Assentor®: an NLP-based Solution to E-mail Monitoring

Chinatsu Aone, Mila Ramos-Santacruz, William J. Niehaus

SRA International, Inc.
4300 Fair Lakes Court
Fairfax, VA 22033
{aonec, mila, niehaus}@verdi.sra.com

Abstract

This paper describes the Natural Language Processing (NLP) component of an e-mail monitoring product called Assentor®. Assentor monitors electronic correspondence for brokerage firms. It uses pattern-matching-based information extraction technology to find and quarantine e-mail messages that indicate, among others, customer complaints, insider trading, stock hyping, hard-pressure sales tactics, and firm preservation issues such as jokes and obscenities. This paper presents a quantitative evaluation of applying pattern matching vs. keyword-based searching to e-mail monitoring. Our evaluation shows that pattern matching performs significantly better than keyword-based searching both in terms of recall (false negatives) and precision (false positives).

Introduction

As e-mail replaces more traditional means of communication in the business arena, firms and institutions need ways to guarantee that their e-mail correspondence is in full compliance with laws and regulations. This is especially true of tightly regulated industries such as the securities industry, where traditionally *compliance officers* review hardcopy correspondence. The securities industry's regulatory bodies (the Securities and Exchange Commission, the National Association of Securities Dealers, and the New York Stock Exchange) have regulations that require a reasonable review of broker e-mail communications (i.e., SEC Rules 17a-3 and 17a-4, NASD Rule 3010 and NYSE Rule 342). The monitoring is designed to prevent violations such as stock hyping, insider trading, and hard-pressure sales tactics, as well as to properly handle customer complaints. Because of the difficulty of monitoring large amounts of e-mail messages, many securities firms have not fully implemented e-mail; others go through tedious manual review procedures to guard against illegal communications. However, manual review of e-mail messages is costly, time-consuming, inconsistent, and simply not feasible for firms with large volumes of e-mail correspondence. Most importantly, manual (pre-)review of e-mail messages defeats the purpose of e-mail, which is immediate delivery.

The problem of e-mail monitoring, therefore, calls for an automated solution. In this paper we describe an e-mail message screening system called Assentor®, which relies on Natural Language Processing (NLP) technology, in particular pattern-matching-based information extraction. The system has been tailored for securities and investment firms. It flags potentially improper communications, such as illegal stock hyping, high-pressure sales tactics, insider trading, and other potentially litigious issues, for human review. In this paper, we describe Assentor's screening capability and present a quantitative evaluation of the NLP-based vs. keyword-based approaches. This evaluation shows that Assentor's NLP-based approach performs significantly better than keyword searching both in terms of precision and recall.

Application Description

Assentor is an e-mail monitoring system tailored for brokerage and investment firms. Assentor sits inside the firm's firewall. As illustrated in Figure 1, the system monitors e-mail messages to and from brokers. The system administrator provides the list of users to monitor and sets the level of monitoring (threshold) for individual users. The system can be configured to apply different levels of monitoring based on seniority and past performance of individual brokers. For instance, senior or highly trusted brokers may be subject to less scrutiny than junior brokers. When Assentor flags a message, it quarantines it for human review. The system can be configured to either let a copy of the message go to its intended recipient or hold the message until a human reviewer approves or rejects it. A compliance officer reviews the message quarantined and can reject it, approve it, or send a warning to the sender.

Assentor integrates a number of Original Equipment Manufacturer (OEM) products. As outlined in the architecture diagram in Figure 2, messages are first decomposed (e.g., attachments are separated) using Integralis MAILsweeper™. Then, messages are checked for viruses using Command Software™ and converted to ASCII text (e.g., Microsoft Word to ASCII text) using INSO Outside In™.

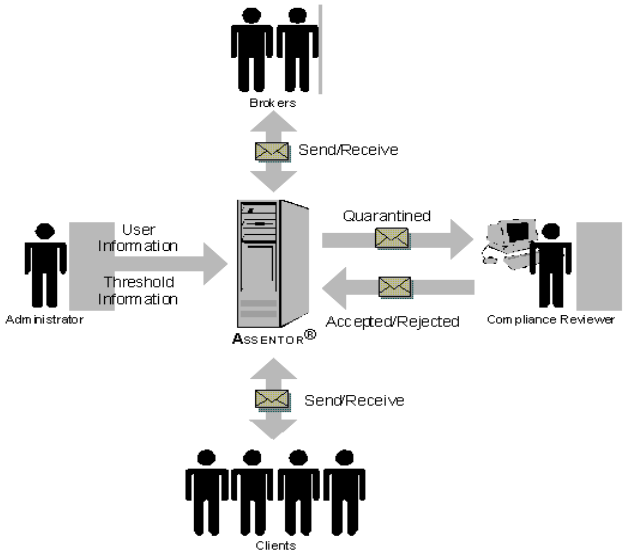

Figure 1: Assentor's Concept of Operation

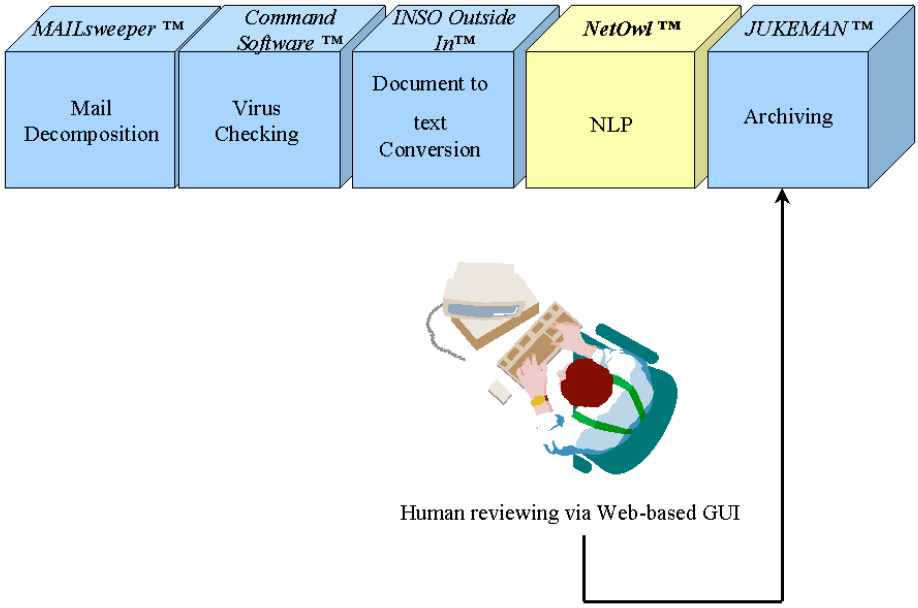

Figure 2: Assentor's Architecture Diagram

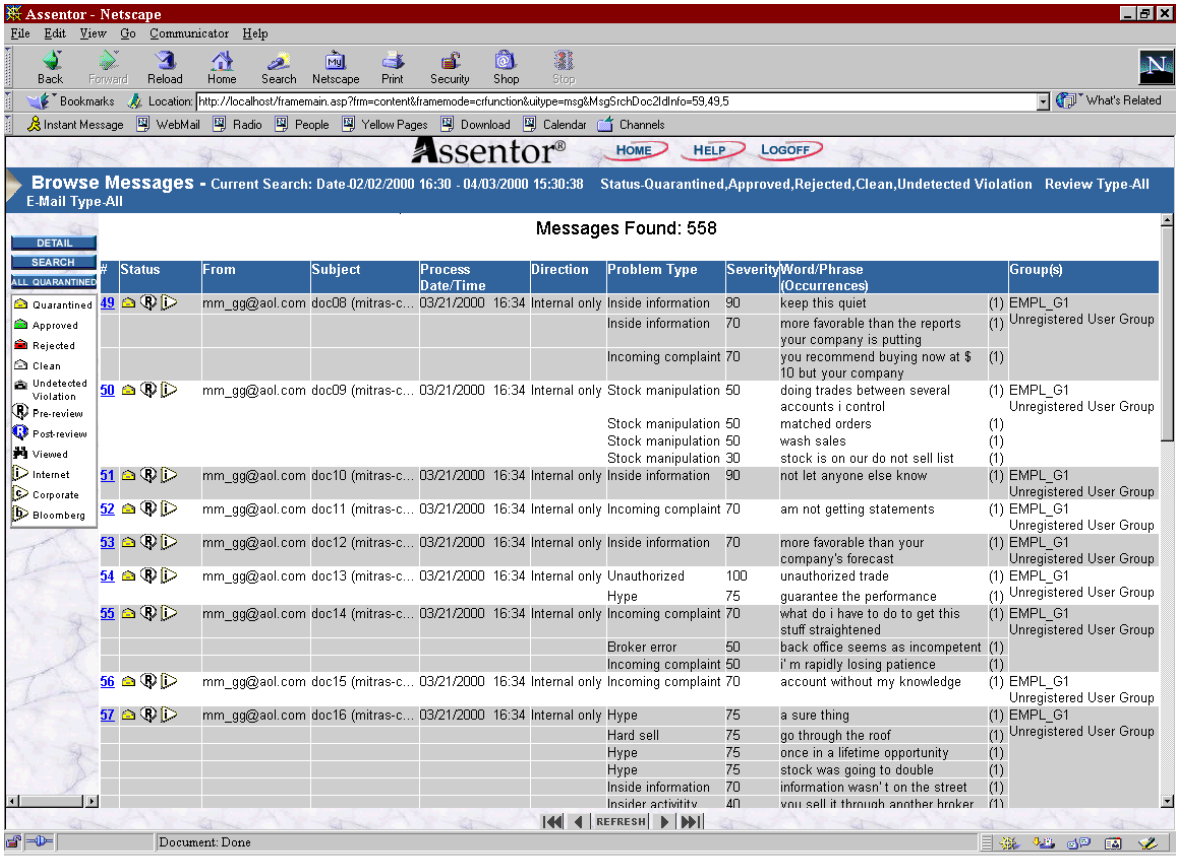

Figure 3: Assentor's "Browse Messages" Screen

At the heart of Assentor is SRA's information extraction engine NetOwl™, which flags potentially improper messages to be quarantined. A compliance reviewer reviews quarantined messages through Assentor's Web-based Graphical User Interface (GUI). Figure 3 shows a list of sample quarantined messages on Assentor's "Browse Messages" screen. Finally, Assentor archives all messages using JUKEMAN™ and generates e-mail management reports. Assentor runs on NT.

In this paper we will concentrate on Assentor's NLP component. Table 1 lists the problem categories tracked by Assentor and the e-mail direction (i.e., inbound or outbound) to which they apply. When Assentor finds an e-mail message containing one of these potential violations, the system quarantines the message.

Category	Direction
Incoming Complaint	In
Legal	In
Money Laundering	In
Stock Manipulation	In
Unauthorized	In
Hard Sell	Out
IPO	Out
Puffery	Out
Reply to Complaint	Out
Restricted Issue	Out
Broker Error	In/Out
Encryption Key	In/Out
Encryption Signature	In/Out
Encryption Message	In/Out
Hype	In/Out
Inappropriate	In/Out
Insider Activity	In/Out
Inside Information	In/Out
Profanity	In/Out
Rumor	In/Out
Terms to Monitor	In/Out

Table 1: Issues Monitored by Assentor

Each problem category (e.g., complaints, rumors, insider trading, etc.) is divided into three sublevels (i.e., high severity, medium severity, and low severity). Each sublevel contains a set of concepts to be monitored. The concepts are organized into these sublevels according to the severity of the content they convey and the ambiguity they may carry. For instance, the high level of the complaint category contains concepts such as complaints about big money losses or broker's recommendations, and requests that an account be terminated. The medium severity includes complaints about unexpected commissions and unsuitable investments. The low severity level is for concepts such as trouble contacting the broker, mild blaming for stock performance, and ambiguous indications of unhappiness by the customer (e.g., "*I'm disappointed*", which could occur in the context of a personal message). The Assentor system administrator at the firm can set up broker groups and turn sublevels on and off depending on the firm's policy for each group. The application can thus be configured to apply different levels of monitoring based on seniority and past performance of individual brokers.

Uses of AI Technology

Assentor relies on NLP technology, in particular information extraction. Information extraction (IE) is emerging as a new technology in commercial solutions. Commercial IE tools include SRA's NetOwl™, IBM's Intelligent Miner™ for Text (Dorre, Gerstl, and Seiffert 1999), and Inxight's Thing Finder™ (www.inxight.com/products/developer/ad_tf.html). To the best of our knowledge, Assentor is the first industry-wide solution that applies pattern-matching-based IE technology to e-mail monitoring and categorization.

Assentor integrates NetOwl and applies generic and domain-specific patterns for the securities industry compliance domain. The generic patterns recognize names of people, companies, locations, e-mail and http addresses, telephone numbers, monetary figures, dates, and time expressions. The domain-specific patterns that we developed specifically for Assentor recognize phrases and sentences relevant to the securities industry compliance domain. Domain knowledge is thus represented in patterns. We worked closely with compliance officers to knowledge engineer the patterns for NetOwl. NetOwl is written in C++ and runs on Windows NT and Sun Solaris.

Various approaches can be adopted for e-mail monitoring from simple keyword-based searching to machine-learning-based techniques. In order to find the most suitable AI technology for e-mail monitoring, we have also evaluated a supervised learning-based approach (nearest neighbor). Our experiments showed than the learning-based categorization approach does not perform well when there is not a large amount of training data available. In our experience only 2% of all messages merit flagging in this e-mail monitoring application for the securities industry. Thus, it is very difficult to collect enough training examples for categories that are rare, such as Stock Manipulation and Insider Trading.

By contrast, since humans can easily generalize a comparatively small set of examples, pattern matching is especially suitable for e-mail monitoring. Moreover, pattern matching has rendered the best results in information extraction, as shown in the series of Message Understanding Conferences (cf. Aone *et al.* 1998).

Some firms use keyword-based searching for e-mail monitoring. A common perception is that keywords flag as many *bad* messages as patterns (or even more), but produce more incorrect hits (*false positives*). To compare the performance of Assentor's NLP-based solution and keyword-based searching, we ran both Assentor and a list of keywords on two sets of unseen messages, or *blind sets*. The first set contains 2000 incoming messages, 163 of which are *bad* messages. The second set contains 1800 outgoing messages, 129 of which are *bad* messages. All our training and blind sets are real-world messages provided by brokerage firms. Our list of keywords was the combination of 3 keyword lists we obtained from 3 brokerage firms (a total of 300 unique keywords).

Our evaluation showed that Assentor's NLP-based approach performs significantly better than keyword searching both in terms of *recall* and *precision.* Recall is the percentage of *bad* messages that were correctly identified as such. Assentor identified 30% more *bad* messages than keyword-based searching. *Precision* is the percentage of flagged messages that were correctly identified as *bad* messages. Assentor's precision was four times better than keyword-based searching. Low precision translates into a higher number of messages that need human review. In other words, the more incorrect hits (*false positives*) a system outputs, the more expensive it is to review the messages.

It is easy to see why pattern matching produces better precision than keyword-based searching. Keywords are typically single words or short phrases. If a document or message contains a keyword, the document is automatically flagged regardless of the keyword's actual meaning. Because of this, keywords generate a high number of false positives. For instance, a keyword like "tip" will not distinguish between a relevant phrase (e.g., "*I'll give you a tip for your investments*") and an irrelevant one (e.g., "*they never tip the waiter*"). Similarly, the keyword "sue" will not distinguish between a relevant phrase (e.g., "*I'm going to sue your company*") and an irrelevant one (e.g., "*Sue called*").

By contrast, Assentor's pattern-matching-based technology incorporates linguistic context, which helps to remove ambiguity. For instance, "*safest*" and other superlatives are stock-hyping keywords, but only when they are predicated of things like "*accounts*" or "*stocks*". Assentor recognizes when superlative terms describe stock phrases. Thus, it flags "*this stock is the best investment*" but not "*an education is the best investment*."

It is very important to note that our NLP-based approach also produces better recall than keyword-based searching. There are a number of reasons for this. First, Assentor's patterns benefit from its dynamic name recognition capability. It dynamically identifies names of companies, people, and locations based on the linguistic properties of these phrases, instead of relying on static lists of names. It also identifies dates and monetary expressions which cannot be listed. Assentor's patterns take advantage of these semantic classes. For instance, a hard-pressure sales tactic pattern for phrases like "you must buy <COMPANY>" will match both known companies as in "*You must buy IBM* " as well as new companies as in "*You must buy XYZ Corp*." Keyword lists, by contrast, are not likely to include a complete list of all companies and are usually not capable of recognizing new company names.

Second, Assentor's NLP-based technology includes the capability to recognize morphological variants, i.e., the different shapes that a single word can take (e.g., "mismanage, mismanagement", "look, looks, looked, looking"). This morphological capability makes its patterns more flexible and robust than a keyword-based system. In a keyword-based system, compliance officers have to think of all possible word variants and combinations of variants. It is easy to forget to list some of these variants or combinations. By contrast, Assentor's morphological analysis capability enables the system to avoid such omissions. For instance, a single pattern for the concept of "mishandling or mismanaging accounts" can match the following phrases:

"You have mishandled my account"
"Mishandling our accounts"
"You mismanaged the account I opened with you"
"Your broker has mishandled my account"
"I see mismanagement of my accounts"
"Mishandling of my account"

Third, relevant concepts are often expressed with sentences consisting only of *common* words such as "*want,*" "*told,*" "*consented,*" "*buy,*" "*fee,*" etc. Keyword lists do not usually include common words because they flag too many messages. Assentor's patterns can recognize expressions involving common words without generating many false positives. For instance, Assentor's patterns can recognize the following types of expressions, which are difficult for the keyword-based approach to recognize without generating many false positives.

Broker Errors:
"I did not want to do this in the first place"
"I thought you were going to buy that Japan Fund"
Unauthorized Activities:
"We never told you to use our margin"
"We had never actually consented to make that purchase"
Complaints:
"We don't think it's fair to be charged these fees"
"This trust has been returning less than I expected"
IPO:
"The firm is looking to do a deal"
"Major systems integration company will be coming to market"
Hard-pressure Sales Tactics:
"You owe it to yourself"
"You know that you deserve it"
Stock Hyping:
"Buy it while you still can"
"Stock is priced like it's going out of business"

In summary, contrary to the common perception that keywords can catch as many relevant messages as more sophisticated techniques, the keyword-based approach not only produces more false positives but also catches fewer relevant messages than Assentor's NLP approach.

Application Development and Deployment

The first step in the development of Assentor was to gather domain knowledge about the securities industry's special needs for e-mail monitoring. During this early stage, securities industry consultants and a group of early adopters and pilot firms provided their insight as well as hardcopy and electronic copy of their correspondence under a strict confidentiality agreement.

As the second step, this correspondence data was split into training and blind sets. With the help of a GUI-based annotation tool, human annotators highlighted questionable messages, and marked and categorized relevant phrases in them.

Computational linguists then performed a data analysis of the training data and wrote patterns. Patterns are generalizations of the examples found in the training corpus. For instance, the example "*I did not receive a confirmation for my IBM trade*" is a complaint about not receiving something. Computational linguists generalize examples using synonyms, domain knowledge, and linguistic constructs. By adding synonyms in the patterns, variations such as "*I did not **get** a confirmation for my IBM trade*" are also captured. Using domain knowledge, other relevant objects are added to the pattern to cover variations such as: "*I did not receive stocks/certificates/information/dividends*," etc. Through linguistic variations (e.g., passivization, contractions, reordering, etc.) of the original example, further variations are captured: "*you did not send me/I didn't get/we never got/we haven't got*" etc.

An automated scoring tool provided feedback about the effectiveness of the rule set on both training and blind sets in terms of standard information extraction metrics: recall, precision, and f-measure. The false positives and false negatives found in the development sets were used to further refine the pattern set.

Most errors in the blind sets are false positives. Assentor's patterns are tailored to brokers' business correspondence, but brokers' correspondence includes a fair amount of non-business related messages. The latter

often contain language that is very similar to the problems that Assentor is searching for. For instance, complaints such as "I'm very unhappy and disappointed" may well occur in the context of a personal message.

Application Use and Payoff

At present, Assentor has been deployed at 77 brokerage and investment firms, with a total of 89,195 seats. Some firms have been using it for as long as two years. Firms' compliance officers use the system on a daily basis. It is estimated that for a firm with 1000 brokers, the complete manual human review solution (randomly selecting 15% of all messages) costs about 5 times as much as the Assentor solution, and that the keyword-based solution costs more than twice as much. As an additional benefit of using Assentor, firms report that it reduces non-business e-mail correspondence. Moreover, the amount of e-mail messages containing inappropriate language (e.g., obscenities and jokes) tends to decrease considerably after the first two weeks of deployment.

Maintenance

Pattern refinement continues on a regular basis as we receive feedback and new sets of sample e-mail messages from firms. As actual examples of violations are scarce, client feedback is an essential part of improving the NLP performance. So far, we have analyzed over 60,000 messages. Pattern updates are provided to the firms quarterly as part of maintenance. Releases are distributed by CD. Patches can be downloaded from the Assentor customer website.

Summary and Future Directions

We have described an e-mail monitoring tool that uses pattern-matching-based information extraction. We presented an evaluation of its performance vs. keyword-based searching. Our evaluation shows that pattern matching performs significantly better than keyword-based searching both in terms of precision and recall.

We are planning to apply Assentor technology to other areas. One area is website monitoring i) to ensure that sensitive or proprietary information of organizations is not publicly disclosed on their webpages, and ii) to allow companies to monitor negative news about themselves. Another exciting application is to use this technology in e-mail response management systems (ERMS) to dramatically improve e-mail-based customer service in e-Business. The NLP technology that underlies Assentor will enable automated or semi-automated response to customer's e-mail messages.

Acknowledgments

Assentor would have not been possible without the contributions of the entire Assentor team and the feedback provided by Assentor's early adopters, pilot firms, and customers. Our thanks to all of them.

References

Aone, Chinatsu, Lauren Halverson, Tom Hampton, and Mila Ramos-Santacruz. 1998. "SRA: Description of the IE^2 System Used for MUC-7." In Proceedings of the 7th Message Understanding Conference (MUC-7). www.muc.saic.com.

Dorre, Jochen, Peter Gerstl, and Roland Seiffert. 1999. "Text Mining: Finding Nuggets in Mountains of Textual Data." In Proceedings of the Knowledge Discovery and Data Mining Conference (KDD-99).