Using Software Agents to Raise Awareness and Lower Information Overload in a Multi-user Collaborative Environment

Conference or Workshop Item

Goddard, David (2019): Using Software Agents to Raise Awareness and Lower Information Overload in a Multi-user Collaborative Environment. In: Proceedings of the 16th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing an the Design of Cooperation Technologies - Doctoral Colloquium Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2019_dc3

Using Software Agents to Raise Awareness and Lower Information Overload in a Multi-user Collaborative Environment

David Goddard
Knowledge Media Institute | Open University
david.goddard@open.ac.uk

Abstract. Ubiquitous devices provide users with notifications that continue to blur the distinction between work and personal activities and can lead to information overload. This research aims to support users of multiple collaborative and social systems who may experience this, by designing and evaluating a personal software agent to support the user and ameliorate the effects of overload. A technique uniting personas and Empathy Map has been applied to model typical user behaviours, goals and pain points, which will inform the design of a solution to manage interruptions and information overload.

Introduction

Most users of social and collaborative software have experienced some form of Information Overload (IO) (Schmitt et al., 2018) and studies on social media platforms (Rodriguez et al., 2014) have shown IO to have a measurable negative impact on users. The intersection between personal and work usage has also become increasingly blurred. We aim to design and evaluate an approach to ameliorate IO arising from frequent interruptions and large volumes of content from diverse sources across work and personal contexts. Many systems - such as Facebook and Twitter - have their own algorithmic approaches to sift and prioritise

large content volumes. We hypothesised that these approaches were flawed[1,2] because of: a) inconsistency of design and implementation across applications; b) mismatch between the priorities of the user and those of the provider; and c) privacy concerns over sharing data with the provider. We further hypothesised that the productivity and well-being of individual users can be improved by shifting more work to autonomous software agents supported by Artificial Intelligence (AI), if these can make content decisions that are consistent with those the user would themselves make.

We use the concept of *awareness* (Metaxas and Markopoulos, 2008) to model the domain in which the agent operates, which we take here to mean the ability of actors[3] to perceive the activities and output of other actors in the system, as influenced by the actor's own activities and goals. We phrase an "awareness problem" - how does a user maintain awareness of relevant information without being so aware of noise that it impairs their function? Seeking to define a relationship between awareness and information overload, we apply a qualifier: *effective* awareness. That is, the actual awareness that an individual is able to maintain taking into account limiting factors such as the human ability to process information. We contend that reduced processing performance due to information overload has a negative impact on an individual's effective awareness; conversely, reducing the effect of information overload on the user can improve their effective awareness.

To test the validity of the hypotheses and provide a basis for solution design and evaluation, we conducted a survey of users of collaborative and social systems. We then developed a set of Personas (Cooper, 1998) to reveal requirements applicable to distinct types of users. These personas will serve as a design and evaluation aide for a novel software system to address the identified needs.

An Awarenesss Agent

The ability of computers to process large quantities of data and learn from user behaviour can help tackle IO. While approaches such as applying simple filters to data can be effective, we investigate how a more sophisticated software agent can act in concert with a human user to enhance their interaction with multiple collaborative systems.

Consider a software agent that acts on behalf of a user to monitor one or more systems, that brings information to their attention and undertakes certain interactions autonomously. We could say this *awareness agent* extends the *focus* of a user of a collaborative system, while also occupying their *nimbus*.

As Benford et al. (1994) summarised it: "The more an object is within your focus, the more aware you are of it. The more an object is within your nimbus, the

[1] https://www.engadget.com/2018/12/18/twitter-chronological-timeline-feature-latest-tweets/

[2] https://www.reviewgeek.com/1328/facebooks-news-feed-algorithm-is-completely-busted/

[3] An actor is not necessarily human, it can also be a software agent or automated emitter of data

more aware it is of you". Adopting social media terminology, we can describe the focus as the people you are *following* and the nimbus as the *followers*.

An agent may also manage a user's own presence as seen by followers, whereby the agent would monitor the output created by its 'owner', and perform actions such as generating push notifications that other actors may receive. The distinction of this mode of operation is that the agent is processing content using rules defined by the *creator* of the content rather than by a consumer. This allows the creator to deliberately promote content items that they themselves judge to be noteworthy.

While many platforms use their own algorithms to select content, the awareness agent is intended to *democratise* this process, handing more control over selection and sharing to the users themselves. The agents and users together can be considered to comprise a Social Machine (Berners-Lee and Fischetti, 1999).

Research Questions and Approach

The following broad research questions are considered:

1. *How should the agent prioritise incoming information and communicate this to the user?*
2. *How should the agent present the user's activities to others?*
3. *How should the user affect and understand the behaviour of the agent?*

The first question considers the *focus* of the user, the second concerns their *nimbus*, while the third examines the relationship between the user and agent itself and considers issues of control, transparency and data privacy.

The domain is intentionally wide: the concept requires that the agent interact with a diverse range of media, applications and individuals to perform its function. However, certain specific cases have been identified, such as the collaboration service Slack, which is widely used, has an extensive API and competes with other media for users' attention. Other cases include email and social media.

An awareness agent should act independently, processing disparate information sources, and should be able to learn via both implicit techniques and explicit training. As well as monitoring *incoming* content, it should also act in an *outward facing* capacity, communicating its owner's activities externally.

A design science methodology is used (Peffers et al., 2007), evaluating theoretical solutions to the problem by developing software implementations of an awareness agent to address specific use cases and then evaluating these solutions.

Survey and Persona Development

User opinions were solicited in a survey, that was advertised with an intention to reach people who may experience IO. This was grouped around five themes:

- Attitudes to interruptions originating from application notifications.
- How well online services understand respondents' preferences and interests.
- Degree of trust and confidence in online services.
- General views on online services, connected applications and smartphones.
- Differentiation between work and personal use of apps and services.

A data-led approach in line with that described by McGinn and Kotamraju (2008) was used to generate the personas, employing a cluster analysis process to map respondent groups to personas (Tu et al., 2010). Persona construction used a hybrid of quantitative and qualitative inputs: the output of the clustering process evidenced the personas, but also supplied some more subjective criteria to enable the creation of a balanced and representative set. The PATHY technique for persona development (Ferreira et al., 2018) was selected for this work because of the improved guidance that it provides to structure users' perceptions and feelings, and to relate software features to personas.

A two-step clustering process was chosen, with individual clusters first generated for each question group, which were then used as a basis for an overall cluster of clusters. This approach was selected because the first analysis suggested that clusters of users tended to respond similarly across a theme. The second level clustering was used to identify commonalities across these theme-based clusters.

Each cluster was assessed for how members had responded to each of the individual questions of the survey (for example, how did Cluster *A* members respond to the question: *I receive so much information online that I often miss things that are important or time critical*?). Chi-squared tests determined in which cases there was a statistically significant relationship between cluster and response.

Having determined the significant combinations, applicable responses to questions were assigned to individual clusters in order to frame archetypes. For example, members of cluster A said that they separated work and personal usage and were not comfortable sharing personal information. These attributes, combined with demographic information that also emerged from the clusters, fed the PATHY technique to derive individual personas. As well as aspects of data-driven development, some subjective input was also used to generate realistic personas and achieve a reasonable balance of types and demographics. The derived set of personas can be found at https://doi.org/10.21954/ou.rd.7700579.

Discussion and Next Steps

The survey confirmed the assumption that many IT users experience IO and have diverse relationships with online services with varying levels of trust in their integrity and competence. The personas developed using the survey data will now be used to inform the design of the agent test platform, and to act as a basis for subsequent evaluation.

Acknowledgments

Many thanks to my supervisors at KMi, Paul Mulholland and Lara Piccolo, whose input has been essential to producing this work.

References

Benford, S., J. Bowers, L. E. Fahlén, and C. Greenhalgh (1994): 'Managing Mutual Awareness in Collaborative Virtual Environments'. *Proceedings of the Conference on Virtual Reality Software And Technology (VRST'94)*, no. August 1994, pp. 223–236.

Berners-Lee, T. and M. Fischetti (1999): *Weaving the Web: the past, present and future of the World Wide Web by its inventor*. New York: Harper Collins.

Cooper, A. (1998): *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*. Indianapolos, Indiana: Sams.

Ferreira, B. M., W. Silva, S. D. Barbosa, and T. Conte (2018): 'Technique for representing requirements using personas: a controlled experiment'. *IET Software*, vol. 12, no. 3, pp. 280–290.

McGinn, J. and N. Kotamraju (2008): 'Data-driven persona development'. *CHI'08*, pp. 1521.

Metaxas, G. and P. Markopoulos (2008): ''Aware of what?' A formal model of awareness systems that extends the focus-nimbus model'. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4940 LNCS, pp. 429–446.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007): 'A Design Science Research Methodology for Information Systems Research'. *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77.

Rodriguez, M. G., K. P. Gummadi, and B. Schoelkopf (2014): 'Quantifying Information Overload in Social Media and its Impact on Social Contagions'. *Eighth International AAAI Conference on Weblogs and Social Media*.

Schmitt, J. B., C. A. Debbelt, and F. M. Schneider (2018): 'Too much information? Predictors of information overload in the context of online news exposure'. *Information Communication and Society*, vol. 21, no. 8, pp. 1151–1167.

Tu, N., X. Dong, P. Rau, and T. Zhang (2010): 'Using cluster analysis in Persona development'. *SCMIS 2010*, pp. 1–5.