A VULNERABILITY ASSESSMENT OF THE U.S. SMALL BUSINESS B2C E-COMMERCE NETWORK SYSTEMS

Jensen J. Zhao
Allen D. Truell
Melody W. Alexander
Sherry A. Woosley

Abstract

Objective: *This study assessed the security vulnerability of the U.S. small companies' business-to-consumer (B2C) e-commerce network systems.* ***Background:*** *As the Internet technologies have been changing the way business is conducted, the U.S. small businesses are investing in such technologies and taking advantage of e-commerce to access global markets and compete with the large companies in their industries. While e-commerce activities have become popular, cyber attacks to the e-commerce sites are also on the rise. Therefore, a need exists for a security vulnerability assessment of the U.S. small companies' e-commerce sites.* ***Method:*** *The study used a combination of three methods—Web content analysis, information security auditing, and computer network security mapping—for data collection and analysis of a sample of 79* Inc. 500 *e-commerce sites.* ***Results:*** *The findings indicate that most e-commerce sites were outsourced to the Internet service companies and had the sites' network information publicly available on the Internet through the Google search. However, these sites had most of their ports closed, filtered, or behind firewalls with very few open ports. Companies in financial services, real estate, marketing, security, construction, education, and transportation were significantly more secure than other companies in protecting their network information.* ***Conclusion and Recommendations:*** *the U.S. small business B2C e-commerce sites were secure on average. But this degree of security is not enough. Therefore, this study provided recommendations such as how to secure network information, how to hide a site's IP address, and how to secure operating systems. In addition, the further research was recommended.*

Introduction

Small businesses in the U.S. represent a majority of employers, create around two thirds of the nation's new jobs, employ about half of the nation's private sector work force, provide half of the nation's nonfarm, private real gross domestic product (GDP), and contribute a significant share of innovations, even in the economic recession (U.S. Small Business Administration, 2003, 2009). As the Internet technologies have been changing the way business is conducted, the U.S. small businesses are investing in such technologies and taking advantage of

Dr. Jensen J. Zhao is a Professor in the Department of Information Systems & Operations Management, Miller College of Business, Ball State University, Muncie, IN.
Dr. Allen D. Truell is a Professor in the Department of Information Systems & Operations Management, Miller College of Business, Ball State University, Muncie, IN.
Dr. Melody W. Alexander is a Professor in the Department of Information Systems & Operations Management, Miller College of Business, Ball State University, Muncie, IN.
Dr. Sherry A. Woosley is an Instructor in the Department of Information Systems & Operations Management, Miller College of Business, Ball State University, Muncie, IN.

e-commerce to access global markets and compete with the large companies in their industries. The shoe retailer Zappos, TV manufacturer Vizio, and Internet-phone service supplier VoIP Supply are just a few examples of small businesses actively involved in e-commerce activities (Brynjolfsson & Smith, 2000; *Inc.*, 2009; U.S. Small Business Administration, 2000).

While e-commerce activities have become popular, cyber attacks to the e-commerce sites are also on the rise. Such attacks would impair or even shut down the e-commerce business completely by damages such as Web site defacement, denial of service, price manipulation, financial fraud, credit-card information thefts, or other data breach (e.g., Greene, 2008; Hovanesian, 2008; Mookhey, 2004). According to Symantec's *Global Internet Security Threat Report*, 90% of all Internet security threats detected by Symantec during 2008 attempted to steal confidential information for financial gains and the Internet became the primary conduit for malicious attack activities (Fossi et al., 2009).

In addition, a 2006 survey of 214 bank Web sites (Hovanesian, 2008) reported that 75% of the sites were vulnerable to hacking, with two big worrisome trends: (a) login boxes were placed on unencrypted Web (http) pages on a bank's domain and (b) the use of third-party services transferred customers to insecure outside pages. According to a cyber security report by NetWitness (Gorman, 2010), from late 2008 to early 2010, hackers gained access to a wide array of data at 2,411 companies, from accessing corporate servers that process credit-card transactions to servers that store large quantities of business data, such as presentations, intellectual property files, contracts, and even upcoming versions of software products.

Since the Internet is now the primary conduit for hacker attacks, it appears necessary to assess how secure the e-commerce sites are to block cyber intrusions and hacker attacks. While small companies in the U.S. account for an important part of the national economy, no nation-wide study of the U.S. small companies' e-commerce security has been identified in the literature. This research gap indicates a need for a security vulnerability assessment of the U.S. small companies' B2C e-commerce sites. We chose the *Inc. 500* companies for this study because these companies represent the U.S. small businesses.

Problem and Purpose of the Study

The problem addressed in this study was to assess the vulnerability status of the *Inc. 500* companies' B2C e-commerce sites. To conduct the study, we raised the following three research questions:

1. What network information of the *Inc. 500* e-commerce sites is publicly available on the Internet?
2. How vulnerable are network systems of the *Inc. 500* e-commerce sites to cyber intrusions and attacks?

3. Are there any significant differences among industry groups of the *Inc. 500* e-commerce sites?

The purpose of the study was to provide the participating companies with the findings that they need for continuous improvement of their e-commerce security. In addition, the findings would enable students specialized in studying Internet security or e-commerce to identify opportunities for internships or jobs at the *Inc. 500* B2C e-commerce sites that need to strengthen or maintain their Internet security.

Methodology

The population of this study consisted of the B2C e-commerce Web sites of the *Inc. 500* companies. A thorough search of the *Inc. 500* corporate Web sites (*Inc.*, 2009) identified 79 B2C e-commerce sites. These 79 sites were all used in the study according to the sample-size requirement (Cochran, 1977). To find out what e-commerce network information of the *Inc. 500* corporations is publicly available on the Internet and how vulnerable their e-commerce portals are to cyber intrusion and attacks, we conducted Google searches for related Web sites and auditing tools. We found three Web sites, *ZoneEdit.com*, *arin.net*, and *insecure.org*, offering the tools.

The *ZoneEdit.com* site is a leading Web site in DNS (Domain Name System) and domain management solutions. It provides a free DNS lookup utility tool, which enables any online user to enter a Web site domain name (e.g., yahoo.com) for searching its IP (Internet Protocol, e.g., 216.115.108.245) address (see at http://www.zoneedit.com/lookup.html).

The *arin.net* (American Registry of Internet Numbers) site provides a free database search service at *ws.arin.net*. The search service allows any online user to find a Web portal's registration information for resources registered with ARIN. The ARIN database contains IP addresses, autonomous system numbers, network name, type, and range, organizations or customers that are associated with these resources, and related points of contact. By entering a portal's IP address into the search tool, any person can get all the registered information of the portal's network systems (see at http://www.arin.net/whois/).

The *insecure.org* site offers a free network port scanning and mapping utility tool, *Nmap*, for network vulnerability assessment. *Nmap* uses raw IP packets to determine what hosts or ports are available on the network, what ports are open, filtered, or closed, what services (application name and version) those hosts are offering, what operating systems (OS) and OS versions they are running, what type of packet filters/firewalls are in use, and many other characteristics (see at http://insecure.org).

To ensure that it would be legal and ethical to use *Nmap* for network port scanning, the study reviewed related literature and could not find federal or state laws that specifically address the issue (e.g., U.S. Department of Justice,

2003). However, in a Georgia District Court case of "Moulton v. VC3," the judge declared a port scan in the case legal because it did not impair the integrity nor availability of the network. The judge found that since the activity performed no damage to the target, it could not be illegal (Jamieson, 2002). The implication of this case is that a port scan is not an attack and usually causes no damage to a target network; the legality and ethics of a port scan depend on whether the intent of a port scan is to cause damage or to improve security. As the purpose of this study was to provide the e-commerce administrators with the findings that they need for continuous improvement of the e-commerce security, using *Nmap* for this study was justified.

Two research assistants were trained to use these three tools to measure the network vulnerability status of each of the 79 corporate B2C e-commerce portals. All the searches and audits of the 79 portals were conducted between October 2009 and February 2010. The results were saved in digital format, and data were recorded and coded. Frequency counts, percentage distributions, and cross-tabulations were prepared for data analysis. Pearson chi-square test was used to determine any significant differences at the .05 alpha level among industry groups in securing network systems. Table 1 illustrates the demographic profile of the 79 *Inc. 500* B2C e-commerce sites.

Table 1

Demographic Profile of Inc. 500 E-Commerce Sites (N=79)

Group	Type of Company Business	No. of Companies	Percentage
Group 1.	Financial services, real estate, marketing, security, construction, education, transportation	33	42%
Group 2.	Food, beverage, retailing	26	33%
Group 3.	Computer, electronics, media, software, and other consumer products	20	25%
	Total	79	100%

Findings

The findings are presented in the sequence of (a) network information available on the Internet, (b) network vulnerability to cyber Intrusion and attack, and (c) significant differences among industry groups.

Network Information Available on the Internet

Research Question 1 asked, "What network information of the *Inc. 500* e-commerce sites is publicly available on the Internet?" The Internet search at *ZoneEdit.com* and *ws.arin.net* identified the IP addresses and network information of the majority of the 79 *Inc. 500* corporate e-commerce sites. As Table 2 shows, 100% of e-commerce sites' IP addresses were publicly available on the Internet.

As a consequence, with these publicly available IP addresses, any online user could go to *ws.arin.net* and enter the IP addresses for identifying a large amount of network information from a majority (76% - 100%) of the e-commerce sites, such as a site's organization address; CIDR (classless inter-domain routing), network name, handle, parent, range, and type; servers' name; registered tech handle, name, phone, and email; organization name and ID. The network information also revealed that the majority (82%) of *Inc. 500* e-commerce sites were outsourced to the Internet service providers; only 18% of the companies ran their sites by themselves (see Table 2).

Table 2

Network Information Availability of Inc. 500 E-Commerce Sites (N=79)

Category	Frequency	Percentage
IP addresses	79	100%
Organization Address (City, State/Province, Country)	79	100%
CIDR (Classless Inter-domain Routing)	79	100%
Network Name	79	100%
Network Handle	79	100%
Network Parent	79	100%
Network Range	78	99%
Network Type	78	99%
Name of Server 1	78	99%
Network Update Information	77	97%
Registered Tech Handle, Name, Phone, Email	72	91%
Organization Name	62	78%
Organization ID	62	78%
Name of Server 2	60	76%
E-commerce Outsourced to Internet Service Companies	65	82%

Network Vulnerability to Cyber Intrusion and Attack

Research Question 2 asked, "How vulnerable are network systems of the *Inc. 500* e-commerce sites to cyber intrusions and attacks?" Network systems connect to the Internet through computer ports. The ports of an Internet-connected computer are classified into the well-known ports, the registered ports, and the dynamic and/or private ports. The numbers of the well-known ports range from 0 to 1,023; those of the registered ports are from 1,024 through 49,151; and those of the dynamic or private ports range from 49,152 to 65,535. If the ports are open on

the Internet without firewalls or filters, they are very vulnerable to cyber intrusions and attacks.

Among the 79 e-commerce sites scanned by using *Nmap*, 32 sites (41%) were detected of running 1,714 Internet ports; 26 sites (33%) running 1,711 ports; and 21 sites (27%) running 1,709 ports. The scan reports revealed that most of these detected ports were closed, filtered, or behind firewalls and only very few ports were detected as open. As Table 3 shows, a majority of sites (54%) had only three or fewer open ports on the Internet, respectively. Nearly one third (30%) of the sites had 4 to 10 open ports at each of these sites, followed by 13% having 12 to 18 open ports, respectively. By contrast, only a tiny minority (3%) had more than 50 open ports each on the Internet.

Table 3

Number of Internet Ports Open at Inc. 500 E-Commerce Sites (N = 79)

		Group	
No. of Open Ports	**No. of Portals**	**Frequency**	**Percentage**
0	5		
1	5		
2	21		
3	12	43	54%
4	7		
5	8		
7	6		
10	3	24	30%
12	3		
14	2		
15	2		
18	3	10	13%
50 or more	2	2	3%
Total	79	79	100%

Table 4 presents the network vulnerability information detected from the open ports of the *Inc. 500* e-commerce sites. The majority (94%) of the 79 sites had Port 80/tcp open for http (hypertext transfer protocol) or World Wide Web services. The common Web servers identified from Port 80/tcp were Apache, Microsoft IIS, and Netscape. Second, 78% of the sites also had Port 443/tcp open for encrypted https services such as personal and institutional Web accounts for business data transactions. In addition, 80% of the sites had their network device type information detected by the network scanner, such as general purpose, authentication server, PDA, and storage. Around one fifth of the sites also revealed the information of their network uptime in days. Finally, only 14% to 15% of the e-commerce sites

had their computer operating systems (OS) information detected by the network scanner, such as running Windows NT/2K/XP/2000/2003/Pocket PC servers and Linux 2.6x servers (see Table 4).

Table 4

Systems Vulnerability Status of Inc. 500 E-Commerce Sites (N = 79)

Category	Frequency	%
Port 80/tcp open; service: http; common servers: Apache, IIS, Netscape	74	94%
Port 443/tcp open; service: https; common servers: Apache, IIS, Netscape	62	78%
Device type information: general purpose, authentication server, PDA, and storage	63	80%
Uptime: in days	18	23%
OS information: e.g., Running: Windows NT/2K/XP/2000/ 2003/Pocket PC	12	15%
OS details: e.g., Windows XP/2000/2003 Servers, and Linux 2.6x	11	14%

Significant Differences Among Industry Groups

Research Question 3 asked, "Are there any significant differences among industry groups of the *Inc. 500* e-commerce sites?" Pearson chi-square test was used to determine whether any significant differences exist among industry groups on network information availability and vulnerability. The 79 participating *Inc. 500* companies were classified into three groups of industries to meet the statistical requirement of group size for comparative analysis (Fraenkel & Wallen, 2006). Group 1 included 33 companies in financial services, real estate, marketing, security, construction, education, and transportation industries. Group 2 had 26 companies in food, beverage, retailing industries. And Group 3 included 20 companies in computer, electronics, media, software, and other consumer products industries.

Overall, just a few significant differences were identified among the three industry groups. Table 5 illustrates only the areas where significant differences existed among the industry groups. Regarding the sites' organization name and ID information, significantly more sites in Group 2 (food, beverage, and retailing; 92%) had their information publicly available in comparison with 78% of the *Inc. 500* 79 e-commerce sites. By contrast, significantly fewer sites in Group 1 (financial services, real estate, marketing, security, construction, education, and transportation; 64%) had such information publicly available on the Internet. In addition, compared with 29% of the *Inc.* 500, significantly more sites in Group 2 (50%) were also detected of their Port 22/tcp open, whereas only 18% of Group 1 had their Port 22/tcp open for ssh (secure shell) communication (see Table 5).

Table 5
Significant Differences Among Industry Groups

Category	% of Inc. 500	vs	Industry Group	%	
Organization Name revealed	78%	vs	Group 2	92%	*
		vs	Group 1	64%	*
Organization ID revealed	78%	vs	Group 2	92%	*
		vs	Group 1	64%	*
Number of Port 22/tcp open	29%	vs	Group 2	50%	*
		vs	Group 1	18%	*

* Significant at .05 level.

Summary and Discussion

The findings of the study indicated that the majority of 79 *Inc. 500* e-commerce sites' network information is publicly available on the Internet through the Google search. The publicly available information indicated that 82% of the *Inc. 500* e-commerce sites were outsourced to the Internet service providers. The available information includes networks' IP address and organization's physical address; CIDR, network name, handle, parent, range, and type; servers' name; network update information, registered tech handle, name, phone, and email address; organization name and ID.

Such publicly available information makes the sites vulnerable to cyber intrusions and hacker attacks. For example, searching for the IP address of a Web site through its Web address (URL) is often the first step for cyber intruders to connect to the server of the site. In addition, the network range and CIDR address reveal the total number of hosts the network posses and the network's higher-level and lower-level routing information. Having put these pieces of information together, a cyber intruder has a full picture of which parts of the network are vulnerable and easy to intrude.

However, the findings of the Internet port scan reports illustrated that the e-commerce sites had their most ports closed, filtered, or behind firewalls; only a very few of ports were detected as open. The majority (94%) of the 79 portals had Port 80/tcp open for http or World Wide Web services and 78% of the portals also had Port 443/tcp open for encrypted https. Although it is common to have Port 80/tcp and Port 443/tcp open for their respective services, such open status is vulnerable to cyber intrusions and attacks because open ports might leak networks server information and operating systems (OS) information, unless the portals use the network address translation (NAT) and the port address translation (PAT) technologies.

For example, 80% of the sites revealed their network device-type information through the network scanner, such as general purpose, authentication server,

PDA, and storage. But only about 15% of the sites had their computer operating systems (OS) detected by the network scanner, which revealed information such as running Windows NT/2K/XP/2000 /2003/Pocket PC servers and Linux 2.6x servers. The findings seem to imply that NAT and PAT technologies were used by many sites.

Finally, the industry group comparison identified that the *Inc. 500* companies in financial services, real estate, marketing, security, construction, education, and transportation were significantly more secure than other *Inc. 500* companies in protecting their network information. This finding of the financial service companies in this study does not support the 2006 survey results of Hovanesian (2008) that 75% of bank Web sites were vulnerable to hacking. The finding indicates that financial services companies have made continuous improvement on the security of their Web sites.

Conclusions and Recommendations

Based on the findings of the study, we conclude that the U.S. small business B2C e-commerce sites were secured by keeping most of their ports closed, filtered, or behind firewalls. But this degree of security is not enough because the publicly available network information of their sites would attract cyber intruders and their few open ports still remain vulnerable to cyber intrusions and attacks.

To further strengthen the sites, we have the following recommendations for the *Inc. 500* and other small companies' e-commerce administrators and developers.

First, consider negotiating with American Registry of Internet Numbers on requiring username and password login for user identity management and access to a Web site's registration information that contains sensitive data such as IP addresses, autonomous system numbers, network name, type, and range. To make the negotiation successful, companies need to form an alliance and conduct collective negotiation with American Registry of Internet Numbers.

Second, to secure the open ports such as Port 80/tcp and Port 443/tcp, consider hiding e-commerce portals' IP addresses and port information by using the network address translation (NAT) and the port address translation (PAT) technologies. These two technologies are usually used together in coordination for two-way communicationNAT is a technique of transceiving network traffic through a router that involves rewriting the source or destination IP addresses and the port numbers of IP packets as they pass through the NAT-enabled router. Therefore, NAT can prevent malicious activity initiated by outside hosts from reaching those local hosts as it disguises the internal network's structure through rewriting and the traffic appears to outside parties as if it originates from the gateway machine. However, using NAT complicates the Internet tunneling protocols because NAT modifies values in the headers which interfere with the integrity checks done by tunneling protocols.

PAT is a device to translate all communications between internal hosts on a private network and external hosts on the Internet. With PAT installed, all

communications sent to or from external hosts contain only the IP address and port information of the PAT device instead of internal host IP addresses or port numbers. PAT translates or replaces IP addresses and ports of its internal hosts; therefore, it effectively hides the true endpoint IP address and port of the internal hosts. External hosts are only aware of the IP address of the PAT device and the particular port being used to communicate on behalf of specific internal hosts. A disadvantage of PAT is that, if many internal hosts on the private network make many connections to the Internet, the PAT device may not have sufficient room in its internal table to keep track of the connections or it may simply run out of unused ports.

Third, a more secure and also more expensive alternative is to use the high anonymity proxy servers. These proxy servers not only hide the portals' original IP addresses but also do not identify themselves as proxy servers, thereby making their portals anonymous on the Internet. Without knowing a portal's original IP address, cyber intruders have difficulties of getting the portal's network information.

If proxy servers and firewall are not used to protect Port 80/tcp, portal administrators have to use intrusion detection software to monitor the port traffic cautiously and detect hackers' fingerprints used in exploitation of Web servers and applications. Regarding the open Port 443/tcp for encrypted https services, user IDs and passwords must be required to grant access to the port and outgoing access to the port from servers should be restricted.

Finally, we recommend that business and information technology (IT) educators consider sharing the findings of the study with their students with the concentration of Internet security or e-commerce and assisting them in identifying opportunities for internships or jobs at the *Inc. 500* B2C e-commerce sites that need to strengthen or maintain their Internet security.

Recommendations for Further Research

We recommend that a further study of this type be conducted in five years among the *Inc. 500* companies for measuring their B2C e-commerce security and identifying opportunities for further improvement.

In addition, we recommend that a comparative study of the security vulnerability of B2C e-commerce sites be undertaken between *Inc. 500* companies and *Fortune 500* companies. The findings of the research will be beneficial not only for big and small companies to learn from each other, but also for business and IT educators to update their relevant courses and programs.

References

Brynjolfsson, E. & Smith, M. D. (2000). *The great equalizer? Consumer choice behavior at Internet shopbots*. Cambridge, MA: Massachusetts Institute of Technology.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley and Sons.

Fossi, M., Johnson, E., Mack, T., Blackbird, J., Low, M. K., Adams. T. et al. (2009, April). *Symantec global Internet security threat report: Trends for 2008* (Vol. 14). Retrieved October 15, 2009, from http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_internet_security_threat_report_xiv_04-2009.en-us.pdf

Fraenkel, J. R. & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). Boston: McGraw Hill.

Greene, T. (2008, August). Business hacks reap money from e-commerce sites. *Network World, 25*(30). Retrieved August 10, 2008, from http://www.networkworld.com/news/2008/ 080808-business-hacks.html

Gorman, S. (2010, February 18). Hackers attack 2.411 firms. *The Wall Street Journal,* p. A3.

Hovanesian, M. D. (2008, August 11). Security holes at the online bank. *Business Week*, 16.

Inc. (2009, September). The Inc. 500 U.S. companies. *Inc.* Retrieved October 1, 2009, from http://www.inc.com/magazine/20090901/index.html

Jamieson, S. (2002). "The Ethics and Legality of Port Scanning", *SANS Institute*. http://www.sans.org/reading_room/whitepapers/legal/the_ethics_and_legality_of_port_scanning_71?show=71.php&cat=legal

Mookhey, K. K. (2004, April 26). Common security vulnerabilities in e-commerce systems. *Security Focus*. Retrieved October 5, 2008, from http://www.securityfocus.com/infocus/1775

U.S. Department of Justice. (2003). "Fraud and Related Activity in Connection with Computers" in the United States Code Annotated Title 18, Chapter 47, Section 1030. Washington, DC: Author, http://www.usdoj.gov/criminal/cybercrime/1030NEW.htm

U.S. Small Business Administration. (2000). Small business expansions in electronic commerce. Retrieved July 2, 2003, from http://www.sba.gov/advo/stats/e_comm2.pdf

U.S. Small Business Administration. (2003). Small business economic indicators for 2002. Retrieved January 7, 2004, from http://www.sba.gov/advo/stats/sbei02.pdf

U.S. Small Business Administration. (2009). Small business economy: A report to the president 2009. Retrieved January 7, 2010, from http://www.sba.gov/advo/research/sbe.html.