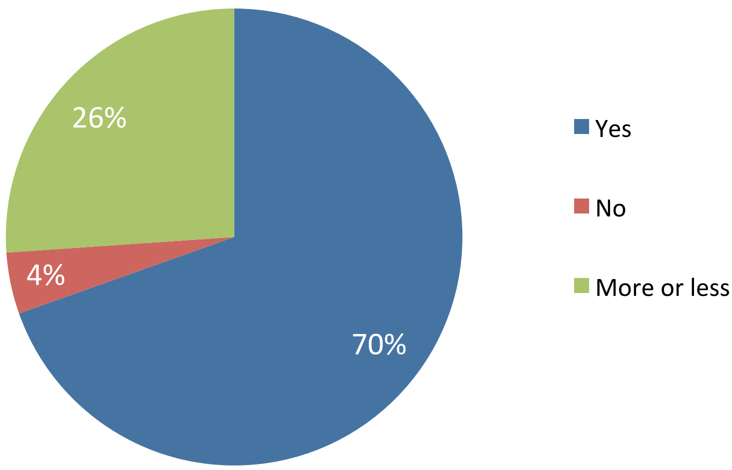


Figure 74 - Users' knowledge of the term soundscape

The second question related to the topic of soundscapes addressed directly the potential of soundscapes as an enhancer of social interrelations (question 4.5). From all the respondents, 4% denied this possibility and 26% were very skeptical about it, however 26% recognize some potential and 44% are sure about it (Figure 75).

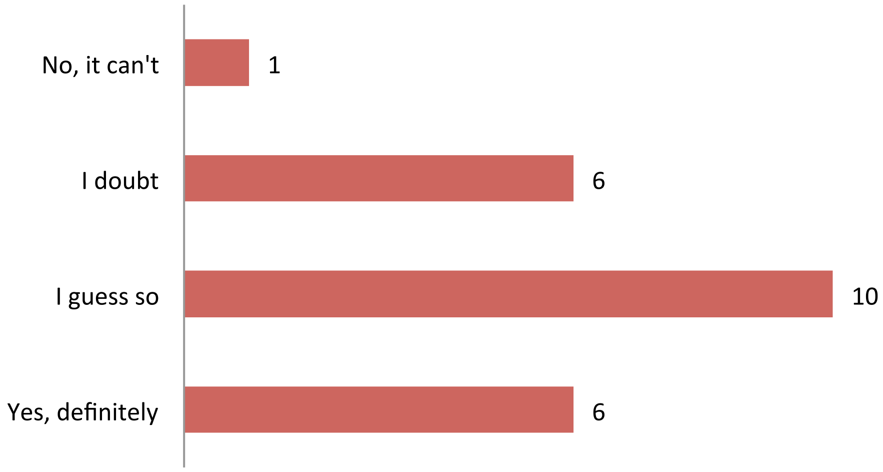


Figure 75 – How subjects evaluate soundscapes as a social enhancer (number of respondents).

Question 4.6 addresses the issue of privacy in the context of OSN. We asked participants to imagine a new version of Hurly-Burly application where the actual soundscape (the sound recording) was shared within the network. The results show that 17% of the users would never use an application like that and 39% would use it in very controlled situations only. 31% said that they would use it with caution and merely 13% evaluated this feature as very appropriate and useful (Figure 76).

Figure 76 - Subjects privacy concerns regarding the sharing of audio clips (number of respondents)

When correlating these results with those obtained in the previous question, one observes a direct relation between the willingness to share sound and the acknowledgment of HB as an ONS enhancer (correlation = 0,6; Figure 77).

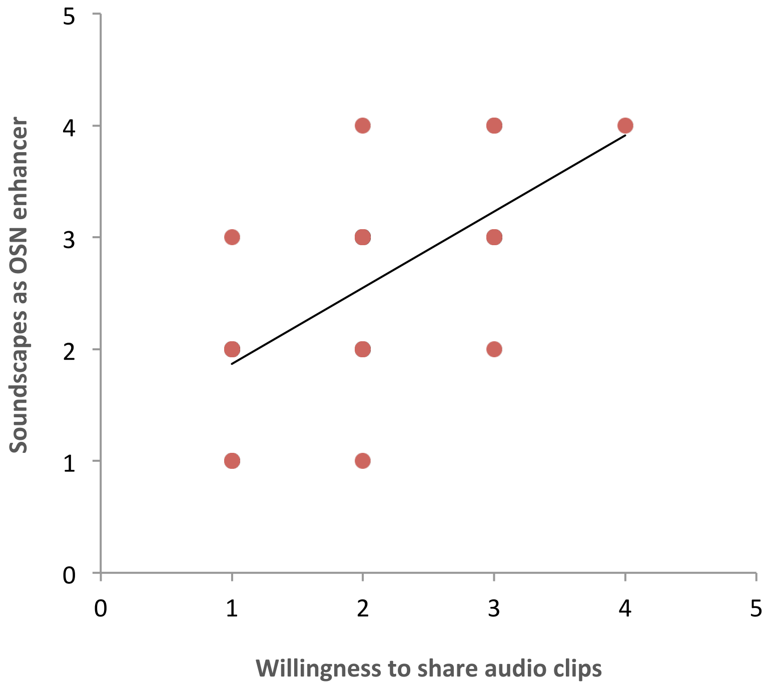


Figure 77 - Correlation between users appreciation of soundscapes as a SN enhancer and their willingness to share audio clips

Question 4.7 addressed the use of HB-like applications as a tool for active participation of citizens in noise control activities, namely by monitoring noise pollution levels. In this matter, 61% of users agree that such applications are useful for this end, 30% believe in that possibility but are not sure and only 9% are skeptical but do not deny it (Figure 78).

Figure 78 - Users opinion about HB-like systems being used as environmental tools

Also in this topic, when we distribute the respondents according to their knowledge about the term soundscape, becomes obvious that those who know the meaning of the word are more inclined to accept HB-like applications as monitoring tools at the service of noise mitigation actions (Figure 79).

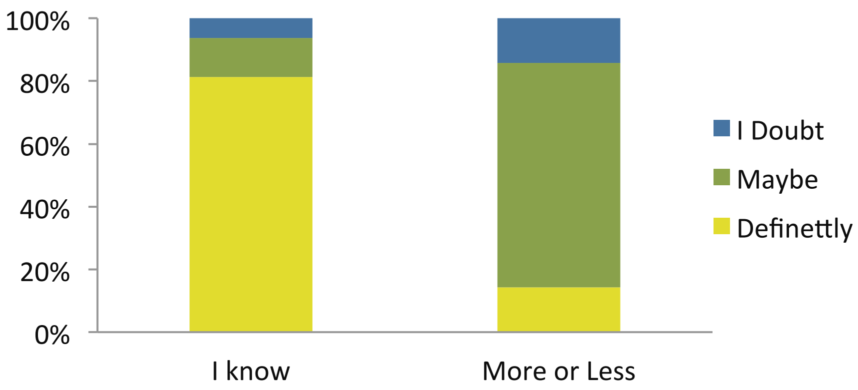


Figure 79 - Distribution of users expectation (number of answers / normalized graph)

Finally, we enquired participants about the plotting of their soundscape history, which they accessed through the web application. The results show that 46% of the users recognized the match between the plot and their activity during the test, although some periods did not match. 46% affirmed that they were able to recognize some parts but not in general and 8% did not find any match (only HB and UCP groups answered this question; Figure 80).

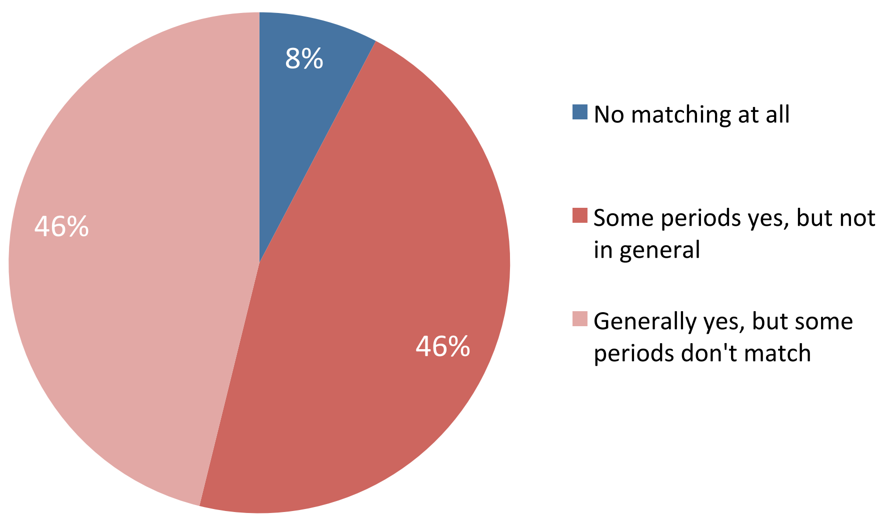


Figure 80 - Subjects evaluation on the matching between their activity and the soundscape history plot

9.3 - Data Analysis – The Soundscape History Plot

In this point we present the analysis of the soundscape history plot, based on the graphs of users that have gathered a relevant amount of data. The soundscape history plot was available to the users through a web application (running in a browser) and is based on the continuous soundscape analysis and movement of the user, monitored by the mobile application. Following, we present three examples that demonstrate the existence of sonic patterns that act as sonic footprints of the users activity.

9.3.1 - Research Office at CITAR

This first example displays data gathered with a device that remained static inside an office at CITAR. This atypical utilization for the application worked as a control test. The experiment ran in optimal conditions, always connected to the Internet and to the power source, allowing a continuous monitoring of the office's soundscape. The graph (Figure 81) displays a long term monitoring from June 11th to June 28th. A zoom out perspective shows the day and night periods, represented by the red and green spots, accordingly. The timespan of this view includes two weekends: 15-16 and 22-23. During these periods the normal activity is dimmed and therefore the concentration of red spikes is much lower. The profile of the sound peaks is an accurate indicator of the room activity. For example, the amount of green/silence on the weekend of 22-23 e larger because it includes the municipal holyday of S. João, and on June 19th a late meeting took place at the office, which is clearly represented by the blues spikes.

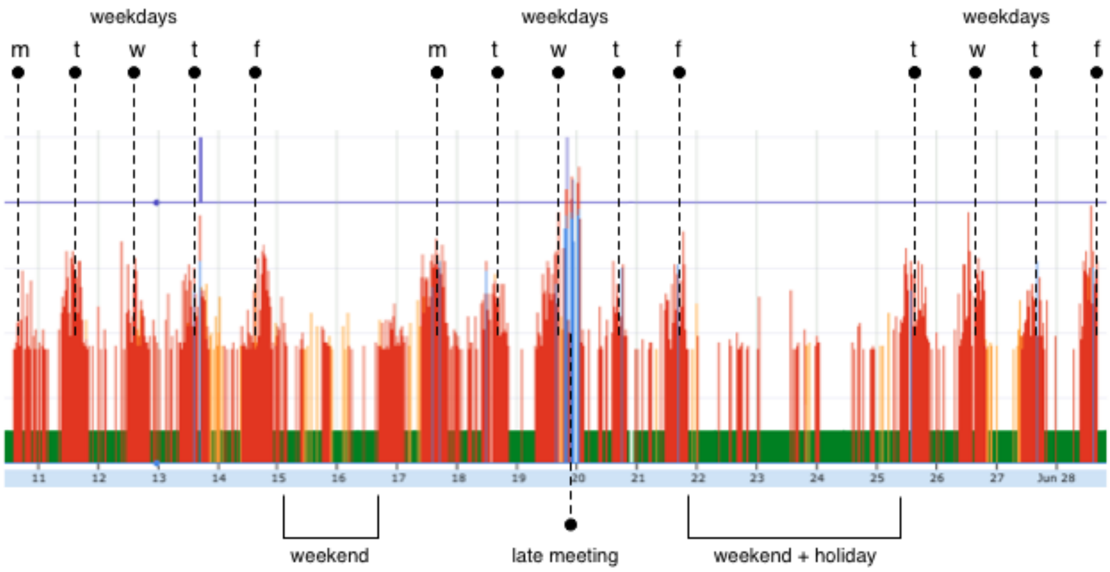


Figure 81 - Soundscape history plot - PhD Students Room (3 weeks)

Zooming-in at the timeframe of the day when the meeting occurred, the activity profile becomes clearer (Figure 82). Around 9am some noises started to emerge, probably arriving from outside the office. These are quite sounds, not continuous in time. At 10am the sound intensity clearly increases, corresponding to the arrival of the researchers sitting in that the room. From 1pm to 2pm it is lunchtime; silence is observed once again, with occasional spikes caused by an internee that remains at the office during lunchtime. From 4:30pm to 5pm there is a data gap; at 5pm everyone leaves the room and returns later for the meeting. At 7pm the meeting starts, the device that is capturing the sound is placed at the center of the table; therefore, the reported sound intensity is louder than the usual sound environment. There are four men in the room participating in the meeting and one participating via videoconference. This fact makes the sound intensity to be louder than usual, since people tend to shout to the computer. At 8pm we interrupt the meeting and went to a restaurant; the device remained at the meeting room. Noises from computers are sufficient to trigger some occasionally red spikes. At 9pm we enter the room and resume the meeting, which lasts until 1am. The sound decreases gradually as some researchers leave the meeting, including that using videoconference. The last part, there are more predominance of red spikes caused by the sound produced by moving the chairs and packing the computers. Around 1:20am everyone abandons the room. Silence emerges. Researchers get back to work at 2pm.

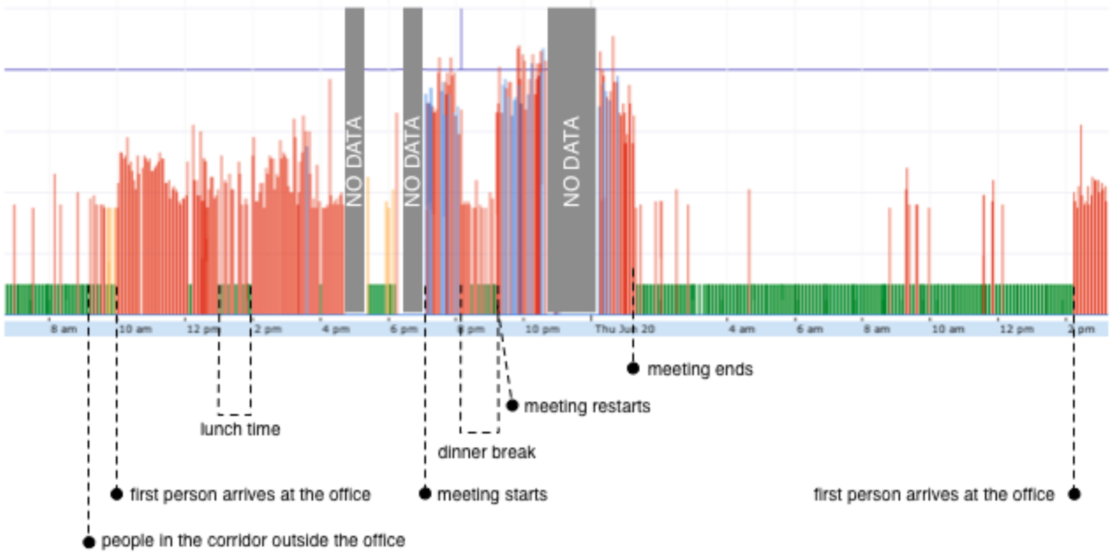


Figure 82 - Soundscape History Plot - PhD students' room (late meeting)

9.3.2 - UCP Participant

In this point we present the analysis of a graph corresponding to one participant of the UCP group. This participant ran the application on an iPad continually during two days, reporting that both remained at home most of this period of time. The iPad was in the living room and the participant developed paper work, mostly in a quiet environment interleaved with light music in background. In fact, the graph (Figure 83) shows an interesting pattern, with a sonic profile almost identical in both days of the experience. At 7:30am the activity starts in the house, showing a steady and slightly above-average sound intensity, arriving from the morning activity (shower, breakfast). Around 10am the louder activity is related with the TV, which was switched-on during a work break. Sleep time is around 24pm, occasionally preceded by a TV zapping.

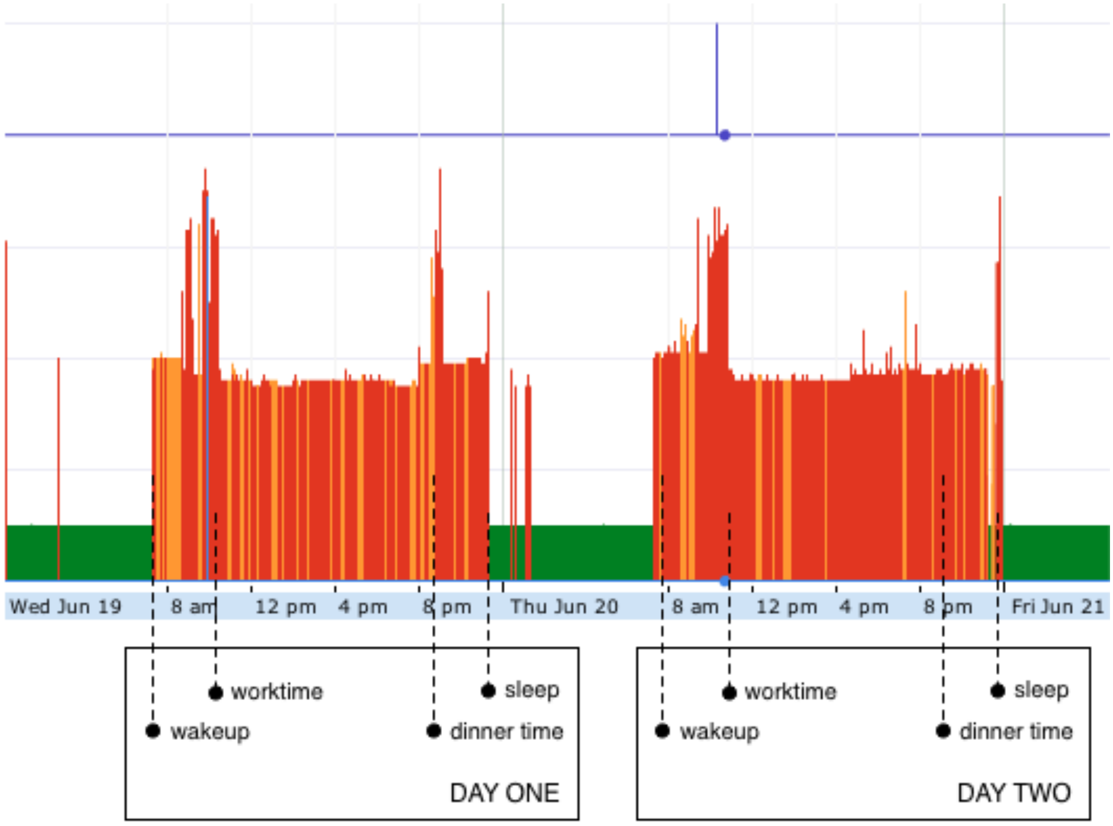


Figure 83 - Soundscape History Plot - UCP member

9.3.3 - USJ Participant

The last example of a soundscape history plot corresponds to a participant from the USJ group, which used an iPod to run the application. The device remained always steady at home, an apartment in a crowded area of Macau. In the excerpt shown (Figure 84) it is possible to identify the night and day periods, clearly noticeable by a change in the sound intensity. From around 1am when the app was started until 9:30am the sound is quieter. At that time, the activity in the house increases; the inhabitants wake up (including two young children) and around 11pm the sound intensity drops when the inhabitants go to sleep. In this example, we are not able to provide a detailed report on the activity of the users since we were not present nor this was asked to the participants. However, the most interesting aspect of this example is the fact that the sound captured is never analyzed as “silence”. This happens because Macau is a very crowded city (the most populated place in the world) and suffers from noise pollution. This apartment is no exception, even during the night, the sound intensity level it is never as low as in the former examples. Some fluctuations are observer during the nigh, sometimes identified as music. We believe this is due to the HVAC systems, which operate during the night creating an orchestration of flat line noises very characteristic of that city. These steady tones (or drones) are frequently identified as music by the application.

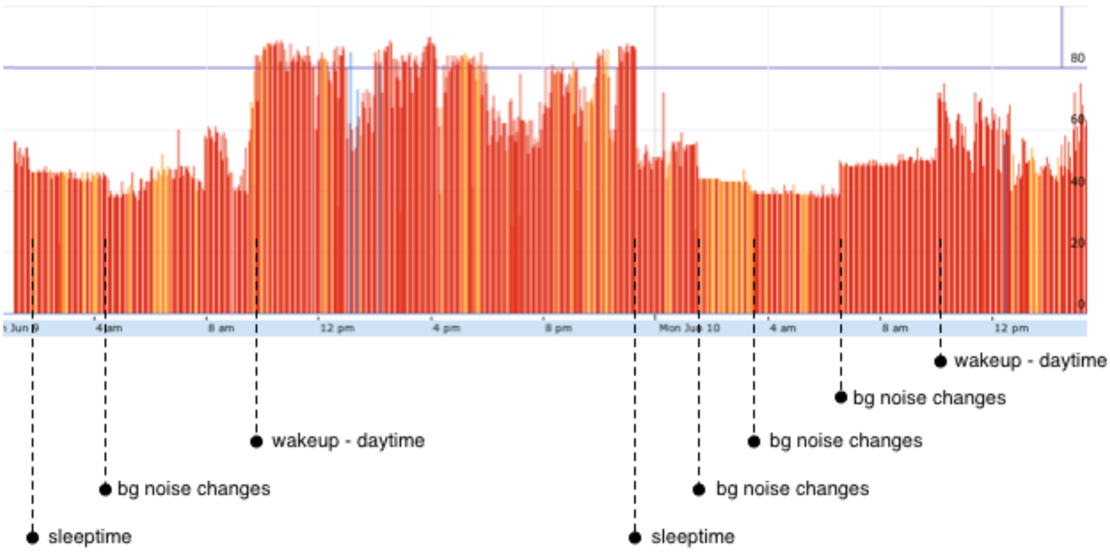


Figure 84 - Soundscape History Plot - Macau member

9.4 - Discussion

In this point, we present an analytical discussion about the data collected during the experimental period. The discussion follows the same structure of the result analysis, starting with the survey results and finishing with the soundscape history plot. It offers our perspective on the results and paves the way for the conclusion of this research project, the facts are examined through the lens of our initial aims and seldom go beyond the topics mentioned up to now. We tried to cover the most significant facets of the results and by no means data are closed to further interpretations; in fact novel approaches are possible and welcomed.

9.4.1 - On the sample

As mention before, we have used a convenience sampling method to gather a subset of individuals to test our hypothesis against and make inferences about a broader population. Next, the discussion will focus on our sample idiosyncrasies.

The first deduction is regarding the engagement of our sample with social media. The fact that 78% of the users were committed with social media helped avoiding bias from the exposure to a new technology. Since most of the users were familiar with social media this variable was not considered as a dependent variable.

We also discovered that in our sample younger groups of students tend to use social media applications as a way of entertainment more than the adult users, which are more inclined towards a utilitarian usage. The groups from UCP and HB elected *information* as the second use for SM, after *socialization*, while USJ group (the younger group) chose entertainment. These results show that our sample is, partially, in tandem with the findings of (Dunne, Lawlor, & Rowley, 2010) and (Cheung et al., 2011); however, extensive studies are lacking in academia regarding the age distribution of users intentionality in SM and OSN. In any case, one may say that the technology behind both uses – entertainment and information - is very similar and is based on powerful sindicalization algorithms, which perform a *gate keeping* task, tailored to the users' interaction and identification profiles.

When users are asked about the things they most like to share/receive in SM, contextual cues are among the least considered. Our study shows that only 8% of the sharing options contemplate information about the context of the user. The two options are *people accompanying the user* (4%) and *geo-localization* (4%). One may argue that text messages and photos - two highly regarded options – may also convey contextual information but the fact is that they are not strictly contextual. This fact raises the discussion about the relevance of soundscape analysis in OSN, since not many users are prone to share contextual cues. Nevertheless, one of the reasons that refrain people of sharing more contextual cues is privacy. A system based on soundscape analysis may overcome these constraints by providing contextual cues without sharing sensitive information.

The answers to the questionnaire also demonstrated that our sample accompanies the overall tendency observed in SM, of moving into the mobile world (<tate of the Media: The Social Media Report< 2012). The dominant device for accessing content in SM is still the laptop (48%), but mobile devices are becoming more prominent. When combining desktop and laptops computers in one category the score reaches 52%, against 48% of both mobile and tablet devices combined. This fact opens wider perspectives for integrating ubiquitous computing features in SM, which is also supported by the increase of computing power in every new generation of mobile devices.

From the direct analysis of the results, we noticed that people have the habit of carrying their mobile devices on a daily basis. Comparing the overall values against the type of device, we found that iPhone users are more likely to bring their devices always with them (67%), wile the majority of iPad and iPod use them most of the time but not always (88% - Figure 85). Although this is not a novelty *per se*, it shows that mobile devices are suitable for real-time and long-term monitoring, at least from the perspective of proximity and intimacy with the user. Moreover, 78% of the participants said they had access to an Internet connection during most of the experience period, which highlights the “always connected” status of contemporary users. This phenomenon was cross-cultural and cross-ages.

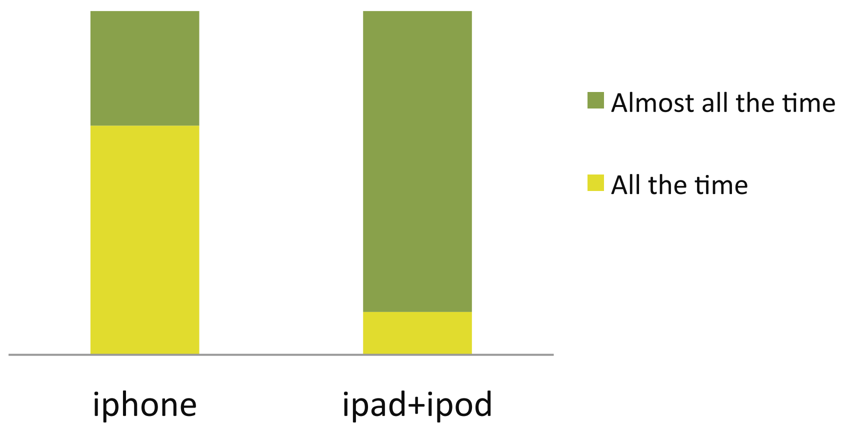


Figure 85 - Distribution of the “user-device closeness” per device type (normalized graph, see percentages above)

9.4.2 - On the user experience

When users were asked to evaluate how much they understood the application’s goals, 91% replied positively. This was somehow surprising since the application goals were not straightforward. We believe that the success in communicating the goals was due to the presentation video describing the overall project and the video tutorial focusing on the application operation. The two videos were highly informative and visually/sonically engaging and could be acceded from a webpage dedicated to the project. Using professional quality material to introduce this research tool helped giving credit to the project and engaging participants, mainly among non-researchers.

Regarding the reasons that would take people to shut down the application, battery issues are the first with 72%. In fact, all users of smartphones and other mobile computing devices struggle against battery consumption. This is a major impediment for long-term background running applications, being one of the reasons why manufactures like Apple present high restrictions to this kind of applications. The second reason elected by users concerns “privacy issues” (21%), showing that privacy is an important element to have in mind when designing monitoring applications, however, the result is not extreme since, we are leaded to conclude, no audio is shared. This conclusion is drawn from answers to question 2.5.

In general, the GUI was well received by users. 69% expressed they liked it and understood the match between sound and visual representations.

9.4.3 - On soundscapes and OSN

Regarding the relevance of soudscapes for OSN, although 70% of the participants provided a positive vote, we found that the results were highly dependent on the application's performance and user's knowledge about the concept of soundscape. Participants who evaluated negatively the application performance ("the application worked with major problems") also evaluated negatively the potential of soundscapes as social networking enhancer. In the other hand, participants who evaluated the application positively also have evaluated positively the potential of using soundscapes in the OSN domain (above 70%; Figure 86). This fact ratifies our methodology: providing a working prototype of a sound-based social network is a convenient way for people to understand the potential of soundscapes in the social domain.

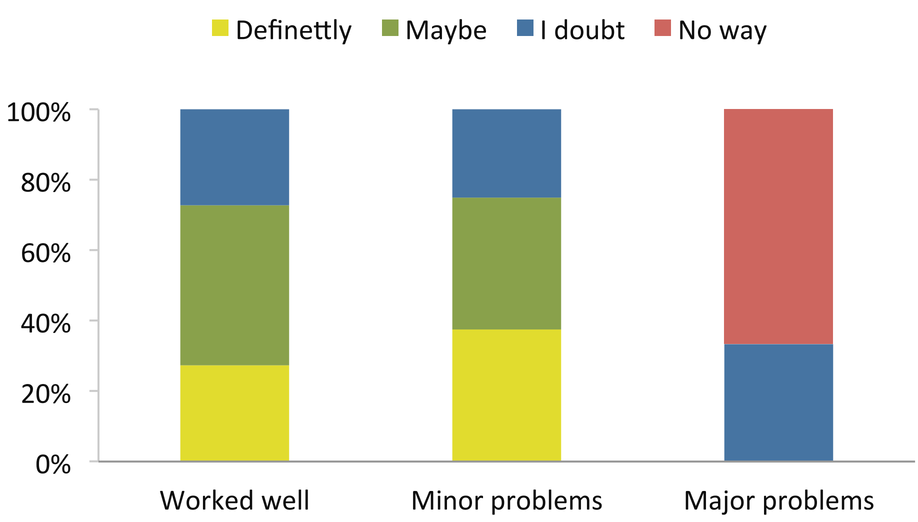


Figure 86 - Distribution of the "soundscape potential for SN" according to the application performance (normalized graph)

Moreover, when comparing the evaluation of soundscapes as social enhancers with the users knowledge of the term soundscape, results confirm that those who have a prior knowledge of the term show a higher tendency to recognize the potential of soundscapes as a social network enhancer (Figure

87). When people are more conscious about the sound phenomenon (soundscapes in particular) they acknowledge more effectively its underlying potential. This conclusion is the driving force of this research: creating awareness for the soundscape phenomenon in order to unveil its full potential and attain a balanced soundscape in contemporary world.

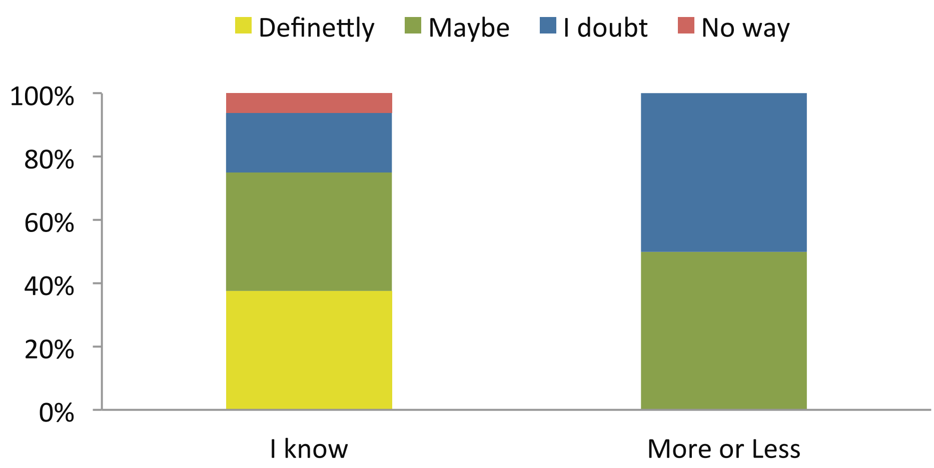


Figure 87 - Distribution of people's opinion on "soudscapes potential for SN" according to their knowledge about the soundscape concept

Regarding the hot topic of privacy, the results gathered by the questionnaire tend to be conservative. When enquired about a version of the application where the audio is shared, 56% say that they would never use it or would do it only in very controlled situations. Only 17% of the participants show no objections to sharing audio in SM. Comparing these results against those from question 4.8, where only 21% of the participants elected privacy issues as the main reason to shut down the application, we conclude that sharing audio is a big step towards privacy disruption. There is a sort of an "uncanny valley" where an excessively realistic representation may refrain people to interact with technology. That confirms our initial hypothesis and validates our design approach. From all the participants, those more willing to share sound are, in general, those who tend to acknowledge the potential of HB in OSN. This suggests that people showing a propensity for this type of applications are more open to share their actual soudscapes, revealing fewer privacy reservations.

9.4.4 - On the soundscape history plot

Referring to the first graph, it is quite clear the match between the room occupancy and the plot. One may argue that this is an *a posteriori* analysis and no explicit information about the room activity can be extract from the mere observation of the graph. In fact, this is our bottom line: extracting high-level information from a soundscape can provide sufficient information for many purposes, without the risk of excessive self-disclosure and privacy issues. Referring to the second plot (Macau) where no silence is ever sensed, it represents a good example of how this kind of applications can be used in environmental noise mitigation actions.

CHAPTER 10 - Conclusions

In this chapter we summarize the most significant outcomes of our research and contributions made to the different fields intersected during the course of the study.

Due to the experimental nature of the project we never aspired to provide a yes-or-no answer, nor an equation able to solve the underlying mechanisms of sound in social networking. Instead, we hope to contribute with a set of guidelines that we consider to be true, repeatable and relevant when approaching soundscapes in the context of social networking.

10.1 - Summary Conclusion

Our initial goal was to assess how social networking could benefit from the use of soundscapes as an integral part of its mechanics. Since the topic of social networking was broad enough to drive us into an immeasurable sea of possibilities, we focused on OSN, narrowing the options and refocusing our attention on the computational and technological aspects of this (relatively) new paradigm. Nevertheless, one should not forget the underlying goal of this research, which was not necessarily related with OSN but instead with creating favorable scenarios to unveil the potential of sound for communicating about the environment.

The first and most important conclusion to draw from our research is that the soundscape can provide useful information regarding our activity and our environment. This aspect has been extensively reported in the literature, namely in the studies of Schafer and Truax. In fact, one may even consider that this feature is what distinguishes and characterizes the soundscape. In this matter, the novelty of our approach relies on the assessment of this phenomenon using portable and personal mobile devices, which were not designed to accomplish such task, originally. We now know that high-end consumer technology in the area of mobile computing are suitable for such long term monitoring tasks, but the technical design is a bottle neck for accomplishing a desirable performance, taking for instance the battery constraints in consumer level mobile devices.

The second conclusion stems from the first. Being proved that soundscapes can be used to provide a valuable indicator about the environment and activity of people and that we are able to monitor this facet on a personal and continuous fashion, then it is possible to use this information successfully in the service of OSN. Broadcasting bits of information regarding users actual status is part of the normal operation of the social networking systems. Using sound on top of this process brings several advantages, namely: sound is computationally cheaper than image, both in capturing, processing and storage, thus allowing less battery consumption and more affordable devices; 2) sound can be sensed omnidirectionally, without demanding the intervention of user, for example facing a camera towards a target; 3) analyzing sound from a non-semantic perspective (after Chion), solves part of the privacy issues regarding self-discloser on OSN.

In particular, this study not only addresses real-time monitoring and broadcasting of sound-related information but also long-term patterns of analyzed sound. Studying the outcomes from these logs, we were able to confirm the existence of a close match between them and the users daily activity. Such profiling can be used in several scopes beyond OSN, as it provides a good opportunity for users to better understand their acoustic environment and draw conclusions regarding their physiological and psychological wellbeing.

The third conclusion we draw from our study is that people who are more aware of their sonic environment, namely by recognizing and/or knowing the concept of soundscape, are also more able to value the approach presented in this study. We are facing a recursive phenomenon were the more people use these type of systems – like Hurly-Burly – the more they learn and value their sonic environment and the more eager they will be to use this type of services. It is a pedagogical snowball towards the vision of Schafer, which can only be attained by an educated population that exert pressure on urban stakeholders to develop more friendly acoustic environments.

Regarding further contributions from this thesis to the field of sound and music computing, we would like to highlight the synthesis in this document of three important research areas that are usually treated separately, without conceptual or practical bonds in between. These areas are: **Sound/Soundscapes**, presented in Chapter 2 - Sound as a Medium, Sound as a Message; **Mobile Computing**, expatiated in Chapter 3 - Mobile Technology, Soundscapes and Context; and **Social Networks**, discussed in Chapter 4 - Being Social in the Information Era. Certainly, much more could have been said regarding each of theses areas in particular, as they represent core subjects in major scientific field as acoustics, computer science and sociology, respectively. Nonetheless, we tried to present these areas from a unique perspective, assuming sound and acoustic communication as the driving force of the endeavor. Those who now aim at following the (narrow) research area of sound-based social networks can find in this thesis a starting point entailing some of the theories and practices that we found central for a good definition of the area. During the writing of this document, we tried hard to avoid any biasing towards our comfort zone, preventing a higher preponderance of the subjects directly related with sound. Since the focus was already pointing to sound, we knew that, in order to create a balanced perspective on the topic, we had to diversify our approach.

By the time we reach the end of this endeavor we recognize that only a small window has been open in this multidisciplinary field. But results are promising: sound can definitely play a more interventional roll in daily interactions within networks of people. Pursuing this goal will lead to an increase of the awareness regarding the sonic phenomenon, in particular the soundscape, which should culminate in more sound-informed, critical and demanding citizens, consequently pushing the bar of the world soundscape towards a more balanced tune.

10.2 - Future Work

Departing from the results achieved insofar, the extension of this research may follow several paths depending on the aspects one wish to focus on. For example, considering the hardware manufacturing, fabricants and researchers may focus on more suitable devices for long-term monitoring, including 1) more approachable and flexible multitasking functions accounting for non-conflicting sound processing, 2) optimized battery consumption algorithms for continuous use and 3) acoustic measuring standards embedded in the devices by default. On the other hand, much work is yet to be done in the field of machine listening, especially regarding the classification of individual and collective environmental sounds. The existing algorithms fail to detect accurately the individual components of a soundscape and their corresponding weightings in the “mix”. Such upgrade on a system like Hurly-Burly would vastly contribute to improve its performance as a social networking enhancer, since the information shared within the network would gain new semantic levels encoded in the presence/absence of the known sound sources.

In the short run, our goals to continue this study include:

1) Embed the sound analysis feature in an existing OSN. This should be the next logical step, since this way users will depart from their pre-established network of friends and be able to use the layer of information obtained from sound in conjunction with other information.

2) Continue with the experimental process, increasing the number of participants in the experience by producing a large-scale experiment using established OSN. Though the conditions of the experiment would be less controlled, the escalation of the num-

ber of participants would provide relevant data to analyze according to the methodology implied insofar.

3) Pursue additional multimodal analysis of contextual cues in order to find correlations between sound and different stimuli gathered in the physical and social environment (temperature, light, geo-location, etc.);

4) Improve the quality of the sound classification algorithm. This endeavor should focus on the premise stated before, of producing more accurate and detailed classification of environmental sounds.

5) Calibrate the sound pressure level measurement for all the devices used in the study, in order to output more coherent results across all participants/devices.

The implementation of these features in future instances of this study will definitely increase the validity of the results achieved and open new understandings regarding the use of sound in the social networking domain.

References

Abowd, G. D., Mynatt, E. D., & Rodden, T. (2002). The human experience. *Pervasive Computing, IEEE*, *1*(1), 48–57. doi:10.1109/MPRV.2002.993144

Addlesee, M., Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A., & Hopper, A. (2001). Implementing a Sentient Computing System. *Computer*, *34*(8), 50–56. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=940013

Akkermans, V., Font, F., Funollet, J., de Jong, B., Roma, G., Togias, S., & Serra, X. (2011). Freesound 2: An improved platform for sharing audio clips. In *Late-braking demo abstract of the Int. Soc. for Music Information Retrieval Conf*. Retrieved from http://www.mtg.upf.es/system/files/publications/freesound_ismir.pdf

Albert, R., Jeong, H., & Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–382. Retrieved from http://www.nature.com/nature/journal/v406/n6794/abs/406378A0.html

Algazi, V. R., Duda, R. O., Thompson, D. M., & Avendano, C. (2001). The CIPIC HRTF Database. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New York: IEEE - Signal Processing Society.

Álvarez, L., Llerena, C., & Alexandre, E. (2011). Application of neural networks to speech/music/noise classification in digital hearing aids. In M. Lazard, A. Buikis, Y. S. Shmaliy, R. Revetria, N. Mastorakis, O. Martin, … G.-R. Gillich (Eds.), *Proceedings of the 11th WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision* (pp. 97–102). Florence: WSEAS Press. Retrieved from http://www.wseas.us/e-library/conferences/2011/Florence/GAVTASC/GAVTASC-15.pdf

Anderson, L. M., Mulligan, B. E., Goodman, L. S., & Regen, H. Z. (1983). Effects of Sounds on Preferences for Outdoor Settings. *Environment and Behavior*, *15*(5), 539–566. doi:10.1177/0013916583155001

Araujo, D., Davids, K., & Passos, P. (2007). Ecological Validity, Representative Design, and Correspondence Between Experimental Task Constraints and Behavioral Setting: Comment on Rogers, Kadar, and Costall (2005). *Journal of Ecological Psychology*, *19*(1), 69–78. doi:10.1080/10407410709336951

Aristotle. (1881). *Aristotelis quae feruntur De coloribus, De audibilibus, Physiognomonica*. B.G. Teubner.

Aristotle. (1987). *Poetics I*. (R. Janko, Ed. & Trans.). Hackett Publishing Company.

Attali, J. (1985). *Noise: The Political Economy of Music*. (B. Massumi, Trans.). University of Minnesota Press. Retrieved from http://books.google.co.uk/books?id=OHe7AAAAIAAJ

Atzmueller, M., Becker, M., Doerfel, S., Hotho, A., Kibanov, M., Macek, B., … Stumme, G. (2012). Ubicon : Observing Physical and Social Activities, 317–324. doi:10.1109/GreenCom.2012.75

Aucouturier, J.-J. J., Defreville, B., & Pachet, F. (2007). The bag-of-frames approach to audio pattern recognition: a sufficient model for urban soundscapes but not for polyphonic music. *The Journal of the Acoustical Society of America*, *122*(2), 881–891. doi:10.1121/1.2750160

Auroux, S., & Weil, Y. (1991). *Dictionnaire des auteurs et des thèmes de la philosophie*. Paris: Hachette Education.

Azizyan, M., Constandache, I., & Roy Choudhury, R. (2009). SurroundSense: mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th annual international conference on Mobile computing and networking* (pp. 261–272). ACM.

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012). Four degrees of separation. In *Proceedings of the 3rd Annual ACM Web Science Conference* (pp. 33–42). New York, NY, USA: ACM. doi:10.1145/2380718.2380723

Ballas, J., & Howard, J. (1987). Interpreting the language of environmental sounds. *Environment and behavior*, *19*(1), 91.

Ballas, J., & Mullins, T. (1991). Effects of context on the identification of everyday sounds. *Human performance*, *4*(3), 199–219.

Barabási, A. L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, *311*(3–4), 590–614. doi:10.1016/S0378-4371(02)00736-7

Barbosa, A. (2005). Public Sound Objects: a shared environment for networked music practice on the Web. *Organised Sound*, *10*(03), 233–242.

Barbosa, A. (2006). *Displaced Soundscapes: Computer Supported Cooperative Work for Music Applications*. *Doctoral Thesis*. Pompeu Fabra University.

Barra, M., Cillo, T., De Santis, A., Petrillo, U. F., Negro, A., Scarano, V., … Maglio, P. P. (2001). Personal WebMelody: Customized Sonification of Web Servers. In *10th International World Wide Web Conference*. Hong Kong.

Barrass, S., & Vickers, P. (2011). Sonification Design and Aesthetics. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 145–172). Berlin: COST Office and Logos Verlag.

Basdogan, C., Ho, C., Srinivasan, M. A., & Slater, M. E. L. (2000). An Experimental Study on the Role of Touch in Shared Virtual Environments. *ACM Trans. Comput.-Hum. Interact.*, *7*(4), 443–460. doi:10.1145/365058.365082

Basili, V. R., Selby, R. D., & Hutchens, D. H. (1986). Experimentation in software engineering. *IEEE Transactions on Sftware Engineering*, *12*(7), 733–743. doi:10.1109/TSE.1986.6312975

Behrendt, F. (2010). *Mobile Sound : Media Art in Hybrid Spaces*. *Doctoral Thesis*. University of Sussex.

Berglund, B, Lindvall, T., & Schwela, D. (Eds.). (1999). *Guidelines for community noise*. Geneva: World Health Organization.

Berglund, Birgitta, & Nilsson, M. E. (2006). On a Tool for Measuring Soundscape Quality in Urban Residential Areas. *Acta Acustica united with Acustica*, *92*(6), 938–944. Retrieved from http://www.ingentaconnect.com/content/dav/aaua/2006/00000092/00000006/art00012

Bergman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Mass.: Bradford Books, MIT Press.

Blesser, B., & Salter, L. R. (2007). *Spaces Speak, Are You Listening?: Experiencing Aural Architecture*. Cambridge: MIT Press.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science (New York, N.Y.)*, *323*(5916), 892–5. doi:10.1126/science.1165821

Boring, R. L. (2002). Human-Computer Interaction as Cognitive Science. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(21), 1767–1771. doi:10.1177/154193120204602103

Boyd, D. M. (2012). Networked Privacy. *Surveillance & Society*, *10*(3), 348–350.

Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x

Brewster, S. A. (2002). Non-Speech Auditory Output. In J. A. Jacko & A. Sears (Eds.), *Human-Computer Interaction Handbook* (pp. 220–239). Mahwah, NJ: Lawrence Erlbaum Associates.

Brown, G., & Cooke, M. (1994). Computational auditory scene analysis. *Computer speech and language*, *8*(4), 287–336. Retrieved from http://staffwww.dcs.shef.ac.uk/people/G.Brown/pdf/csl.pdf

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments* (2nd ed.). Berkeley: University of California Press.

Cage, J. (2011). *Silence: Lectures and Writings* (50th Anniv.). Middletown: Wesleyan University press.

Camurri, A., Volpe, G., Vinet, H., Bresin, R., Fabiani, M., Dubus, G., … Seppanen, J. (2010). User-Centric Context-Aware Mobile Applications for Embodied Music Listening. In P. Daras & O. M. Ibarra (Eds.), *User Centric Media SE - 3* (Vol. 40, pp. 21–30). Springer Berlin Heidelberg. doi:10.1007/978-3-642-12630-7_3

Caplow, T. (1968). *Two against one: Coalitions in triads* (4th Ed.). Prentice Hall.

Castells, M. (2007). *A Sociedade em Rede. A Era da Informação: Economia, Sociedade e Cultura* (3ª ed., Vol. I). Lisboa: Fundação Calouste Gulbenkian.

Chadabe, J. (1996). *Electric Sound: The Past and Promise of Electronic Music*. Upper Saddle River: Pearson Education.

Chen, W., Du, J., Liu, Y., Liu, H., & Mai, L. (2011). Energy-efficient Ambient Sound Sensing and Classification Using Smart Phones. London: Imperial College of London. Retrieved October 18, 2012, from http://www.doc.ic.ac.uk/~lm111/articles/Ambient_Technical Report.pdf

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. doi:10.1121/1.1907229

Cheung, C., Chiu, P.-Y., & Lee, M. (2011). Online social networks: Why do students use facebook? *Computers in Human Behavior*, *27*(4), 1337–1343. doi:10.1016/j.chb.2010.07.028

Chion, M. (1994). *Audio-Vision: Sound On Screen*. (C. Gorbman, Ed.). New York: Columbia University Press.

Chuengsatiansup, K. (1999). Sense, symbol, and soma: illness experience in the soundscape of everyday life. *Cult Med Psychiatry*, *23*(3), 273–301.

Clarke, R. (1994). The digital persona and its application to data surveillance. *The Information Society*, *10*(2), 77–92.

Collins, K. (2008). *Game Sound - An Introduction to the History, Theory, and Practice of Video game Music and Sound Design*. Cambridge: The MIT Press.

Comission of the European Communities. (2009). *Progress Report On The Single European Electronic Communications Market* (Vol. 1). Brussels. Retrieved from http://ec.europa.eu/information_society/policy/ecomm/doc/implementation_enforcement/annualreports/14threport/Vol1Part1_30072009.pdf

Cordeiro, J. (2011). Murky Shooting : em busca da primazia sonora nos videojogos. In A. R. Fernandes, B. Oliveira, V. Orvalho, & A. A. de Sousa (Eds.), *Acta da 4ª Conferência anual em Ciência e Artes dos Videojogos* (pp. 165 – 171). Porto: Sociedade Portuguesa de Ciências dos Videojogos.

Cordeiro, J., Baltazar, A., & Barbosa, A. (2012). Murky shooting: the use of auditory (non-speech) feedback on mobile audiogames. In *Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound. Corfu - Greece* (pp. 40–43). New York: ACM. doi:10.1145/2371456.2371462

Cordeiro, J., Barbosa, A., & Santos, A. (2013). A Soundscape Assessment Tool Based On A Massive Multichannel Loudspeaker Setup. In *Proceeding of Echopolis-Days of Sound 2013: Sounds, noise and music for rethinking sustainable city and econeighborhood*. Athina: SDMed.

Cosenza, V. (2013). World Map of Social Networks. Retrieved August 08, 2013, from http://vincos.it/world-map-of-social-networks/

Couvreur, C., Gaunard, P., Mubikangiey, C. G., & Fontaine, V. (1998). Automatic classification of environmental noise events by hidden markov models. *Applied Acoustics*, *54*(3), 187–206. doi:10.1016/S0003-682X(97)00105-9

Cross, S. (2007). Community and Media. In (G. Ritzer, Ed.)*Blackwell Encyclopedia of Sociology Online*. Wiley-Blackwell Publishing. doi:10.1111/b.9781405124331.2007.x

Daft, R., & Lengel, R. (1983). *Information richness. A new approach to managerial behavior and organization design*. *Organizations As Information Processing Systems - Office of Naval Research Report Series*. College Station, Texas. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA128980

De Vos, P., & Van Beek, A. (2011). Environmental Noise. In *Reference Module in Earth Systems and Environmental Sciences, from Encyclopedia of Environmental Health* (pp. 476–488). Burlington: Elsevier. doi:10.1016/B978-0-444-52272-6.00252-X

Dey, A. (2001). Understanding and using context. *Personal and ubiquitous computing*. Retrieved from http://www.springerlink.com/index/1d9grxkjvquhpwkw.pdf

Dick, F., Bussiere, J., & Saygin, A. (2002). The effects of linguistic mediation on the identification of environmental sounds. *Center for Research in Language Newsletter*, *14*(3), 3–9.

Donovan, J. (1975). Community noise criteria. *Journal of American Industrial Hygiene Association*, *36*(11), 849–854.

Dunne, Á., Lawlor, M.-A., & Rowley, J. (2010). Young people's use of online social networking sites – a uses and gratifications perspective. *Journal of Research in Interactive Marketing*, *4*, 46–58. doi:10.1108/17505931011033551

Durlach, N., & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators & Virtual Environments*, *9*(2), 214–217.

Edworthy, J. (1998). Does sound help us to work better with machines? A commentary on Rauterberg's paper about the importance of auditory alarms during the operation of a plant. *Interacting with Computers*, *10*, 401–409.

Eggen, B., Mensvoort, K. van, Menting, D., Vegt, E., Widdershoven, W., & Zimmermann, R. (2008). Soundscapes at Workspace Zero–Design Explorations into the Use of Sound in a Shared Environment. In *Proceedings of the 6th International Conference on Pervasive Computing*.

Entertainment Software Association. (2011). *Essential Facts About the Computer and Video Game Industry*. *Annual Report*. Washington D.C.

Erlmann, V. (Ed.). (2004). *Hearing Cultures: Essays on Sound, Listening and Modernity*. Oxford: Berg Publishers.

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the internet topology. *ACM SIGCOMM Computer Communication Review*, *29*(4), 251–262.

Ferscha, A., Vogl, S., & Beer, W. (2002). Ubiquitous Context Sensing in Wireless Environments. In P. Kacsuk, D. Kranzlmüller, Z. Németh, & J. Volkert (Eds.), *Distributed and Parallel Systems* (Vol. 706, pp. 98–106). Springer US. doi:10.1007/978-1-4615-1167-0_12

Fields, J. M., & Walker, J. G. (1982a). Comparing the relationships between noise level and annoyance in different surveys: A railway noise vs. aircraft and road traffic comparison. *Journal of Sound and Vibration*, *81*(1), 51–80. Retrieved from http://www.sciencedirect.com/science/article/pii/0022460X82901778

Fields, J. M., & Walker, J. G. (1982b). The response to railway noise in residential areas in Great Britain. *Journal of Sound and Vibration*, *85*(2), 177–255.

Fink, A., Mechtley, B., Wichern, G., Liu, J., & Thornburg, H. (2010). Re-Sonification Of Geographic Sound Activity Using Acoustic, Semantic, And Social Information. In *Proceedings of the 16th International Conference on Auditory Display, Washington D.C., 9-15 June* (pp. 305–312).

Fishwick, P. (Ed.). (2008). *Aesthetic computing*. Cambridge, MA, USA: MIT Press.

Freni, D., Vicente, C. R., & Mascetti, S. (2010). Preserving location and absence privacy in geo-social networks. *Proceedings of the 19th ACM international conference on Information and knowledge management*.

Friberg, J., & Friberg, J. (2004). Audio Games: New perspectives on game audio. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 148–154). ACM. doi:10.1.1.108.5572

Gärdenfors, D. (2003). Designing sound-based computer games. *Digital Creativity*, *14*(2), 111–114. doi:10.1076/digc.14.2.111.27863

Gauvain, J.-L., & Lamel, L. (2000). Large-vocabulary continuous speech recognition: advances and applications. *Proceedings of the IEEE*, *88*(8), 1181–1200. doi:10.1109/5.880079

Gaver, B., Dunne, T., & Pacenti, E. (1999). Cultural Probes. (J. Blickstein, Ed.)*Interactions*, *6*(1), 21–29. doi:10.1145/291224.291235

Gaver, W. (1993). What in the World Do We Hear? An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, *5*, 1–29.

Gaver, W., Smith, R., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKOLA simulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 85–90). New York: ACM. doi:10.1145/108844.108857

Gaye, L., Mazé, R., & Holmquist, L. (2003). Sonic City: the urban environment as a musical interface. *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*, 109–115.

Ghannam, J. (2011). *Social Media in the Arab World: Leading up to the Uprisings of 2011. A Report to the Center for International Media Assistance*. Center for International Media Assistance. Retrieved from http://www.edots.ps/internews/userfiles/CIMA-Arab_Social_Media-Report_1.pdf

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin Harcourt (HMH).

Gilfix, M., & Couch, A. (2000). Peep (The Network Auralizer): Monitoring Your Network with Sound. In *Proceedings of the 14th USENIX conference on System administration* (pp. 109–118). Berkeley, CA: USENIX Association.

Gjerdingen, R. O., & Perrott, D. (2008). Scanning the Dial: The Rapid Recognition of Music Genres. *Journal of New Music Research*, *37*(2), 93–100. doi:10.1080/09298210802479268

Greenfield, A. (2006). *Everyware: The Dawning Age of Ubiquitous Computing*. Berkeley, CA: New Riders Publishing.

Greif, I. (1988). *Computer-Supported Cooperative Work: A Book of Readings*. San Mateo, CA: Morgan Kaufmann.

Guastavino, C., Katz, B. F. G., Polack, J., & Levitin, D. J. (2005). Ecological Validity of Soundscape Reproduction. *Acta Acustica united with Acustica*, *91*(September 2004), 333–341.

Guillén, J., & Barrio, I. L. (2007). The soundscape experience. In *Comunicaciones del XIX Congreso Internacional de Acústica, (ICA2007), Madrid, 2-7 Sep. 2007* (pp. 43–66). Sociedad Española de Acústica (SEA).

Guo, G., & Li, S. Z. (2003). Content-based audio classification and retrieval by support vector machines. *IEEE transactions on neural networks*, *14*(1), 209–15. doi:10.1109/TNN.2002.806626

Guski, R. (2001). Environmental Stress and Health. In (Neil J. Smelser & P. B. Baltes, Eds.) *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Pergamon.

Hage, P., & Harary, F. (1983). *Structural Models in Anthropology*. Cambridge, UK: Cambridge University Press.

Hammond, K. (1993). Naturalistic decision making from a Brunswikian viewpoint: Its past, present, future. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 205–227). Norwood, New Jersey: Ablex Publishing Corporation.

Hammond, K. (1998). Ecological Validity: Then and Now. Brunswik Society. Retrieved July 24, 2013, from http://www.albany.edu/cpr/brunswik/notes/essay2.html

Han, X., & Tan, Q. (2010). Dynamical behavior of computer virus on Internet. *Journal of Applied Mathematics and Computation*, *217*(6), 2520–2526. doi:10.1016/j.amc.2010.07.064

Hansen, P., & Jurs, P. (1988). Chemical applications of graph theory. Part I. Fundamentals and topological indices. *Journal of Chemical Education*, *65*(7), 574.

Heidemann, J., Klier, M., & Probst, F. (2012). Online social networks: A survey of a global phenomenon. *Computer Networks*, *56*(18), 3866–3878. doi:10.1016/j.comnet.2012.08.009

Helmholtz, H. Von. (1913). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig, Germany: Vieweg & Sohn.

Henri, F., & Pudelko, B. (2003). Understanding and analysing activity and learning in virtual communities. *Journal of Computer Assisted Learning*, *19*(4), 474–487.

Hermann, T. (2011). Model-Based Sonification. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 399–427). Berlin, Germany: Logos Publishing House.

Hermann, T., Hunt, A., & Neuhoff, J. (2011a). *The Sonification Handbook*. (T. Hermann, A. Hunt, & J. G. Neuhoff, Eds.) (pp. 1–586). Berlin, Germany: Logos Publishing House. Retrieved from http://sonification.de/handbook

Hermann, T., Hunt, A., & Neuhoff, J. G. (2011b). Introduction. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 1–6). Berlin, Germany: Logos Publishing House.

Hermann, T., & Ritter, H. (1999). Listen to your data: Model-based sonification for data analysis. *Advances in intelligent computing and multimedia systems*, *8*, 189–194.

Hethcote, H. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*(4), 599–653. doi:10.1137/S0036144500371907

Hiltz, S. R., & Turoff, M. (1993). *The Network Nation: Human Communication Vira Computer*. Cambridge, MA: Mit Press.

Hoschka, P. (1998). CSCW research at GMD-FIT: from basic groupware to the social Web. *ACM SIGGROUP Bulletin*, *19*(2), 5–9. doi:10.1145/290575.290576

Hosokawa, S. (1984). The Walkman Effect. *Popular Music*, *4*(1984), 165–180. doi:10.2307/853362

Imielinski, T. (1996). *Mobile Computing*. (H. F. Korth, Ed.). Norwell, MA, USA: Kluwer Academic Publishers.

Jensen, K. (1999). *Timbre models of musical sounds. (Doctoral Dissertation)*. Department of Computer Science, University of Copenhagen, Copenhagen, Denmark.

Kang, J. (2011). Noise Management: Soundscape Approach. In J. O. Nriagu (Ed.), *Encyclopedia of Environmental Health* (pp. 174–184). Burlington: Elsevier. doi:10.1016/B978-0-444-52272-6.00260-9

Kanjo, E. (2009). NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping. *Mobile Networks and Applications*, *15*(4), 562–574. doi:10.1007/s11036-009-0217-y

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. doi:10.1016/j.bushor.2009.09.003

Kay, A. (1972). A personal computer for children of all ages. In *Proceedings of the ACM Annual Conference - Volume 1*. New York, NY, USA: ACM. doi:10.1145/800193.1971922

Kerber, S., & Fastl, H. (2008). Prediction of perceptibility of vehicle exterior noise in background noise. In *Fortschritte der Akustik DEGA 2008* (pp. 623–624). Berlin, Germany: DEGA.

Kilander, F., & Lönnqvist, P. (2001). A weakly intrusive ambient soundscape for intuitive state perception. *Communication presented at Continuity in future computing systems - I3 Spring Days Workshop, 23-24 April, Oporto*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.27.9719&rep=rep1&type=pdf

Kilander, F., & Lönnqvist, P. (2002). A whisper in the woods-an ambient soundscape for peripheral awareness of remote processes. In *Proceedings of the 8th International Conference on Auditory Display (ICAD 2002)*. Kyoto, Japan: Advanced Telecommunications Research Institute.

Kilgour, F. G. (1963). Vitruvius and the Early History of Wave Theory. *Technology and Culture*, *4*(3), 282–286.

Kim, H. G., Moreau, N., & Sikora, T. (2005). *MPEG-7 audio and beyond: Audio content indexing and retrieval*. Chichester, West Sussex, UK: Wiley.

Kim, W., Jeong, O.-R., & Lee, S.-W. (2010). On social Web sites. *Information Systems*, *35*(2), 215–236. doi:10.1016/j.is.2009.08.003

Kirsty, B., & Samuel, F. (2009). An Interface and Framework Design for Interactive Aesthetic Sonification. In *Proceedings of the 15th International Conference on Auditory Display*. Copenhagen, DK: Re:New – Digital Arts Forum.

Kjeldskov, J., & Graham, C. (2003). A Review of Mobile HCI Research Methods. In L. Chittaro (Ed.), *Proceedings of the 5th International Symposium, Mobile HCI 2003 - Human-Computer Interaction with Mobile Devices and Services* (Vol. 2795, pp. 317–335). Springer Berlin Heidelberg. doi:10.1007/978-3-540-45233-1_23

Klæboe, R. (2011). Noise and Health: Annoyance and Interference. In (Jerome O. Nriagu, Ed.)*Encyclopedia of Environmental Health*. Burlington, Canada: Elsevier. doi:10.1016/B978-0-444-52272-6.00242-7

Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech: a review. *International Journal of Speech Technology*, *15*(2), 99–117. doi:10.1007/s10772-011-9125-1

Koskinen, I., & Battarbee, K. (2006). Thinking about Sound in Mobile Multimedia. *Communication presented at the Pervasive Image Capture and Sharing Workshop in the 8th International Conference on Ubiquitous Computing (UbiComp 2006), Orange County, USA*. Retrieved from http://www2.uiah.fi/~ikoskine/recentpapers/mobile_multimedia/thinking_about_sound.pdf

Kraitchik, M. (1953). *Mathematical Recreations* (2n Ed.). Mineola, NY, USA: Dover Publications.

Krishnamurthy, B., & Wills, C. (2008). Characterizing privacy in online social networks. In *Proceedings of the first workshop on online social networks* (pp. 37–42). New York, NY, USA. doi:10.1145/1397735.1397744

Lane, D. R., & Lane, R. A. (1994). *Transistor Radios: A Collector's Encyclopedia and Price Guide*. Radnor, PA, USA: Wallace-Homestead Book Co.

Lasen, A. (2004). Affective technologies—emotions and mobile phones. *Receiver - Connecting to the future*, *11*, 1–8.

Laske, O. (1974). Musical Acoustics (Sonology): A Questionable Science Reconsidered. *Numus West*. Retrieved from http://scholar.google.co.uk/scholar?cluster=15308735449473169125&hl=en&as_sdt=2005&sciodt=0,5#0

Lee, J., & Tan, D. (2006). Using a low-cost electroencephalograph for task classification in HCI research. In *Proceedings of the 19th annual ACM symposium on User interface software and technology* (pp. 81–90). New York, NY, USA: ACM. doi:10.1145/1166253.1166268

Lercher, P., & Schulte-Fortkamp, B. (2003). The relevance of soundscape research to the assessment of noise annoyance at the community level. In G. D. J. Ronald (Ed.), *Proceedings of the 8th International Congress on Noise as a Public Health Problem*. International Commission on Biological Effects of Noise.

Lu, H., Pan, W., Lane, N. D., Choudhury, T., & Campbell, A. T. (2009). SoundSense: scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th international conference on Mobile systems, applications, and services* (pp. 165–178). New York, NY, USA: ACM. doi:10.1145/1555816.1555834

Lu, L., Jiang, H., & Zhang, H. H. (2001). A robust audio classification and segmentation method. In *Proceedings of the ninth ACM international conference on Multimedia - MULTIMEDIA '01* (pp. 203–211). New York, New York, USA: ACM Press. doi:10.1145/500170.500173

Lukowicz, P., Ward, J., Junker, H., Stäger, M., Tröster, G., Atrash, A., & Starner, T. (2004). Recognizing workshop activity using body worn microphones and accelerometers. In A. Ferscha & F. Mattern (Eds.), *Pervasive Computing - Second International Conference, PERVASIVE 2004* (Vol. 3001, pp. 18–32). Springer Berlin Heidelberg. doi:10.1007/978-3-540-24646-6_2

Macaulay, C., & Crerar, A. (1998). Observing the Workplace Soundscape: Ethnography and Auditory Interface Design. In S. Brewster (Ed.), *Proceedings of ICAD*. London, UK: British Computer Society.

Maisonneuve, N., & Stevens, M. (2009). Citizen noise pollution monitoring. In *Proceedings of the 10th …* (pp. 96–103). Retrieved from http://dl.acm.org/citation.cfm?id=1556198

Manach, J.-M. (2009). La vie privée, un problème de vieux cons? *InternetACTU.net*. Retrieved July 10, 2013, from http://www.internetactu.net/2009/03/12/la-vie-privee-un-probleme-de-vieux-cons/

Maranto, G., & Barton, M. (2010). Paradox and Promise: MySpace, Facebook, and the Sociopolitics of Social Networking in the Writing Classroom. *Computers and Composition*, *27*(1), 36–47. doi:10.1016/j.compcom.2009.11.003

Martins, L. G. (2008). *A Computational Framework for Sound Segregation in Music Signals*. *(Doctoral Dissertation)*. Faculdade de Engenharia da Universidade do Porto, Oporto, Portugal. Retrieved from http://repositorio-aberto.up.pt/handle/10216/58570

Mathison, S. (2012). The history of telenet and the commercialization of packet switching in the US. *Communications Magazine, IEEE*, *50*(5), 28–45. doi:10.1109/MCOM.2012.6194380

Matthews, T., Carter, S., Pai, C., Fong, J., & Mankoff, J. (2006). Scribe4Me: Evaluating a Mobile Sound Transcription Tool for the Deaf. In P. Dourish & A. Friday (Eds.), *UbiComp 2006: Ubiquitous Computing* (Vol. 4206, pp. 159–176). Springer Berlin Heidelberg. doi:10.1007/11853565_10

Maur, P., Costa, P. M., Pitt, J., Galvão, T., & e Cunha, J. (2013). Assessing contextual mood in public transport: a pilot study. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services* (pp. 498–503). New York, NY, USA: ACM. doi:10.1145/2493190.2494429

McAdams, S., & Bigand, E. (Eds.). (1993). *Thinking in Sound: The Cognitive Psychology of Human Audition*. Oxford, UK: Oxford University Press.

McLuhan, M. (1962). *The Gutenberg Galaxy: The Making of Typographic Man*. Toronto, Canada: University of Toronto Press.

McLuhan, M., & Lapham, L. H. (1994). *Understanding Media: The Extensions of Man* (p. 392). The MIT Press. Retrieved from http://www.amazon.com/Understanding-Media-The-Extensions-Man/dp/0262631598

Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, *84*(4), 857–870. doi:10.1037/0022-3514.84.4.857

Middlebrooks, J. J. C., & Green, D. M. D. (1991). Sound localization by human listeners. *Annual review of psychology*, *42*, 135–159. doi:10.1146/annurev.ps.42.020191.001031

Midolo, C. L. (2008). SoundFishing. In *Proceedings of the 5th International Mobile Music Workshop, Vienna, 13-15 May* (pp. 41–44). Vienna: University of Applied Arts, Vienna.

Milgram, S. (1967). The small world problem. *Psychology Today*, *1*(1), 61–67.

Miller, C., & Wortham, J. (2010, August 29). Technology Aside, Most People Still Decline to Be Located. *The New York Times*. New York. Retrieved from http://www.nytimes.com/2010/08/30/technology/30location.html?pagewanted=1&_r=2&

Miluzzo, E., Cornelius, C. T., Ramaswamy, A., Choudhury, T., Liu, Z., & Campbell, A. T. (2010). Darwin phones: the evolution of sensing and inference on mobile phones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services* (pp. 5–20). New York, NY, USA: ACM. doi:10.1145/1814433.1814437

Miluzzo, E., Lane, N., Eisenman, S., & Campbell, A. (2007). CenceMe – Injecting Sensing Presence into Social Networking Applications. In G. Kortuem, J. Finney, R. Lea, & V. Sundramoorthy (Eds.), *Smart Sensing and Context - Proceedings of the 2nd European Conference, EuroSSC 2007, Kendal, England, 23-25 october* (Vol. 4793, pp. 1–28). Springer Berlin Heidelberg. doi:10.1007/978-3-540-75696-5_1

Miluzzo, E., Papandrea, M., Lane, N. D., Sarroff, A. M., Giordano, S., & Campbell, A. T. (2011). Tapping into the vibe of the city using vibn, a continuous sensing application for smartphones. In *Proceedings of 1st international symposium on From digital footprints to social and community intelligence* (pp. 13–18). New York, NY, USA: ACM. doi:10.1145/2030066.2030071

Mitrovic, D., Zeppelzauer, M., & Breiteneder, C. (2010). Features for Content-Based Audio Retrieval. In *Advances in Computers: Improving the Web* (Vol. Vol. 78, pp. 71–150). Elsevier. doi:10.1016/S0065-2458(10)78003-7

Mohan, P., Padmanabhan, V. N., & Ramjee, R. (2008). Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM conference on embedded network sensor systems* (pp. 323–336). New York, NY, USA: ACM. doi:10.1145/1460412.1460444

Møller, H., Hammershøi, D., Johnson, C. B., & Sørensen, M. F. (1999). Evaluation of Artificial Heads in Listening Tests. *Journal of the Audio Engineering Society*, *47*(3), 83–100.

Monk, A., & Gilbert, G. N. (Eds.). (1995). *Perspectives on HCI: diverse approaches (Computers and People)*. Michigan, USA: Academic Press.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*(20). doi:10.1103/PhysRevLett.89.208701

Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review Letters*, *67*(2). doi:10.1103/PhysRevE.67.026126

Norman, D. (1999). *The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Solution*. Cambridge, MA, USA: The MIT Press.

Oksanen, S., Kleimola, J., & Valimaki, V. (Eds.). (2009). *D5.2 First Evaluation Report*. *SAME Sound and Music for Everyone Everyday Everywhere Every Way*.

Olguín, D., & Pentland, A. (2008). Social Sensors for Automatic Data Collection. In I. Benbasat & A. R. Montazemi (Eds.), *Proceedings of the 14th Americas Conference on Information Systems. Learning from the past: charting the future of the discipline.* Illinois, USA: Association for Information Systems.

Olson, H. F. (1967). *Music, Physics and Engineering*. Mineola, NY, USA: Dover Publications.

Pachet, F., & Roy, P. (2009). Analytical Features: A Knowledge-Based Approach to Audio Feature Generation. *EURASIP Journal on Audio, Speech, and Music Processing*, *2009*, 1–23. doi:10.1155/2009/153017

Paelke, V., Oppermann, L., & Reimann, C. (2008). Mobile Location-Based Gaming. In L. Meng, A. Zipf, & S. Winter (Eds.), *Map-based Mobile Services - Design, Interaction and Usability* (pp. 310–334). Springer Berlin Heidelberg. doi:10.1007/978-3-540-37110-6_15

Palfrey, J., Sacco, D., & Boyd, D. (Eds.). (2008). *Enhancing Child Safety & Online Technologies*. *Final Report of the Internet Safety Technical Task Force to the Multi-State Working Group on Social Networking of State Attorneys General of the United States*. Durham, North Carolina, USA: Carolina Academic Press. Retrieved from http://www.cap-press.com/pdf/1997.pdf

Pathan, K. T., & Reiff-Marganiec, S. (2009). Towards Activity Context using Software Sensors. *Electronic Proceedings in Theoretical Computer Science*, *2*, 27–35. doi:10.4204/EPTCS.2.3

Peeters, G. (2004). *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*. *CUIDADO I.S.T. Project Report*. IRCAM.

Peltonen, V., Tuomi, J., Klapuri, A., Huopaniemi, J., & Sorsa, T. (2002). Computational auditory scene recognition. In *Proceeding of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 2, pp. II–1941–II–1944). IEEE. doi:10.1109/ICASSP.2002.5745009

Peppers, D., & Rogers, M. (2009). The Social Benefits of Data Sharing. Retrieved July 10, 2013, from https://www.privacyassociation.org/publications/2009_01_the_social_benefits_of_data_sharing

Pilcher, E. J., Newman, P., & Manning, R. E. (2009). Understanding and managing experiential aspects of soundscapes at Muir woods national monument. *Environ Manage*, *43*(3), 425–435. doi:10.1007/s00267-008-9224-1

Poole, A., & Ball, L. (2006). Eye tracking in HCI and usability research. In (C. Ghaoui, Ed.)*Encyclopedia of human computer interaction*. Hershey, PA, USA: Idea Group Reference.

Postman, N. (1993). *Technopoly: The Surrender of Culture to Technology*. New York, NY, USA: Vintage.

Prensky, M. (2001). Digital natives, digital immigrants Part 1. *On the Horizon*, *9*(5), 1–6.

Rachuri, K. K., Musolesi, M., Mascolo, C., Rentfrow, P. J., Longworth, C., & Aucinas, A. (2010). EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM international conference on Ubiquitous computing* (pp. 281–290). New York, NY, USA: ACM. doi:10.1145/1864349.1864393

Rana, R. K., Chou, C. T., Kanhere, S. S., Bulusu, N., & Hu, W. (2010). Ear-phone: an end-to-end participatory urban noise mapping system. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks* (pp. 105–116). New York, NY, USA: ACM. doi:10.1145/1791212.1791226

Rao, B., & Minakakis, L. (2003). Evolution of mobile location-based services. *Communications of the ACM*, *46*(12), 61–65. Retrieved from http://dl.acm.org/citation.cfm?id=953490

Rayleigh, J. W. S. (1877). *The Theory of Sound* (Vol. 1). London, UK: Macmillan and co.

Rayleigh, J. W. S. (1878). *The Theory of Sound* (Vol. 2). London, UK: Macmillan and co.

Reydellet, J. (1996). Pierre Schaeffer, 1910-1995: The Founder of "Musique Concrète." *Computer Music Journal*, *20*(2), 10–11.

Rheingold, H. (1993). *The Virtual Community: Finding Commection in a Computerized World*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Rheingold, H. (2001). Mobile virtual communities. *Receiver - The mobile self*, *6*, 1–4.

Ruben, M. K. (2010, January). Radio Activity: The 100th Anniversary of Public Broadcasting. *Smithsonian Magazine*. Retrieved from http://www.smithsonianmag.com/history-archaeology/Radio-Activity-The-100th-Anniversary-of-Public-Broadcasting.html

Ruiz Vicente, C., Freni, D., Bettini, C., & Jensen, C. S. (2011). Location-Related Privacy in Geo-Social Networks. *IEEE Internet Computing*, *15*(3), 20–27. doi:10.1109/MIC.2011.29

Sá, C. (2010). *O que é o interface? Da entificação á identificação do interface enquanto complexo mediador*. *(Doctoral Dissertation)*. Universidade Nova de Lisboa, Lisbon, Portugal. Retrieved from http://run.unl.pt/handle/10362/5582

Sagan, C. (1990). Why We Need To Understand Science. *Skeptical Inquirer*, *14*(13).

Salber, D., Dey, A. K., & Abowd, G. D. (1999). The context toolkit: aiding the development of context-enabled applications. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (pp. 434–441). New York, NY, USA: ACM. doi:10.1145/302979.303126

Saldanha, A. (2009). Soundscapes. In (R. Kitchin & N. Thrift, Eds.)*International Encyclopedia of Human Geography*. Oxford, UK: Elsevier. doi:10.1016/B978-008044910-4.00979-2

Sallai, G. (2012). Defining infocommunications and related terms. *Acta Polytechnica Hungarica*, *9*(6), 5–15.

Saussure, F. de. (2011). *Course in General Linguistics*. (W. Baskin, Trans., P. Meisel & H. Saussy, Eds.). Columbia University Press.

Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). Distance matters: geo-social metrics for online social networks. In *Proceedings of the 3rd conference on Online social networks*. Berkeley, CA, USA: USENIX Association. Retrieved from http://dl.acm.org/citation.cfm?id=1863190.1863198

Schafer, R. M. (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books.

Schilit, Bill, Adams, N., & Want, R. (1994). Context-Aware Computing Application. In *First Workshop on Mobile Computing System and Applications*. Santa Cruz, Ca.

Schilit, BN, & Theimer, M. (1994). Disseminating active map information to mobile hosts. *Network, IEEE*. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=313011

Schmidt, A. (2012). Context-Aware Computing: Context-Awareness, Context-Aware User Interfaces, and Implicit Interaction. In (M. Soegaard & R. F. Dam, Eds.)*Encyclopedia of Human-Computer Interaction*. Aarhus, Denmark: The Interaction Design Foundation. Retrieved from http://www.interaction-design.org/encyclopedia/context-aware_computing.html

Schmidt, A., Beigl, M., & Gellersen, H. (1999). There is more to context than location. *Computers & Graphics*. Retrieved from http://www.sciencedirect.com/science/article/pii/S009784939900120X

Schuler, D. (1994). Social computing. *Communications of the ACM*. Retrieved from http://dl.acm.org/citation.cfm?id=175223

Schulte-Fortkamp, B. (2002). The Meaning of Annoyance in Relation to the Quality of Acoustic Environments. *Noise & Health Journal*, *4*(15), 13–18.

Schulte-Fortkamp, B., Brooks, B. M., & Bray, W. R. (2007). Soundscape: An Approach to Rely on Human Perception and Expertise in the Post-Modern Community Noise Era. *Acoustics Today*, *3*(1), 7–15. doi:10.1121/1.2961148

Schwartz, T. (1974). *The Responsive Chord*. Garden City, NY, USA: Anchor Press.

Serra, X., Leman, M., & Widmer, G. (Eds.). (2007). *A Road Map For Sound And Music Computing* (1st ed., pp. 1–167). S2S2 Consortium.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Hoboken, NJ, USA: Wiley.

Simmel, G. (1908). *Soziologie: Untersuchungen über die formen der vergesellschaftung*. Berlin, Germany: Duncker & Humblot.

Simpson, J., & Weiner, E. (1989). *The Oxford English Dictionary* (2nd Ed., Vol. 1). Oxford, UK: Oxford University Press.

Smith, D., Ma, L., & Ryan, N. (2005). Acoustic environment as an indicator of social and physical context. *Personal and Ubiquitous Computing*, *10*(4), 241–254. doi:10.1007/s00779-005-0045-4

Smith, E. (2010). *iPhone applications & privacy issues: An analysis of application transmission of iPhone unique device identifiers (UDIDs)*. PreSet Kill Limit. Retrieved from http://pskl.us/wp/wp-content/uploads/2010/09/iPhone-Applications-Privacy-Issues.pdf

Smith, W. P., & Kidder, D. L. (2010). You've been tagged! (Then again, maybe not): Employers and Facebook. *Business Horizons*, *53*(5), 491–499. doi:10.1016/j.bushor.2010.04.004

Spooner, S. (Ed.). (1922, December). Marconi's Wireless Telegraph Co., Ltd. *Flight*, *XIV*(50), 752.

Sporns, O. (2003). Graph theory methods for the analysis of neural connectivity patterns. *Neuroscience Databases*. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4615-1079-6_12

Stansfeld, S., & Clark, C. (2011). Mental Health Effects of Noise. In (J. O. Nriagu, Ed.)*Encyclopedia of Environmental Health*. Burlington, MA, USA: Elsevier. doi:10.1016/B978-0-444-52272-6.00248-8

Steffens, J. (2003). Realism and Ecological Validity of Sound Quality Experiments on Household Appliances Laboratory and Reality Methodical approach. In *Proceedings of AIA-DAGA 2013 International Conference on Acoustics, Merano, 18-21 March 2013, Merano, Italy* (pp. 1473–1476). Berlin: German Acoustical Society (DEGA).

Subrahmanyam, K., & Reich, S. (2008). Online and offline social networks: Use of social networking sites by emerging adults. *Journal of Applied Developmental Psychology*, *29*(6), 420–433. doi:10.1016/j.appdev.2008.07.003

Suhanek, M., & Djurek, I. (2012). Perception and evaluation of sudden loudness changes in familiar soundscapes. In T. Bucak & K. Jambrošiæ (Eds.), *Proceedings of the 5th Congress of the Alps Adria Acoustics Association*. Zagreb, Croatia: Acoustic Society of Croatia (HAD).

Suied, C., & Viaud-Delmon, I. (2009). Auditory-visual object recognition time suggests specific processing for animal sounds. *PLoS ONE*, *4*(4). doi:10.1371/journal.pone.0005256

Szeremeta, B., & Zannin, P. H. T. (2009). Analysis and evaluation of soundscapes in public parks through interviews and measurement of noise. *The Science of the Total Environment*, *407*(24), 6143–9. doi:10.1016/j.scitotenv.2009.08.039

Taruskin, R. (2009). *Music in the Late Twentieth Century: The Oxford History of Western Music*. New York, NY, USA: Oxford University Press.

The Nielsen Company. (2012). *State of the Media: The Social Media Report*. Retrieved from http://www.nielsen.com/us/en/reports.html

Theile, G. (2000). Multichannel natural recording based on psychoacoustic principles. In *Proceedings of the 108th AES Convention* (p. 5156). AES. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=9182

Tractinsky, N., & Eytam, E. (2012). Considering the Aesthetics of Ubiquitous Displays. In A. Krüger & T. Kuflik (Eds.), *Ubiquitous Display Environments* (pp. 89–104). Springer Berlin Heidelberg. doi:10.1007/978-3-642-27663-7_6

Trepte, S. (2006). Social identity theory. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. 255–271). London, USA: Routledge.

Truax, B. (2001). *Acoustic Communication* (2nd ed.). Santa Barbara, CA, USA: Greenwood Publishing.

Tuan, Y.-F. (1979). Space and place: humanistic perspective. In S. Gale & G. Olsson (Eds.), *Philosophy in Geography* (pp. 387–427). Springer Netherlands. doi:10.1007/978-94-009-9394-5_19

Tzanetakis, G., & Cook, P. (1999). Multifeature audio segmentation for browsing and annotation. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99 (Cat. No.99TH8452)* (pp. 103–106). IEEE. doi:10.1109/ASPAA.1999.810860

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, *10*(5), 293–302. doi:10.1109/TSA.2002.800560

VanDerveer, N. J. (1979). *Ecological Acoustics: Human Perception of Environmental Sounds. (Doctoral Dissertation)*. Cornell University, Ithaca, NY, USA.

Verne, J. (1889, February). In the Year 2889. (L. S. Metcalf, Ed.)*The Forum*, *VI*(6), 662–677.

Vickers, P. (2011). Sonification for Process Monitoring. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 455–491). Berlin, Germany: Logos Publishing House.

Vickers, P., & Hogg, B. (2006). Sonification abstraite/sonification concrete: An "Æsthetic perspective space" for classifying auditory displays in the ars musica domain. In T. Stockman, L. V. Nickerson, C. Frauenberger, A. D. N. Edwards, & D. Brock (Eds.), *Proceedings of the 12nd International Conference on Auditory Display*. London, UK: Department of Computer Science, Queen Mary, University of London.

Viollon, S., Lavandier, C., & Drake, C. (2002). Influence of visual setting on sound ratings in an urban environment. *Applied Acoustics*, *63*(5), 493–511. doi:10.1016/S0003-682X(01)00053-6

Vitale, R., & Bresin, R. (2008). Emotional cues in knocking sounds. In *Proceedings of the 10th International Conference on Music Perception and Cognition* (pp. 276–281). Aichi, Japan: The Japanese Society for Music Perception and Cognition.

Vitruvius. (1931). *Vitruvius: On Architecture, Volume I, Books 1-5 (Loeb Classical Library No. 251)*. Loeb Classical Library. Retrieved from http://www.amazon.com/Vitruvius-Architecture-Volume-Classical-Library/dp/0674992776

Walker, B., & Kramer, G. (1996). Human Factors and the Acoustic Ecology: Considerations for Multimedia Audio Design. In *Audio Engineering Society Convention 101*. AES. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=7436

Wang, G. (2009). Designing Smule's Ocarina: The iPhone's Magic Flute. In *Proceedings of the 9th International Conference on New Interfaces for Music Expression* (pp. 303–307). Pittsburgh, Pennsylvania, USA: Carnegie Mellon University.

Warren, W., & Verbrugge, R. (1984). Auditory perception of breaking and bouncing events: a case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 704–712.

Watts, D. J., & Strogatz, S. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. doi:10.1038/30918

Weinreich, F. (1997, February). Establishing a point of view toward virtual communities. *Computer-Mediated Communication Magazine*. Retrieved from http://www.december.com/cmc/mag/1997/feb/weinon.html

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *3*(3), 3–11. doi:10.1145/329124.329126

Weiser, M. (1993). Hot Topics - Ubiquitous computing. (R. D. Williams, Ed.)*Computer*, *26*(10), 71–72. doi:10.1109/2.237456

Weiser, M., & Brown, J. (1996a). Designing calm technology. *PowerGrid Journal*, *1*(1), 75–85.

Weiser, M., & Brown, J. (1996b). The coming age of calm technology. In P. J. Denning & R. M. Metcalfe (Eds.), *Beyond calculation - The Next Fifty Years of Computing* (pp. 75 – 85). Springer New York. doi:10.1007/978-1-4612-0685-9_6

Weiser, M., Gold, R., & Brown, J. S. (1999). The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal*. doi:10.1147/sj.384.0693

Weisstein, E. W. (n.d.). Königsberg Bridge Problem. Wolfram Research, Inc. Retrieved August 08, 2013, from http://mathworld.wolfram.com/KoenigsbergBridgeProblem.html

Westerkamp, H. (2002). Linking Soundscape Composition and Acoustic Ecology. *Organised Sound*, *7*(1), 51–56. doi:10.1017/S1355771802001085

Wolf, K. (2011). *The sociology of Georg Simmel*. Charleston, South Carolina, USA: Nabu Press.

Woszczyk, W., Bech, S., & Hansen, V. (1995). Interaction Between Audio-Visual Factors in a Home Theater System: Definition of Subjective Attributes. In *Audio Engineering Society Convention 99*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=7633

Yu, L., & Kang, J. (2009). Modeling subjective evaluation of soundscape quality in urban open spaces: An artificial neural network approach. *J Acoust Soc Am*, *126*(3), 1163–1174. doi:10.1121/1.3183377

Zhang, S., Jiang, H., & Carroll, J. M. (2010). Social Identity in Facebook Community Life. *International Journal of Virtual Communities and Social Networking*, *2*(4), 64–76. doi:10.4018/jvcsn.2010100105

APPENDIXES

Appendix A

Questionnaire used in Murky Shooting research

Avaliação Jogo Murky Shooting 07/2012

Teste 1: Iphone

Já tinha jogado o Murky? Pontuações

__________ 1º________ 2º________ 3º________ 4º________

Comentários (acerca da jogabilidade, do controlo, do som. Sugestões):

____

____

____

Teste 2: Versão computador – ipod via OSC - 2 versões

Pontuação: A) ____ B) ____

Notou alguma diferença? Sim____ Não____

Se sim, qual?

____

____

Qual versão prefere? A) ___ B) ___ NSNR _____

Porquê:

____

____

Teste 3: interfaces – versão computador – Rato vs ipod

Pontuação: Rato ____ Ipod ____

Qual interface prefere?

A) rato
B) ipod / iphone

Porquê?

____

____

____

Appendix B

Questionnaire used in MMucS research

Questionário de Investigação Científica sobre
Percepção de Paisagens Sonoras

Investigadores:
João Cordeiro e Adrian Santos

---

1. Caracterização do utilizador

Nome: ____

Idade: _____ Sexo: m [ ] f [ ] Ocupação: ________________

2. Caracterização das experiências

2.1 Do ponto de vista meramente acústico, qual das situações corresponde melhor ao ambiente sonoro do local?

[a] Com Auscultadores
[b] Sem Auscultadores
[c] Igual nas duas
[d] Ns/Nr

2.2 Em termos gerais, em qual das duas situações sentiu uma maior aproximação à experiência real sentida no local?

[a] Com Auscultadores
[b] Sem Auscultadores
[c] Igual nas duas
[d] Ns/Nr

2.3 Porquê?

[a] Por não ter que usar auscultadores (Se respondeu *Sem Auscultadores*)
[b] Melhor sensação de espaço
[c] Mais detalhe nos eventos sonoros
[d] Mais imersivo
[e] Menos artificial
[f] Melhor distinção na posição das fontes sonoras
[g] Maior qualidade de reprodução sonora

3. Relação do utilizador com o local da experiência

3.1 Reconhece o local retratado nas experiências? [a] Sim [b] Não

3.2 Se conhece, de que local se trata? ________________________________

3.3 Com que frequência utiliza o local?

[a] 4-5 dias por semana
[b] 1-3 dias por semana
[c] 3-4 vezes por mês
[d] Menos que 3-4 vezes por mês

3.4 Normalmente, em que altura do dia frequenta este local:

[a] É indiferente, tanto de tarde como de manhã
[b] Mais da parte da tarde
[c] Mais da parte da manhã
[d] Entre a manhã e a tarde (almoço)

3.5 Quais a sua principal motivação para se deslocar ao local (admite múltipla escolha):

[a] Comer/beber
[b] Relaxar
[c] Conviver
[d] Jogar
[e] Trabalhar
[f] Estudar/Ler
[g] Estar na companhia de pessoas (mesmo que desconhecidas)

3.6 Quando vai ao local, leva em mente:

[a] Demorar o menor tempo possível
[b] Demorar o tempo que for necessário
[c] Demorar o maior tempo possível, pois gosta de lá estar

3.7 Em termos **gerais**, com classifica a sua experiência no local:

Péssima [1] [2] [3] [4] [5] Excelente

3.8 Em termos **acústicos**, que qualidades associa ao local (admite múltipla escolha):

[a] Barulhento
[b] Desagradável
[c] Agradável
[d] Silencioso
[e] Muito reverberante ("ecos")
[f] Dificulta a comunicação
[g] Bom para conversar
[h] Bom para ouvir música
[i] Má acústica
[j] Boa acústica

4. Relação com a utilização de auscultadores

4.1 Em termos gerais, gosta de usar auscultadores?

Detesto [1] [2] [3] [4] [5] Adoro

4.2 Qual a utilização que faz de auscultadores?

[a] Quase nunca, só mesmo quando necessário
[b] Usa uma vez por semana
[c] Quase todos os dias

4.3 Seleccione os 3 principais pontos que caracterizam a sua utilização de auscultadores

[a] Por motivos profissionais
[b] Para ouvir música
[c] Para comunicar (intercomunicadores, auriculares...)
[d] Para ouvir/ver filmes, vídeos, séries, tv
[e] Para jogar videojogos
[f] Outro, (qual: ________________________)

4.4 Porque utiliza auscultadores como meio de escuta?

[a] Para me isolar do ambiente exterior
[b] Para ter melhor qualidade de reprodução
[c] Para não incomodar as pessoas à minha volta
[d] Para aquecer as orelhas

4.5 Pense nas situações em que normalmente utiliza os auscultadores. Em geral, se tivesse uma alternativa viável ao seu uso (p.ex.: colunas direccionais), continuava a utilizá-los?

[a] Sim
[b] Não
[c] Talvez
[d] Ns/Nr

****
NÃO RESPONDA DE IMEDIATO À PRÓXIMA PERGUNTA.
ESPERE PELA PRÓXIMA EXPERIÊNCIA
****

5. Soundscape Mixer:

5.1 Considera que este tipo de exercício pode melhorar a sua capacidade de compreender o ambiente sonoro que o/a rodeia?

Não contribui [1] [2] [3] [4] [5] Sim, contribui

FIM

Obrigado pelo seu tempo! ☺

Appendix C

Questionnaire used in Hurly Burly research

Sound Based Social Networks

This survey is part of a research project on the topic of Sound-Based Social Networks. It is intended for all who used the application Hurly-Burly (www.hurlyburly.eu). There is no "right or wrong" questions.It takes about 5 minutes to complete (less then 20 questions).All data is confidential and will be use for scientific purposes only.
* Required

Respondent Identification
Name *
Age*
Nationality *
Gender *

1. Mobile Devices

1.1 Which device did you use to run the application? *
Mark only one oval.
- ▯ iPhone
- ▯ iPad
- ▯ iPod
- ▯ I don't know
- ▯ I didn't run the application

1.2 During the experiment, was the device always with you? *
Mark only one oval.
- ▯ Yes, I carried the device always with me
- ▯ Most of the time, but not always
- ▯ No, the device and I were in separate places

1.3 During the experiment, how often was your device connected to the Internet? *
Mark only one oval.
- ▯ Always
- ▯ Most of the time, but not always
- ▯ 50% of the day
- ▯ Occasionally
- ▯ Never

1.4 During the experiment, how did you connect your device to the Internet? *
Mark only one oval.
- ▯ Through my 3G/4G plan
- ▯ Through Wi-Fi
- ▯ Both (3G/4G + Wi-Fi)

- I didn't connect the device
- Other ________

9. **1.5 When do you generally use sound in your device? (Top 3) ***
Mark only one oval per row.

- During calls
- Listening to music
- Recording Sound
- Watching videos
- Playing Games
- Musical/Sound App
- Other ________

10. **1.6 How often did you use de app during the test period?**
Mark only one oval.
Always

- Most of the time
- Half of the time
- A few times
- I tried it once or twice
- Never
- I'm not sure

2. Social Media

2.1 What is your favorite social media application? *
Mark only one oval.

- Facebook
- Tweeter
- Sina
- Weibo
- Other ________

2.2 Why do you use social media applications? *
Check all that apply.

- Socialize (chat, change messages...)
- Leisure (movies, music, cartoons...)
- Work (advertise your products, contact clients...)
- Information (news, politics, sports, technology, culture...)
- Other ________

2.3 What is your MAIN goal when using social networking applications? *
Mark only one oval.

- Meet and find new people/make new friends
- Keep in touch with people I met in the real world

- Both options
- Other

2.4 Regarding social networking, what information do you find more useful to share and receive? (CHOOSE THREE) *

Check all that apply.

- Textual messages (thought, moods, opinions, status...)
- Where I am (geo location, check in...)
- Things I like (artists, places, food...)
- People I am with, at the moment
- What music I'm listening
- Videos
- Photos
- Audio
- Other

2.5 How often do you use social media applications? *

Check all that apply.

- Everyday
- Twice a week
- Once a week
- Once a month
- Never
- I'm no sure
- Other

2.6 How do you prefer to use social media applications? *

Mark only one oval.

- Mobile devices (smartphones, iPods...)
- Tablets
- Laptops
- Desktops
- Other

3. Soundscapes

3.1 Do you know what a soundscape is? *

Mark only one oval.

- Yes
- More or less
- No

4. Hurly Burly Application

4.1 Did you understand the main goal of the application? *

Mark only one oval.
Not really 1 2 3 4 Yes, I understood completely

4.2 Was it easy to use? *
Mark only one oval.
Very easy 1 2 3 4 Very difficult

4.3 Did the application worked without problems? *
Mark only one oval.
- Yes, the application worked fine
- The application worked fine, with minor problems
- The application worked with major problems
- No, the application didn't work at all
- I don't know

4.4 From your experience using the application, what were the main problems you detected?

4.5 Do you think that knowing your friends soundscapes may enhance social interactions? * *Mark only one oval.*
- Yes, definitely
- I guess so
- I doubt
- No, it can't
- Other

4.6 Imagine that instead of sharing sound information, you were sharing the actual sound. What do you think of that? *
Mark only one oval.
- I would use it, I think it increases the sense of realism of the social experience
- I might use it with caution
- I would use it only in very controlled situations
- I would never use an application like that, it violates my privacy
- I'm not sure

4.7 Noise problems affect the lives of numerous people in urban and industrial centers. Do you think that applications like Hurly-Burly can be used as tools to monitor and raise awareness for this problem? *
Mark only one oval.
- Yes, definitely
- I guess so
- I doubt
- No, it can’t
- I don’t know

- Other

4.8 What would be the first reason to quit the app?
Mark only one oval.

- Low battery
- Privacy issues
- Keeps my device slow
- I don't like the red bar on top of the screen
- I'm not sure
- Other

4.9 What do you think about the graphic representation of the soud-scapes (waveforms)?
Mark only one oval.

- I did'n understand the match
- I understood the match but didn't like it
- I understood the match and I liked it
- I think it was too obvious
- I'm not sure
- Other

Suggestions to improve the application:

5. By looking into your soundscape history graph, can you identify a match between the graphic representation of the sound and your activity during the experiment?
Check your e-mail for a link to the soundscape history graph
Mark only one oval.

- Yes, definitely, a perfect match
- Generally yes, but some periods don't match
- Some periods yes, but not in general
- No matching at all
- I didn’t collect sufficient data to give a proper answer
- I'm not sure
- Other