Workspace Awareness without Overload: Contextual Filtering of Social Interactions

Adrien Joly [♠,♣], Pierre Maret [♦], Johann Daigremont [♠]
{ adrien.joly, johann.daigremont } @alcatel-lucent.com
pierre.maret@univ-st-etienne.fr

[♠] *Alcatel-Lucent Bell Labs France*
[♣] *Université de Lyon, LIRIS UMR 5205, INSA de Lyon*
[♦] *Université de Lyon, Laboratoire Hubert Curien, UMR CNRS 5516*

Abstract. At a time when social media revolutionize the exchange of information between users of the Web 2.0, many enterprises are still applying older collaboration paradigms. Following the explosion of blogs and social sharing services on the Internet, some companies are beginning to catch up with these trends, but not all employees are familiar with these new practices which require manual contribution, even though they became usual for Web 2.0 users. Inspired from previous works and currents trends in Social Networking and Context-Awareness, we propose the iterative development of an experimental framework for contextual sharing and filtering of social updates. The intension is to trigger more opportunities for communication in a collaborative environment, while preventing information overload. In this paper, we present some preliminary results on synthetizing contextual clouds reflecting current working activity of users, based on the extraction of descriptive tags from web documents that they are currently accessing.

Keywords. Enterprise knowledge, User activity, Context similarity, Web 2.0, Collective intelligence, Notifications, Tag clouds

1. Introduction

On the so-called Web 2.0, numerous Social Networking Sites (SNS) sush as Facebook[1] , Twitter[2] or even YouTube[3] have known a tremendous success as they brought new ways to communicate, interact, learn about people, and share content with them. But, more importantly, they can easily become addictive because they are a great way to expose one's identity through self-reflection and they provide a sense of awareness about people that matter.

Following this trend, some companies are beginning to catch up with the Social Networking paradigm by proposing platforms for blogging, video sharing and microblogging inside the enterprise. Besides improving informal social ties [20], transposing Awareness to the enterprise is very promising for promoting collaboration.

[1] http://www.facebook.com
[2] http://www.twitter.com
[3] http://www.youtube.com

Today, communication inside usual contemporary enterprises relies on intranets, multi-recipient emails and several collaborative spaces or tools for more specific purposes. Intranets are usually complex and overly formal information spaces that are managed and redacted by a transversal communication department close to the strategic layers of the enterprise. This implies that, instead of actually helping collaboration, the information published on intranets and newsletters is often extensive, general updates about the enterprise; thus employees have to find out other ways to exchange and orchestrate their work on a large scale. This often leads to the situation of employees who have no clue of what their next-office neighbor is working on, and mostly rely on occasional short breaks around the watercooler to be Aware of what is going on.

In [11], we identified that conversations on Social Networking Sites rely on contexts that are relevant (or inspiring) to both interacting parties (i.e. the author and the commenter). As the amount of digital information continually grows, filtering becomes a necessity for reducing productivity loss. In this paper, we propose a context-aware [4] mechanism to promote relevant communication (and thus, collaboration) opportunities to office workers. Before describing the proposed research in section 3, we present some previous works on the Ambient Awareness domain in section 2, then we will explain the research methodology we intend to apply in section 4, before concluding this paper with open issues in section 5.

2. Background

Ambient Awareness[4] supports informal and lightweight social communication, to help people maintain peripheral awareness of others. According to Liechti et al. [13], this wide research domain can be divided in four categories: group awareness, workspace awareness, contextual awareness and peripheral awareness. Several experiments have been carried out for collaborative bug tracking [7] and configuration management of collaborative software development projects [17,1], in order to promote communication, reduce overlapping work, and improve ability to detect and resolve conflicts. Dourish et al. [5] propose the shared feedback approach, consisting of tracking users' actions on documents, and allowing them to comment them, in order to provide useful context information and avoid work duplication. Erickson et al. [9] experimented the social proxy as an abstract representation of activity in an enterprise communication space: babble. In these semi-private asynchronous discussion spaces, users are aware of the level of participation of their colleagues on several ongoing entitled discussions, and can decide opportunistically to join one of those.

Damian et al. reported [3] that information overload must be avoided by filtering updates by relevance, in order to prevent users from ignoring too frequent updates potentially containing crucial information for the success of a project. Machine-interpretable contextual information can support awareness systems to identify unplanned interactions (e.g. such as changes in the team members involved, which are very frequent in collaborative development projects [3]), and thus dynamically adapt the dispatching and presentation (e.g. notification modalities) of awareness information to users according to their current context [13].

Other awareness efforts outside the scientific community include Social Networking Sites (SNS), web-based platforms that rely on manually shared content and

[4] http://www.awareness.id.tue.nl/

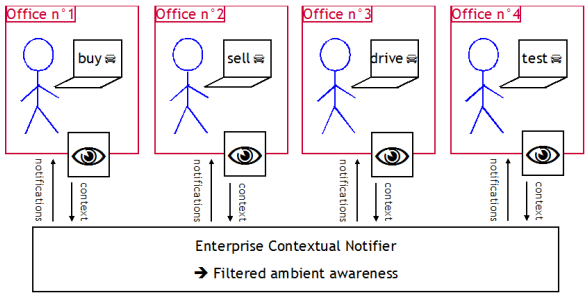

Figure 1: Awareness of co-workers' activities

status messages for triggering communication. Amongst the most popular informal SNS, Facebook[5] intends to help friends keeping in touch by providing a personal and dynamic news feed that contains updates about the content they shared and the comments they wrote, as further explained in [11]. Twitter[6] proposes a simpler approach: invite its users to spread 140 characters long textual messages to share concise updates about their activities, current interests and opinions, with their followers. Whereas Twitter is a good tool to improve visibility, Erickson et al. [8] deplore that is it too poor for proper awareness, as it does not provide a feeling of accountability to its users.

Similarly to Budzik et al.'s approach [2], we believe that opportunistic communication is key for awareness, and that the similarity of current working contexts (i.e. based on common concepts of documents/resources being currently manipulated) is a great opportunity to propose such communication. In this paper, we propose an extension of this approach that relies on description of resources instead of the resource itself, in order to widen recommendations of relevant contacts while improving privacy protection.

3. Digital resources as contextual cues, approach and scenario

In this study, we support the hypothesis that people carrying out activities in a given context are interested in resources and social updates from colleagues that are currently in a similar context. Indeed, when people lose visibility of their common context (e.g. when working silently on digital documents) and that they don't explicitly communicate about their activities, they miss collaboration opportunities.

By running a software agent watching users' activities (e.g. which resources/documents they are manipulating), as depicted on figure 1, their work context can be known. From there, users' actions can be selectively notified to other users which work context is similar, giving them opportunities for relevant communication. Indeed, common context between two people is an important factor to trigger interaction [19] (e.g. sign of respect, conversation, or exchange).

The first application of our research is to develop software to enable Ambient Awareness between employees, based on the description of documents that they consult and manipulate, or more generally their computer-based activities. This software aims to leverage numerous enterprise information sources (e.g. user directory,

[5] http://www.facebook.com
[6] http://www.twitter.com

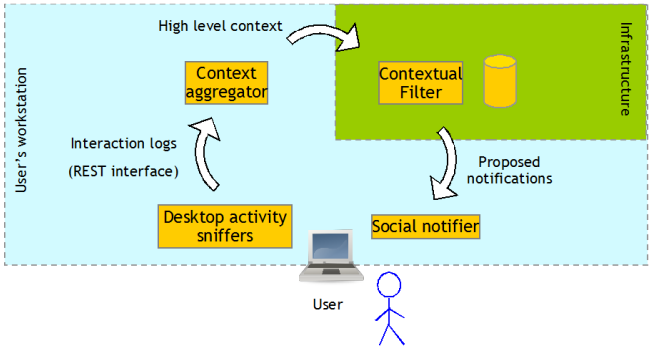

Figure 2. Modular overview of the prototype.

knowledge management/sharing platforms, shared calendars, document archives, digital communication channels) and workstation interaction logs (e.g. currently opened documents and windows, input patterns on mouse and keyboard, and actions on web applications) in order to consolidate a rich enterprise social network in which updates are spread through real-time notifications depending on relevance with the users' current context. Here, what we call context is a synthetic representation of a user's current activity and the description of this ongoing activity (e.g. the name of a document, of a project, a topic), visually summarized as a tag cloud (similarly to the Library Mirror [12]),.

To illustrate the workflow of envisioned Ambient Awareness system in the enterprise, we propose the following use case. Alice and Bob are employees of a same company, but they don't know each other. Bob is currently developing a prototype in Java, but he experiences difficulties to use a module M1 that was developed internally by team T1. The system recognizes this context and proposes him to make it visible to other employees, he accepts. Alice is one of the developers which were involved in the development of this module M1. As she is currently developing in Java, like Bob, she sees a notification saying that Bob is experiencing difficulties while trying to integrate module M1 in his prototype. She decides to help him out by initiating a voice call. In this use case, the contextual notification system proposed a communication opportunity which was relevant to both users, thanks to the knowledge of their common context. Without this system, Bob would have probably lost too much time (at least compared to the short time Alice took to solve the problem, thanks to her experience on the subject), and Alice would have lost an opportunity to get to know Bob and feel helpful to her colleague.

4. Contextual Notification Framework

In order to prove that the similarity between digital resources being manipulated by two users gives a relevant opportunity for communication, we are developing a framework for contextual filtering of social notifications. As depicted on figure 2, our current prototype is composed of four modules:

- Desktop Activity sniffers track the content being produced, manipulated and consulted by the user, and actions undertaken by this user;

Interaction/action	Intended context provider
Reading a web page about a specific subject	Firefox Extension: tabs watcher
Browsing a web site that has a specific purpose (e.g. giving information about a subject, providing a service)	Firefox Extension: tabs watcher
Posting/replying on a microblogging platform	HTTP request analysis (through Firefox Extension or proxy)
	RSS feed of the microblogging platform
Posting/sharing content on a site (KM or social platform)	HTTP request analysis (through Firefox Extension or proxy)
	RSS feed of the site
Annotating content on a site (rating or commenting)	HTTP request analysis (through Firefox Extension or proxy)
	RSS feed of the site
Attending a project event/meeting with colleagues	Personal Information Manager (e.g. Outlook, other calendar)
Writing a document for a project	Windows tracking + extract document name from title
Chatting with colleagues about a given subject	Jabber/XMPP server to extend and deploy
Logging on workstation / Resuming from screensaver	OS-specific event listener (software daemon)
Logging off workstation / Starting screensaver	OS-specific event listener (software daemon)

Table 1. Intended context providers.

- a local context aggregator gathers interaction logs from the sniffers to synthetize the current user's context as a contextual clouds (inspired from tag clouds, aka folksonomies);
- a centralized contextual filter evaluates the similarity of current contextual clouds of every user to filter notifications and propose contact when relevant;
- and a social notifier displays these notifications and contact proposals on the users' terminal.

Similarly to the TaskTracer project [6], activity sniffers can be implemented as plug-ins for resourceful applications (e.g. document editors), as daemons hooking on interaction events of the wokstation's operating system and/or applications which expose an API (i.e. to figure out which files are currently being manipulated), or even as proxies (e.g. for HTTP, email traffic). The point of activity sniffers is to extract meaningful cues that can help to synthetize the context, e.g. known symbols from file/directory names, metadata. In table 1, we propose a list of sniffers that could track interesting actions and context information. Additional sniffers could also leverage static information from other databases such as user profiles, project team lists (e.g. people who subscribed to the mailing list of a same project), in order to infer more contextual information around the people involved directly or indirectly with user's current activities and content.

For customization and privacy protection purposes, the context aggregator must allow the user to consult, manipulate and annotate the synthetized contexts before confirming (or not) their transmission to the centralized contextual filter. This could be implemented as on-screen notifications that would pop-up when the current contextual cloud had significant changes, waiting for user feedback before transmission or disappearing after a while.

The contextual filter dispatches the activities (updates) of a user to another user if their current contextual clouds are similar, as a measurement of relevance. This similarity function brings several interesting research questions about the knowledge representation model and algorithm to use, and the temporal dynamics of the contextual cloud (e.g. how to handle topic sustain and deprecation).

On the social notifier side, the notifications must be concise, and might use illustrations to give a quick clue of its topic without being too intrusive on the user's

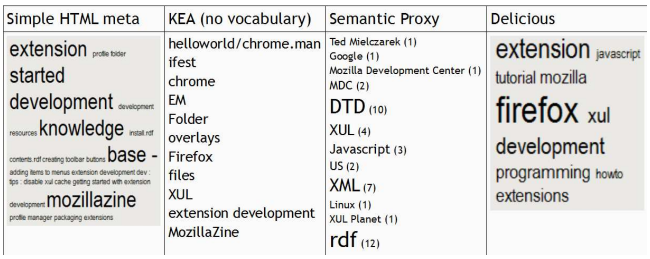

Figure 3. Tag clouds.

screen. This notification must also provide browseable links (i.a. like on web pages) so that the user can consult more information about the entities and symbols involved (e.g. the author, the topics), annotate it, and communicate with the author.

5. First results with context sniffers

As a first class of documents to analyze for context synthesis, we chose to focus on web browsing. Each opened web page is considered as a document that is manipulated by the user, assuming that this web page is relevant with the user's current activity. In this section, we will present our first experimental results of document analysis towards contextual cloud synthesis.

As a first experiment with web pages, we developed an extension for Mozilla Firefox[7] that tracks when web pages are being opened and closed, and logs these events and the corresponding URLs to the local context aggregator. For every update, the context aggregator updates the user's current contextual cloud with four different enhancers which turn URLs into tag clouds (i.e. a set of weighted keywords):

- a simple HTML metadata reader that extracts keywords from the title, description and keywords embedded in the page's HTML elements;
- a KEA-based (Keyphrase Extraction Algorithm)[8] keyword extractor that
- can summarize natural language content into a list of keyphrases, after a training phase;
- a semantic keyword extractor based on the OpenCalais Semantic Proxy web service[9], returning the textual representation of semantic instances that were recognized in the page;
- and a last one that returns the tags that del.icio.us[10] users manually entered to describe that page, the weight of each tag being the number of users who used this tag to describe the page.

In order to evaluate the quality of these enhancers for synthetizing a resembling and consistent contextual cloud based on the web pages that the user is browsing, we generated a contextual cloud with each enhancer at a time on a same page and

[7] http://www.mozilla.com/firefox/
[8] http://www.nzdl.org/Kea/
[9] http://semanticproxy.com/
[10] http://www.delicious.com/, a web-based social bookmarking service

compared the results. We ran this experiment on a page entitled "Getting started with extension development"[11], which is about Mozilla Firefox extension programming.

From the resulting contextual tags depicted on figure 5, we can draw the following conclusions:

- The KEA-based enhancer returned several relevant keywords (e.g. firefox, xul), including higher-level concepts (e.g. extension development), but also too many inconsistent keywords (e.g. ifest, EM, files).
- The Semantic Proxy-based enhancer picks up well-recognized knowledge such as standards. Indeed, RDF, DTD, XML and XUL technologies are very relevant keywords for the given page that are correctly emphasized by their high weight (based on the number of occurrences in the content). Lower weighed keywords are also relevant with the content of the page, but they do not really reflect its topic.
- The Simple HTML meta enhancer mostly emphasized keywords that were found in the title of the page. In our case, the page is quite well described by its title, that is why the results are satisfying. Nevertheless, this is not the case on all web pages.
- Finally, the del.icio.us-based enhancer returned fewer keywords but they better describe the topic of the page, and their weights are more heterogeneous, emphasizing the most descriptive keywords, such as: firefox, extension, development. Even lower weighted keywords are quite well describing the topic and the utility of the page, thanks to common sense emerging from the crowd del.icio.us' users: programming, tutorial, xul.

We conclude that the results from the del.icio.us-based enhancer are the most suitable for describing current user's context based on his/her web browsing session. Nevertheless, we must keep in mind that not every web page is tagged on this service. Therefore, we decided to combine the results of del.icio.us-based enhancers with the simple HTML meta enhancer, so that any web page can generate keywords that can be represented in the contextual cloud.

We would like to insist on the fact that these experimental results are preliminary and provided here as a first proof that contextual clouds can effectively be generated from a list of currently browsed web pages, enabling our future experiments of contextual cloud similarity.

6. Conclusion

In this paper, we proposed a Contextual Notification Framework to promote relevant communication and collaboration opportunities to enterprise workers, based on work context similarity. We compared our first results of context synthesis from user-browsed web pages using four different modules.

In the next steps of our research, we intend to develop a similarity algorithm that could be applied to our contextual clouds for recommending relevant contacts and social updates, according to their current context. For the iterative development and

[11] http://kb.mozillazine.org/, "Getting started with extension development"

experimentation of this platform, we will firstly rely on the enterprise environment, and then generalize our approach to mobile terminals with physical context sensors, also leveraging information from social networking sites.

7. References

[1] J. T. Biehl, M. Czerwinski, G. Smith, and G. G. Robertson. Fastdash: a visual dashboard for fostering awareness in software teams. In CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 1313–1322, New York, NY, USA, 2007. ACM.

[2] J. Budzik, X. Fu, and K. Hammond. Facilitating opportunistic communication by tracking the documents people use. In CSCW 2000 Workshop on Awareness and the WWW. Retrieved January, volume 30, page 2005, 2000.

[3] D. Damian, L. Izquierdo, J. Singer, and I. Kwan. Awareness in the wild: Why communication breakdowns occur. Global Software Engineering, 2007. ICGSE 2007. Second IEEE International Conference on, pages 81–90, Aug. 2007.

[4] A. K. Dey. Understanding and using context. Personal and Ubiquitous Computing, 5(1):4–7, 2001.

[5] P. Dourish and V. Bellotti. Awareness and coordination in shared workspaces. In CSCW'92: Proceedings of the 1992 ACM conference on Computer-supported cooperative work, pages 107–114, New York, NY, USA, 1992. ACM.

[6] A. N. Dragunov, T. G. Dietterich, K. Johnsrude, M. McLaughlin, L. Li, and J. L. Herlocker. Tasktracer: a desktop environment to support multi-tasking knowledge workers. In IUI'05: Proceedings of the 10th international conference on Intelligent user interfaces, pages 75–82, New York, NY, USA, 2005. ACM.

[7] J. B. Ellis, S. Wahid, C. Danis, and W. A. Kellogg. Task and social visualization in software development: evaluation of a prototype. In CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 577–586, New York, NY, USA, 2007. ACM.

[8] I. Erickson. The translucence of twitter. In EPIC 2008, Ethnographic Praxis in Industry Conference, pages 58–72, 2008.

[9] T. Erickson and W. A. Kellogg. Social translucence: an approach to designing systems that support social processes. ACM Trans. Comput.-Hum. Interact., 7(1):59–83, 2000.

[11] A. Joly, P. Maret, and J. Daigremont. Context-awareness, the missing block of social networking. International Journal of Computer Science and Applications, 4(2), February 2009.

[12] M. Koch. Supporting Community Awareness with Public Shared Displays. In Proc. Bled eCommerce Conference. 2005. Bled, Slowenia, 2005.

[13] O. Liechti. Awareness and the www: An overview. ACM SIGGROUP Bulletin, 21:3–12, 2000.

[16] C. Plaue, T. Miller, and J. T. Stasko. Is a picture worth a thousand words? an evaluation of information awareness displays. In Graphics Interface, pages 117–126, 2004.

[17] A. Sarma, D. Redmiles, and A. van der Hoek. Empirical evidence of the benefits of workspace awareness in software configuration management. In SIGSOFT '08/FSE-16: Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of software engineering, pages 113–123, New York, NY, USA, 2008. ACM.

[19] Y. Vogiazou, M. Dzbor, J. Komzak, and M. Eisenstadt. Buddy space: Large scale presence for communities at work and play. In Proc. ECSCW03 Workshop, 2003.

[20] D. Vyas, M. van de Watering, A. Elins, and G. van der Veer. Universal Access in HumanComputer Interaction. Ambient Interaction, chapter Engineering Social Awareness in Work Environments. 2007.