웹 게시판에서 스팸 게시물 탐지 및 블러킹 방안

조유형*, 민성기
*고려대학교 컴퓨터정보통신대학원
e-mail : morningstar@korea.ac.kr ,sgmin@korea.ac.kr

Spam post detection and blocking mechanism on web bulletin board

Yu Hyoung Cho*, SungKi Min
Dpt. Information Technology, Graduate of School Computer and Information Technology,
College of Information and Communications, Korea University.

요 약

웹 게시판 서비스에서의 비정상행위 탐지 및 블러킹 방안 제시는 서비스를 제공하는 주체에게는 대량의 광고성 게시글로부터 안정적 서비스 운영이 가능하게끔 하고, 서비스를 이용하는 사용자에게는 원하지 않는 게시글로부터 블러킹 방안을 제공 받아 깨끗한 웹 게시판 서비스를 가능한 방안으로 인터넷 환경이 제공되면서 필터링 기술 발전 수준이 가장 높은 스팸 메일 필터링 기술을 응용하여 웹 게시판 서비스에 적용하여 필터링 효과 수준을 측정하고 다른 웹 서비스 등에 활용할 수 있는 방안을 제시한다.

1. 서론

정보 통신 기술의 발달로 수많은 사용자들은 자신의 웹 사이트를 생성하여 인터넷에 연결하거나 블로그나 홈페이지 서비스를 제공하는 사업자의 서비스를 활용하여 자신이 가지고 있는 정보를 게시한다. 이러한 서비스 증가로 인해 악의적으로 광고 및 타 사이트 유도등 스팸성 글이 게시되어 해당 게시판이 광고, 유해 사이트 홍보에 이용되는 사례가 증가하고 있으며, 서비스 제공 사업자측면에서는 대량의 스팸 글로 인한 서비스 안정성 저하 및 불필요한 트래픽 발생으로 이를 유지하기 위한 추가 비용 지불이 필요한 악순환이 지속되고 있다.

본 논문에서는 게시판 형태의 웹사이트에서 메일 서비스에서 주로 사용되는 안티스팸 기술을 응용하여 웹 게시판 형태의 서비스에서 효과적인 광고성 글에 대한 탐지 및 블러킹하는 방안을 제안한다. 이 방법은 웹 사이트에 접근하는 사용자 아이피, 유저아이디 게시글 제목등을 메모리 데이타베이스에 저장하여 사이트 접속 할 때마다 검사하여 비교하는 방식으로 특정 행태가 다수 발생하면 필터링하는 방법이다.

본 논문의 구성은 다음과 같다. 2 장에서는 인터넷 서비스에서의 스팸 기술과 그에 대응하는 안티스팸 필터링 방법에 대해 살펴보며, 3 장에서는 본 논문에서 중점적으로 다루고 해결하고자 하는 게시판 형태의 웹서비스에서의 스팸 기술과 필터링에 대해 설명하고 본 연구의 목적을 제시하고 4 장에서는 실제 적용 방법과 사례를 제시한다. 마지막으로 5 장에서 결론과 향후 연구 방향에 대해 언급한다.

2. 관련연구

2.1 메일 서비스에서의 스팸과 필터링 기술

WWW(World Wide Web)에서의 대표적 스팸기술이 메일 서비스 스팸이며, 스팸 메일의 정의는 발송자가 재화나 서비스 판매 목적으로 수신자의 동의 없이 불특정 다수에게 대량으로 전송되는 이메일을 말한다[1]. 안티스팸 필터링 기술에는 메일 클라이언트에서 차단, 메일 서버에서 차단, 메일 프로토콜을 사용하여 차단하는 방법등이 있으며, 대표적으로 메일 서버에서 차단하는 방법에는 메일 서비스 업체에서 메일 송신 서버 주소를 등록하여 해당 서버 이외의 주소에서 발송되는 메일을 필터링하는 White-List DB 관리방법과 메일 문서내 단어들을 대상으로 통계적 방법의 SVM(Support Vector Machine)을 이용하여 스팸 메일 필터링하는 방법등이 있다.[2]

2.2 검색엔진의 순위조작을 위한 스팸과 필터링 기술

웹 검색은 대다수의 검색엔진에 질의어를 입력하고 입력된 질의어와 가장 관련성이 높은 웹 페이지를 찾아내기 위한 검색엔진 자체의 쿼리엔진 알고리즘을 가지고 있다. 검색 쿼리엔진 알고리즘의 가장 큰 문제점은 질의어에 대한 결과가 사용자의 요구와 일치하기 어렵고 엔진 자체의 부정확함과 검색엔진의 높은 우선순위 획득을 위해 악의적 작성된 웹 페이지로 인해 웹 페이지 스패밍(Spamming)이 발생한다. 위의 문제는

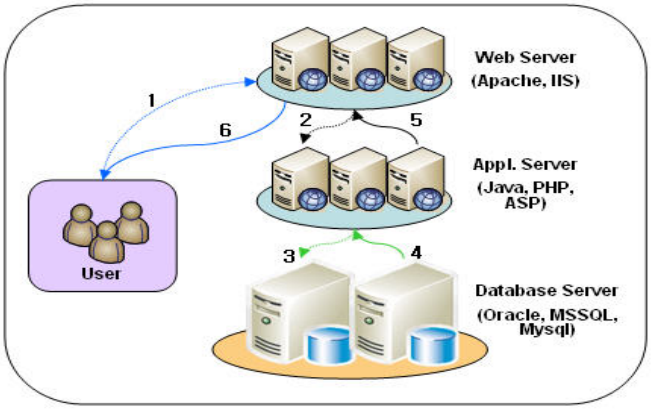
2001 년에 처음 발생하였으며, Google Bombing 과 같은 사건이 검색엔진 알고리즘의 약점을 이용한 것이다[3]. 구글 검색 엔진이 페이지 랭크 기반 순위 방식으로 인해 특정 키워드에 대해 웹 페이지가 과다하게 링크되게 함으로써 구글 검색의 우선순위를 조작하는 방식이다. 위와 같이 링크기반 검색 방식에서의 스패밍을 줄이기 위해 웹 페이지의 컨텐츠 구조기반 접근법을 사용한다.[4] 또한 link context 를 이용하여 목적지 페이지의 중요도를 상대적으로 조절하여 각 페이지의 링크 순위를 계산하는 방법을 이용하여 스패밍 페이지 필터링을 하는 방법등이 제시되고 있다.[5]

2.3 웹 게시판에서의 스팸과 필터링 기술

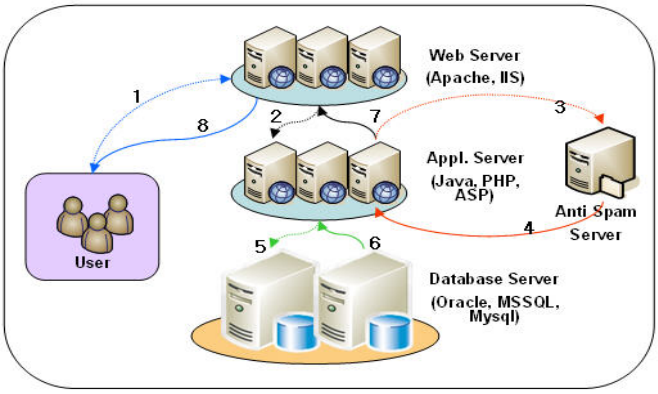
최근의 웹 서비스 사용자 형태가 메일에서 블로그, 개인 홈피 이용률 증가와 함께 메일에서의 안티 스팸 기술의 발달로 인해 블로그나 개인 홈피 게시판에 대한 스팸 공격이 증가하고 있으며, 또한 스팸 공격 방법이 메일의 경우는 사용자의 개인 정보 취득이나, 유해성 싸이트 홍보등에서 특정 웹 싸이트 홍보나 방문을 유도하는 스팸글 공격 형태를 나타낸다. 이러한 스팸 공격에 대한 유해는 메일이나, 검색에서의 스팸과 다른 형태로써 스팸 글이 데이터 베이스에 저장이 되므로 대량의 트래픽 유입으로 서비스 불안정성 가중과 이에 대한 사업자의 안정적 서비스 운영에 대한 비용 증가를 가져 오고 블로그나 개인 홈피 사용자 입장에서는 불필요한 게시글로 인한 깨끗한 인터넷 환경의 저하를 가져올 수 있다. 이에 대한 필터링 방안으로 시간과 접속 리퀘스트 조절하여 효과적으로 방어가 가능한 알고리즘을 응용하여 웹 게시판에 적용하여 필터링 하는 방법을 사용하고자 한다.

3. 웹 게시판에서 스팸글에 대한 필터링 방법

일반적인 웹 게시판은 서비스 관리자가 게시판을 생성하고 방문자가 남긴 글을 읽거나 질문사항에 답변하는 기능으로 주로 이용되어 왔다. 최근에는 다양한 계층의 웹 서비스 접근이 보편화 됨 으로써 게시판의 비속어 증가나 상품 광고, 특정 싸이트 방문 유도등 가상 공간의 익명성을 활용한 게시판 스팸이 증가 하고 있다.



[그림 1] 일반적인 웹 게시판 서비스 구조도

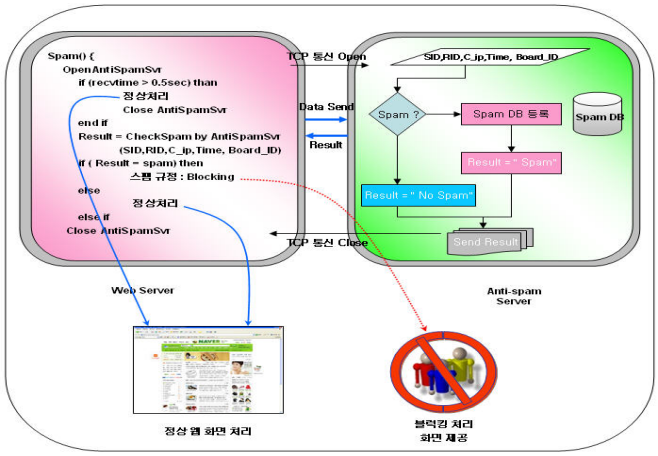


[그림 2] 스팸 필터기능을 포함한 웹 게시판 구조도

스팸 필터 기능을 포함한 구조는 스팸에 대한 정보를 보관하는 서버를 별도로 구성하여 사용자의 요청 정보가 데이터 베이스에 도달하기 전에 판단하기 위해 [그림 2]와 같이 구성한다. 또한 위와 같이 스팸 관리가 가능한 웹 서비스는 NCSA httpd v2.0 기반의 아파치 웹 서비스에 mod_security 모듈[6]을 추가하여 텍스트 기반의 비속어나 특정 단어 매칭을 통한 필터링이 가능하나[7], 이 방법은 한정된 단어와 관리자의 개입이 필수적이며 필터 대상을 추가 할 때는 서비스 중지가 필요하여 대용량의 서비스를 운영하는 곳에서는 불합리한 점이 발생할 수 있다.

위와 같이 안티 스팸 기능은 웹 서비스에 추가 기능으로 제공하는 것과 개별 상황에 적절하게 구축하는 것이 있는데 본 논문에서는 서비스 상황에 적절하게 개발하여 구축하는 것이 대규모 온라인 서비스를 하는 곳에서는 필요하다. 이것은 특정 스팸 사용자가 동일한 형태의 스팸 공격을 하기 때문에 서비스 운영 중 스팸룰 관리나 즉각적 대응으로 안정적 서비스를 목표로 하는 곳에서는 바람직한 구성이다.

일반적인 웹 시스템 구성도에서 추가적으로 안티 스팸이 들어가 있어 사용자의 응답지연 현상이 발생할 수 있는 부분에 대해서는 웹 소스에서 안티 스팸 모듈을 호출할 때 특정 시간 내에 응답이 없을 경우에는 정상처리가 되도록 구성을 함으로써 안티 스팸 서버의 장애나 응답 지연으로 인한 서비스 지연이 발생하지 않도록 구현이 되어야 한다.



[그림 3] 안티 스팸 서버의 내부 흐름도

안티 스팸 서버는 다음과 같은 알고리즘으로 구성되어 있다. 멀티 서비스를 운영하기 위해 서비스 필드와 각 서비스의 정책을 관리하는 룰과 이 룰의 정보를 보관하는 카테고리로 구성되어 있다.

[표 1] 안티 스팸 DB 구성 알고리즘

서비스 (Service)	SID	Sdate	Stime	Edata	Etime	Desc.
룰 (Rule)	SID	RID	R-Type	Block On/Off	Category Count	Desc.
카테고리 (Category)	RID	CID	C-Type	Count	Time	Desc.

위와 같은 안티 스팸 알고리즘으로 상위 주체에 대해 하위 항목을 관리하는 방향으로 구성되어 있다. 또한 안티 스팸에 대해 실 서비스 적용 방안은 다음과 같이 일반 사용자와 스팸을 발생시키는 사용자를 구분하기 위해 룰을 선택한다.

게시판 서비스에 안티 스팸을 적용하기 위한 방법은 다음과 같다. 안티 스팸 서비스에 적용할 대상 게시판에 고유한 SID (Service Identification)을 부여한다. 이 SID 를 기준으로 룰 항목의 RID 와 카테고리 항목의 CID 값이 동일하게 하여 일관성을 갖는 구조로 구현한다. 스팸 구성 알고리즘에서 각각의 항목에 대해 설명을 하면 서비스 항목에서 Rule Count 값은 하나의 서비스에 다중을 Rule 선택이 가능하므로 이에 대한 개수 관리가 필요하며, 룰 항목에서 R-Type 은 해당 서비스에서 로그인 사용자 또는 비 로그인 사용일 경우에는 클라이언트 IP Address 등이 될 수 있으며 또한 룰 항목에서 다중의 카테고리를 선택 할 수 있으므로 카테고리 카운트 항목이 필요하다. 카테고리 항목에서의 C-Type 은 스팸 관리가 특정 시간에 대한 관리(Time)와 특정 트래픽에 대한 관리로 구분 될 수 있다. 또한 룰에 대한 사용자 구분은 로그인 사용자는 로그인 아이디, 비 로그인 사용자는 클라이언트 아이피로 선택 할 수 있으며 대상 항목은 대상 게시판 ID 나 특정 게시물 번호 등이 될 수 있다. 또한 블러킹 대상 시간과 횟수를 선택 가능함으로써 대량의 스팸 트래픽이 유입될 경우에 대한 대응도 가능하다. 위와 같은 방법에 대해 정리를 하면 아래의 표와 같다.

[표 2] 게시판 안티 스팸 서비스 적용 조건

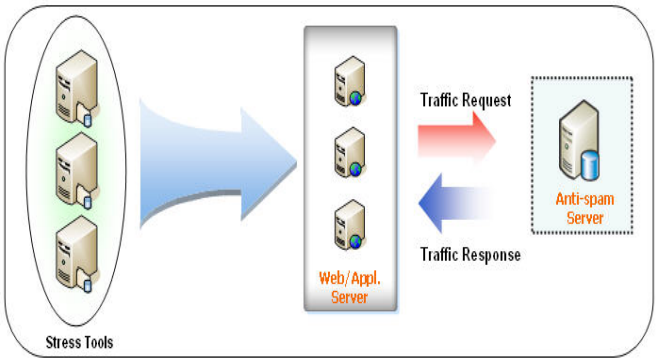
사용자 구분	대상항목	블러킹 항목	구분
로그인 ID, 사용자 IP Address	게시판 ID, 특정 게시물 번호	시간, 횟수	스팸으로 판단할 경우 블러킹 시간

[표 2]와 같은 조건들을 다양한 방법으로 조합을 하면 게시판 서비스에서 정교한 스팸 필터링이 가능하다 예를 들어 특정 로그인 사용자가 특정 게시판에 1 분에 10 회 이상 게시판을 읽거나 쓰는 행위가 발생하면 하루동안 그 게시판에 접근 불가능하게 함. 또는 게시판의 내용 중에 유해단어가 포함된 글을 쓰는 것을 방지 하는 것도 가능하며 제한된 사용자에 대한 블러킹 시간을 유연하게도 가능하다.

위의 서비스 적용 조건에 대한 예를 들면 10 분 동안에 특정 블로그 게시판에 대한 리퀘스트가 50 회 이상이고 20 분내에 동일 클라이언트 IPaddress 에서의 리퀘스트가 20 회 이상이면 해당 블로그 게시판에 대해 해당 클라이언트 IP 주소에 대해 2 시간 동안 블러킹 할 수 있다.

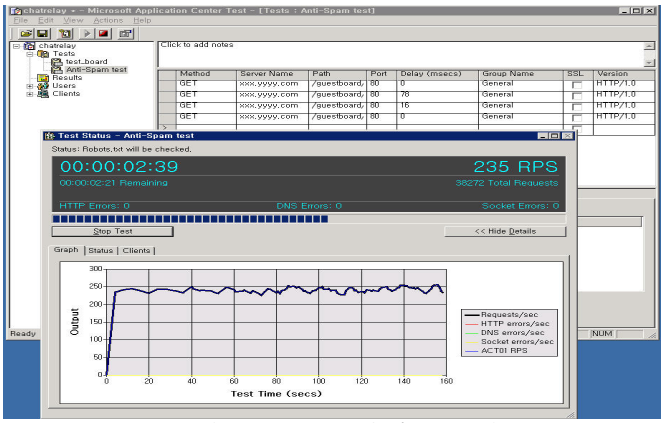
4. 검증

본 논문에서 제시하고자 하는 게시판에서의 안티 스팸 적용은 다음과 같은 테스트를 통하여 안티 스팸 서버의 부하를 측정하며 위의 결과를 토대로 서비스 적용 가능 여부 및 범위를 수립한다.



[그림 4] 부하 테스트 구성도

위의 테스트 구성환경에서 부하테스트 툴은 마이크로 소프트 사의 Application Center Tools(이하 ACT)를 사용한다.[8] ACT 는 웹 서버의 스트레스를 테스트함으로써 가상의 사용자를 임의적으로 부여하여 응용프로그램에 액세스 할 때 웹 서버의 반응을 확인할 수 있다.



[그림 5] ACT 부하테스트 화면

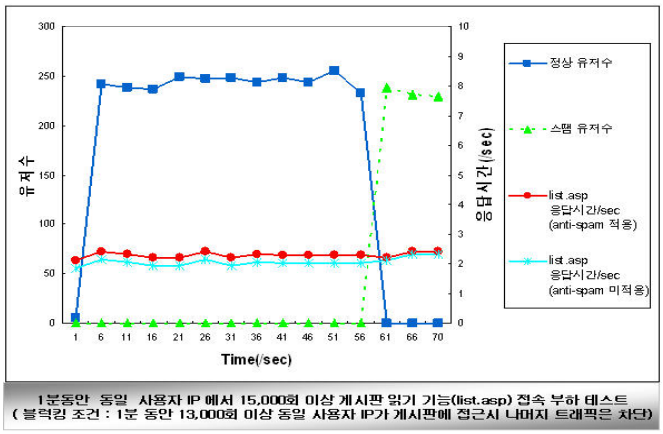
스트레스 툴 서버는 Windows2003 Server 에 ACT 를 설치하고 웹 서버는 Windows 2003 Server 에 IIS 6.0 으로 구성하며, 안티 스팸 서버는 Linux Kernel 2.6 + Berkeley Memory DB 로 구성된다.

테스트는 게시판의 읽기 기능과 쓰기 기능으로 분류하여 단위 시간당 허용시간, 허용 트래픽, 블록킹 및 안티 스팸 서버의 부하로 인해 통과되는 리퀘스트의 량을 측정하여 안티 스팸 서버의 기능 구현과 성

능을 파악할 수 있다.

테스트를 수행하기 전에 안티 스팸 DB 에 서비스, 룰, 카테고리를 등록하여야 한다. 추가 해야 하는 항목은 스팸 디비 알고리즘에 맞추어 다음과 같이 할 수 있다.

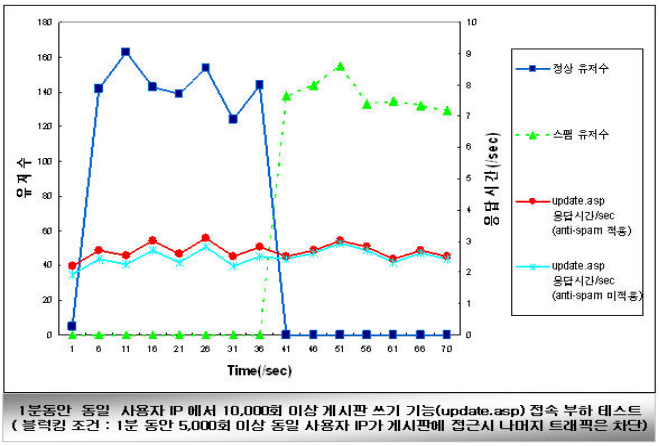
[list.asp : 게시판 읽기 기능에 대한 테스트]
Service : ADD, 10005, 20070701, 20070901, 0, 24, Read
Rule : ADD, 10005, 20005, Client_IP, 0, 1
Category : ADD, 20005, 30005, Time, 0, 0



[그림 6] 게시판 읽기 기능에 대한 안티 스팸 부하 테스트

위의 [그림 6]에서와 같이 안티 스팸 기능을 포함한 게시판에서의 부하 테스트 결과는 블러킹 조건에 만족하면 더 이상 정상 프로세스 처리가 아닌 스팸으로 처리하여 블러킹 카운트가 증가되며 서비스 또한 블러킹 오류 페이지로 전환된다. 또한 해당 읽기 기능을 처리하는 list.asp 파일에 대한 응답 시간은 안티 스팸 적용전에는 평균 2.32 초 이며 안티 스팸 기능을 추가하면 2.57 초로 기능의 유무에 대한 차이는 0.4 초 이내며 스팸 확인후 해당파일의 응답시간은 안티 스팸 서버가 없는 상태의 응답시간에 수렴하며, 전체적인 서비스 지연현상이 미미 하다.

[update.asp : 게시판 쓰기 기능에 대한 테스트]
Service : ADD, 10006, 20070701, 20070901, 0, 24, Write
Rule : ADD, 10006, 20006, Client_IP, 0, 1
Category : ADD, 20006, 30006, Time, 0, 0



[그림 7] 게시판 쓰기 기능에 대한 안티 스팸 부하 테스트

위의[그림 7]은 게시판 쓰기 기능에 대한 부하 테스트이며 update.asp 에 대한 응답시간은 list.asp 보다 약 0.4 초이상 높은 2.71 초이며 스팸 기능 추가전 2.49 초보다 0.22 초 증가되었으며 스팸이 발생하는 부하 테스트 40 초 이후에는 안티 스팸 적용전 응답시간에 수렴하는 형태를 보이며 전체적인 웹 서비스에 대한 지연시간에는 큰 변화는 없다.

5. 결론

본 연구에서는 최근에 많이 사용하고 있는 웹 게시판에서의 스팸에 대한 대응 방안으로 안티 스팸 서버를 사용함으로써 서비스 지연현상을 최소화 하면서 게시판에서의 스팸을 필터링 하는 방안을 제안하였다. 사용자의 IP 주소나 특정 트래픽에 대한 블러킹으로 일반적인 패턴에 대한 방어는 가능하나 프락시 서버를 사용하는 기업이나, 관공서 사용자에 대한 스팸머 구분에 대한 세부적인 방안에 대해서는 향후 연구 과제로 제시되어야 하며, 또한 일반적인 웹 서비스에서의 스팸 블러킹에 대한 적중율을 높일 수 있는 알고리즘에 대한 연구가 필요가 있다.

참고문헌

[1] 김자경, 이광수, "스팸 메일 차단 방법론 비교,분석", 한국정보과학회 2004 년 추계학술대회, 2004

[2] 이신영, 길아라, 김명원, "링크구조분석을 이용한 스팸메일 분류", 한국정보과학회 소프트웨어 및 응용 제 34 권 제 1 호, 2007

[3] Ken "Caesar" Fisher, Google Bombing heating up, Ars Technica Newsdesk, January 22th 2004

[4] 신광섭, 이우기, 강석호, "Web Structure Management 기법을 이용한 Spamming page filtering algorithm", 한국정보과학회 2004 년 춘계학술대회, 2004

[5] 이우기, 신광섭, 강석호, "링크내역을 이용한 페이지점수법 알고리즘", 한국정보과학회 데이터베이스 제 33 권 제 7 호, 2006

[6] Open Secure Web Application Firewall, www.modsecurity.org

[7] 김은정, 조동욱, 이성환, "웹 게시판 유해 단어 현황, 문제점과 해결 방안의 제시", 한국컨텐츠학회/한국통신학회 2003 년 추계학술대회, 2003

[8] Microsoft Co., Microsoft Application Center Test, http://support.microsoft.com/kb/231282, http://www.microsoft.com/downloads/details.aspx?FamilyID=e2c0585a-062a-439e-a67d-75a89aa36495&DisplayLang=en