2002년도 대한전자공학회/대한전기학회 시스템 및 제어분야 합동 추계학술대회 논문집(2002.11.30)

패킷 마이닝 기법을 사용한 인터넷 서버 프로파일의 자동생성 연구

곽 미 라, 조 동 섭
이화여자대학교 과학기술대학원 컴퓨터학과
전화 : 02-3277-2309 / 휴대폰 : 016-741-6870

Generation of Internet Server Profile Using Packet Mining

Mira Kwak, Dong-sub Cho
Dept. of Computer Science and Engineering EIST Ewha Womans University
E-mail: mirakwak@ieee.org

Abstract

Management of internal Internet servers is increasingly becoming an important task. According to meet this requirement, they use service log analysis tools and network monitoring tools. But these are not enough to produce advanced management information considering contents of Internet services. Therefor we propose a system and let it detect Internet server systems existing in internal network and individuate those systems with providing profile. Internet server profile includes system's basic information, network traffic information, and Internet service usage information.

I. 서론

네트워 및 인터넷의 사용이 활발해지면서, 기업이나 학교 등 한 조직의 내부 네트워에서 인터넷을 통한 서비스를 제공하는 서버 시스템들의 수가 증가해왔다. 이에 따라 조직에서 내부 네트워의 서버 시스템들의 활용도 및 이상 사용을 발견하고 그 현황을 관리하는 방법이 필요하게 되었고, 그 방법으로 시스템의 서비스 로그 분석과 네트워 모니터링이 널리 사용되고 있다. 서버 시스템에 기록되는 로그 데이터는 그 서비스 사용 현황을 파악하기 위한 분석에 기본적인 자료로 널리 활용되고 있지만, 그 내용이 제한적이고 서버의 운용환경에 따라 다른 형식으로 기록되므로 이를 분석하는데 번거로움이 따른다. 네트워 모니터링은 관리자로 하여금 네트워의 내부에서 혹은 내부와 외부 사이에서 오고가는 모든 네트워 트래픽을 실시간으로 관찰하고 시간에 따른 변화추이를 살필 수 있게 하여 네트워 사용 분석에 유용하지만, 서비스의 내용을 고려한 고급 분석에 미흡한 점이 많다. 이에 내부 네트워에 존재하는 서버들에 대한 네트워 트래픽 현황 및 그 서비스 사용 현황을 관찰하고 분석, 기록하는 발전된 시스템이 요구된다.

본 연구에서는 이러한 요구를 해결할 수 있도록 내부 네트워 분석 기능을 수행할 뿐 아니라, 내부 네트워에 존재하는 서버들에 대한 정보를 사람에 의해 직접 입력받지 않고 각 서버들에 대해 수집된 정보를 바탕으로 하여 자동으로 파악하는 시스템을 설계하였다. 이를 통해 조직의 내부 네트워의 인터넷 서버들을 스스로 감지하고 그 사용 현황을 분석하는 시스템이 가능하게 된다. 이를 위해 본 연구에서는 IP 패킷들을 수집하여 이에 데이터마이닝 알고리즘을 적용하는 방법을 사용하였고, 이를 패킷 마이닝이라 이름지었다.

II. 인터넷 서버 프로파일의 생성

표 2 인터넷 서버 프로파일의 내용

항목	내용
서버의 기본정보	호스트이름, IP주소, MAC주소, 컴퓨터이름
제공 서비스	FTP, HTTP, Mail, Telnet 서비스 제공 여부
활동이 활발한 시간대	전체 네트웍 트래픽 수준, 프로토콜별 및 서비스별 분석
데이터 송수신 현황	전체적인 송수신 현황, 프로토콜별 및 서비스별 송수신 현황
트래픽 추이 분석	전체 네트웍 트래픽 수준, 프로토콜별 및 서비스별 분석
기간 분석	트래픽 추이를 월/요일/일/시간을 기준으로 통계적 분석

2.1 인터넷 서버 프로파일의 설계

각 인터넷 서버의 프로파일은 표1과 같은 내용을 가진다.

표 1에서 보인 것과 같은 프로파일을 생성하기 위하여, 2.2 절의 방법으로 수집된 기본적인 네트웍 트래픽 데이터는 두 번의 분석 과정을 거친다. 첫 번째 분석을 통하여 네트웍 트래픽 데이터는 종합적인 트래픽 데이터와 프로토콜별 트래픽 데이터, 그리고 WWW, FTP, Mail, Telnet 등의 서비스별 트래픽 데이터로 구분되어 각각 저장된다. 이 과정을 2.2절에서 설명한다. 두 번째 분석단계에서는 이렇게 분리 저장된 데이터로부터 각 서버에 대한 정보를 구성하게 된다. 두 번째 과정을 2.3절에서 설명한다.

2.2 IP 패킷 수집과 정보의 추출

패킷의 수집은 libpcap이라는 라이브러리를 사용하여 이루어진다. 패킷 수집 작업을 수행하는 모듈은 내부 네트웍 상의 모든 패킷 데이터를 수집하기 위하여 여러 호스트 시스템들이 연결되어 있는 허브들보다 상부에 위치한 시스템에서 동작하도록 하였다.

수집된 패킷 데이터는 트래픽을 프로토콜별 정보와 서비스별 정보로 구분하여 저장하는 과정을 거친다.

2.3 인터넷 서버 프로파일의 생성

2.2절의 결과 데이터베이스로부터 내부 네트웍에 존재하는 인터넷 서버들이 자동으로 파악되며, 각 서버별 프로파일이 생성되어 저장된다.

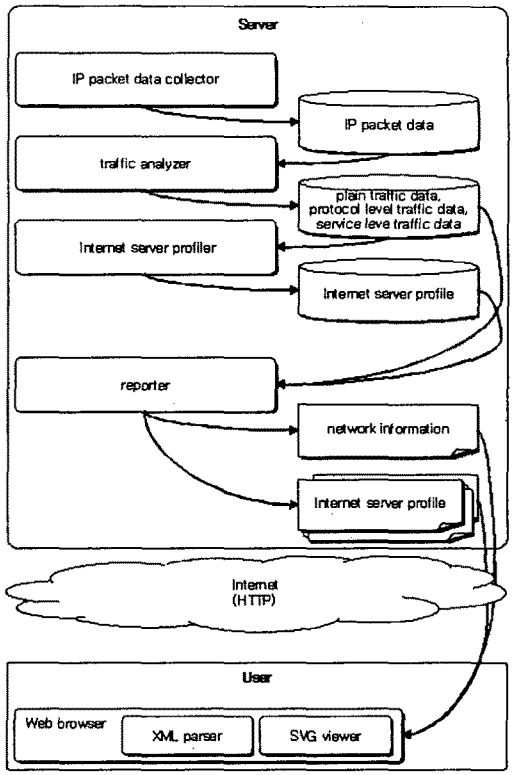

그림 1 인터넷 서버 프로파일 생성 시스템의 구조도

트래픽 데이터의 흐름 정보로부터 트래픽의 출발지와 도착지의 IP 주소와 포트 번호, 패킷의 조합을 통해 식별된 전달 내용을 알 수 있다. 이를 바탕으로 내부 네트웍에 존재하는 여러 호스트 시스템들로부터 인터넷 서비스를 받는 단말들과 인터넷 서비스를 제공하는 서버 시스템들을 구분할 수 있다.

발견된 인터넷 서버 시스템들 각각에 대하여 그 기본적인 정보, 제공하고 있는 서비스들, 활동이 활발한 시간대, 데이터의 송수신 현황, 트래픽의 추이와 그 기간별 분석 내용이 데이터베이스에 저장된다.

이후 보고서 생성부에 의해, 내부 네트웍에 대한 개괄적인 자료와 각 인터넷 서버에 대해 그 프로파일이 표 1의 내용을 포함하도록 설계된 XML 스키마를 따르는 XML 파일로 생성되며, 그 내용을 가시화하기 위하여 SVG파일로 그래프가 생성된다. 사용자의 요구에 따라 간추린 종합적인 정보, 상세한 각 서버별 정보 등을 출력하는데 유연함을 가지기 위하여 XSL을 작성하여 XML 형식의 결과 파일을 다양하게 포맷팅하도록 하였다.

III. 인터넷 서버 프로파일 생성 시스템의 설계와 구현

본 논문에서 제안하는 인터넷 서버 프로파일 생성 시스템의 전체적인 구조는 그림 1에서 보는 바와 같다. 시스템은 크게 네 개의 부분으로 구성되며 그 내용은 다음과 같다.:

- IP 패킷 수집부: 네트웍 상에 오가는 모든 패킷을 수집하여 적재
- 트래픽 분석부: 수집된 패킷 데이터로부터 트래픽을 전체적 트래픽, 프로토콜별 및 서비스별 트래픽으로 나누어 저장
- 인터넷 서버별 프로파일 작업부: 트래픽 분석부의 결과를 바탕으로 각 인터넷 서버를 찾아내고 그에 대한 프로파일을 생성
- 보고서 생성부: 생성된 프로파일을 웹 인터페이스를 통해 사용자의 요청에 따라 전달

보고서 생성부의 출력은 XML 형식이며, 트래픽 및 여러 통계적인 분석 내용의 시각화를 위하여 XML 그래픽 표준인 SVG를 사용한다.

IV. 결론

본 연구를 통하여 웹 기반으로 내부 네트웍의 여러 인터넷 서버들을 스스로 감지하고 그에 대한 프로파일을 구성하는 시스템을 제안하였다.

현재, 제안된 시스템은 수집된 데이터를 바탕으로 분석 작업을 배치 방식으로 실행하고 있다. 이를 개선하여 실시간으로 적재되는 대량의 데이터를 세련된 방법으로 처리하고, 이를 바탕으로 인터넷 서버 프로파일을 주기적으로 갱신할 수 있도록 할 것이다.

또한 네트웍과 관련한 정보들 외에 서버의 시스템 자원 사용 현황 정보를 포함하도록 인터넷 서버 프로파일의 설계를 확장하고자 한다.

참고문헌

[1] Aloke Gupta, "Performance aspects of computers with graphical user interfaces", University of Illinois at Urbana-Champaign, 1993.

[2] Luca Deri, Stefano Suin, "Ntop: beyond Ping and Traceroute", Proceedings of DSOM '99, Zürich, Switzerland, October 1999.

[3] Seong Jin Ahn, Seung Keun Yoo, Jin Wook Chung, "Design and Implementation of a Web-based Internet Performance Management System Using SNMP MIB-II", International journal of network management 9, pp.309~321, 1999.

[4] Jia-yu Pan, Srinivasan Seshan, Christos Faloutsos, "FastCARS: Fast, Correlation-Aware Sampling for Network Data Mining", Proceedings of IEEE GlobeCOM 2002, Nov. 17-21, 2002.