

CHARACTERIZING APPLICATION ATTENTIVENESS TO ITS USERS: A METHOD AND POSSIBLE USE CASES

MILAN Z. BJELICA

*Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovica 6,
Novi Sad, 21000, Serbia
milan.bjelica@rt-rk.com
<http://www.ftn.uns.ac.rs>*

NIKOLA TESLIĆ

*Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovica 6,
Novi Sad, 21000, Serbia
nikola.teslic@rt-rk.com
<http://www.ftn.uns.ac.rs>*

Mitigating risks of rejection by end users should be the ultimate goal for any computer-based system or application. Latest researches have shown that with the growth of wearable and mobile computer-based products, the obtrusiveness of user applications has become significant. User time and his attention should be regarded as resources, as important as processing power or consumed energy. In this paper, we propose a novel method to characterize undesired interference between application usage and habitual activities of users, to what we refer as *attentive interference*. We argue that this interference is inversely proportional with application usability. We also present a set of heuristics that can be followed in order to increase application attentiveness, a case study for a commercial product and an overview of an ongoing implementation of presented characterizations to increase efficiency of context-aware systems.

Keywords: attentiveness; awareness; usability.

1. Introduction

Context-aware functions have become a very important in design and development of new software applications. Final judgment on the quality and usability of an application is therefore given by its acceptance by end-users. The most sophisticated software designed to be used by people appears worthless if they do not want to use it. Malhotra et al. (2004) state two basic reasons of failure for a new user-oriented system: users are not motivated to use functions provided by the system; system makes it difficult to perform the functions which users are motivated to use. In this regard, new applications should not invent functions that go beyond users' common interests. They should also try to keep all the functions as intuitive as possible. Noticeably, software is likely not to be used if additional training or effort is needed to start using it, even if we provide such training, or detailed tutorials and instructions.

The fact that users are able to use the software is not enough – they must want to use it (Carr (2003)). A statement that users will accept new software if it performs well is simply false. The study regarding this problem has been done by Markus et. al (1994).

It appears that consumer electronics market is saturated with the number of “useful” devices and applications that are always on users’ disposal, many of which are carried as regular outfit (e.g. cellular phone, portable music player, photo camera). Since all needed functions are often distributed to several devices and/or applications, users’ attention becomes fragmented. Each minute, cell phones interrupt the user with signaling tones regardless of his current activities. Music playing through the headphones diverts user’s attention from the busy traffic. Office desk has also become a source for torrents of different information requiring immediate attention. Several software programs on user’s PC fight for user’s gaze: chat clients, e-mail clients, different reminders and sticky notes, social websites. Different applications running in the background begin flashing taskbar buttons or playing sounds to confirm that previous processing has been completed. On the other hand, issuing commands to devices has been sped up to the extent that transforms users into multitasking “machines”. All sums up as additional stress, fatigue and tiredness. We can expect in a near future the users to start discarding existing devices and applications instead of adopting new ones – they are likely to start building a wall towards any innovative gadget, under the excuse of health hazards.

User attention and his time must be regarded as resources, if we mean to succeed with another user-oriented application. Vertegaal (2003) states that solution lies in the use of an Attentive User Interface (AUI). By that paradigm, decision on whether to divert user’s attention to the application depends on current state that user is in. That state can be determined in different ways. Sensors of presence and speech capture inputs on user’s whereabouts. They use the additional intelligent algorithms to help the application decide on any action. For example, device should postpone or cancel the notification if it concludes that will upset the user or interrupt him. This mostly addresses the issues with software applications that output their result in an asynchronous manner, such as communication software. Given the all growing trends in achieving pervasiveness and ubiquity in computing, communication and sharing information between devices and between devices and people are not possible to avoid. Therefore, possible intrusiveness and obtuseness must be carefully assessed for any new application and the quality in that regard must be designed in.

Other authors also discuss users, their motivation to use systems, their privacy and attention. Jaimes (2006) puts the user as a central point for the design of every new multimedia device, emphasizing the cultural diversity of users. Baker (2006) argues that any interactivity should not disrupt regular daily activities of users.

This paper introduces attentiveness as the key enabler when assessing application usability. Application should be attentive. This means that it should seek means to minimize interference with habitual activities of users, but still remain effective. For that purpose we will present several metrics and heuristics for defining how much the application is attentive. In the scope of this paper, attentive application is the one that seeks gaps in user behavior and his daily habits. It combines sociological constraints with

technology possibilities to assure good starting ground that mitigates risks of rejection by users. In the end, we will present a case study to illustrate how metrics can be used to characterize attentiveness for an exemplary user-oriented application.

2. A Concept of Attentive Interference

The goal of this section is to define a new term of attentive interference and to present means for characterizing it.

Let us first define a term of *user-system interference*. User-system interference is the amount of possible intersection in terms of time between the use of the software application, with the use of other devices and software applications or with the habitual activities of people as users. Attentive interference, as defined in this paper, is the value of user-system interference below certain predefined threshold that systems should thrive not to exceed. This threshold should be extracted from a sample of real usage activities that resulted in a quality user experience assessed by a user survey. Nevertheless, the system, device or application that has higher percentage value of user-system interference, we can regard as less attentive than the system with lower percentage of this interference. Interference is generally not desired, but sometimes it can be defined as a requirement. For example, it is supposed that user keeps his cell phone on in order to even consider using FM receiver that is embedded into phone. We define band of interference for this case, where upper bound provides attentiveness threshold, while lower bound provides entry usage threshold.

Interference can be estimated by using probability theory. We need sets of equations, which allow us to calculate:

- (1) Probability that users perform activities during a specific period of time that use of the system relates to (*related habitual activities*);
- (2) Probability that users would need to use functions of the system in given moment (*application/system need*);
- (3) *User-system interference*, to measure how well the system fits to users' habitual activities and therefore their daily living environment.

2.1. Related habitual activities

Listing all related habitual activities is an uneasy task. It is advised that this list needs to be as narrow as possible. We should consider only activities that can obviously and unconditionally impact the use of the device. For example, for the e-mail client on a mobile device, we can elicit three crucial related activities: (1) being out of office; (2) having a need to check mailing lists; (3) having a need to write longer e-mails. Generally, we can regard more than one user, if our application is a shared resource, or if habitual activities can be performed by more than one person.

For every related habitual activity we define, we select among several equations proposed in this paper to calculate probability that the activity will be performed. We define (a) *static* and (b) *dynamic* probability equations. Static equations give probability disregarding time variable, while dynamic equations give probability functions over time.

Additionally, we propose a refinement scheme based on which it is possible to adapt/refine the equations based on real-time inputs. This refinement makes the approach viable even for individual users, rather than for an average population.

2.2. Static related habitual activity probability equations

First we define total number of people participating in the activity with the Eq. 1.

$$r_s = \sum_{i=1}^q n_i \quad (1)$$

In this relation, q presents the number of distinct user groups (e.g. men, women, elderly, students etc.), and n_i defines the number of people in a group. It is not important the numbers to be exact – they just should correlate with one another in correct proportion.

Probability, that at least one member of a group k participates in a related habitual activity is given with the Eq. 2.

$$P_k^\alpha = 1 - \frac{\sum_{j=1}^{r_s} \left(\sum_{i=1}^{k-1} n_i + \sum_{i=k+1}^q n_i \right) \binom{r_s}{j}}{\sum_{j=1}^{r_s} \binom{r_s}{j}} \quad (2)$$

where n_i , $i \neq k$ is the number of people in a group other than k . Additionally, we can derive the probability, that only members of group k participate in the related habitual activity (Eq. 3).

$$P_k^\beta = \frac{\sum_{i=1}^{n_k} \binom{n_k}{i}}{\sum_{j=1}^{r_s} \binom{r_s}{j}} \quad (3)$$

where n_k is the number of people in the group k . Further we can define probabilities that the related habitual activity is performed by one user only (Eq. 4) or by several users (at least two) (Eq. 5).

$$P^1 = \frac{r_s}{\sum_{j=1}^{r_s} \binom{r_s}{j}} \quad (4)$$

$$P^n = 1 - P^1 \quad (5)$$

2.3. Dynamic related habitual activity probability equations

Related habitual activities can also be regarded with respect to time. This paper defines *P-models*, which are polynomial functions of time used to define the probability that the related habitual activity is going to be performed in some particular moment. Note that *P-model* is not a statistical distribution function, since it integrates to a sum of total time, not total probability. *P-model* is a polynomial of the form given by Eq. 6.

$$P(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \dots + \alpha_n t^n \quad (6)$$

α coefficients should be obtained by linear regression performed upon a set of collected sample data.

To include differences related to specific users (age, gender etc), we define *P-model variance* that further adjusts the probability. This probability is added to or subtracted from the *P-model* at the exact time (Eq. 7).

$$P_{\text{var}}(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n \quad (7)$$

where β coefficients are also obtained by linear regression, and x denotes the property for which variance is calculated (e.g. user age). Variance is used to fine-tune the model depending on various user properties of interest.

Let us assume there are no specific user properties that are respected. We can define *general related habitual activity time* T_{rha} (Eq. 8), to calculate time users dedicate to performing a related activity,

$$T_{rha} = \int_a^b P(t) dt \quad (8)$$

where a is start, and b is end hour, $a, b \in [0, 24]$ $a, b \in \mathfrak{R}$. Let us include differences between user groups in calculation. Then we can derive Eq. 9

$$T_{rha} = \int_a^b \sum_{i=1}^r \left(\left(P_i^\alpha - \sum_{j=1}^{i-1} P_j^\beta - \sum_{j=i+1}^q P_j^\beta \right) \cdot P(t) \right) dt \quad (9)$$

where r is the number of user groups for which the time is calculated.

Let us include additional properties (e.g. age or gender). Then we can define *P-model addition*, which defines addition in probability that depends on those properties. General formula would be as shown in Eq. 10.

$$P_{\text{add}}(t) = \frac{P(t)}{2} \left(\frac{1}{(a_{\text{up}} - a_{\text{low}})} \int_{a_{\text{low}}}^{a_{\text{up}}} P_{\text{var}_m}(x) dx + \frac{1}{(b_{\text{up}} - b_{\text{low}})} \int_{b_{\text{low}}}^{b_{\text{up}}} P_{\text{var}_f}(x) dx \right) \quad (10)$$

This formula calculates addition in probability of the related activity, for males and females of property (e.g. age) defined in ranges $[a_{\text{up}}, a_{\text{low}}]$ and $[b_{\text{up}}, b_{\text{low}}]$, respectively.

Related activity time including *P-model addition* related to additional properties, now produces (Eq. 11)

$$T_{rha} = \int_a^b \sum_{i=1}^r (P(t) + P_{add}^i(t)) dt \quad (11)$$

where $P_{add}^i(t)$ is a probability addition depending on a group i for the moment t .

To calculate related habitual activity time when static probability equations are included, we derive the Eq. 12.

$$T_{rha} = \int_a^b \sum_{i=1}^r \left(\left(P_i^\alpha - \sum_{j=1}^{i-1} P_j^\beta - \sum_{j=i+1}^q P_j^\beta \right) \cdot (P(t) + P_{add}^i(t)) \right) dt \quad (12)$$

2.4. Application (system) need

It is necessary to define a new function $p(t)$, that would give probability that our system would be used in a specific moment. To simplify further equations, we will use a discrete set of probability values:

$$p_i, i \in [0,23], i \in \mathbb{N}_0 \text{ and } t_i \in [0,1], t_i \in \mathbb{R}$$

where i corresponds to a specific timeslot. In general case, we define one-hour timeslots (e.g. for $i=10$, time interval that will be considered is 10:00-11:00 AM). This probability set should be obtained by experiments, or estimated by importing data related to use of other, similar systems. We use discrete set, since this probability should be made easy to input to calculus. This is important and makes possible one-time definition of related habitual activity probabilities and their reuse for new products. Additionally, we can define the time of system need as (Eq. 13):

$$T_{sn} = \sum_{i=a}^{b-1} p_i t_i \quad (13)$$

2.5. User-system interference

Finally, we can define *user-system interference time* by the Eq. 14.

$$T_{if} = \sum_{i=a}^{b-1} \left\{ p_i \int_i^{i+1} \sum_{j=1}^r \left(\left(P_j^\alpha - \sum_{k=1}^{j-1} P_k^\beta - \sum_{k=j+1}^q P_k^\beta \right) \cdot (P(t) + P_{add}^j(t)) \right) dt \right\} \quad (14)$$

where a is the start hour in a day, b is end hour $[0,24]$. r is the number of groups ($r > 1$) that participate in the related activity, while q being the total number of user groups.

Several simplifications can be introduced by the following equations. If there are no multiple groups involved in the related habitual activity, we can use (Eq. 15):

$$T_{if} = \sum_{i=a}^{b-1} \left\{ p_i \int_i^{i+1} P(t) dt \right\} \quad (15)$$

On a shared system, single user can interfere with the system in the amount defined by (Eq. 16):

$$T_{if} = \sum_{i=a}^{b-1} \left\{ p_i \int_i^{i+1} (P^1 \cdot P(t)) dt \right\} \quad (16)$$

More than one user interfere the system in the amount defined by (Eq. 17).

$$T_{if} = \sum_{i=a}^{b-1} \left\{ p_i \int_i^{i+1} (P^n \cdot P(t)) dt \right\} \quad (17)$$

If system need is triggered by the related habitual activity, or vice-versa, then we use (Eq. 18).

$$T_{if} = \sum_{i=a}^{b-1} \left\{ \int_i^{i+1} \max \left[\sum_{j=1}^r \left(\left(P_j^\alpha - \sum_{k=1}^{j-1} P_j^\beta - \sum_{k=j+1}^q P_j^\beta \right) \cdot p_i \right), p_i \right] dt \right\} \quad (18)$$

2.6. Attentive interference

We propose a new category for usability/quality measurement for a user-oriented system. Depending of the system (photo camera, multimedia player, set-top box etc) several related habitual activities would need to be defined, as in the example presented later in this paper. Quality would be measured for the average household, with gender and age equally distributed (e.g. 2 young (20), 2 middle-age (40), 2 elder (60)). The application would be considered attentive, if for n regarded related activities with user-system interference times $T_{if,n}$ the following is true (Eq. 19):

$$\frac{1}{n} \sum_{i=1}^n T_{if,n} < \frac{1}{\psi} \cdot T_{sn} \quad (19)$$

In this equation, ψ is the constant that should be obtained through an inverse problem for systems that we can classify as attentive. The best way to determine attentiveness is to conduct a user survey, aiming to assess user satisfaction. Emphasis in the survey should be the users' feeling of intrusiveness. This and other possible factors that can help solving the inverse problem are given as *attentiveness heuristics* in the Section 3. Although more research is needed to properly examine the variance of ψ , first results suggest the mean value of $\psi \approx 2$.

2.7. Models refinement scheme

Latest research results on the topic of attentive interference and system intrusiveness (Bjelica et. al (2010)) suggest refinements of P-models and therefore the final decision on system attentiveness tailored to suit the specific context. Consider having P-models and application/system need probabilities initially set by the statistical data, as described previously. Then, based on the acquisition of different context parameters, models are subject to continuous refinement. Models refinement scheme depends on two input vectors: Awareness Input Vector (V_{aw}) and Rules Vector (V_r) (Eq. 20).

$$V_{aw}(t) = \begin{bmatrix} v_{r_1}(t) \\ v_{r_2}(t) \\ \dots \\ v_{r_n}(t) \end{bmatrix} \quad V_r(t) = \begin{bmatrix} f_{r_1}(V_{aw}(t), V_{aw}(t - \Delta t_1)) \\ f_{r_2}(V_{aw}(t), V_{aw}(t - \Delta t_2)) \\ \dots \\ f_{r_m}(V_{aw}(t), V_{aw}(t - \Delta t_m)) \end{bmatrix} \quad (20)$$

Awareness Input Vector contains variables that define current context. Elements in the Rules Vector are all different functions of V_{aw} obtained in present moment, and, possibly, including data obtained earlier. These functions are defined separately depending on the type of awareness input variables. Based on each V_{aw} , new rules vector V_r elements are calculated. Each element within the rules vector is assigned a weight ω_i that defines how much the vector element influences the refinement of used P-model. Then, addition value of P-model that is going to be used for the refinement is obtained by scalar vector product (Eq. 21):

$$V_{\omega} = [\omega_1 \quad \omega_2 \quad \dots \quad \omega_m] \quad P_{ref}(t) = V_{\omega} \cdot V_r(t) \quad (21)$$

Finally, we calculate new P-model value at given t (Eq. 22).

$$P'(t) = P(t) + P_{ref}(t) \quad (22)$$

$P'(t)$ is used instead of $P(t)$ for calculating attentive interference in particular moment t . System use data is refined similarly, with the difference that refined value of p_i is stored and used in the future. This way, system use is profiled for the actual user ecosystem. The example for calculating feedback schemes is given in Section 4, while real world application is presented in Section 5.

3. Attentiveness heuristics

In wider sense attentiveness can also be achieved/measured by following a set of heuristic requirements that we here provide. Application is considered attentive, if those heuristics are met. There are several psychological and sociological constraints that are being considered. These constraints are subject to change depending on the functionality of the application, and are often elicited by an ethnography study. For example, O'Brien et al. (1999) conducted an ethnography study on the use of TV receiver in an average household, to capture behavior related to this device.

The list of proposed heuristics follows.

A. Nurturing related activities. Users might align their daily activities with the use of the application. If there is a need for functionality, its provision should be straightforward and quick. For example, user might need to take a photograph or answer a call, which are all functions that could be provided by the aforementioned application. Effort needed should be minimized, or the purpose of the application can be jeopardized (e.g. a short moment must be captured by photo camera, or an important call must be quickly made).

B. Ownership. If the application is dedicated to single user only, then users might struggle for ownership. Users can be conflicted in this manner.

C. Socializing. If an individual using the application becomes separated from a group he belongs to, the application is not considered attentive.

D. Privacy. Application should facilitate concept of privacy in a multi user environment. Data created or accessed by the application must be treated with respect to privacy – a mechanism to protect valuable memories or private content must be provided.

E. Control. User must have the feeling he is in control over the application, and not vice-versa. This implies that application must decrease automatics with functions such as content sharing, auto-answer to calls, etc.

F. Trust. Application should be trusted by the user. For applications on embedded devices, user interface should not resemble too much to PC user interface. It is essential that data processed within the application remains safe and that it cannot be accidentally lost.

G. Familiarity. It should not take too long for users to get familiar with application functions. User interface look and feel should therefore associate the user to the way of use they are already accustomed to. It is not attentive to introduce revolutionary changes in user interface between versions.

H. Interface. This category is further split into:

a. Device-proven graphical elements. This means that all GUI elements should be reused from similar concepts for the target device.

b. Comprehension time. Time for learning to use GUI should not take too long. The expected level of knowledge for the average user of a multimedia system is much lower than for PC users. Target group is much wider.

c. Data input. The way data are entered must be well tailored to suit the physical input device. E.g. if the keypad is the only means of input, device should facilitate using arrows for navigation, and T9 or multi tap way of text entering.

d. Visibility. Is the device regarded from the distance or from the vicinity? Data should read clearly and concept of wizards (Next-Back) should be facilitated.

e. Ease. The way the device is to be controlled should be aligned with its purpose. If the device targets relaxation and entertainment (e.g. multimedia player, TV receiver), controls should be made extremely simple. The device could also perform tasks related to data processing, or any other operation that can irreversibly make changes on the user content. When performing such operations it is suggested that user inputs more data, and less data should be assumed or suggested to user. This is to prevent errors, and, more importantly, keep the user appear in control.

4. Case study

A case study has been conducted in order to test metrics proposed in this paper and demonstrate their efficiency. Usability assessment with respect to attentiveness has been done for the case of a set-top box (TV set) upgraded with a multimedia software application. The application enables users to make and receive phone calls, send and receive short text messages and browse multimedia contents on their mobile phones via *Bluetooth*. Application provides a connection mechanism to Skype application on a nearby PC, providing Skype functions on the TV set. Name used for the application in the study is therefore *SkypeTV*. We selected this specific device and application, because we believe that TV set and concepts behind TV program watching have been set up for a long period of time (dating back to the first half of 20th century). Users are very affiliated to TV set. It became a central place in a household. *SkypeTV* multimedia application is intended to extend main functionality of TV set. It should be attentive to all common concepts related to television as a global phenomenon of today.

Let us first apply the concept of attentive interference to the case we selected. Inevitable related habitual activity for this case is **TV watching**. Since *SkypeTV* software is run on TV, it is important to calculate interference with regular usage of TV. Using statistical data from Australian Government Research and Statistics (2009) and UK Office for National Statistics (2003) we can use linear regression to derive dynamic *P-model* for TV watching. Best fit to data can be achieved by using a polynomial of 14th order (Eq. 23):

$$P_{TV}(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \dots + \alpha_{14} t^{14} \quad (23)$$

Coefficients for this polynomial are given in Table 1. Model gives probability in percentage that TV is watched in a household in a given time moment during a day [0-24h]. Model plot is given in Figure 1.

P-model variance, that includes gender and age, is also obtained by the regression and results in a 6th order polynomial for males (Eq. 24) and 7th order polynomial for females (Eq. 25). Given polynomial orders represent best fit to used statistical data.

In *P-model* variance x denotes the age for which the variance is sought.

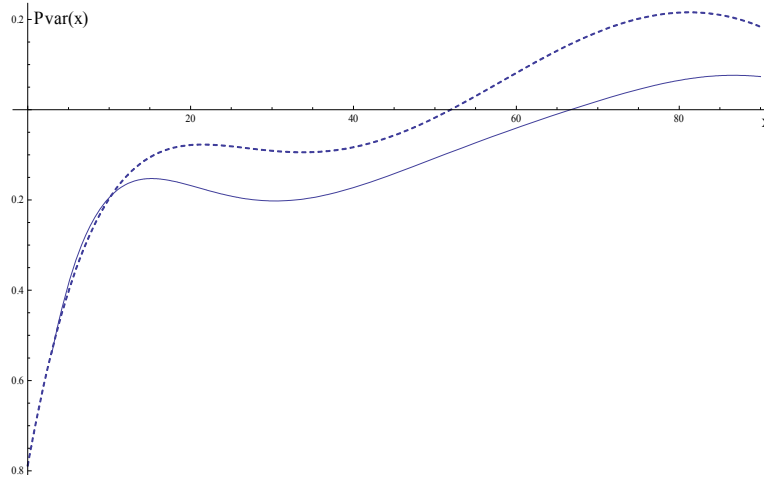


Fig. 2: P-model variance for males (dashed) and females (solid)

$$P_{m,TV}(x) = \beta_{m,0} + \beta_{m,1}x + \beta_{m,2}x^2 + \dots + \beta_{m,6}x^6 \quad (24)$$

$$P_{f,TV}(x) = \beta_{f,0} + \beta_{f,1}x + \beta_{f,2}x^2 + \dots + \beta_{f,7}x^7 \quad (25)$$

Coefficients for these two polynomials are given in Table 1.

P-model variance gives the probability in percentage that needs to be added to or subtracted from the P-model probability for any given age [0-90 years]. P-model variance plots for both males and females are given in Figure 2.

We picked a hypothetic household to define static TV watching probability.

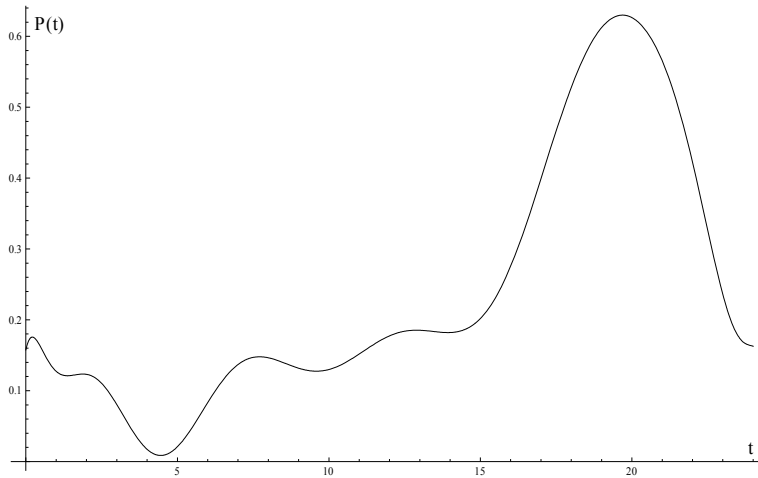


Fig. 1: P-model plot for TV watching

Table 1. P-Models coefficients obtained by linear regression

P-model for TV watching			
Coeff.	Value	Coeff.	Value
α_0	452535470 2873103700	α_8	-16997689 36465652900
α_1	108593841 551482700	α_9	41800111 1278186895300
α_2	-358845822 537223700	α_{10}	-15857461 9881752890500
α_3	675571380 889537100	α_{11}	384694 7128252379500
α_4	-490198798 1114908500	α_{12}	-14755 12445665612200
α_5	91033027 610764900	α_{13}	3285 214308148717900
α_6	-138132775 4294183200	α_{14}	-269 3040371986555200
α_7	35051583 7527201700		
P-model variance for males and females by their age			
Coeff.	Value	Coeff.	Value
$\beta_{m,0}$	-784170653 994662400	$\beta_{f,0}$	-113081463 128331100
$\beta_{m,1}$	215200202 2141194900	$\beta_{f,1}$	120741694 840324500
$\beta_{m,2}$	-21358235 4037422700	$\beta_{f,2}$	-67090167 6277493900
$\beta_{m,3}$	11060057 85091570300	$\beta_{f,3}$	18972706 49089967500
$\beta_{m,4}$	-752600 476893512900	$\beta_{f,4}$	-2030778 267011954900
$\beta_{m,5}$	20269 2156245618300	$\beta_{f,5}$	247594 2940048349100
$\beta_{m,6}$	-2758 122715942192500	$\beta_{f,6}$	-133131 69563828600000
		$\beta_{f,7}$	906 759825057232900

Household members, their names, gender and age are given in Table 2. Our target group in this household consists of four members (Roger, Claire, John and Jane), as it is expected that they will be the most frequent users of the application. This influences static related activity probability equation.

Definition of related activity is: TV watching by at least one person from our target group. We assume remote controller, as the only means of input to a TV set, would be shared among all the people watching TV at a specific moment. By using Eq. 2 we get:

$$P_1^\alpha = 1 - \frac{\sum_{i=1}^8 \binom{4}{i}}{\sum_{j=1}^8 \binom{8}{j}} \approx 0.94$$

This means that there is 94% chance that one member of our target group is among the audience in front of the TV in any given moment. Further on, we assume that related habitual activity time is the time users are awake (8 AM – 24 PM).

Table 3. Hourly probability estimates for Skype TV use

p_0	0.23	p_6	0.25	p_{12}	0.36	p_{18}	0.37
p_1	0.25	p_7	0.29	p_{13}	0.37	p_{19}	0.37
p_2	0.26	p_8	0.33	p_{14}	0.38	p_{20}	0.37
p_3	0.26	p_9	0.37	p_{15}	0.39	p_{21}	0.32
p_4	0.25	p_{10}	0.36	p_{16}	0.38	p_{22}	0.27
p_5	0.24	p_{11}	0.35	p_{17}	0.37	p_{23}	0.25

Now, according to Equation 12 we get:

$$T_{rha} = P_1^\alpha \int_8^{24} P_{TV}(t) \left(1 + \frac{1}{4} \begin{pmatrix} P_{m,TV}(14)+ \\ P_{m,TV}(34)+ \\ P_{f,TV}(18)+ \\ P_{f,TV}(32) \end{pmatrix} \right) dt \approx 4.0112h$$

This means that, on average, around 4 hours each day TV is being watched by at least one member of our target group. We also need a discrete set of probabilities, that multimedia application will be needed in a given hour. We used statistical data related to use of PC Skype application, to estimate system need for *SkypeTV*. The duration of use in that hour would be approximately 10 min ($t=0.17$), according to Kuan-Ta et al. (2006). By using hourly online users chart from SkypeStats.com (2008) and estimation of total number of real Skype users from Ckipe.com (2009) we can estimate a set of discrete probabilities as shown in Table 3.

Next we can calculate system need, as the total time during a day that *SkypeTV* application is going to be needed:

$$T_{sn} = 0.17 \sum_{i=8}^{23} p_i = 0.9537h$$

User-system interference time is, according to Equation 14:

$$\begin{aligned} T_{if} &= 0.037_{[8,9]} + 0.038_{[9,10]} + 0.04_{[10,11]} + 0.047_{[11,12]} + \\ &+ 0.053_{[12,13]} + 0.055_{[13,14]} + 0.058_{[14,15]} + 0.073_{[15,16]} + \\ &+ 0.676_{[17,21]} + 0.129_{[21,22]} + 0.071_{[22,23]} + 0.037_{[23,24]} = \\ &= 1.314h > 0.477h = \frac{1}{2} T_{sn} \end{aligned}$$

Roughly an hour and a half during a day will be spent in “interference”, meaning that *SkypeTV* is not going to be available because TV program is being watched, and vice-

Table 2. Hypothetic household members

Jimmy Doe	Male	3 years	John Doe	Male	34 years
Sarah Doe	Female	4 years	Jane Doe	Female	32 years
Roger Doe	Male	14 years	Arthur Doe	Male	70 years
Claire Doe	Female	18 years	Amanda Doe	Female	65 years

versa. According to the hypothesis from Equation 19, this system is not attentive in terms of interference to users' habitual activity.

Based on the module refinement scheme presented in Section 2.7, we will now illustrate how the refinement can be used in our current example, to take into account user feedback loop. Let us assume that an optical camera and a microphone are positioned above the TV screen, recording environment in front of the TV. A face detection algorithm, provided with the camera, would detect if there is a person watching TV (facing the screen) or not. The more faces are detected, the more probable TV watching activity is. Also, the longer a face is detected, the more probable TV watching activity is. TV watching activity would be less probable if people are speaking (what is detected by the microphone). Based on this definition, we can define model refinement scheme:

$$V_r(t) = \begin{bmatrix} v_{fcount}(t) \\ V_{r1}(t) \cdot v_{fcount}(t-10) \\ V_{r2}(t) \cdot v_{fcount}(t-30) \\ -3 \cdot v_{speaking}(t) \cdot v_{speaking}(t-10) \end{bmatrix}$$

$$V_{aw}(t) = \begin{bmatrix} v_{fcount}(t) \\ v_{speaking}(t) \end{bmatrix} \quad V_w = [0.002 \quad 0.004 \quad 0.02 \quad 0.01]$$

Awareness Input Vector consists of two variables, v_{fcount} , reporting the number of currently detected faces, as well as $v_{speaking}$, that can have the following values:

$$v_{speaking}(t) = \begin{cases} 1, & \text{speaking} \\ 0, & \text{not speaking} \end{cases}$$

Rules Vector is defined to provide positive increments for TV watching probability, when there are faces detected. Rules Vector enhances the certainty that TV is being watched, if faces are also detected 10 seconds ago, as well as 30 seconds ago. Probability is decreased (therefore the minus sign) if speaking is detected, but only if it is detected for at least last 10 seconds. Weights are given to every parameter, so the output can be fine tuned later without the need to change the rules kernel.

Now, if the family has TV ser on before they go to work / school to hear the news and weather forecast, each morning between 8 and 9 AM, but are drinking coffee at the same time and discussing the show, we could easily get:

$$\begin{aligned}
V_{aw}(8h10m15s) &= \begin{bmatrix} 3 \\ 1 \end{bmatrix} \\
V_{aw}(8h10m30s) &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\
V_{aw}(8h10m38s) &= \begin{bmatrix} 3 \\ 1 \end{bmatrix}
\end{aligned}
\quad
V_r(8h10m38s) = \begin{bmatrix} 3 \\ 3 \cdot 2 \\ 6 \cdot 3 \\ -3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 18 \\ -3 \end{bmatrix}$$

We calculate refined $P_{TV}(t)$:

$$P_{TV}^{-1}(8h10m38s) = P_{TV}(8h10m38s) + V_r(8h10m38s) \cdot V_w = P_{TV}(8.1772) + 3 \cdot 0.002 + 6 \cdot 0.004 + 18 \cdot 0.02 - 3 \cdot 0.01 = 0.1445 + 0.36 = 0.5045$$

that is much more accurate than the original value based on general statistics.

Next we check proposed attentiveness heuristics. Here we will discuss some aspects related to the nature of TV and its users, and try to answer to heuristic requirements. In latter experiment, we will present the results of questionnaires filled in by users that had hands-on experience with the multimedia application.

A. Nurturing related activities. It is noted that TV viewers tend to organize their daily activities in alignment with TV program scheme. For example, TV can be used to watch a football match, and therefore other planned activities can be postponed until the match is finished. Very sensitive period is around prime-time (20-22 PM). P-model for TV watching and results obtained for attentive interference indicate that this heuristic requirement may not be met.

B. Ownership. There is only one remote control device to operate the TV set. Moreover, using *SkypeTV* requires the possession of the remote control. TV is also watched potentially by more than one person. Using *SkypeTV* may appear intrusive to other viewers, especially if user interface covers a lot of the screen surface. This may be a problem if there is only one application on one TV set. However, when browsing multimedia contents, it can be desired that more people are watching e.g. photographs. This heuristic requirement is likely not to be met with communication part of *SkypeTV*. Meeting the requirement with content-browsing part depends on the content being browsed. If the content is private then ownership concerns are justified. If the intention is to present content to multiple viewers, then ownership might not be an issue.

C. Socializing. People are gathered when watching the same TV program, so in this manner TV set socializes people. On the other hand, if someone is watching TV program while there are others in the same room (e.g. house guests) and a conversation is on, it is often requested that TV is off. Again, if a communication over *SkypeTV* is needed, it would require TV set to be turned on. This heuristic requirement has a slight chance not to be met for the communication part of the application. However, browsing multimedia contents can be a socializing activity, if the intention is that content is presented to house guests.

D. Privacy. For TV set, privacy can be questioned since there is more then one person watching, and screen is visible from a broad viewing angle. With *SkypeTV* application,

this is definitely the issue, both with communication, and with private multimedia contents.

E. Control. Whether user feels being in control depends on the multimedia application only. Only users of the application can give an opinion in this regard.

F. Trust. Trust should not be the problem, since the application does not provide any content storage or processing, nor is it connected to internet that all could be the reasons to question safety of data being browsed.

G. Familiarity. TV users are used to TV-like user interface (extensive use of vertical menus, simple and infrequent use of the remote controller). *SkypeTV* must be tailored to give look and feel of the traditional TV user interface. Users give opinion in this regard.

H. Interface. Several authors discussed interactivity with respect to user interface in their studies. Obrist et al. (2008) base their study on the interactive television service. Lekakos et al. (2001) analyze potential use of interactive commercials. Related to TV receiver, we can redefine the following:

- a. Application should use TV-like graphical elements (e.g. vertical menus);
- b. Expected level of knowledge for an average TV user is far below that of the average PC user.
- c. The way of entering data must suit remote controller as the input device.
- d. Device is regarded from the distance, and possibly screen resolution is low. Application should use larger screen graphical elements and bigger text.
- e. Watching TV program should remain an entertaining, relaxing activity suitable for the living room. User posture is also relaxed, laid back, so this can represent an additional constraint.

To support the theory with evidence, an experiment was conducted. Structure of users participating was similar to the structure from the Table 2. There were 16 people involved: 2 very young ones (5-8 years), 6 youths (12-18), 6 middle-aged ones (24-50) and 2 elderly ones (~70). Gender balance was also met (8 males, 8 females).

Experiment consisted of two separate phases: Phase I lasted for two days: one arbitrary weekday and a Saturday. Each participant was asked to note every instance when he wanted to make a call, or when the call was incoming: (a) whether the TV was on, (b) what was the number of people watching, and (c) was he close to the TV (was he watching). Depending on each answer, at the end of the experiment, each user noted whether he would or would not prefer to use *SkypeTV* over the conventional Skype or mobile phone usage. In this manner experimental results consisted of a set of true/false statements, for every participant.

Table 4. Results of the Phase I

	A Skype-like Messages, calls	B Multimedia content browsing
Would use TV	26%	63%
Would not use TV	74%	37%

Phase II of the experiment involved real equipment with *SkypeTV* software set up. TV platform used was *VGCB* chip based on *MIPS* processor from *Trident Microsystems*. Chip supports interfacing to LCD panels through HDMI, LVDS and RGB as well as graphic processing, so it could support showing graphics on screen. Bluetooth communication to Skype or mobile phone was provided by *Bluegiga WT11* chip, interfaced to *VGCB* via UART (Universal Asynchronous Receiver/Transmitter). To capture voice from the speaker a microphone array was used, based on *SEA²M* technology for audio processing (Papp et al. (2007)). More details on *SkypeTV* are given by Lakobrija et al. (2007) and Bjelica et al. (2008). The look of the user interface of *SkypeTV* is shown in Figure 3.

Each participant commented on all the heuristics defined in this paper, by using 10-level Likert scale, with some help in understanding the meaning of each requirement. To better assess the results, two use-cases were regarded – one for using Skype-like functions (calls and messages), and the other for browsing multimedia contents from mobile phone (handling files, image preview, slideshow).

Table 4 gives an overview of the results obtained by the Phase I. Percentage shows the amount of TRUE statements (would use the *SkypeTV* in the particular case) as opposed to FALSE (would not use the TV). Column A stands for the use case of Skype-like

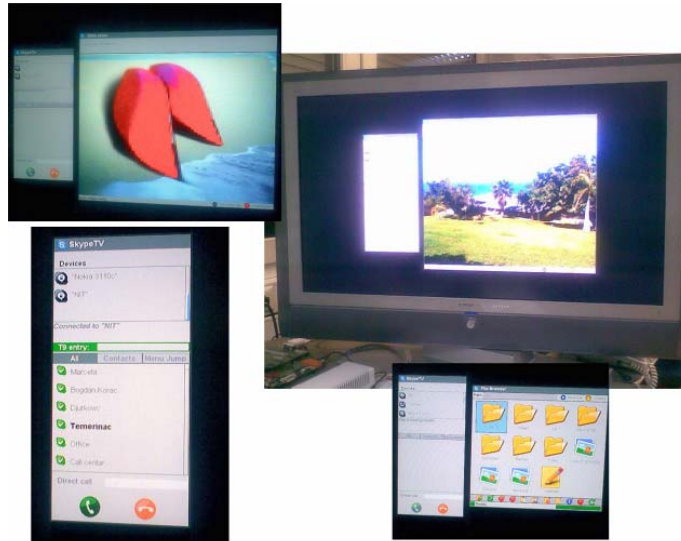


Fig. 3. SkypeTV

functions, while column B stands for multimedia contents browsing.

Table 5 gives an overview of heuristics assessment by users in the experiment. Percentage shows the amount of agreement according to Likert scale (0% - strongly disagree, 100 % - fully agree). Higher percentage reflects the opinion that a heuristic requirement is more likely to be met.

We see that totals from Table 4 and Table 5 do correlate to some extent. Pondering results on privacy heuristics would make totals even closer, since it is clear that privacy is where use-case A scored the worst.

Interference factor calculated earlier gives us a reasonable amount of doubt in the success of *SkypeTV* with users (attentive interference threshold is over by more than 50%). Heuristics support the doubt, especially when the case A is considered. Those two inputs are apparently very valuable in making a decision, since Phase I yielded results that are not totally in favor of *SkypeTV*. However, in the section 5 we present the way to use attentive interference model to suggest operational profile to *SkypeTV* in order to increase user satisfaction.

5. Application in User Awareness Kit (UAK)

Possible applications of the attentiveness characterizations presented in this paper are most certainly within user awareness systems. Here we present an ongoing research project that resulted in a prototype system called *User Awareness Kit* (UAK) (Bjelica et al (2010)). UAK aims to be a future off-the-shelf solution that should be able to provide attentiveness towards users, for any consumer device or application that is subject of such an upgrade. The main intention of UAK is to integrate to the host device, and it can do so in several possible ways: as an add-on system on a chip (SoC) communicating with the host via an intra-processor serial protocol; as a simple software extension to the target system; or, as a network attached device (Network User Awareness Kit – NUAK). UAK tries to provide not only user awareness, but goes further in giving advices of an operational mode, or the current intrusiveness level (IL). Intrusiveness level is calculated as the amount of user-system interference (between the habitual activity models, and system need model), as presented in this paper, upon a Δt interval surrounding the

Table 5. Results of the Phase 2

Heur.	A Skype-like Messages, calls	B Multimedia content browsing	Heur.	A Skype-like Messages, calls	B Multimedia content browsing
A	14%	37%	H.a	60%	85%
B	12%	53%	H.b	53%	55%
C	22%	44%	H.c	70%	95%
D	11%	33%	H.d	62%	78%
E	60%	71%	H.e	32%	85%
F	82%	74%			
G	44%	84%			
Total A-G:	35%	57%	Total H:	55%	80%
			TOTAL:	40%	61%

moment of inquiry (Figure 4). Host device therefore can interact with UAK in a minimal fashion – it would be enough just to poll IL or to seek advice, and based on these data to further tailor GUI appearance, delay notifications or in any other way adapt their behavior towards becoming attentive. The interaction between host device and UAK is done by simple API calls.

UAK works upon two essential pillars: one being attentive interference model, and the other physical sensor data based on which the interference model is refined. Prototype UAK system was connected to the following sensors:

- 3D camera to obtain information on proximity for up to 5 people in front of the camera;
- Optical camera (Logitech QuickCam Pro 9000) to detect presence/motion for up to 5 people in front of the camera;
- Microphone array (5 microphones) based on SEA²M technology to detect speech activity and the position of the current speaker in the 6m radius;

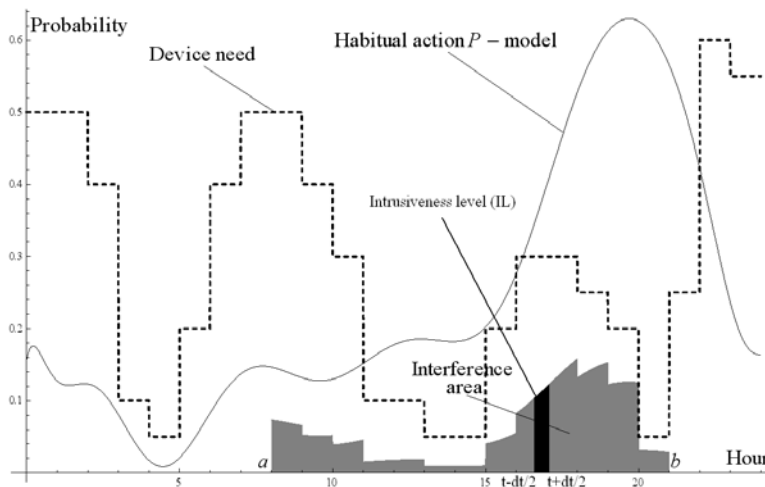


Fig. 4. Interference area and UAK intrusiveness level

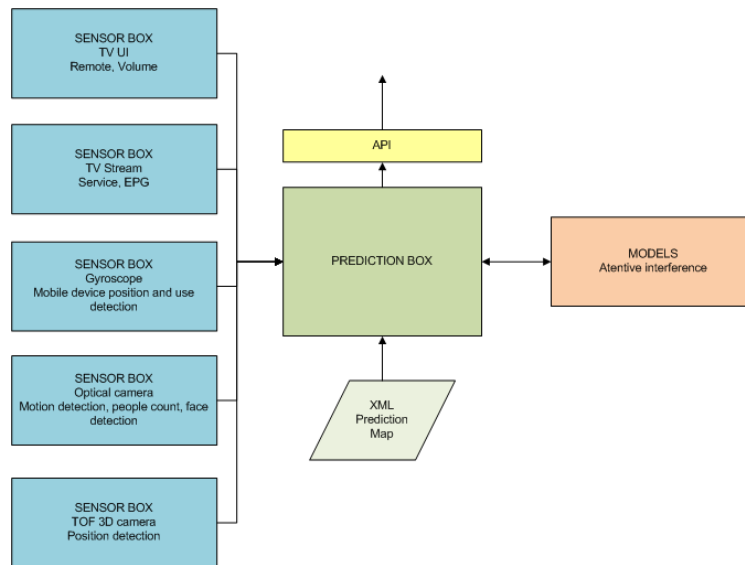


Fig. 5. UAK software architecture

- Accelerometer, attached to cell phone and to a remote controller to gain knowledge on the position/use of these devices.

UAK has the ability of configuring with different habitual activity and system usage models. However, these models are backed by the sensor readings and a mapping mechanism. Mapping block acquires sensor inputs, and based on if-then-else rules and sensor events timing, performs an update action on models. Therefore, models are alive and adapt to the specific environment. The software architecture of UAK is presented in Figure 5.

5.1. SkypeTV Use Case

Prototype UAK system was used for a TV set as a host extended with a *SkypeTV* application. By having a *P-model* that describes TV-watching as a possible interfering habit (Figure 1) and based on sensor measurements, UAK helped the application to choose a mode to be in.

In *Silent (Off)* mode, all alerts are not reported to users: UAK suggests this mode when there is a high probability of concentrated TV watching. Once the probability becomes lower, application shows an indication in the corner of the screen informing the user of all pending alerts. This way user can decide whether to see detailed info or not (this mode is called *Passive*). When the probability of interference is low, mode becomes *Active*, and alerts are made both visually and with appropriate sound. When there are no users in front of the TV, all the alerts are delayed until someone arrives to watch (Figure 6).

Before integration with UAK, *SkypeTV* was always in *Active* mode with alerts obvious to users at all times. We conducted an experiment, aiming to prove that with UAK, users that watch TV do not feel interrupted with incoming alerts. We selected a family, consisting of four members (husband, wife, two teenage children - a girl and a boy). The experiment lasted for two days – Saturday (setup without UAK) and Sunday (setup with UAK). All family members were asked to press green button on the remote controller, if they saw the *SkypeTV* alert and felt comfortable about it. Yellow button should have been pressed if the alert was moderately intrusive. Red button press meant that the notification



Fig. 6. SkypeTV with UAK in Active, Passive and Off modes

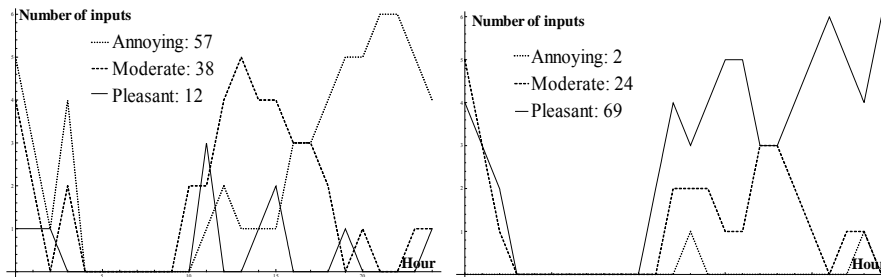


Fig. 7. Experimental results without UAK (left) and with UAK (right)

was annoying. All alerts were pre-programmed to follow similar showing pattern, and TV software logged the results for us to analyze. Figure 7 shows the results.

6. Related works

To the best of our knowledge, in recent researches usability assessment and attentiveness of the system have not been regarded as a whole, nor there are any studies trying to provide a tangible characterization. Many authors address usability of user interfaces and provide guidelines for creating user-oriented UIs. Juristo et al. (2007) propose a method to characterize user interface parameters, and propose a list of parameters to be assessed. Davis (1989) introduced a technology acceptance model (TAM), emphasizing perceived usefulness and perceived ease-of-use as most important enablers. Intille (2002) suggests a concept of interactivity for the embedded devices, where messages should not be aggressive towards users and should not be presented unless absolutely necessary. Concepts of privacy have been analyzed in a study done by Beckwith (2003). Attentive user interface (AUI) paradigm was introduced by Vertegaal (2003). Effects of usability to user trust and satisfaction were examined by Casalo et al. (2008).

Work on how multimedia applications affect people has been done already by other authors. Makela (2005) investigates how multimedia affects people culturally. Jaimes (2006) in his study also states that user should be a central point for each new multimedia system, emphasizing cultural background of users.

Usable inputs to the topic of this paper were provided by several other researches. Koskela et al. (2004) in their ethnography study, analyze the use of a light control application on TV receiver by a young married couple. Wonneberger et al. (2009) examine dynamics of individual television viewing behavior.

7. Conclusion

In this paper, we presented metrics and heuristics for characterizing application attentiveness to its users, which we believe to be the next crucial enabler for the success of new user-oriented technologies. We introduced notions of *system attentiveness* and *attentive interference*, as new deliverables for the final system quality. System

attentiveness and system respect to its users should always be a milestone to reach in terms of quality. No quality assessment can be complete for any user application, if system attentiveness is not assessed with care.

This paper gives a solid start for future work on profiling attentiveness metrics in more detail, with respect to the impact of different applications to users' everyday life. It provides a new angle and a new way of thinking prior to bringing decisions on new applications – it is necessary not only to provide functions, but to provide them with care.

References

- Australian Government Research and Statistics (2009). *Get the Picture, Free-to-air TV, Audiences (viewing patterns)*, <http://www.screenaustralia.gov.au/gtp/wftvviewtimeofday.html>
- Baker, K. (2006). *Intrusive Interactivity Is Not an Ambient Experience*, IEEE Multimedia, Vol. 13, No. 2, pp. 4-7.
- Beckwith, R. (2003). *Designing for ubiquity: the perception of privacy*, IEEE Pervasive Computing, Vol. 2, No. 2, pp. 40-46.
- Bjelica, M. Z.; Teslic, N. (2009). *A concept of usability assessment for user-centered multimedia applications*, in proceedings of International Multiconference on Computer Science and Information Technology, 12-14 Oct. 2009, pp. 443-450.
- Bjelica, M. Z.; Teslic, N. (2009). *A Concept of System Usability Assessment: System Attentiveness as the Measure of Quality*, in proceedings of First IEEE Eastern-European Regional Conference on the Engineering of Computer Based Systems, 7-8 Sept. 2009, pp. 144-145.
- Bjelica, M. Z.; Teslic, N. (2010). *Multi-Purpose User Awareness Kit for Consumer Electronic Devices*, in proceedings of International Conference on Consumer Electronics, 11-13 Jan. 2010
- Bjelica, M.; Savic, M.; Aleksic, T. (2008). *One solution of integrated file browser for connected Bluetooth device as an application for TV receiver*, in proceedings of 52nd ETRAN conference, 8-12 June 2008, TE2.2-1-4
- Bluegiga WT11 Module (2009). *Products, Bluetooth modules, WT11 - Class 1 Bluetooth® 2.1+EDR Module*, http://www.bluegiga.com/WT11_Class_1_Bluetooth_Module
- Carr, N. (2003). *IT Doesn't Matter*, Harvard Business Review, Vol. 81, No. 5, pp. 41-49.
- Casalo, L. V.; Cisneros, J. (2008). *An Empirical Test of the Multiplicative Effect of Usability on Consumer Trust and Satisfaction*, in proceedings of 19th International Conference on Database and Expert Systems Application, 1-5 Sept. 2008, pp. 439-443.
- Ckipe.com (2009). *Skype Statistics: CKIPE, the borderless communicator, Recent Skype Data (concurrent users online)*, <http://ckipe.com/borderless>
- Davis, F. D. (1989). *Perceived Usefulness, Perceived Ease Of Use, And User Acceptance Of Information Technology*, MIS Quarterly, Vol. 13, No. 3, pp. 318-340.
- Intille, S. (2002). *Designing a Home of the Future*, IEEE Pervasive Computing, Vol. 1, No. 2, pp. 76-82.
- Jaimes, A. (2006). *Human-Centered Multimedia: Culture, Deployment, and Access*, IEEE Multimedia, Vol. 13, No. 1, pp. 12-19.
- Juristo, N.; Moreno, A. M.; Sanchez-Segura, M.-I. (2007). *Guidelines for Eliciting Usability Functionalities*, IEEE Transactions on Software Engineering, Vol. 33, No. 11, pp. 744-758.
- Koskela, T.; Vaananen-Vainio-Matilla, K. (2004). *Evolution towards smart home environments: empirical evaluation of three user interfaces*, Personal and Ubiquitous Computing, Vol. 8, No. 3-4, pp. 234-240.
- Kuan-Ta, C.; Chun-Ying, H.; Huang, P.; Chin-Laung, L. (2006). *Quantifying Skype user satisfaction*, in proceedings of the 2006 conference on Application, technologies, architectures and protocols for computer communications, pp. 399-410.

- Lakobrija, R.; Smiljkovic, N.; Papp, I.; Kukulj, D.; Savic, M. (2007). *One implementation for extension of TV device with Skype application*, in proceedings of 15th conference TELFOR, 20-22 Nov. 2007, pp. 549-552
- Lekakos, G.; Chorianopoulos, K.; Spinellis, D. (2001). *Information systems in the living room: A case study of personalized interactive TV design*, in proceedings of the 9th European Conference on Information Systems, 27-29 June 2001, pp. 319-329.
- Makela, T. (2005). *Multimedia Software as Culture: Towards Critical Interaction Design*, IEEE Multimedia, Vol. 12, No. 1, pp. 14-15.
- Malhotra, Y.; Galetta, D. F. (2004). *Building systems that users want to use*, Communications of the ACM, Vol. 47, No. 12, pp. 88-94.
- Markus, L.; Keil, M. (1994). If we build it, they will come: Designing information systems that people want to use, MIT Sloan Management Review, Vol. 35, No. 4, pp. 11-25.
- MESA Imaging (2009). *SwissRanger™ SR4000 Overview*, <http://www.mesa-imaging.ch/prodview4k.php>
- O'Brien, J.; Rodden, T.; Rouncefield, M.; Hughes, J. (1999). *At home with the technology: an ethnographic study of a set-top-box trial*, ACM Transactions on Computer-Human Interaction, Vol. 6, No 3, pp. 282-308.
- Obrist, M.; Bernhaupt, R.; Tscheligi, M. (2008). *Interactive TV for the Home: An Ethnographic Study on Users' Requirements and Experiences*, International Journal on Human-Computer Interaction, Vol. 24, No. 2, pp. 174-196.
- Papp, I.; Saric, Z.; Jovicic, S.; Teslic, N. (2007). *Adaptive Microphone Array for unknown desired speaker's transfer function*, Journal of the Acoustic Society of America, Vol. 122, No. 2, EL44-EL49.
- SkypeStats.com (2008). *Daily Skype Users Online Worldwide*, <http://www.skypestats.com/>
- Trident Microsystems (2009). *HiDTV™ digital TV SoC Video Processor Family (VGCB is discontinued)*, <http://www.tridentmicro.com/product.asp#3>
- UK Office for National Statistics (2003). *The UK 2000 Time Use Survey, TV Viewing*, http://www.statistics.gov.uk/timeuse/summary_results/activities_age_gender.asp
- Vertegaal, R. (2003). *Attentive User Interfaces*, Communications of the ACM, Vol. 46, No. 3, pp 30-33.
- Wonneberger, A.; Schoenback, K.; Meurs, L. (2009). *Dynamics of Individual Television Viewing Behavior: Models, Empirical Evidence, and a Research Program*, Communication Studies, Vol. 60, No. 3, pp. 235-252.