Informeta, L.L.C.

Mentys™ Solution for Letter-of-Credit AML

Version 1.0, 16 August 2006

Matthew Johnson
matthew.johnson@informeta.net

Sumanth Sridhar
sumanth.sridhar@informeta.net

1 Civic Center Plaza, Suite 540
Poughkeepsie, NY 12601
www.informeta.net

1. Introduction

In today's business environment, increasingly stringent regulatory requirements are forcing organizations to monitor pricing of goods on Letter of Credit (L/C) transactions. Federal regulators are holding the financial institutions who participate in these transactions liable for any gross pricing irregularities on these instruments.

Each year, financial institutions are paying heavy fines for failing to comply with government regulations. By maintaining systems that are capable of detecting fraudulent L/C transactions, businesses can demonstrate a commitment to meeting their legal and moral obligations.

Common techniques for laundering money through L/C include misrepresentation of price, quantity, or quality of imports or exports.[1] Until now, no system has been capable of spotting these kinds of transactions.

Mentys™ is an artificially intelligent software engine developed by Informeta that learns complex relationships and trends directly from numeric and/or text data, with or without additional expert knowledge, to produce a single, unified knowledge base. Informeta has used the capabilities of Mentys successfully in a number of different applications, including detection of data anomalies, prediction, and error correction.[2]

Informeta proposes an innovative solution built around Mentys that will spot pricing anomalies and report possible occurrences of L/C fraud or money laundering. A financial institution implementing our solution will be taking proactive steps toward identifying and stopping these fraudulent transactions. This solution will maintain an audit trail of all processed L/Cs, which can be reported to the federal authorities on a regular basis.

2. Mentys solution

Our automated solution will use Mentys' predictive features, combined with prior knowledge from application experts, to validate new L/C transaction details against historical data and current pricing extracted from the Web.

The solution will use three distinct sources of information to establish the credibility of a new L/C.

1. Historical L/Cs – details of past transactions can be extracted from a known repository, such as a customer's database, and used to extrapolate a plausible range of current-day prices from recent trends.
2. Current pricing information – retail, wholesale, or manufacturer's prices for goods can be found using external sources (e.g., the Web) and used to determine a plausible price range.
3. Expert prior knowledge – URLs for Web sites that provide pricing information, markup percentages for back-converting retail/wholesale to manufacturer's price, etc.

It is important to note that the universe of products that may be traded via L/C is effectively unlimited. Historical data may be insufficient or irregularly occurring, making it difficult to discern meaningful trends. Similarly, the availability on the Web of accurate prices for a particular product is not

1 Refer to “Trade-Based Money Laundering,” Financial Action Task Force, June 2006.
2 See Informeta's application white papers at http://www.informeta.net/index.php?page=products.

guaranteed. We believe it is important to combine all three sources of knowledge mentioned above to produce a reliable determination.

In previous work, we demonstrated that occurrences of specific keywords extracted from text articles on the Web can be combined with numeric data to enhance the predictive performance of Mentys. We believe that the combination of information from the historical databases and the Web, including both text and numeric fields, will allow us to successfully evaluate a considerable percentage of L/Cs. The success rate of this solution should improve as we accumulate more awareness and knowledge over time.

2.1 Historical trends

Existing historical databases will contain years' worth of past L/C transactions. We will search the available historical data for L/Cs that have similar product descriptions and terms of sale. Using Mentys to analyze matching transactions for pricing trends, we will predict the expected range for present-day pricing. We will also evaluate the credibility of other L/C details, such as brand, quantity, freight costs, and country.

2.2 Web spider

A Web spider is a software package that has the ability to navigate through web pages automatically and extract desired information. We propose the development of a series of spiders capable of finding prices for goods or services from multiple sources on the Web.

These spiders will search predetermined Web sources for current prices that match the goods or services description given on the L/C. These sources can be structured pages from seller's websites or shopping portals (e.g., Froogle, Bizrate). We could also perform a search of unstructured text from free or subscription online publications, though this would entail another layer of complexity.

2.3 Prior knowledge

In our experience, many important relationships cannot be gleaned from the data alone for one reason or another. The data samples may not be sufficiently large, or some relationships may depend on factors external to the known data. Application experts who understand some of these relationships can provide additional input to Mentys in the form of prior knowledge.

As we noted above, we may find retail or wholesale prices for goods, which require conversion back to the original manufacturer's price. Knowledge of a seller or distributor's profit margin can help to determine reasonable markup percentages.

Prior knowledge such as a list of brand names, unit conversions, and standard industry terminology can help to interpret and normalize L/C product descriptions. Other useful lists could include URLs for trusted sources of online pricing information.

It is not necessary that all relevant prior knowledge be available up front. Additional information or changes can be incorporated into the system over time, so that the solution will continue to improve and adapt.

2.4 Reporting

The primary output of the solution will be a report for each new L/C transaction. This will include reference numbers and other pertinent details, matching historical L/C prices, the most reliable current prices from the Web (along with links to the source pages), and a recommendation of whether the new price is acceptable. In case a determination cannot be made due to insufficient or unreliable data, these conditions will be reported.

We will ensure that the solution also keeps an audit trail of each processed L/C, even in cases where we are unable to validate the L/C price. This will allow an L/C expert to pick up where the Mentys solution leaves off in cases of insufficient supporting data. The audit trail will include any and all details required by the customer or government regulations.

2.5 Other considerations

The nature of L/C transactions poses a number of specific challenges for any solution. In particular, the product description field contains unstructured text that may include details such as quantity, units, brand name, freight (e.g., "FOB"), and country of origin/manufacture. Whether searching the historical L/C data or general Web content, it could be that an exact match for the complete product description cannot be found.

For example, a current L/C may have the following description[3]:

```
100 Sets 'ABC' Brand Pneumatic Tools, 1/2″ drive,
complete with hose and quick couplings, CIF Sunny Port
```

It is likely that no historical record exists that matches that description exactly. It is similarly likely that a search of the Web for the complete description will not produce any results. Hence the best we can achieve will be partial matches.

Our solution will use a sophisticated set of rules to determine what constitutes a "good match" versus a "bad match." This will require parsing the description into distinct semantic components that can be prioritized. The spider will then search for various combinations of one or more components of the description and effectively rank the matches that it finds.

In the example description above, examples of semantic components are:

- Base product – e.g. "`Pneumatic Tools`"
- Brand name – e.g. "`'ABC' Brand`"
- Quantity & units – e.g. "`100 Sets`"
- Freight/insurance costs – e.g. "CIF Sunny Port"
- Country of manufacture/origin – *not applicable in this example*
- Other – e.g. "`1/2″ drive`" or "`hose`" or "`quick couplings`"

We note also that, provided we have the necessary information, we could potentially validate the buyer and seller of each transaction. In this way, the solution might be used to identify false businesses that may serve as a front for illicit activities.

3 Excerpt of sample Documentary Credit opened by means of full text cable (in SWIFT format), courtesy of EXPORT911, http://www.export911.com/e911/export/docLC2.htm.

3. Conclusions

Continued growth in global trade and the increasing speed of information transfer provide ever more attractive opportunities for money launderers, terrorist financiers, and other criminals. Careful analysis of Letter of Credit transactions must be an important part of any comprehensive solution.

Our previous work has proved that Mentys can be used as a very effective and robust engine in applications that need to handle high-volume data in real time. We have used Mentys to develop custom applications in areas where traditional software systems have been ineffective.

With the ability to merge text, numeric data, and prior knowledge, we believe that Mentys has a unique capacity to handle the problem of Letter of Credit AML, which no other known software package has yet addressed.