

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/301404126>

Intelligent Text Mining Based Financial Risk Early Warning System

Conference Paper · April 2015

DOI: 10.1109/ICISCE.2015.68

CITATIONS

2

READS

106

9 authors, including:



Qikun Wu

University of Pittsburgh

1 PUBLICATION 2 CITATIONS

[SEE PROFILE](#)



Hongyu Mao

The Hong Kong University of Science and Technology

2 PUBLICATIONS 5 CITATIONS

[SEE PROFILE](#)



Mubing Zhou

Southern University of Science and Technology

2 PUBLICATIONS 3 CITATIONS

[SEE PROFILE](#)



Huaqing Wang

Southern University of Science and Technology

158 PUBLICATIONS 3,564 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Text Mining Technology for Financial Markets [View project](#)

Intelligent Text Mining Based Financial Risk Early Warning System

KQ Wang, QK Wu, HY Mao, MB Zhou, K Jiang, XP Zhu, L Yang, T Wang, HQ Wang

Dept. of Financial Mathematics and Financial Engineering
South University of Science and Technology of China
Shenzhen, China
wang.hq@sustc.edu.cn

Abstract—With increasing risks faced by financial firms and markets, text mining is playing more and more important role. However, because of the complexity between a text and the consequences caused by the text, it is hard for traditional text mining techniques to meet the needs. In order to solve this problem, we focus on the employment of intelligent agents to enhance the capability of text analysis. In our system, various classes of intelligent agents are proposed for financial risk early warning. The text mining agent gives influence measurement, the relationship agent generates relationship measurement between the text and the target, and the reliability agent measures whether the text is trustable or not. Collaboration within such agents is able to produce more effective and accurate early warning messages.

Keywords—Knowledge Engineering; Intelligent Systems; Text Mining

I. INTRODUCTION

Financial experts manage daily financial activities mainly by researching on the news, blogs, annual reports, research reports and other information sources. Normally, they focus on the quantitative expressions of such text information. However, with the rapid development of internet, global information makes financial market change dramatically time to time. It is hard for financial experts to acquire useful information quickly from such global information resources by their manual work. Existing text analyzing software basically cannot identify the relationships between important texts and financial tools.

Financial news and stocks research reports can rapidly reflect the investment tendency of investors and can influence them both directly and indirectly. Using this information wisely can provide us with a new perspective of understanding financial markets.

II. ARCHITECTURAL DEVELOPMENT

Informally, intelligent agents can be seen as software agents with intelligent behavior. Formally, intelligent agents are used to denote a software-based computer system that is capable of autonomous action in order to meet its design objectives [1]. Our financial information service system based on intelligent text mining technology is a multi-agent system,

which consists of a text collecting agent, a text mining agent, a text relationship analyzing agent, a reliability analyzing agent and an user interface agent. Our proposed system can complete the whole analyzing process within mini-second level, which is much fast than manual analyzing process.

A. Theoretical Foundation: Simon's Decision Making Model

According to the Nobel-prize winner professor H.A. Simon's [2] decision making process model, the decision making process can be divided into four stages as follows:

- 1) *Intelligence*: Collect and pre-process.
- 2) *Design*: Analysis of information.
- 3) *Choice*: Select the best scheme from all candidates.
- 4) *Review*: Evaluate solutions.

According to Simon's theory, the knowledge, information, experience and ability of decision makers are limited. It is hard to consider all the possible situations and then find the best solution. So from Simon's view, whether or not a decision is rational should be based on decision maker's psychological limitation, the knowledge limitation, the motivation limitation and the interaction limitation.

B. The Multi-agent Framework

In our multi-agent framework, each agent is independent and can solve their subtasks. Multi-agent can reduce the complex of both the system and problem solving process dramatically, making it to be organized. Our multi-agent system breaks the limitation of a single expert system. Using different agents for solving one problem greatly can enhance the capability of the system.

Each agent in our system provides different information service. They communicate with each other and transfer necessary information. And the running of our system is divided into several stages. In each stage, there will be some agents that receive tasks, react and solve the tasks cooperatively. Following steps are needed in a financial information service system: collection, pre-process, text analysis, and transferring. We introduce the function of each agent below:

- Text collection agent: Collect information from internet and preprocess them.

- Text mining Agent: Text mining.
- Relation analysis Agent: Analyze the relationship between the acquired text and objects.
- Reliability analysis Agent: Analyze the reliability of the information acquired.
- User Interface Agent: Give a final process to all the information and display them to users.

There are system layer, application layer and service layer in our system. Figure 1 is the SOA framework of our system , while Figure 2 is the data flow chat.

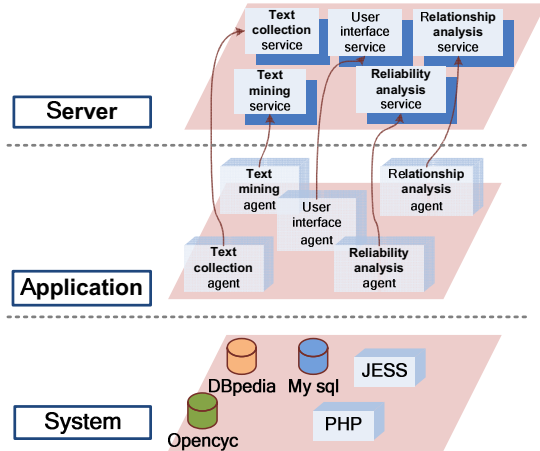


Fig. 1. SOA Architecture

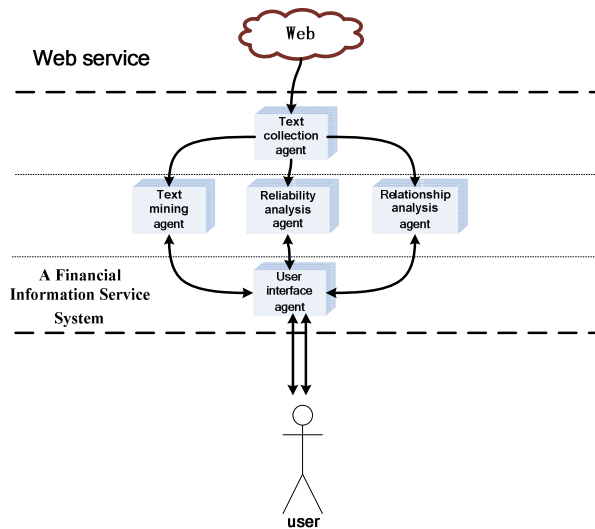


Fig. 2. System flow chart

In the text collection agent, a Web spider is used in this agent to search related information set by users and restore them from internet. Then a text filter is employed to filter repeated and useless text (which is judged by factors like time).

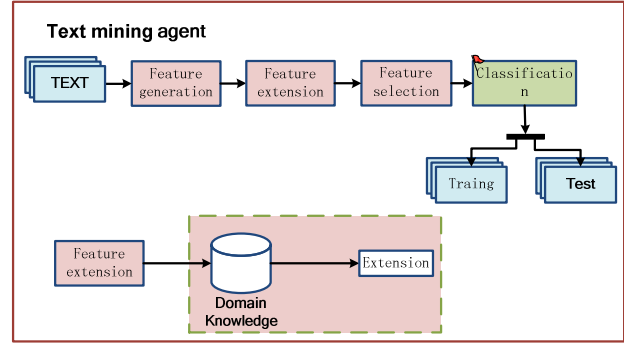


Fig. 3. Text mining agent

The architecture of the text mining agent is shown in the figure 3. The main difference between our text mining agent and traditional text mining is that there is an inference module in our agent. By adding the inference capability, this agent is able to produce better outputs, comparing with traditional text mining technologies.

The architecture of the relationship analysis agent is shown in the figure 4. It is able to discover the relationships between the text and the targets which the user sets up.

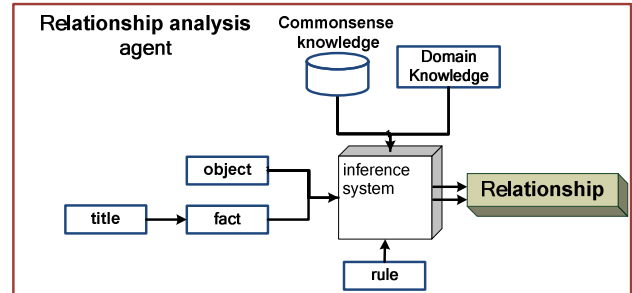


Fig. 4. Relationship analysis agent

In order to perform the relationship analyzing tasks, this agent needs know the world, i.e. it should have a huge intensive knowledge, including “common-sense knowledge”. This is one of the hardest problems that the artificial intelligence researchers face. In order to solve it, John McCarthy proposed a formal logic approach to program with common sense [3]. Marvin Minsky illustrated a more comprehensive approach to simulate common-sense reasoning [4]. In our system, the knowledge is divided into two categories: domain knowledge and common-sense knowledge. Two famous existing common-sense knowledge bases, OpenCyc [5] and DBpedia [6] have been used in our system. By using our own domain knowledge base, i.e. JESS (Java Expert System Shell) facts and rules [7], as well as OpenCyc and DBpedia commonsense knowledge bases, this agent is able to reason the connection between the text and the objects.

The architecture of the Reliability analysis agent is shown in the figure 5. This agent is developed to decide whether the text is reliable or not. As shown in the figure 5, the reliability is based on the reliability of the text source, i.e. the publisher,

the reliability of the author(s), and the times of cited. A set of JESS rules have been developed in this agent to perform such tasks.

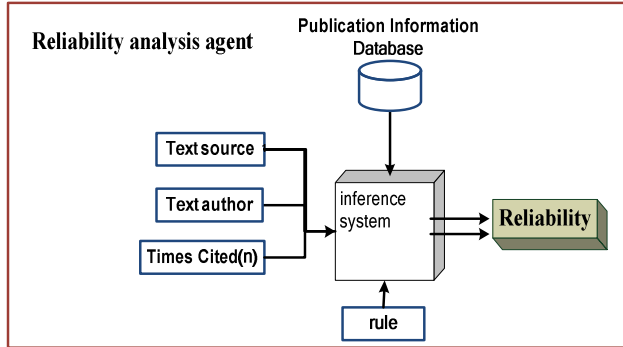


Fig. 5. Reliability analysis agent

The outputs from the text mining agent, the relationship analysis agent and the reliability analysis agent will go to the user interface agent for final processing. These outputs are in the following three categories:

- Influence measurement from the text mining agent: a score of the text, i.e. whether this text can have important influence.
- Relationship measurement from the relationship agent: a score to represent the relationship between the text and an object, which the user is interested.
- Reliability measurement: the score of the reliability of the text.

Based on these three scores, this agent will do the following tasks. Firstly, it will determine whether the result needs to be passed to the user or not. If needed, this agent will summarize generate a co-responding warning message to pass to the user.

III. SYSTEM OPERATION

A time sequence diagram of the system operation is shown in the Figure 6. The text collection agent collects data from Internet, based on the user's specification and send such data to Text mining agent, Reliability analysis agent and Relationship analysis agent for further processing. After receiving the analyzing results from these three agents, the User interface agent will summarize them and send final results to the users.

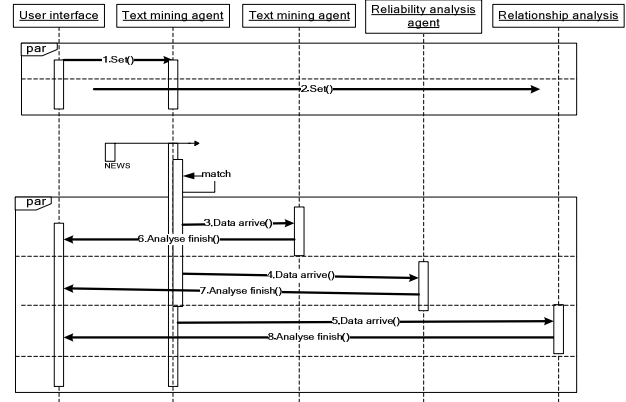


Fig. 6. Time sequence diagram

IV. CONCLUSIONS

In this paper, a novel Intelligent Text Mining Based Financial Risk Early Warning System has been presented. Novel multi-agents architecture has been designed. In our system, the text mining agent gives influence measurement, the relationship agent generates relationship measurement between the text and the target, and the reliability agent measures whether the text is trustable or not. Collaboration within such agents is able to produce more effective and accurate early warning messages.

V. ACKNOWLEDGMENT

This is a joint research project with Shenzhen PROLTO Supply Chain Management Co. Ltd. Many thanks to Mr. Suzhi Chen, the Chairman, and Mr. Sui Shi, the Chief Financial Officer from PROLTO.

REFERENCES

- [1] H. Wang, C. Wang, "Intelligent Agents in the Nuclear Industry," IEEE Computer, vol 30, pp. 28-34, 1997.
- [2] H.A. Simon, The new science of management decision. 1977.
- [3] J. McCarthy, "Some Expert Systems Need Common-sense," Annals of the New York Academy of Sciences, vol 426.1, pp.129-137, 1984.
- [4] M. Minsky, The Society of Mind. Simon & Schuster, 1988.
- [5] D.B. Lenat and R.V. Guha, Building large knowledge-based systems; representation and inference in the Cyc Project. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1989.
- [6] C. Becker and C. Bizer, DBpedia Mobile: A Location-Enabled Linked Data Browser. 1st International Workshop about Linked Data on the Web, Beijing, China, 2008.
- [7] E. Friedman-Hill, Jess in Action: Java Rule-Based Systems. Greenwich, CT: Manning, 2003.