

Fabio Paternò,
Antonio Giovanni
Schiavone
(CNR-ISTI, Human
Interfaces in
Information Systems
Laboratory)

Valutazione dell'usabilità di applicazioni web su dispositivi mobili tramite individuazione di “bad usability smells”

Evaluation of web applications' usability on mobile devices through “bad usability smells” detection

Sommario: Il Web è lo strumento più diffuso e pervasivo per comunicare informazioni, stimolare la crescita di comunità o fornire servizi interattivi. Oggigiorno, i dispositivi più comuni per la navigazione Web sono gli smartphone, che ormai da tempo hanno superato i personal computer nelle statistiche relative all'accesso in rete. Conseguentemente, nell'ambito dello sviluppo di applicazioni e siti web è diventato fondamentale garantire una buona usabilità per l'accesso tramite dispositivi mobile. Nel corso degli anni sono stati proposti vari strumenti automatici per la valutazione dell'usabilità di applicazioni web, molti dei quali non consentono di effettuare un'analisi dell'usabilità basata sul comportamento degli utenti nei loro contesti quotidiani.

Il nostro laboratorio ha sviluppato un nuovo metodo di valutazione dell'usabilità, basato sull'analisi delle interazioni degli utenti, allo scopo di identificare specifici comportamenti degli utenti (detti “bad usability smells”), che spesso vengono manifestati per gestire situazioni problematiche nell'uso di una applicazione. Il metodo consiste nell'analizzare i log delle interazioni degli utenti ed identificare la presenza di queste interazioni problematiche, che sono descritte e formalizzate tramite un linguaggio specifico realizzato ad hoc. Tale metodo è stato applicato, tramite il relativo strumento automatico di supporto, in un caso di studio relativo ad una applicazione Web italiana di largo uso: attraverso tale approccio è stato possibile identificare alcune parti dell'interfaccia utente dell'applicazione che erano problematiche dal punto di vista dell'usabilità.

Abstract: The Web is the most widespread and pervasive way for communicating information, stimulating community growth or providing interactive services. Nowadays, the most common devices for Web browsing are smartphones, which have long since surpassed personal computers in network access statistics. Consequently, in the development of applications and websites it has become essential to guarantee good usability for access via mobile devices. Over the years, various automatic tools have been proposed for the evaluation of the usability of web applications, many of which do not allow usability analysis based on user behavior in their daily contexts.

Our laboratory has developed a new usability evaluation method, based on the analysis of user interactions, in order to identify specific

user behaviors (called "bad usability smells"), which are often manifested to handle problematic situations during the use of a web application. The method consists in analyzing the log of user interactions and identifying the presence of these problematic interactions, which are described and formalized through a specific language created for this purpose. This method has been applied, through the relative automatic support tool, in a case study of a widely used Italian Web application: through this approach it was possible to identify some parts of the user interface of the application that were problematic from the usability point of view.

1. Introduzione

Il World Wide Web è un mezzo globale indispensabile di comunicazione per persone, aziende e pubbliche organizzazioni. Lo sviluppo di applicazioni Web di semplice uso è diventato ormai un elemento cruciale per chiunque voglia promuovere servizi o trasmettere informazioni. Questa necessità è resa ancora più stringente dall'uso diffuso di dispositivi mobili, i quali oggi sono le piattaforme più utilizzate per svolgere attività di svago (come effettuare ricerche online[1] o usufruire di contenuti multimediali) e il loro utilizzo nelle attività professionali è in costante crescita[2].

Difatti, l'uso di tali dispositivi mobili ha cambiato il modo in cui le persone navigano nel Web ed usano i servizi online: lo svolgimento di tali attività non è più relegato in ambienti definiti quali ad esempio uffici o case, ma può essere effettuato praticamente in qualunque luogo, come ad esempio sui mezzi di trasporto pubblico, in ambienti all'aperto, all'interno di locali pubblici, etc).

Per diversi anni, i ricercatori hanno condotto studi relativi all'analisi e al miglioramento dell'usabilità delle applicazioni Web[3][4], proponendo diversi strumenti, metodologie e tecniche per questo scopo. In particolare, gli studi si sono concentrati sullo sviluppo di strumenti di valutazione automatica dell'usabilità [5], il cui utilizzo consente di ridurre i tempi e i costi necessari all'analisi dell'usabilità, liberare i valutatori da compiti ripetitivi e noiosi, e che le valutazioni siano possibili anche per applicazioni complesse senza aumentare i costi di valutazione eccessivi. Gli strumenti di valutazione automatica dell'usabilità possono essere classificati in due gruppi principali: quelli che usano il codice sorgente delle pagine Web (cioè la loro struttura e / o contenuto) come fonte di dati per il rilevamento di problemi di usabilità, e quelli che si concentrano sull'analisi dei dati di interazione dell'utente reale. Il primo gruppo include alcuni strumenti commerciali come Google Mobile Friendly Test Tool [6] o Bing's Mobile Friendliness Test Tool [7]. A partire dalla struttura della pagina Web, questi strumenti cercano di valutarne l'usabilità nello specifico contesto della navigazione web effettuata tramite dispositivi mobili. Nel secondo gruppo, invece, i dati di utilizzo possono essere recuperati dai log del server (contenente principalmente le sequenze cronologiche delle pagine web visitate) o dalla registrazione lato client delle attività degli utenti mentre stanno navigando (registrando così sia la sequenza di pagine Web visitate che interazioni

infra-pagina, come ad esempio selezioni con il mouse, uso della barra di scorrimento, ecc.).

L'ascesa di dispositivi come smartphone e tablet ha portato all'ampia adozione di tipi di interazioni utente che sono significativamente diversi da quelli sui dispositivi desktop. Le differenze nascono dai molti possibili contesti d'uso, dalle limitazioni tecniche dei dispositivi mobili (ad es. Connettività non sempre ottimale, dimensioni dello schermo minori, con risoluzioni diverse, capacità di elaborazione talvolta limitata), e dal modo in cui gli utenti interagiscono con loro (ad esempio, alcuni utenti preferiscono interagire con gli smartphone con entrambe le mani, gli altri preferiscono interagire con una sola mano [8]). Variazioni in ciascuno di questi fattori (ad esempio il cambiamento della dimensione dello schermo [9]) può quindi portare a diverse percezioni di usabilità. Al fine di comprendere e analizzare meglio questi tipi di interazioni degli utenti è quindi necessario definire criteri specifici e sviluppare nuovi strumenti di valutazione per garantire poi un'adeguata valutazione dell'usabilità anche in contesti mobili. Per venire incontro a tale esigenza, presso l'HIIS Laboratory dell'Istituto di Scienza e Tecnologie del Consiglio Nazionale delle Ricerche è stato elaborato una metodologia basata sul rilevamento automatico di indicatori di problemi di usabilità (definiti "Bad Usability Smells") nei dispositivi mobili durante l'accesso alle applicazioni Web. Per mettere in pratica tale metodologia è stato inoltre sviluppata una piattaforma di valutazione dell'usabilità, chiamata M.U.S.E. (Mobile Usability Smell Evaluator)[10], che è in grado di registrare l'utente comportamento durante l'interazione con qualsiasi applicazione Web attraverso qualsiasi tipo di dispositivo abilitato per il browser. Le interazioni utente così ottenute vengono elaborate da un algoritmo per l'identificazione di specifici pattern di interazione che indicano la presenza potenziale di problemi di usabilità.

2. Formalizzazione dei Bad Usability Smell

I Bad Usability Smell rappresentano degli indicatori della potenziale presenza di qualche problema di usabilità all'interno di un'applicazione web.

Nello sviluppo di un approccio basato sull'individuazione di tali indicatori, un ruolo fondamentale è ricoperto dagli utenti di tali applicazioni, che, tramite le loro interazioni con l'applicazione, sono la fonte dei dati utili al rilevamento di tali criticità dal punto di vista dell'usabilità. Al fine quindi di garantire la qualità dei dati acquisiti, il rilevamento delle interazioni dovrebbe essere effettuato in modo non invadente, in modo da non influire sul comportamento dell'utente durante l'esecuzione dei suoi compiti. Le strategie adottate per svolgere compiti possono variare da utente per utente in base a vari aspetti, ad esempio contesti d'uso diversi, dispositivi diversi, diversa personalità, ecc. Anche di fronte allo stesso problema di usabilità, gli utenti possono adottare varie strategie (e di conseguenza, comportamenti) volte a superare eventuali situazioni problematiche. Nonostante tale variabilità, è possibile definire un piccolo sottoinsieme di comportamenti più

frequentemente adottati dagli utenti per far fronte a specifici problemi di usabilità. Questo sottoinsieme è quindi la base per la definizione dei Bad Usability Smell. Vari ricercatori hanno condotto negli anni alcuni studi volti a catturare tali “comportamenti più frequenti”, o in relazione a particolari tipologie di dispositivi (ad esempio nei confronti di applicazioni desktop) oppure rivolti a diverse finalità (ad esempio il supporto all'adattamento automatico delle pagine web). A causa delle limitate dimensioni dello schermo, della scarsa precisione di puntamento del tocco basata sulle dita umane, gli utenti, navigando tramite dispositivi mobili, possono incontrarne delle difficoltà se la progettazione dell'interfaccia utente non tiene conto delle peculiarità di tali dispositivi.

Una prima fase del nostro lavoro è stato definire un insieme di problemi di usabilità che possono essere rivelati tramite l'analisi del comportamento dell'utente. Per identificare questo set ci siamo basati sui lavori presenti in letteratura scientifica (ad es. [11]), informazioni fornite da software commerciali, come ad esempio [12], importanti studi relativi all'usabilità mobile [13] e nostre analisi sullo sviluppo di vari siti Web e di come gli utenti interagiscono con essi in ambiente mobile. Alla fine di questa fase di studio, abbiamo identificato sei diversi problemi di usabilità:

- **Elementi troppo piccoli o vicini:** questo bad smell è caratterizzato da elementi interattivi nella pagina Web che sono eccessivamente piccoli o vicini. (ad es. Figura 1)
- **Links troppo vicini:** corrisponde a una selezione dell'utente che implica il caricamento di una pagina sbagliata.
- **Contenuto distante:** quando l'utente è forzato ad eseguire molti scroll in alto ed in basso.
- **Sezione Troppo Piccola:** una sezione troppo piccola richiede specifiche azioni di ingrandimento.
- **Cattiva leggibilità:** quando vi sono testi con font troppo piccoli o spaziature troppo ridotte (ad es. Figura 2 – lato sinistro)
- **Form lunghe:** quando vi è un numero di elementi in una form che può essere considerato eccessivo per una buona usabilità su dispositivo mobile (ad es. Figura 2 – lato destro)

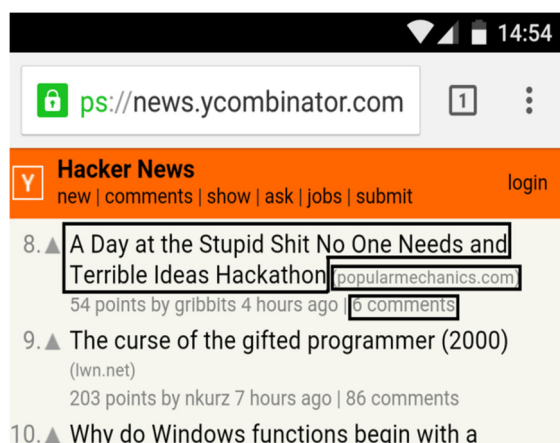


Figura 1. Esempio di elementi troppo vicini

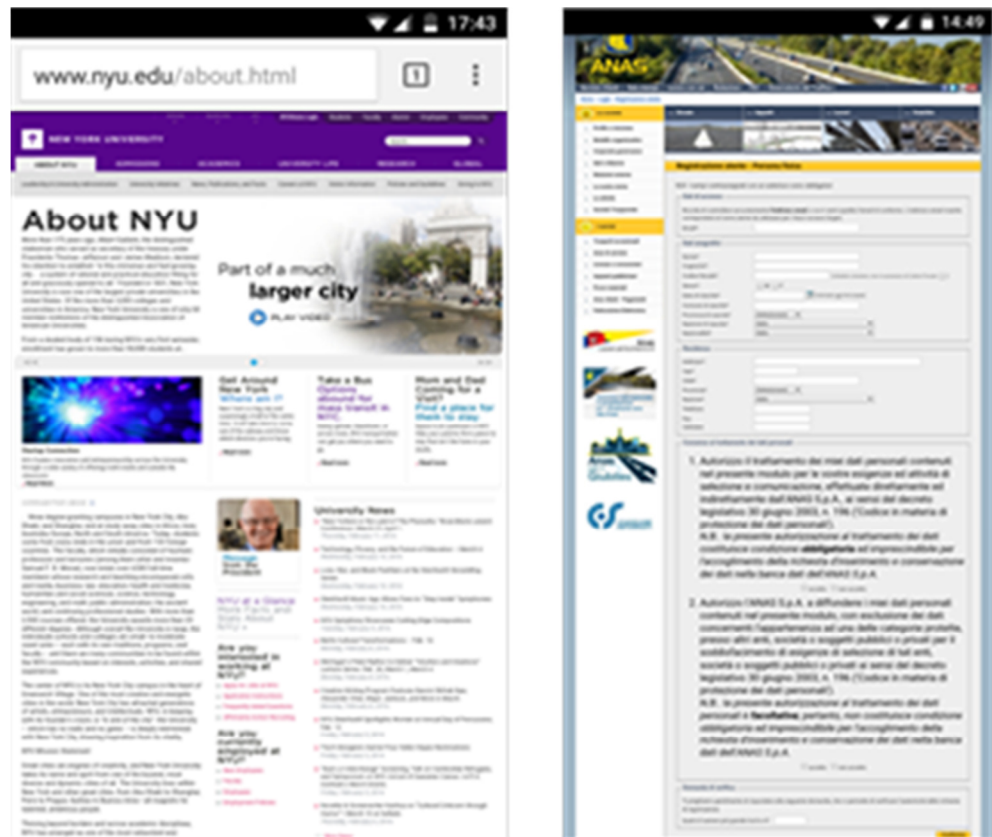


Figura 2.. Esempio di cattiva leggibilità (a sinistra) e di form lunghi (a destra)

Un passo necessario al fine di sviluppare un sistema automatico per il rilevamento dei bad smell è stato lo sviluppo di un linguaggio utile per formalizzare tali indicatori. Poiché le sessioni utente sono registrate come sequenze di eventi, anche tali indicatori sono stati formalizzati come pattern di eventi. Quindi, il rilevamento di bad usability smell si ottiene controllando, all'interno delle sequenze di eventi registrate, la presenza di pattern di eventi che rappresentano uno o più cattivi odori di usabilità. Durante la fase di sviluppo di tale approccio, abbiamo notato come sia molto difficile individuare sequenze di eventi esattamente precise per rilevarli. Infatti, i comportamenti degli utenti possono variare in molti modi, anche minimi, e quindi è impossibile associare un problema di usabilità con una sequenza di eventi esatta. Questo aspetto ha portato all' introduzione di una serie di operatori per definire le tipologie di sequenze di eventi rilevanti, e quindi destinati a facilitare il rilevamento di sequenze simili tra loro, anche se non esattamente identiche, in modo da introdurre una certa flessibilità nel meccanismo di rilevamento. Ad esempio abbiamo introdotto:

- Operatori per indicare il numero di ripetizioni di un evento (o per indicare la presenza di nessun limite di ripetizioni, indicate tramite il carattere speciale *).
- Operatori per indicare una direzione negli eventi per cui l'informazione ha senso (ad esempio lo scroll), oppure lasciare indefinita tale informazione (indicate tramite il carattere speciale \$).
- Operatori per definire un'intervallo temporale, inteso come tempo massimo che può trascorrere tra l'evento precedente e quello corrente.

In Tabella 1 è possibile osservare alcuni esempi di formalizzazione astratta dei bad smell. Per esempio, nel primo caso abbiamo una sequenza di eventi di pinch, seguita da una sequenza di eventi pan in qualunque direzione, e conclusa da un evento tap su un elemento con conseguente cambio di focus.

Bad Usability Smell	Pattern di Eventi
Too Small or Close Elements	[*] Pinch(out) + [*]Pan(\$) + Tap + Focus(in)
Too Close Links	[*]Tap + Beforeunload + Pageview + Beforeunload + Pageview
Distant Content	[5]Pan(down)

Tabella 1. Pattern comportamentali per alcuni bad smell

Per l'implementazione concreta del linguaggio definito, utile per consentirne l'applicazione in un sistema automatico, seguendo l'esempio di numerosi altri lavori in letteratura (ad esempio [14]), abbiamo utilizzato un semplice linguaggio XML. I vantaggi nell'uso di XML per formalizzare tali pattern sono che la loro descrizione è facilmente comprensibile sia da esseri umani che da sistemi informatici, può essere facilmente estesa e/o modificata e può essere convalidata definendo un appropriato Schema XSD.

3. Il validatore di usabilità M.U.S.E.

Al fine di implementare concretamente l'approccio proposto, abbiamo sviluppato uno strumento automatico di valutazione dell'usabilità web chiamato M.U.S.E. (Mobile Usability Smell Evaluator), la cui architettura è illustrata in Figura 3.

Tale strumento è basato su di un proxy che è in grado di registrare il file con gli eventi generati durante da un utente durante l'interazione con qualsiasi Web applicazione tramite dispositivi desktop o mobili.

I dati sul comportamento degli utenti sono raccolti attraverso un Logger JavaScript iniettato nella pagina Web dal proxy: lo strumento è quindi in grado di registrare le interazioni dell'utente con qualsiasi sito Web, e quindi senza la necessità per il proprietario del sito Web di installare manualmente gli script di registrazione dei dati.

Architettura di M.U.S.E. (Mobile Usability Smell Evaluator)

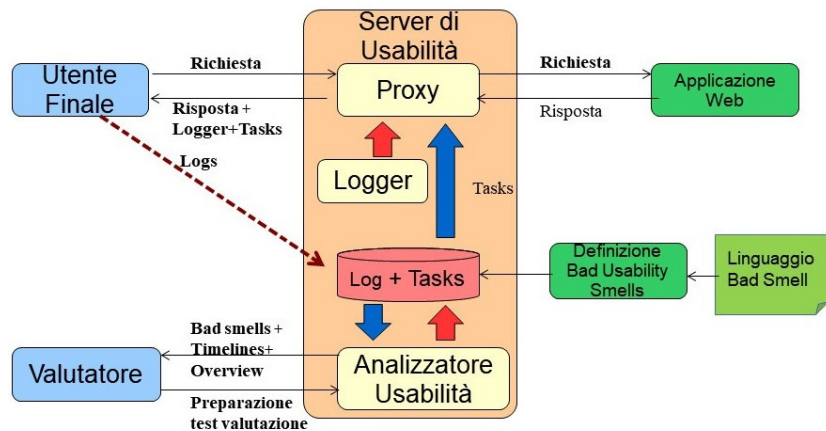


Figura 3 . L'architettura di M.U.S.E.

M.U.S.E. include anche un pannello di gestione attraverso il quale un esperto di usabilità, oltre ad associare la piattaforma ad un particolare sito, può definire dei compiti specifici che gli utenti possono completare al fine di testare una particolare pagina o una particolare funzionalità di un'applicazione web. Ogni interazione dell'utente è registrata come sequenza di eventi, che include sia quelli generati direttamente dall'utente (ad es. click, tap, movimento del mouse) che quelli generati dal browser in risposta alle azioni dell'utente (ad esempio ridimensionamento della pagina, cambiamento di orientamento di un dispositivo mobile).

Una volta catturati tali sequenze di eventi, il validatore consente agli esperti di usabilità di analizzare le interazioni dell'utente, graficamente rappresentate tramite delle timeline [15]: ognuna di esse può essere fatta scorrere temporalmente in avanti ed indietro, ed inoltre una coppia di timeline può essere sovrapposta per consentire di confrontare visivamente il comportamento di due distinti utenti nell'esecuzione dello stesso compito. Sempre tramite lo stesso pannello, l'esperto può attivare la funzionalità di riconoscimento automatico dei bad smell, precedentemente formalizzati e caricati sulla piattaforma. La Figura 4 mostra una timeline in cui è stata rilevata la presenza di un sottoinsieme di eventi corrispondente ad un bad smell, evidenziato in rosso.

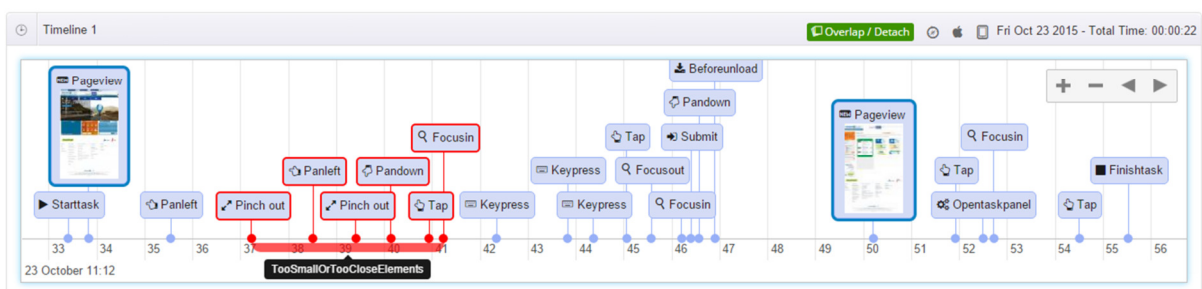


Figura 4. Esempio di timeline all'interno della quale è stato rilevato un bad smell

4. Test Utenti

Al fine di testare il nostro approccio, fu creata una sessione di valutazione dell'usabilità, composta da quattro diversi compiti, da svolgersi sulla versione inglese di un sito relativo alle autostrade italiane, che fornisce servizi di pubblica utilità e che presentava problematiche di usabilità. A tale test hanno partecipato circa 40 utenti (età media 28) di entrambi i sessi, che hanno eseguito i compiti previsti dalla sessione di test tramite il loro smartphone senza vincoli di luogo e di tempo. L'esecuzione di tale test ha portato alla generazione di più di 14000 eventi, in cui MUSE ha trovato 51 istanze di bad usability smells.

I risultati dei test evidenziarono alcuni problemi di usabilità riguardanti la struttura dell'interfaccia utente, avente troppi elementi e non bene organizzati: in particolare, i più comuni bad usability smells rilevati nell'applicazione considerata furono "Elementi Troppo Piccoli o Vicini" e "Contenuto Distante".

5. Conclusioni

Abbiamo presentato un nuovo approccio per la valutazione dell'usabilità delle applicazioni web, con particolare riferimento all'usabilità in ambito mobile. Tale metodologia, basata sulla raccolta di dati relativi all'interazione di alcuni utenti all'interno di una applicazione web e sul successivo rilevamento di bad usability smell, è stata implementata uno strumento automatico di valutazione, chiamato M.U.S.E.

La soluzione proposta consente la valutazione di applicazioni Web mobile senza obbligare il possessore dell'applicazione a modificarla manualmente per seguire tali test. Inoltre, secondo l'approccio proposto, gli utenti che partecipano al test possono operare "in the wild", ossia senza l'obbligo dell'uso di dispositivi specifici, di svolgere i test in luoghi o in tempi predefiniti.

L'insieme dei bad usability smells, definiti all'interno della nostra metodologia, è stata formalizzata tramite linguaggio XML, in modo che tale insieme sia facilmente modificabile ed espandibile, senza richiedere cambiamenti nell'implementazione del validatore di usabilità. Pianifichiamo ulteriori studi per migliorare le funzionalità di M.U.S.E. e di effettuare ulteriori studi volti a migliorare e raffinare le definizioni dei bad usability smells.

Bibliografia

- [1] Google's AdWords official blog (May 2015)
<http://adwords.blogspot.com/2015/05/building-fornext-moment.html>.
- [2] ComScore's Global Mobile Report (July 14, 2015).
<http://www.comscore.com/Insights/Presentations-andWhitepapers/2015/The-Global-Mobile-Report>
- [3] J. Grigera, A. Garrido, and J. M. Rivero "A tool for detecting bad usability smells in an automatic way" in Web Engineering, ser. Lecture Notes in Computer Science. vol. 8541 pp. 490–493, 2014.
- [4] V. F. de Santana and M. C. Calani Baranauskas "WELFIT: A Remote Evaluation Tool for Identifying Web Usage Patterns through Client-Side Logging" in International Journal of Human-Computer Studies, vol. 76 no. C pp 40-49, 2015.
- [5] M. Y. Ivory and M. A. Hearst "The state of the art in automating usability evaluation of user interfaces" in ACM Computing Surveys (CSUR) vol. 33 no. 4 pp. 470-516, 2001.
- [6] Google Mobile Friendly Test Tool.
<https://www.google.com/webmasters/tools/mobilefriendly/>
- [7] Bing's Mobile Friendliness Test Tool.
<https://www.bing.com/webmaster/tools/mobilefriendliness>
- [8] S. Boring, D. Ledo, X. Chen, N. Marquardt, A. Tang and S. Greenberg, "The fat thumb: using the thumb's contact size for single-handed mobile interaction" in Proc. 14th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'12), 2012, pp. 39-48.
- [9] D. Raptis, N. Tselios, J. Kjeldskov and M. Skov "Does size matter? investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency" in Proc. 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI 2013), 2013, pp. 127-136).
- [10] F. Paternò, A. G. Schiavone, A. Conti: "Customizable Automatic Detection of Bad Usability Smells in Mobile Accessed Web Applications", Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2017)
- [11] J. Grigera, A. Garrido, and J. M. Rivero "A tool for detecting bad usability smells in an automatic way" in Web Engineering, ser. Lecture Notes in Computer Science. vol. 8541 pp. 490–493, 2014.
- [12] Google Search Console. Retrieved February 6, 2017 from
<https://www.google.com/webmasters/tools/>
- [13] J. Nielsen and R. Budiu. Mobile Usability. New Riders (2013)

- [14] A. G. Schiavone, F. Paternò: "An extensible environment for guidelinebased accessibility evaluation of dynamic Web-applications", *Universal Access in the Information Society* v. 14, no. 1
- [15] F. Paternò, A. G. Schiavone, P. Pitardi. "Timelines for Mobile Web Usability Evaluation", in *Proc. of the International Working Conference on Advanced Visual Interfaces (AVI 2016)*, 2016, pp. 88-91.