Information Sciences 358–359 (2016) 1–17

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Assessing e-mail intent and tasks in e-mail messages

M. Sappelli[a,b,*], G. Pasi[c,*], S. Verberne[b], M. de Boer[a,b], W. Kraaij[a,b]

[a] TNO, Anna van Buerenplein 1, Den Haag, The Netherlands
[b] Radboud University, Toernooiveld 300, Nijmegen, The Netherlands
[c] Universit di Milano Bicocca, Italy

ARTICLE INFO

Article history:
Received 28 July 2015
Revised 4 February 2016
Accepted 4 March 2016
Available online 12 March 2016

Keywords:
E-mail intent
E-mail annotation scheme
Task-based e-mail categorization
Human annotation

ABSTRACT

In this paper, we analyze corporate e-mail messages as a medium to convey work tasks. Research indicates that categorization of e-mail could alleviate the common problem of information overload. Although e-mail clients provide possibilities of e-mail categorization, not many users spend effort on proper e-mail management. Since e-mail clients are often used for task management, we argue that intent- and task-based categorizations might be what is missing from current systems.

We propose a taxonomy of tasks that are expressed through e-mail messages. With this taxonomy, we manually annotated two e-mail datasets (Enron and Avocado), and evaluated the validity of the dimensions in the taxonomy. Furthermore, we investigated the potential for automatic e-mail classification in a machine learning experiment.

We found that approximately half of the corporate e-mail messages contain at least one task, mostly informational or procedural in nature. We show that automatic detection of the number of tasks in an e-mail message is possible with 71% accuracy. One important finding is that it is possible to use the e-mails from one company to train a classifier to classify e-mails from another company. Detecting how many tasks a message contains, whether a reply is expected, or what the spatial and time sensitivity of such a task is, can help in providing a more detailed priority estimation of the message for the recipient. Such a priority-based categorization can support knowledge workers in their battle against e-mail overload.

1. Introduction

In the project SWELL[1] we aim to develop ICT applications that minimize the risk of burn-out and improve the well-being of employees. A large source of stress at work originates from information overload, and more specifically e-mail overload [2,13]. Whittaker and Sidner [34] believe that this is caused by the misuse of the original purpose of the e-mail system. The authors state that although e-mail was originally developed for the purpose of asynchronous communication, it is currently being used for task management, scheduling and personal archiving as well. This causes cluttered in-boxes and information getting lost in archives.

* Corresponding author at: TNO, Anna van Buerenplein 1, Den Haag, The Netherlands.
E-mail addresses: maya.sappelli@tno.nl (M. Sappelli), pasi@disco.unimib.it (G. Pasi), s.verberne@cs.ru.nl (S. Verberne), maaike.deboer@tno.nl (M. de Boer), kraaijw@acm.org (W. Kraaij).

[1] http://www.swell-project.net

http://dx.doi.org/10.1016/j.ins.2016.03.002
0020-0255/

Attempts to improve the organization of inboxes include the automatic detection of spam [28], message categorization [3,8,15,22,27,30] and priority estimation [1,11,29]. Complete agents exist that help the user file messages into folders [31]. However, not many of these automated techniques are adopted in current systems and many users do not even use category folders at all [15,22]. The most likely actions users make are splitting personal and work-related e-mail by using separate mailboxes [7] and cleaning e-mails at the end of the day [19]. In general, not many users spend effort on general e-mail management (deleting, moving, flagging) [17]. Nevertheless, research indicates that proper categorizations could alleviate the problem of feeling overloaded [2,4,34]. The fact that categorizations are not used suggests that there may not be a full understanding of what type of categorization is needed to properly support users in the way they use e-mail.

Since e-mail clients are often used for task management [7,34], we believe that intent- and task-based categorizations might be what is missing from current systems. This paper studies the realization of tasks in e-mail messages to better understand what the intent is behind an e-mail. In order to do so we annotate e-mail messages with both their e-mail intent and task intent. By e-mail intent we mean the intent of the sender; why did a person send the message. In that case, the intent refers to a message as a whole. Then, within a message the sender has (either implicitly or explicitly) possibly specified one or more tasks to be undertaken by the receiver. This latter aspect is referred to as the task(s) in the message. In this paper we investigate both the intent of the message and the tasks that are conveyed in the message. Additionally, we investigate how an e-mail conversation between two individuals evolves over time.

This paper makes four contributions compared to the previous literature: First, a taxonomy of tasks that can be found in e-mail messages is proposed. Second, a manual annotation of two datasets is provided; these annotations will be available to the research community to provide new opportunities for the development of (automated) e-mail support systems. Third, we present an initial analysis of how senders convey tasks in e-mail messages. And finally we present some initial results on automated classification for one of the dimensions. We address the following research questions:

1. To what extent do corporate e-mail messages contain tasks?
2. What are the characteristics of tasks in e-mail messages?
3. How do work-related e-mail conversations evolve?
4. Can we determine the number of tasks in a message automatically?

We start this paper with an overview of literature on the analysis of e-mail message content. Then we present the results of a pilot study where we developed our e-mail classification scheme in Section 3. In this study we determine which dimensions of content analysis are reliable for annotation. Additionally, we assess the validity of using annotations by independent assessors. In Section 4, we present the results of a larger-scale annotation study, where we annotate messages from the Enron and Avocado datasets. These datasets originate from a company setting and are likely to be representative of how tasks are conveyed in a work environment. In Section 5 we demonstrate the possibility to use the collected data for automated classification.

2. Background literature

In this section we shortly report the literature related to the analysis of e-mail message content. A limitation of the research that addresses the analysis of e-mail messages is the limited availability of public datasets of e-mail messages. The most used publicly available dataset of e-mail messages is the Enron dataset. This is a set of messages that was made public during a legal investigation of the Enron company [20]. It contains over 200,000 messages. Many researchers, however, make use of their own privately collected sets of e-mail messages [1,10,11,21].

Some research into the content of e-mail messages has been directed at the timing of communication. In an interview study, Tyler and Tang [32] investigate the concept of the *responsiveness image* of a person in order to understand what information is conveyed by the timing of email responses. They distinguish *response expectation* (the implicit time the sender gives to the recipient to respond) from *breakdown perception* (the initiation of a follow-up action that occurs when the response expectation time has ended).

This responsiveness image could be seen as a request for attention. Hanrahan et al. [17] analyze responsiveness in a 2-week study by logging user interactions with e-mail and compared these interactions to diary entries of the participants. The authors propose that e-mails can be categorized into 4 groups of requests for attention: ignore, accountable non-answer (engage with message but do not reply), postponed reply and immediate reply. This categorization provides insight in both the timing as well as the type of response that is expected.

Kooti et al. [21] add that the request of attention is not solely based on the contents of a message. They note that there is an effect of load on the replying behavior of people. As users receive more e-mail messages in a day, they will reply to a smaller fraction of messages.

Besides the research centered on replying behavior, another line of research that addresses the issue of identifying e-mail intent is finalized at an analysis of the content of messages. Gains [12] focuses on the language that is used in messages, in particular on aspects related to the pattern and style of a text. He has analyzed messages in a commercial and in an academic setting and found that commercial e-mail messages tend to follow standard written business English, while messages in an academic setting follow a more pseudo-conversational pattern where for example the salutation is absent.

To analyze the message style, Gains [12] uses a classification scheme from business communication described by Ghadessy and Webster [14]. The authors state that there are roughly three types of business communication: informative