

An Integrated Agent System for E-Mail Coordination using Jade

K. P. Thooyamani^{1*}, V. Khanaa² and R. Udayakumar³

¹Professor of School of Computing, Bharath University, Chennai-600073; thooyamani@hotmail.com

²Dean, Information Technology, Bharath University, Chennai-600073 India; drvkannan62@yahoo.com

³Associate Professor, Information Technology, Bharath University, Chennai-600073; rsukumar2007@hotmail.com

Abstract

With growing usage of computer applications there is growing dependence of business over computer aided services like e-mails. The CRM or E-Mail based marketing companies are becoming pervasive. There is always a greater need to prioritize the e-mails which has direct impact on the ROI - Return on Investment to run these companies in a way that to customers satisfied. In this paper, a method is proposed to prioritize the unread E-Mails based on users 'interest and priority. This is implemented in this system using agent technology. The Agent mechanism is guided using JADE Middleware and it's underlying architecture. The Intelligent agent used here can learn the interest and priority of the user and use this knowledge to rearrange the unread mails. The Text corpus in the mail would undergo tokenization and each token would be matched with knowledgebase by the agent. The 90 percent precision of the prioritization. Thereby, the E-Mails are ranked according to interest and priority of the user.

Keywords: Intelligent Agent, Tokenization, E-Commerce, Artificial Intelligence.

1. Introduction

E-mail is the world's most preferred form of communication. Billions of E-mails traverse the globe daily. A person, who E-mails to inwards and outwards, would daily be dealing with a number of E-mails. It becomes a difficult task for the user to e-mails. The wise email client would sort of the in the user's inbox. This can be achieved using an Intelligent Agent. The Intelligent Agent will be the responsible to rank each E-Mail according to its, and display the ranks in the user's inbox.

2. Related Work

Email marketing is a form of direct marketing using email as a means of communicating commercial or fund-raising

messages to an audience. In its broadest sense, every email sent to a potential or current customer could be considered email marketing. However, the term refers to:

- (1) Sending email messages with the purpose of enhancing the relationship of a merchant with its current or previous customers, to encourage customer loyalty and repeat business.
- (2) Sending email messages with the purpose of acquiring new customers or convincing current customers to purchase something immediately.
- (3) Adding advertisements to email messages sent by other companies to their customers.
- (4) Sending email messages over the Internet, as exists outside the Internet e.g., network email.

**Corresponding author:*

K. P. Thooyamani (thooyamani@hotmail.com)

3. Challenges of Mail Prioritization

Personalized email prioritization (PEP) is an ordinal regression problem, which is different from conventional text classification where for each category; there are only two levels, true or false. Users rated their importance from one to five or in the alternative from not important at all to very important, resulted in ordinal regression problem. Given limited amount of time, users may want to selectively read important emails or take action to certain important messages levels. The personalized email prioritization entails two main research challenges: (I) the sparse training data and (II) one's own priority definition. First of all, unlike spam filtering, we could not share training data among different users because of privacy issues and different interests. People hesitate to share their very personal labeling information except spam emails. Even though there are users who are willing to share the very personal labeling information in practice it cannot be shared.

For instance, the importance of a grant proposal email could be extremely important to the principal investigator but it could be marginally important or not important to the person who's not actively working on the proposal. Second, one's own priority definition could lead to diverse way of defining priority. In that case, the assumption of the current state-of-the-art ordinal regression such as Support Vector Ordinal Regression (SVOR) [1] might not be good enough. For instance, regression-based approaches assume one weight vector to model all levels of email priorities from the lowest priority level to the highest priority level, resulted in all parallel decision boundaries. Since the email text is very high dimensional space, it is not easy to visualize and check whether regression-based approach assumption will be applicable or not [2–4]. Therefore, we have to adopt forms of empirical evaluation to conform the best approaches. There are some Customer Relationship Management Tools that come with integrated support for emails. They may discover crucial information for decision making by tracking the e-mail conversations with customers. They also enable Centralization of all customer information and email correspondence inside CRM Association of customer's email conversation with CRM data. For driving such business deals it is mandatory that the emails are properly prioritized according to the purpose and should these preferences carry out the workflow. This type of functionality is used the Integrated Agent System for Email Coordination (IASEC) integrated with the existing mail system [5–8].

4. System Architecture & Related Technology

The Figure 1 shows that the 2-tier Architecture of an Email Prioritizing Agent in Jade. The architecture is explained in detail as below. The Agent works in association with the email - client which may in turn be associated to any Mail Server. The Agent takes up the e-mail morphology and tokenizes the entire content. The logic for the inference mechanism is stored in the inference engine which works in association with the Knowledge Base, which stores the list of words along with its weights, which are directly proportional to the interests of the user. The weight of the words in the database is updated by a web spider based upon the interests and activities of the user. The spider accomplishes the above task by visiting the web pages and parsing it into words following which it performs morphological analysis to update the weight of the words present in the database. It also takes into account the various factors such as the sender of the E-mail, the periodicity of the mail and the users' interests and the user's web browser history. The inference Engine would resolve the degree of importance of the mails in the inbox working with the knowledge base. This is summarized as a prioritized list of content on the email client consequently. The Agent Technology is a rapidly developing area of research. Agent Technology is to

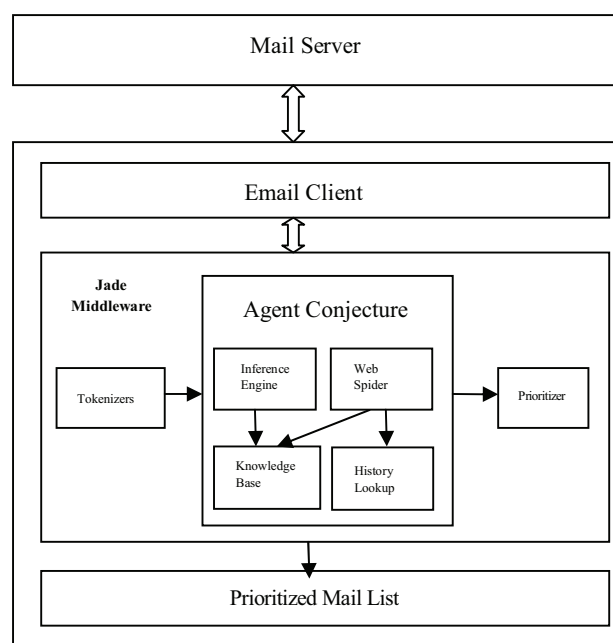


Figure 1. IASEC Architecture.

shortly commence the development phase of its multi-protocol network. Intelligent Agent is an entity which observes and manipulates an environment (E-mail client in our case) and directs its activity toward achieving its goal. Intelligent Agent will learn and use its knowledge to achieve its goal. An agent program can be defined briefly as a mathematical function of an agent, which maps all the possible sequence of percepts to every action that the agent can perform either to a coefficient or feedback elements or even to a function or constant all of which affects eventual actions.

5. Prioritized E-mail Coordination

The powerful basis for understanding how agent technology can help people work together more effectively will result from a better understanding of the nature of coordination. The coordination of distributed processing is of great interest to a number of research communities. In particular, Coordination technology is a way of designing and studying innovative computer systems that help people work together in small or large groups. The coordination process must organize the flow of information in such a way that communication between the managers of an agent proves to be most effective. JADF (Java Agent Development Framework) and it is referred in Figure 2. [9–10] is a software Framework fully implemented in Java language. It is a middleware developed by TILAB that simplifies the development of applications. .

Framework
The Inference process
Mail Analysis
Tokenization
The Inference process

For the inference process, a database is maintained containing a list of words, which are directly proportional to the interests of the user. During inception, the database contains the words considered universal for any user. The content of the knowledge base are updated by a web spider

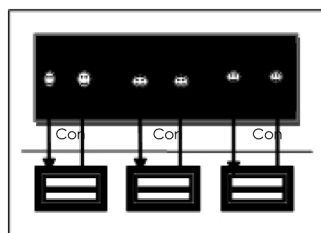


Figure 2. JADE Architecture.

(Bot) based upon the interests and activities of the user. Mail Analysis In order to declare that a particular E-Mail.

Important to the user the agent considers some of the factors such as the sender of the E-mail, the periodicity of the mail and the users' interests. The sender and the periodicity can be directly obtained from the mail, whereas the interest of the user can be obtained by studying the user's web browser history.

6. Internal Functionality of the Agent

In this section the internal working of the IASEC is explained in detail. The system retrieves email from mailbox and prioritizes them using an agent along with the Bots, which are programs capable of performing an automated task over the internet, which in turn uses Web Spider, a kind of Bots which fetches and analyzes the information available in the web pages.

The agent analyzes the sender and periodicity of the mail in the following way:

6.1 Sender

For a given mail the agent checks if the sender of the E-Mail is in the contact list of the user. If the sender is found in the contact list, then preference is given to an E-mail obtained from a known source over an E-Mail from an unknown source. Moreover if the user gets an E-Mail from a person who is in the contact list and is an active member (friend or family, Member) then, the corresponding E-Mail is provided more preference by the Agent. That's referred in Table 1 and Figure 3.

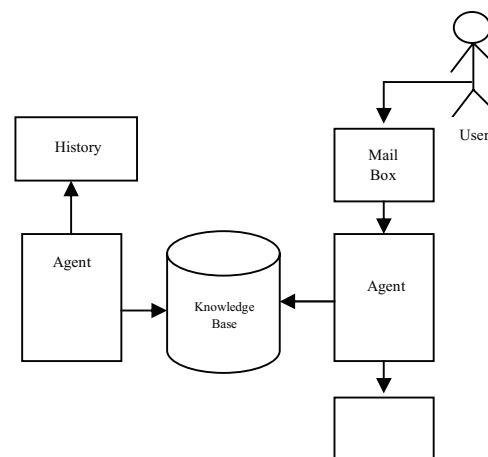


Figure 3. Evaluation of metrics is cited.

Table 1. Evaluation of metrics is cited

Evaluation	Education Sports	Job	Metric
Precision	0.5	0.5	0.5
Recall	0.6	0.8	0.4
F-Measure	17.14	17.17	0.16
Accuracy	0.5	0.5	0.5
Error Rate	0.5	0.5	0.5

6.2 Periodicity

The agent analyzes the periodicity of the sender which means that the agent will check if the user gets an E-Mail from a particular contact on a fixed interval of time (as in the case of newsletters). If the agent finds such a kind of mail, the agent awards less priority to that particular E-mail.

7. Conclusion

An approach to arrange E-mails according to the priority of the user using an intelligent agent has been presented in this paper. The Intelligent agent can be enhanced with consideration of the attached content also in email by the inference engine.

8. Acknowledgement

The support, encouragement and the guidance received by the author from Dr. K. L. Shunmuganathan, his research supervisor at Bharath University has been very motivating. The author is thankful to his students Mr. John Royceton,

Mr. Jobin George and Mr.Vinothkumar of RGCE for their support.

9. References

1. Shen Y, and Jiang J (2003). Improving the performance of Naive Bayes for text classification, CS224N Spring Report.
2. Koller D, and Sahami M (1997). Hierarchically classifying documents using few words, Machine Learning International Workshop, Technical Report.
3. McCallum A, and Nigam K (1998). Comparison of event classification for event models for text classification, AAAI-98 Workshop on Learning for Text Categorization.
4. Woolridge M, and Jennings N R (1995). Intelligent agents: theory and practice, Knowledge Engineering Review, vol 10, No. 2.
5. Sahami M, Dumais S et al. (1998). A bayesian approach to filtering junk E-mail, AAAI-98 Workshop on Learning for Text Categorization.
6. Russell S J, and Norvig P (2003). Artificial intelligence: a modern approach, 2nd Edn., Prentice Hall, 375–398.
7. Grefensette G, and Tapaninian P (1994). What is a word, what is a Sentence? Problem of tokenization, Rank Xerox Research Center, Grenoble Laboratory, France.
8. Shen W, Hao Q et al. (2006). Applications of agent based systems in intelligent manufacturing: an overview, Advanced Engineering Informatics, vol 20(4), 415–431.
9. Caire G, Lavarone G et al. (2011). Jade Tutorial, Jade Programming for Android, Available From: <http://jade.tilab.com>
10. Leonardo (2001). Intermedia in Electronic Images, MIT Press Journal, vol 34(1), 49–54.