COMPETITOR ANALYSIS AND ITS DEFENSES IN THE E-MARKETPLACE

A firm’s presence on the Internet opens a valuable channel for its competitors to collect and analyze its business information—to the firm’s competitive disadvantage—unless it devises a defensive strategy.

Information is a critical factor in achieving any form of business success. Plans for defending against competitive analysis systems (CASs) [4] include guidelines for preventing rival firms from getting such information. Here, we explore how the Internet, intelligent software agents, and data mining tools aid competitor analysis and how firms are increasingly exposed to their competitors through the network. We also propose a framework for deploying a plan to limit the competitive threat resulting from that exposure.

The Internet has become a competitive, as well as operational, necessity for all kinds of business. On the one hand, it gives firms an effective channel for conducting business 24 hours a day; on the other, it serves as a new channel for their competitors to automatically and invisibly collect important business information for the purpose of competitor analysis. Unfortunately, it has become increasingly effective and efficient for competitors to identify, collect, codify, and mine information. Such effectiveness is due to the availability of intelligent agents (IAs), peer-to-peer computing, network connections, large databases, and advanced data mining tools.

All firms are aware of the security issues (such as denial of service and malicious penetrations) threatening their Internet-based systems. While many academic papers, practical articles, and books detail coping methods [4, 7], many firms still ignore legitimate probing from their competitors or, worse, may be ignorant of these activities. Firms may not realize that their Web sites and online retail stores could be valuable resources for competitors’ CASs, as well as ideal entry points for probes or even attacks.

Firms use competitor analysis to attempt to define and develop a deeper understanding of their industries and identify and target existing, as well as potential, competitors, determining their strengths and weaknesses and anticipating their strategic and tactical moves [11]. To be useful, a CAS must provide trustworthy information about the goals and plans of a firm’s competitors, as well as their ability to achieve them. An important step in competitor analysis is collecting and analyzing a competitor’s data [3]. Due largely to the recent development of IAs and other
advanced information technologies, it is easier and quicker than ever [1, 2, 8]. IAs are programmed to seek out competitive intelligence on behalf of their owners, operating without constant human control or communication—autonomously, proactively, reactively, and sociably [9]. For example, IAs can be programmed to automatically search the Internet for exactly what their owners are looking for, evaluate the reams of information being retrieved, and deliver the refined information back to their owners [2, 8, 9]. Such functionality supports significant new threats to firms doing business on the Internet.

The Internet is increasingly used as a channel to communicate with customers, although this communication is also an entry point through which competitors might obtain a firm’s customer information and study how it conducts its business. For example, many express-service shipping companies allow customers to use package-tracking numbers (see Figure 1) to locate their packages in transit. When a customer enters a tracking number on the company’s Web site, the delivery route for the package is spelled out on the company’s Web page. Figure 2 includes such information from two major U.S. express-service shipping companies; for each of them, this way of tracking packages is a dramatic improvement in terms of customer service and cost saving over traditional telephone-based methods. Customers access delivery information through mouse clicks. Phone calls, which could be costly for the companies, are not needed.

However, from the point of view of competitor analysis, this means of customer service makes it relatively convenient for competitors to also collect the company’s business information. For instance, a competitor could simply program a number of IAs to enter an express service company’s Web site to query delivery information of all the packages delivered by, say, sequentially generating the tracking numbers from 00...00 to 99...99, then send the retrieved delivery information back to the competitor’s CAS. Using this data, competitors might ascertain how the company routes packages and which geographic areas represent a significant portion of the company’s business, along with possibly the identities of its big accounts. Based on this information, these competitors might then be able to develop better ways to provide express service to the customers, perhaps luring them away.

In another example, price aggregators (such as mysimon.com and pricescan.com) contract with vendors to publish price information associated with the merchandise on their Web sites, providing shoppers with attractive prices [8]. They make money through advertising and other means. However, competitors of the aggregators could program their IAs to collect price data directly from the aggregators’ Web sites without going through the vendors. IAs easily retrieve price data from the original aggregators, reformat it,
and provide it to their customers. In this scenario, the original price aggregators do all the work (getting the price data from the vendors in the first place), while their competitors harvest the information and, worse, undermine their customer bases and revenue streams. This means of exploiting data is definitely not what the original price aggregators want to see (see the sidebar “Aggregating Data Through IAs”).

### Competitor Information

Firms increasingly allow their customers to order products or services online anytime from anywhere. However, whatever is disclosed to customers on the Internet is also available to competitors’ IAs. With IAs, competitors can monitor the firm’s online pricing information, discount rates, shipping charges, warranties, and in-stock/out-of-stock information, all in the interest of developing competing strategies. These strategies (as well as tactics) can also be carried out automatically and autonomously by IAs at the speed of electronic communication. For example, if the firm lowers the price of a product by $5, a competitor’s IAs (monitoring this particular product) could know the price move instantly, then alert it to lower the price of its own competing product(s) by, say, $2 more to maintain its price advantage.

In addition to pricing information, IAs, combined with sophisticated data mining tools, can help identify other valuable business information (such as hot-selling products, whether or not a firm is able to maintain reasonable stock levels, the new features of new products, and trends in the firm’s product development effort). With all this information, competitors are better able to develop more focused and effective competing strategies.

Public online sources (such as filings to the U.S. Securities and Exchange Commission and the online versions of newspapers, magazines, industrial surveys, and individual firms’ own press releases) make the collection and analysis of a firm’s financial information through IAs more convenient than ever. Advanced Web page searching algorithms (such as Google’s PageRank) and data mining tools are also more able than ever to locate and analyze a competing firm’s valuable information [1, 2, 4].

As business-to-business Web sites and online auction and negotiation systems have become more popular for implementing e-procurement [6], a great deal of transaction information has also become available online. Accessing it through IAs, firms study and monitor their competitors’ cost structures, production cycles and schedules, procurement and distribution channels, inventory timelines, and turnover rates.

### False Information

In addition to obtaining valuable information from target firms, competitors can in some instances also use IAs to communicate false information to target firms. Many firms collect information about visitors to their Web sites (such as their browsing and buying habits and patterns), then use it to analyze and predict their personal preferences. The results can guide the firms to later provide better services and products, as well as marketing programs, to their customers through, say, accurate and timely recommendations, customized discount rates, and person-

---

**Aggregating Data Through IAs**

Running for about 10 minutes each morning, my IAs would regularly request the prices of approximately 450 books from eight price aggregator sites. Each aggregator employs its own programs to collect price data directly from online retailers selling books. The data they aggregate is what makes it possible for them to compete against other aggregators. All have invested time and money in their data-collection systems. To protect that investment, they want to prevent other organizations from accessing their price data and relaying it to users in other formats. Most price aggregators make money either through advertising or click-through agreements with vendors; if aggregator A dynamically crawls aggregator B’s Web site and relays the data to a browser in aggregator A’s format, then aggregator B has just been taken out of the compensation loop—even though it did most of the work.

The aggregator sites being studied as part of a project at Tulane University examining how shopbots tend to represent markets do not know I am an academic researcher. They know only that they receive regular hits from my system requesting prices for books. After I had collected data for about two months (March 2002 to May 2002), one of them stopped responding to requests from my IP address. I spoke to the site’s administrator who said he routinely terminates access for automated agents accessing his site, believing almost all automated traffic to his site is attributable to competitor behavior. When he found out I was an academic researcher, he allowed my agents to continue running against the site.

**Gove Allen** (gallen@tulane.edu) is an assistant professor in the Department of Information and Operations Management in the A.B. Freeman School of Business at Tulane University, New Orleans, LA.
alized Web pages. This method is increasingly popular and important among firms selling online, along with customer relationship management systems. Its success depends on the accuracy of the information being collected, because only accurate information yields accurate predictions that lead to or reinforce customer loyalty.

However, if a firm’s competitors were to reduce the accuracy of the firm’s prediction systems, hurting its ability to understand customer preferences, the competitors could simply develop a large number of shopping IAs to randomly browse the firm’s Web site or establish a large number of shopping baskets with random items. These activities could generate a quantity of nonsense Web browsing patterns and nonsense merchandise combinations in shopping baskets. However, because the Web server would have difficulty discriminating among patterns generated by people and patterns generated by IAs, random patterns and merchandise combinations, mixed with the patterns generated by people, are all collected in the customer patterns database by the firm’s Web server. The nonsense data floods the firm’s customer-patterns database and lowers the accuracy of the information that serves as the source of the prediction and recommendation system, later leading to a high error rate in recommendations and predictions, eventually making the systems useless.

Before the publicly accessible Internet era, competitor analysis relied largely on humans [3, 10, 11] but is now primarily conducted by computers at the speed of electronic communication. Firms must be aware of these new threats and adopt preemptive, as well as reactive, defensive methods.

A FRAMEWORK FOR DEFENSE

One view [4] of how to design an effective defensive framework involves five steps—identify current and potential competitors; determine which information they would want to know; assess the availability of that information; determine the value of a firm’s information; and impede the competitors’ CAS—detailed here:

Identify key information. An effective defense typically begins by identifying competitors; however, identifying existing competitors is relatively easy, while potential competitors could be anywhere and are usually invisible [4]. Many business research organizations are themselves not direct competitors of a particular firm but could be hired by the firm’s competitors to analyze its business situation [3]. Thus, it is quite difficult for a firm to enumerate the information a competitor might want to obtain and deploy a defensive mechanism around it. The firm needs to place itself in the position of its competitors to determine what they most want to know [4].

A firm should also want to determine what inappropriate information its competitors might want to send in its direction to confuse it as to their true intentions; an example is the flooding with false data of a firm’s database that records customer browsing patterns and merchandise combinations in shopping baskets. A firm should identify such information and the channels through which its competitors might pass it along, then develop preventive strategies. We strongly encourage research in this area.

Evaluate the accessibility of key information. Public sources of competitor data include industry norms (such as D&B and Standard & Poor’s Industry Survey) [10], but information available on the Internet has become as immediate as it is valuable. Moreover, because most of this information is in text/ASCII format, it is relatively easy for IAs to collect and analyze. Information collectors program IAs and send them out to desired Web sites (such as a target firm’s Web site, third-party online databases, including Compu-
stat and Edgar, and major newspaper Web sites) to collect relevant information and codify it automatically [1, 2]. Moreover, XML specifications make the data on Web sites more understandable, because XML tags have semantic meanings that make clear the data’s use and why it was collected in the first place [5]. This additional information is a great help to IAs trying to understand, refine, and analyze data.

Firms should be aware of how their most important business information can be accessed, along with its consequences, then determine whether protections (such as access passwords) are necessary.

**Defensive Methods**

Once a firm determines what information its competitors would want and how they might get it, as well as what false information they might try to pass along and how they might communicate it, it can begin to develop the defensive methods outlined here:

- **Routine analysis and monitor the system logs of information system activities.** When probing information through IAs, the IAs usually leave patterns of their activity (such as dramatically increased numbers of database queries or of newly established shopping baskets). Upon identification of a suspicious pattern, system managers should activate counterspying procedures (such as blocking IA access) (see Figure 3).

- **Eliminate unnecessary conveniences on a firm’s Web sites.** Analyzing its own sites, a firm might identify functional conveniences it could do without (if, say, site visitors are people) but can be exploited by IAs to greatly increase the efficiency of their information probing. For example, some firms allow visitors to query product or service information in batches. Visitors can input hundreds of product or tracking numbers at a time, instead of only one or a few at a time; the system then searches them all at once. This feature is rarely used by human visitors but could be exploited by IAs to speed their information-collection activities. Eliminating or limiting them could help the firm inhibit competitors’ information-probing goals.

- **Limit the information lifetime on Web sites.** Firms are enthusiastic about putting information on their sites but are practically indolent about removing it. Plenty of information on the Internet may be outdated for customers but isn’t for competitors. Firms should thus periodically identify and remove it; for example, there is no need to retain parcel delivery information beyond, say, a month after a parcel is delivered.

- **Make competitors’ IAs work longer.** Many firms’ product and tracking numbers have only a few digit combinations; IAs quickly enumerate all possible combinations. Extending the length of product and tracking numbers, mixing letters with the numbers, or using complex algorithms to generate the numbers in the first place can greatly increase the time and complexity required for enumeration. IAs thus need significantly more time to conduct a complete information sweep if they do it through enumeration.

- **Put false information in a firm’s own databases.** Firms could feed competitors’ IAs false information; for example, they could make up false product numbers, concoct false or misleading information under the numbers, and put them in the same database with real product numbers. If competitors’ IAs collect information by guessing product numbers, they will also likely get to these fake products and access the false information. The false information could thus lower the accuracy of the information collected through probing activities, possibly leading competitors to devise ineffective competitive tactics and strategies. The firm itself should be able to identify and control this false information.

- **Publish more soft and less hard data.** Soft data usually has no meaning or incorrect meaning or is easily misinterpreted without understanding the assumptions or context surrounding its use. IAs need more time and intelligence to understand and codify such data but are sometimes unable to do so; for example, saying, “This product’s sales increased by 12%” is safer than saying, “This product’s sales are $1.34 billion.”

- **Backup data more frequently.** Competitors not only want a firm’s valuable information, they might also want to deliberately send false information to the firm for their own strategic or tactical reasons. Monitoring system logs to discover and stop such activities is use-
ful. However, when the databases are flooded, archival data can be used to cleanse them of incorrect information, making them complete and accurate again.

**Conclusion**

Firms can recognize and prevent illegal penetrations and malicious attacks on their information infrastructures but often ignore or are unaware of legitimate (though invisible and undetectable) information probing from their competitors. Many still don’t realize that competitor analysis is carried out automatically and autonomously by IAs crawling the Internet at the speed of electronics. This ignorance could result in severe information leakage and represents a significant self-imposed business threat.

In order to defend itself, a firm’s IT professionals must be business savvy, knowledgeable about the business information that is important to the firm, and then develop defensive systems on the information level, not just on the technology level; for example, setting up a network firewall to stop malicious and illegal network intruders is a solid security practice. At least as important is looking out for the activities that collect valuable business information via the Internet (such as IAs emulating customer queries sent to collect package delivery information).

Top business managers must be technology savvy, aware that Internet technology not only opens new market channels but brings new threats to the business. They must emphasize information protection while enhancing Internet-based communications with customers and business partners. This attention to protection demands close working relationships among business and technology managers.

Meanwhile, more laws need to be enacted or amended to define the legal boundaries of information probing. Although the information on the Internet, especially when unprotected, is assumed to be freely available to the public, boundaries still exist between information probing and information poaching. The Internet is still a relatively new public network, and many laws are vague and ambiguous when applied to its activities. Improvement is necessary. But no matter how the legal system addresses these issues, business managers must be alert to them. Otherwise, competitive business damage and worse might already be on the way.

---

**References**