(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년01월25일
(11) 등록번호 10-1009482
(24) 등록일자 2011년01월12일

(51) Int. Cl.
G06F 21/20 (2006.01) *H04L 12/22* (2006.01)
G06F 11/30 (2006.01) *G06F 17/10* (2006.01)
(21) 출원번호 10-2009-0013412
(22) 출원일자 2009년02월18일
심사청구일자 2009년02월18일
(65) 공개번호 10-2010-0049471
(43) 공개일자 2010년05월12일
(30) 우선권주장
1020080108352 2008년11월03일 대한민국(KR)
(56) 선행기술조사문헌
KR100745613 B1*
KR1020070050726 A
한국컴퓨터종합학술대회2005 박현도 외1인, "랜덤성을 이용한 알려지지 않은 신종 웜에 대한 탐지 기법, 논문집 Vol.32. No.1 pp. 133-135.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
고려대학교 산학협력단

(72) 발명자
이희조

박현도

(74) 대리인
유미특허법인

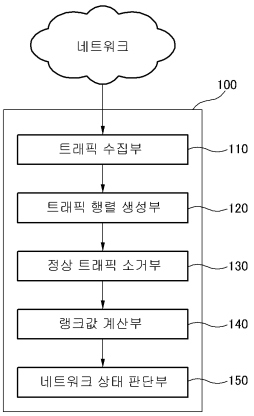
전체 청구항 수 : 총 16 항 심사관 : 양찬호

(54) 웜 탐지 방법 및 장치

(57) 요 약

네트워크 상에서 인터넷 웜을 탐지하는 방법 및 장치로서, 패턴 매칭 기반의 웜 탐지 방법이나 행위 기반의 웜 탐지 방법을 이용하지 않고도 네트워크 트래픽에 발생한 랜덤성의 변화를 기반으로 적은 연산량만으로도 새로운 유형의 신종 인터넷 웜을 조기에 탐지할 수 있는 웜 탐지 방법 및 장치가 개시된다.

대 표 도 - 도1



특허청구의 범위

청구항 1

네트워크로 유입되는 트래픽을 수집하는 트래픽 수집단계와;

제1 시각에 개시되는 제1 시간 영역에서의 상기 트래픽의 특성을 나타낸 제1 트래픽 행렬과, 상기 제1 시간 영역의 종료시각인 제2 시각에 개시되는 제2 시간 영역에서의 상기 트래픽의 특성을 나타낸 제2 트래픽 행렬과, 상기 제2 시간 영역의 종료시각인 제3 시각에 개시되는 제3 시간 영역에서의 상기 트래픽의 특성을 나타낸 제3 트래픽 행렬을 생성하는 트래픽 행렬 생성단계와;

상기 제1 시간 영역, 상기 제2 시간 영역 및 상기 제3 시간 영역에서의 정상적인 플로우에 대응되는 행렬 성분을 소거하되, 상기 정상적인 플로우 중 상기 제2 시각의 정각에 종료되거나 개시된 플로우 및 상기 제3 시각에 종료된 플로우에 대응되는 행렬 성분도 소거하여 정상 트래픽 소거 행렬을 생성하는 정상 트래픽 소거 단계와;

상기 정상 트래픽 소거 행렬의 랭크값을 계산하는 랭크값 계산 단계 및;

상기 랭크값을 바탕으로 상기 네트워크의 상태를 판단하는 네트워크 상태 판단 단계를 포함하는 웜 탐지 방법.

청구항 2

제1항에 있어서,

상기 정상 트래픽 소거 단계는,

상기 제2 트래픽 행렬과 상기 제3 트래픽 행렬을 배타적 논리합하여 제1 행렬을 얻는 단계와;

상기 제1 행렬과 상기 제1 트래픽 행렬을 논리곱하여 제2 행렬을 얻는 단계를 더 포함하며,이 때 상기 정상 트래픽 소거 행렬은 상기 제1 행렬과 상기 제2 행렬을 배타적 논리합한 행렬인 웜탐지방법.

청구항 3

제1항에 있어서,

상기 트래픽 행렬 생성 단계에서, 상기 트래픽 수집단계에서 특정 플로우가 수집된 경우, 상기 제1 트래픽 행렬, 상기 제2 트래픽 행렬 및 상기 제3 트래픽 행렬의 상기 플로우의 IP 주소에 대응되는 행렬 성분이 0으로부터 1로 변경되는 웜 탐지 방법.

청구항 4

제1항에 있어서,

상기 랭크값 계산 단계에서, 상기 랭크값은 상기 정상 트래픽 소거 행렬에 가우스 소거법을 적용하여 얻어지는 0이 아닌 행의 개수인 웜 탐지 방법.

청구항 5

제4항에 있어서,

상기 랭크값 계산 단계에서, 상기 랭크값을 r, 상기 정상 트래픽 소거 행렬이 랜덤 행렬이 아닐 확률을 P, 상기 정상 트래픽 소거 행렬의 행의 크기 및 열의 크기를 각각 m 및 n이라고 할 때, 상기 랭크값은

$$P = 2^{r(n+m-r)-nm} \prod_{i=0}^{r-1} \frac{(1-2^{i-n})(1-2^{i-m})}{(1-2^{i-r})} \quad \text{where } r = 1, 2, \ldots, \min(m, n)$$

에 의하여 얻어지는 웜 탐지 방법.

청구항 6

제1항에 있어서,

상기 네트워크 상태 판단단계에서, 상기 랭크값이 소정의 기준값보다 큰 경우 상기 네트워크는 웜에 감염된 것

으로 판단되는 웜 탐지 방법.

청구항 7

제1항에 있어서,

상기 네트워크 상태 판단단계에서, 상기 랭크값이 0에 근접한 값으로 급격하게 감소하는 경우, 상기 네트워크는 웜에 감염된 것으로 판단되는 웜 탐지 방법.

청구항 8

제6항에 있어서,

상기 기준값은 상기 정상 트래픽 소거 행렬이 랜덤 행렬이 아닐 특정 확률에 대한 최대 랭크값인 웜 탐지 방법.

청구항 9

네트워크로 유입되는 트래픽을 수집하는 트래픽 수집부와;

제1 시각에 개시되는 제1 시간 영역에서의 상기 트래픽의 특성을 나타낸 제1 트래픽 행렬과, 상기 제1 시간 영역의 종료시각인 제2 시각에 개시되는 제2 시간 영역에서의 상기 트래픽의 특성을 나타낸 제2 트래픽 행렬과, 상기 제2 시간 영역의 종료시각인 제3 시각에 개시되는 제3 시간 영역에서의 상기 트래픽의 특성을 나타낸 제3 트래픽 행렬을 생성하는 트래픽 행렬 생성부와;

상기 제1 시간 영역, 상기 제2 시간 영역 및 상기 제3 시간 영역에서의 정상적인 플로우에 대응되는 행렬 성분을 소거하되, 상기 정상적인 플로우 중 상기 제2 시각에 종료되거나 개시된 플로우 및 상기 제3 시각에 종료된 플로우에 대응되는 행렬 성분도 소거하여 정상 트래픽 소거 행렬을 생성하는 정상 트래픽 소거부와;

상기 정상 트래픽 소거 행렬의 랭크값을 계산하는 랭크값 계산부 및;

상기 랭크값을 바탕으로 상기 네트워크의 상태를 판단하는 네트워크 상태 판단부를 포함하는 웜 탐지 장치.

청구항 10

제9항에 있어서,

상기 정상 트래픽 소거부는,

상기 제2 트래픽 행렬과 상기 제3 트래픽 행렬을 배타적 논리합하여 제1 행렬을 얻고, 상기 제1 행렬과 상기 제1 트래픽 행렬을 논리곱하여 제2 행렬을 얻은 후,

상기 제1 행렬과 상기 제2 행렬을 배타적 논리합하여 상기 정상 트래픽 소거 행렬을 얻는 웜 탐지 장치.

청구항 11

제9항에 있어서,

상기 트래픽 행렬 생성부는, 상기 트래픽 수집단계에서 특정 플로우가 수집된 경우, 상기 제1 트래픽 행렬, 상기 제2 트래픽 행렬 및 상기 제3 트래픽 행렬의 상기 플로우의 IP 주소에 대응되는 행렬 성분을 0으로부터 1로 변경하는 웜 탐지 장치.

청구항 12

제9항에 있어서,

상기 랭크값 계산부는, 상기 정상 트래픽 소거 행렬에 가우스 소거법을 적용하여 얻어지는 0이 아닌 행의 개수를 상기 랭크값으로 계산하는 웜 탐지 장치.

청구항 13

제12항에 있어서,

상기 랭크값 계산부에서, 상기 랭크값을 r, 상기 정상 트래픽 소거 행렬이 랜덤 행렬이 아닐 확률을 P, 상기 정상 트래픽 소거 행렬의 행의 크기 및 열의 크기를 각각 m 및 n이라고 할 때, 상기 랭크값은,

$$P = 2^{r(n+m-r)-nm} \prod_{i=0}^{r-1} \frac{(1-2^{i-n})(1-2^{i-m})}{(1-2^{i-r})} \quad \text{where } r = 1, 2, \ldots, \min(m, n)$$

에 의하여 얻어지는 웜 탐지 장치.

청구항 14

제9항에 있어서,

상기 네트워크 상태 판단부는, 상기 랭크값이 소정의 기준값보다 큰 경우상기 네트워크가 웜에 감염된 것으로 판단하는 웜 탐지 장치.

청구항 15

제9항에 있어서,

상기 네트워크 상태 판단부는, 상기 랭크값이 급속히 0으로 수렴하는 경우, 상기 네트워크가 웜에 감염된 것으로 판단하는 웜 탐지 장치.

청구항 16

제14항에 있어서, 상기 기준값은 상기 정상 트래픽 소거 행렬이 랜덤 행렬이 아닐 확률에 대한 최대 랭크값인 웜 탐지 장치.

명 세 서

발명의 상세한 설명

기 술 분 야

[0001] 본 발명은 네트워크 상에서 인터넷 웜을 탐지하는 방법에 관한 것으로서, 보다 상세하게는 적은 연산량만으로 새로운 유형의 신종 인터넷 웜을 조기에 탐지할 수 있는 웜 탐지 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 네트워크에서 웜을 탐지하기 위한 백신 프로그램이나 IDS, IPS, Firewall과 같은 기술들은 많다.

[0003] 그러나, 대규모 네트워크에서 적은 연산량과 작은 메모리 공간을 사용하여 웜을 탐지하는 데에는 여전히 부족함이 많다. 웜을 탐지하고, 검출된 웜에 적절한 대응조치를 행하는 대부분의 보안 기술들은 많은 연산량과 거대한 메모리 공간을 필요로 한다.

[0004] 한편, 웜을 탐지하는 기존의 방식을 크게 나누면, 패턴 매칭을 기반으로 웜을 탐지하는 기법과 웜의 행위를 기반으로 웜을 탐지하는 기법들이 있다. 이들 각각의 단점은 다음과 같다.

[0005] 먼저, 패턴 매칭을 통한 웜 탐지 기법들은, 알려지지 않은 신종 웜을 탐지하지 못한다는 단점이 있다.

[0006] 다음으로, 웜의 행위를 기반으로 웜을 탐지하는 기법들은 오탐률이 높으며, 또한 웜을 탐지하기 위하여 네트워크의 많은 정보들을 이용하여야 하므로 탐지를 위하여 많은 연산량을 필요로 한다는 단점이 있다.

[0007] 한편, 행위 기반 웜 탐지 기법 중 한 가지 유형으로서 네트워크의 엔트로피를 이용하는 기법들이 있지만, 엔트로피를 이용하는 기법들 또한 많은 연산량을 필요로 하므로 대용량, 고속의 네트워크, 예컨대 백본 네트워크에 적용하기 어렵다.

[0008] 이처럼, 현재의 웜 탐지 기법들은 대규모 네트워크에서 효율적으로 신종 웜을 탐지하지 못하는 문제점이 있었다.

발명의 내용

해결 하고자하는 과제

[0009] 본 발명은 상기와 같은 문제점을 해결하기 위하여 착안된 것으로서, 상세하게는 적은 연산량만으로도 대규모 네트워크에서 새로운 유형의 신종 인터넷 웜을 조기에 탐지할 수 있는 웜 탐지 방법 및 장치를 제공하는 것을 목적으로 한다.

과제 해결수단

[0010] 상기와 같은 목적을 달성하기 위하여, 본 발명에 의한 웜 탐지 방법은, 네트워크로 유입되는 트래픽을 수집하는 트래픽 수집단계와; 제1 시각에 개시되는 제1 시간 영역에서의 상기 트래픽의 특성을 나타낸 제1 트래픽 행렬과, 상기 제1 시간 영역의 종료시각인 제2 시각에 개시되는 제2 시간 영역에서의 상기 트래픽의 특성을 나타낸 제2 트래픽 행렬과, 상기 제2 시간 영역의 종료시각인 제3 시각에 개시되는 제3 시간 영역에서의 상기 트래픽의 특성을 나타낸 제3 트래픽 행렬을 생성하는 트래픽 행렬 생성단계와; 상기 제1 시간 영역, 상기 제2 시간 영역 및 상기 제3 시간 영역에서의 정상적인 플로우에 대응되는 성분을 소거하되, 상기 정상적인 플로우 중 상기 제2 시각에 종료되거나 개시된 것에 대응되는 성분과 상기 제1 시각에 개시되어 상기 제3 시각에 종료된 성분도 소거하여 정상 트래픽 소거 행렬을 생성하는 정상 트래픽 소거 단계와; 상기 정상 트래픽 소거 행렬의 랭크값을 계산하는 랭크값 계산 단계 및; 상기 랭크값을 바탕으로 상기 네트워크의 상태를 판단하는 네트워크 상태 판단 단계를 포함한다.

[0011] 또한, 본 발명의 다른 측면에 의한 웜 탐지 장치는, 네트워크로 유입되는 트래픽을 수집하는 트래픽 수집부와; 제1 시각에 개시되는 제1 시간 영역에서의 상기 트래픽의 특성을 나타낸 제1 트래픽 행렬과, 상기 제1 시간 영역의 종료시각인 제2 시각에 개시되는 제2 시간 영역에서의 상기 트래픽의 특성을 나타낸 제2 트래픽 행렬과, 상기 제2 시간 영역의 종료시각인 제3 시각에 개시되는 제3 시간 영역에서의 상기 트래픽의 특성을 나타낸 제3 트래픽 행렬을 생성하는 트래픽 행렬 생성부와; 상기 제1 시간 영역과 상기 제2 시간 영역 및 상기 제3 시간 영역에서의 정상적인 플로우에 대응되는 성분을 소거하되, 상기 정상적인 플로우 중 상기 제2 시각에 종료되거나 개시된 것에 대응되는 성분과 상기 제 1시각에 개시되어 상기 제 3시각에 종료된 성분도 소거하여 정상 트래픽 소거 행렬을 생성하는 정상 트래픽 소거부와; 상기 정상 트래픽 소거 행렬의 랭크값을 계산하는 랭크값 계산부 및; 상기 랭크값을 바탕으로 상기 네트워크의 상태를 판단하는 네트워크 상태 판단부를 포함한다.

효 과

[0012] 본 발명에 의한 웜 탐지 방법 및 장치를 이용하면, 적은 연산량만으로도 대규모 네트워크에서 새로운 유형의 신종 인터넷 웜을 더욱 향상된 정확도로 조기에 탐지할 수 있는 효과가 있다.

[0013] 또한, 본 발명에 의한 웜 탐지 방법 및 장치를 이용하면, 패턴 매칭 기반의 웜 탐지 방법이나 행위 기반의 웜 탐지 방법을 이용하지 않고도 네트워크 트래픽에 발생한 랜덤성의 변화를 기반으로 웜을 탐지하는 새롭고 정확성이 한층 향상된 방법을 이용할 수 있는 효과가 있다.

발명의 실시를 위한 구체적인 내용

[0014] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0015] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "…부", "…기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0016] 이제 본 발명의 다양한 실시예에 따른 웜 탐지 방법 및 장치를 이하의 도면들을 참조하여 상세하게 설명한다.

[0017] **Ⅰ. 웜 탐지 장치의 구성**

[0018] 도 1은 본 발명의 실시예에 따른 웜 탐지 장치를 도시한 블록도이다.

[0019] 도 1에 나타낸 바와 같이, 본 발명의 실시예에 따른 웜 탐지 장치는, 제1 네트워크(10)와 연동하며, 트래픽 수

집부(110), 트래픽 행렬 생성부(120), 정상 트래픽 소거부(130), 랭크값 계산부(140) 및 네트워크 상태 판단부(150)를 포함한다.

[0020] 웜 탐지 장치의 각각의 구성요소의 역할을 간단히 설명하면 다음과 같다.

[0021] 먼저, 트래픽 수집부(110)는 네트워크(10)로 유입되는 트래픽을 수집한다.

[0022] 다음으로, 트래픽 행렬 생성부(120)는 트래픽 수집부(110)가 수집한 트래픽의 특성을 행렬로 표현한다. 이 때, 트래픽 행렬 생성부(120)는 트래픽에 포함된 각각의 플로우(패킷)의 목적지 IP(Internet Protocol) 주소를 바탕으로 트래픽 행렬 내에서 해당 플로우가 배치되는 성분을 결정한다.

[0023] 정상 트래픽 소거부(130)는 인터넷 웜의 전파로 인하여 발생한 트래픽이 아닌 정상적인 트래픽을 트래픽 행렬에서 소거한다.

[0024] 랭크값 계산부(140)는 정상 트래픽 소거부에서 정상적인 트래픽을 소거한 결과인 정상 트래픽 소거 행렬의 랭크값(rank)을 계산함으로써 랜덤성 정도를 측정한다.

[0025] 마지막으로, 네트워크 상태 판단부(150)에서는 랭크값 계산부(140)에서 산출된 랭크값을 바탕으로 해당 네트워크가 인터넷 웜의 침입을 받았는지 여부 등의 네트워크 상태를 판단한다.

[0026] 이하에서는 본 발명의 실시예에 따른 웜 탐지 장치에서 이루어지는 웜 탐지 방법을 순차적으로 더욱 상세히 설명하기로 한다.

[0027] **II. 트래픽 행렬 생성부에서의 트래픽 행렬의 구축**

[0028] 트래픽에 무작위성 체크를 적용하기 위하여, 트래픽을 행렬의 형태로 변환하게 된다. 특히, IP 주소가 공격 트래픽의 내에서 무작위성을 가지는 행렬에 표현될 필요가 있다. 또한, 웜은 계속하여 무작위로 목표물을 선택하므로, 인터넷 웜의 활동성을 포착하기 위하여, 경유하는 트래픽의 목적지 IP 주소를 보아야 한다. 오늘날의 인터넷 웜은 난수발생기를 사용하여 전파의 속도를 최대화함과 동시에 검출을 회피하고 있다.

[0029] 트래픽 행렬 구축 설계 단계에서는, 활동중인 웜이 채택한 특정 전파전략에 따라 IP 주소 내의 4개의 옥텟들의 양태가 다르게 됨에 주목하여야 한다. 인터넷 웜은 상이한 스캐닝 전략들을 사용하여 왔다. 네트워크 내의 트래픽에 있어서 IP 주소의 각각의 옥텟을 **IP1, IP2, IP3, IP4**라고 할 때, IP 주소는 아래 식과 같이 나타낼 수 있다.

수학식 1

[0030] **IP=IP1.IP2.IP3.IP4**

[0031] 도 2는 균일 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면이다.

[0032] 슬래머 웜과 코드레드 웜은 랜덤 스캔 기법을 사용하는데, 이는 4개의 옥텟(**IP1 ~ IP4**) 모두를 무작위로 선택한다. 이러한 기법은 “균일 스캔(uniform scan)”이라고도 일컫는다.

[0033] 도 3은 서브넷 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면이다.

[0034] 코드레드II와 같은 인터넷 웜의 경우에는 “서브넷 스캔(subnet scan)”이라고도 하는 부분적 선호도 기법을 사용한다. 이 웜은 랜덤 스캔을 1/8의 확률로 수행한다. 그리고, 1/2의 확률로 동일한 IP1을 유지한다. 또한, 3/8의 확률로 동일한 **IP1.IP2**를 유지한다.

[0035] 따라서, 코드레드II의 스캐닝 기법에서는 **IP3**과 **IP4**가 전부 무작위적이지만, **IP1**과 **IP2**에서는 부분적으로만 무작위적이다.

[0036] 도 4는 순차적 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면이다.

[0037] 순차적 스캔은 예컨대 블라스터 웜이 채용하고 있는 스캔 기법이다.

[0038] 블라스터 웜은 하나의 **IP1.IP2**만을 무작위로 선택하고, 이후의 타겟은 클래스 B 네트워크 내에서 다른 타겟 네

트워크를 선택할 때까지 순차적으로 스캔을 수행한다. 따라서 IP1과 IP2는 무작위적이지만, IP3와 IP4의 분포는 순차적이다.

[0039] 그러나, 공격당한 네트워크에서 주목할 점은, 스캔한 트래픽의 목적지 IP주소의 분포는 무작위적이지 않고 순차적일 수 있다는 것이다.

[0040] 이하에서는 편의상 "균일 스캔" 및 "서브넷 스캔"을 단순히 "랜덤 스캔"이라고 하기로 한다. 이들 모두 IP 주소의 일부를 무작위로 가지기 때문이다. 랜덤" 유형과 "순차적" 유형을 아래의 인터넷 웜의 스캐닝 기법을 나타내기 위하여 사용하기로 한다.

[0041] 본 발명의 실시예에서, 트래픽 행렬은 256개의 행과 256개의 열로 되어있으며, 트래픽 행렬의 각 성분은 1비트의 크기를 갖는다.

[0042] 한편, 이처럼 트래픽 행렬이 비교적 대규모인 256 ×256 이진 행렬이 되는 경우에도, 컴퓨터에서 트래픽 행렬을 처리하는 데에는 단지 8Kbyte의 작은 메모리 공간만을 필요로 한다.

[0043] 256 ×256 트래픽 행렬은 최대 65,536개의 구별되는 목적지 IP 주소를 가질 수 있다. 이 때, 웜에 의하여 사용되는 IP 주소의 랜덤성을 유지하기 위해서는, 해당 트래픽에 포함된 각각의 플로우가 가지는 목적지 IP 주소를 트래픽 행렬의 성분에 적절히 대응시켜야 한다.

[0044] i,j를 각각 행 인덱스와 열 인덱스라고 할 때, 트래픽에 포함된 각각의 플로우의 목적지 IP 주소와 트래픽 행렬의 성분과의 대응관계는 아래와 식과 같이 나타낼 수 있다.

수학식 2

[0045] $$i = \mathrm{IP}_1 \oplus \mathrm{IP}_3, \quad j = \mathrm{IP}_2 \oplus \mathrm{IP}_4$$

[0046] 연산기호 $\oplus$ 는 배타적 논리합(XOR: exclusive OR) 연산을 의미한다.

[0047] 한편, 인터넷 웜에서 실제로 사용되는 스캐닝 기법에서는 랜덤성이 높게 유지된다는 점에 주목할 필요가 있다.

[0048] 순차적 스캔 웜의 경우에는, 매핑 함수가 XOR에 따른 치환(permutation)을 수행하지만, 일단 매핑이 각각의 행(XOR이 수행된 IP3)을 선택하고 나면 치환된 순서대로 해당 행에 1을 순서대로 채워 넣는다(XOR이 수행된 IP4). 결국, 해당 행은 모두 1인 엔트리를 가지며, 시간이 흐름에 따라 웜의 스캔 트래픽은 증가하게 됨으로써 모두 1인 엔트리를 가지는 행은 순차적으로 많아진다.

[0049] 이러한 행들은 랭크값을 계산하는 랭크값계산부에서 수행하는 가우시안소거법(Gaussian Elimination)에 의해 모두 0으로 가게 됨으로서 랭크값은 극단적으로 줄어들어 결국 랭크값은 0에 가깝게 되는 결과를 초래한다.

[0050] 랜덤 스캔 웜의 경우에는, 트래픽 행렬에 랜덤한 엔트리들이 증가하게 됨으로써 랜덤성 정도가 유지된다.

[0051] 만약 랜덤 스캔 웜과 순차적 스캔 웜이 복합적으로 발생하는 경우에는 순차적 스캔 웜의 경우와 같이 랭크값이 0에 가깝게 극단적으로 줄어든다. 이유는 본 기술이 트래픽 행렬을 생성하는 트래픽 행렬 생성부(120)에서는 매핑 함수가 트래픽 행렬의 엔트리를 기록할 때 중복기록(overwrite)을 허용하기 때문에 웜에 감염된 호스트가 늘어남에 따라 행렬에는 순차적 스캔 웜에 의한 트래픽이 기록되기 때문에 순차적 스캔 웜에 의한 랭크값이 극단적으로 0에 가깝게 되는 것과 동일한 결과를 초래한다.

[0052] 일단 매핑이 정의되면, 웜 탐지 장치(100)의 트래픽 행렬 생성부(120)에서 이루어지는 트래픽 행렬의 구축은 용이하다.

[0053] 각 단위 시간 영역이 개시되는 시점에, 트래픽 행렬은 0으로 채워진다. 그 후 유입되는 플로우(패킷)의 각각의 목적지 IP 주소에 대하여 대응되는 엔트리가 1로 변경된다. 행렬은 이런 식으로 단위 시간 영역이 종료할 때까지 계속하여 채워진다.

[0054] 이 때, 단위 시간 영역의 지속시간은 작업환경에 좌우되나, 전형적인 값은 초 단위로서, 예컨대 1초, 10초 등이 된다. 본 실시예에서는 단위 시간 영역을 1초로 설정하고 있는데, 이는 인터넷 웜에 감염됨으로써 발생하는 비

정상적인 플로우의 지속시간이 대개 1초 미만이지만, 정상적인 플로우의 지속시간은 대개 1초보다 크다는 점에 착안한 것이다.

[0055] **III. 트래픽 필터링 행렬의 동작**

[0056] 단위 시간 영역에 대한 트래픽 행렬의 구축이 완료되면, 웜 탐지 장치(100)의 정상 트래픽 소거부(130)에서는 트래픽 내의 정상적인 플로우가 가능한 한 많이 제거됨으로써, 침입 탐지의 정확성 향상을 꾀하게 된다.

[0057] 정상 트래픽 소거부(130)에서는 트래픽 행력에 대하여 아래에서 설명하는 바와 같은 연산을 수행함으로써 정상적인 트래픽을 필터링하게 된다.

[0058] 특히, 정상 트래픽 소거부(130)에서는 연속적인 시간 단위의 두 트래픽 행렬에 대한 비트 당 XOR 동작을 행함으로써, 연속적인 두 트래픽 행렬에 존재하던 정상적인 플로우의 대부분을 제거하고, 가장 의심되는 트래픽만을 결과 행렬에 남긴다. 또한, 지속시간이 긴 정상적인 트래픽에 대해서 2개의 연속되는 행렬 상에서 비트 당 AND 동작을 수행하게 된다.

[0059] 본 출원인은 적은 연산량만으로도 대규모 네트워크에서 새로운 유형의 신종 인터넷 웜을 조기에 탐지할 수 있는 웜 탐지 방법을 출원하여 등록받은 바 있다(등록 제745613호 “네트워크 감시를 수행하는 장치 및 프로그램이 저장된 기록매체”).

[0060] 이에 따르면, 네트워크를 통한 유입 트래픽과 유출 트래픽의 특성이 반영된 트래픽 행렬의 랭크값을 이용하여 네트워크의 구체적인 상태를 파악하도록 하고 있다. 즉, 트래픽 행렬의 랭크값이 소정의 정상 범위를 초과하는 경우에는 인터넷 웜의 공격을 받고 있는 것으로 판단할 수 있으므로, 조기에 웜의 공격을 발견하여 대응할 수 있는 것이다.

[0061] 본 출원인의 기존 발명에 따른 웜 탐지 방법은 상당한 정확도로써 웜의 네트워크 침입을 조기 탐지할 수 있는 효과가 있었다.

[0062] 그러나, 본 출원인의 기존 발명에서는, 정상적인 플로우가 시각 (t-1)에 개시되어 시각 t에 종료되는 특정 시구간의 개시 시점과 동시(시각 (t-1) 정각)에 새로이 시작되거나, 또는 해당 시구간의 종료 시점(시각 t 정각)에 종료되는 경우에는, 이를 행렬에서 제거하지 못하는 문제점이 있었다.

[0063] 이에, 본 발명의 실시에에 의한 웜 탐지 장치(100)에서는 정상 트래픽 소거부(130)에서 정상적인 플로우를 제거하는 필터링을 수행함에 있어서, 정상적인 플로우 중 기존 발명에서 제거하지 못했던 플로우도 제거할 수 있는 새로운 필터링 메커니즘을 적용하여 웜 탐지의 정확도를 제고하였다.

[0064] 도 5는 본 발명의 실시예에 의한 웜 탐지 방법에서 사용되는 필터링 및 랭크값 측정의 메커니즘을 나타낸 도면이다.

[0065] M(t)를 시각t부터 시각(t+1) 까지의 시간 영역에서 수집된 트래픽으로 구축된 트래픽 행렬로 정의하기로 한다.

[0066] 이 때, M(t-1)은 시각(t-1)과 시각(t) 사이의 시간 영역에서 수집된 트래픽에 대응되는 트래픽 행렬이 되며, 마찬가지로 M(t-2)는 시각 (t-2)와 시각(t-1) 사이의 시간 영역에서 수집된 트래픽에 대응되는 트래픽 행렬이 된다.

[0067] 또한, $M(t) \oplus M(t-1)$ 은 M(t)와 M(t-1) 내의 대응되는 성분(entry)에 대하여 배타적 논리합(XOR) 연산이 행해진 후의 행렬을 나타낸다.

[0068] XOR 연산은 중복 엔트리를 제거하기 위한 것인데, 중복 엔트리는 일반적으로 문제가 되지는 않는다. 예컨대, 4개의 정상적인 플로우가 있고, 웜에 의하여 단위시간당 1개의 패킷이 생성되는 경우에, XOR 동작이 단순화된 4×4 행렬에 적용된 필터링의 결과는 아래 수학식과 같다.

수학식 3

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

[0069]

[0070] 즉, 수학식 3에서 보는 바와 같이, 정상적인 플로우는 일반적으로 지속시간이 1초보다 길기 때문에, 인접 시간 영역의 두 행렬 내의 동일한 위치에서 모두 1로 나타나게 된다. 따라서, 배타적 논리합 연산의 수행 결과로 남은 1의 성분은 대부분 인터넷 웜의 전파로 인하여 발생한 비정상적인 패킷으로 볼 수 있다.

[0071] 네트워크가 정상적인 상태에서는, 정상적인 신규 플로우가 시각 t에 급격히 네트워크로 쇄도하지 않는 이상, 행렬 $M(t) \oplus M(t-1)$ 의 랭크값이 M(t)의 랭크값보다 작게 나타난다.

[0072] 반면, 인터넷 웜이 네트워크를 통하여 전파됨으로써 랜덤 스캔 트래픽이 증가하는 경우에는, $M(t) \oplus M(t-1)$ 의 랭크값이 M(t)의 랭크값보다 한층 높게 나타난다.

[0073] 한편, 배타적 논리곱 연산의 수행을 통하여 해당 시간 영역 내의 정상적인 플로우가 대부분 소거된 이후에도, 정상적인 플로우 중 일부는 트래픽 행렬 내에 잔존할 수 있다.

[0074] 예컨대, 시간 영역과 시간 영역이 구분되는 특정 시각에 새로이 개시되거나 종료되는 플로우의 경우가 그러하다.

[0075] 이러한 플로우까지도 제거하기 위하여, 본 발명의 실시예에서는 정상 트래픽소거부(130)에서 아래 수학식과 같은 행렬 연산이 수행됨으로써 정상적인 플로우를 더욱 효과적으로 소거하게 된다.

수학식 4

$$M'(t) = M_{XOR}(t) \oplus (M_{XOR}(t) \bullet M(t-2))$$

[0076]

[0077] 위 수학식에서 $M'(t)$를 정상 트래픽 소거 행렬이라 하기로 한다.

[0078] 이 때, $M_{XOR}(t)$ 는 아래 수학식과 같은 두 개의 연속적인 행렬에 의한 XOR 동작을 나타낸다.

수학식 5

$$M_{XOR}(t) = M(t) \oplus M(t-1)$$

[0079]

[0080] 어떤 정상적인 트래픽이 시각(t-1) 또는 (t)에서 종료되는 경우, 해당 트래픽은 $(M_{XOR}(t) \bullet M(t-2))$ 에 포함될 수 있지만, $M'(t)$에서 제거된다.

[0081] 도 5에서, 둥근 점은 정상적인 트래픽을, 십자기호는 네트워크에 침입한 웜에 의하여 발생한 비정상적인 트래픽을 각각 나타낸다.

[0082] 도 5의 최하단에서 나타낸 바와 같이, 본 발명의 실시예에 따른 필터링 수행 후의 트래픽 행렬인 정상 트래픽 소거 행렬에는 비정상적인 플로우만이 남게 된다. 정상 트래픽 소거 행렬은 이후 해당 시구간에 대한 랭크값을 측정하는 데에 사용된다.

[0083] 도 6은 본 발명의 실시예에 따른 필터링의 원리와 트래픽 행렬의 관계를 나타낸 벤다이어그램이다.

[0084] 도 6에서, 시각 t에서 개시되는 접속이 정상적인 플로우인지 아닌지 여부는 시각 (t+1)이 될 때까지는 판단할 수 없으므로, M'(t)에서는 그 플로우가 삭제되어서는 안된다. 또한, 정상적인 플로우일 가능성이 높은 정각(t-1) 및 정각(t)에서의 플로우가 제거된 형태를 확인할 수 있다.

[0085] **Ⅳ. 랭크값의 측정**

[0086] 랭크값 계산부(140)에서는 정상 트래픽 소거부에서 정상적인 트래픽을 소거한 결과인 정상 트래픽 소거 행렬의 랭크값(rank)을 계산함으로써 랜덤성 정도가 측정된다.

[0087] 랜덤한 m×n 이진 행렬의 랭크값은 다음 수학식 6에 따르는 확률분포를 가진다.

수학식 6

[0088] $$P = 2^{r(n+m-r)-nm} \prod_{i=0}^{r-1} \frac{(1-2^{i-n})(1-2^{i-m})}{(1-2^{i-r})} \quad \text{where } r = 1, 2, \ldots, \min(m, n)$$

[0089] 수학식 6에 2를 밑으로 하는 로그를 취하여 랭크값을 계산하면, 수학식 7과 같은 관계식을 얻게 된다.

수학식 7

[0090] $$(m-r)^2 > \log_2 \frac{1}{P}$$

[0091] 즉, 확률(P)을 거의 0에 가까운 값인 0.001%라고 가정할 때, 256×256 이진행렬에서는 252가 가장 큰 랭크값이 된다.

[0092] 다시 말해서, 기준값은 정상 트래픽 소거 행렬이 랜덤 행렬이 아닐 특정 확률(P)에 대한 최대의 랭크값이 된다.

[0093] 예컨대, 어떠한 256×256 이진행렬이 252 이상의 랭크값을 가지는 경우,이 행렬은 99.999%의 확률로 랜덤 행렬임을 알 수 있다.

[0094] 따라서, 웜 탐지 장치(100)의 네트워크 상태 판단부(150)는 랭크값 계산부(140)에서 계산된 정상 트래픽 소거 행렬의 랭크값이 소정의 값을 넘는 경우, 예컨대 256×256 의 경우에 랭크값이 252보다 큰 경우에는 해당 네크워크가 인터넷 웜의 침입을 받기 시작한 것으로 네트워크의 상태를 판단하게 된다.

[0095] **Ⅴ. 인터넷 웜의 탐지 방법**

[0096] 네트워크에 인터넷 웜 중 랜덤 스캔 타입의 웜이 침입하는 경우를 예로 들어 설명한다. 또한 이 웜이 전파되는 규모가 슬래머(Slammer) 웜의 전파에 필적한다고 가정한다.

[0097] 이 때 N을 특정 웜에 취약한 집단의 크기, L을 모니터링되는 네트워크 크기 (IP 주소의 수 기준)로, α는 감염률이라고 한다. 이에 따라, αN은 취약 집단 내의 감염된 호스트의 수를 가리킨다. 또한, 웜 스캔율 β는 1초당 1개의 웜에 대하여 수행되는 스캔 회수를 가리킨다.

[0098] 시뮬레이션을 위하여, $N=10^6$, $\beta = 3\times10^2$, 6×10^2, 10^3 $L=2^{16}$으로 설정하였다. 참고로, 슬래머웜의 스캔율이 최대 26,000스캔/초이고, 평균 약 4,000스캔/초 · 웜이었으며, 과거 코드레드II 웜의 변종 중의 하나가 탐지를 회피하기 위해 수행했던 느린 스캐닝의 스캔율이 300스캔/초이었던 점을 고려하면 β값은 적절한 수준으로 설정되었다. 또한, β의 값을 세가지(3×10^2, 6×10^2, 10^3)의 서로 다른 값으로 정한 것은 본 기술이 느리게 스캐닝을 수행하는 웜 또한 조기에 탐지 가능함을 보여주기 위해서이다.

[0099] 도 7은 시뮬레이션의 수행 결과를 감염된 호스트 수와 랭크값의 관계로 나타낸 그래프이다.

[0100] 도 7에서 나타낸 바와 같이, 감염된 호스트의 수효가 증가함에 따라 본 발명의 실시예에 따른 정상 트래픽 소거 행렬의 랭크값은 단시간 내에 매우 급격하게 증가한다. 웜의 스캔율이 1000스캔/초의 경우 단지 인터넷 전체에서의 3%의 취약 호스트가 감염된 경우에도, 랭크값은 이미 252를 넘어서서 최대값인 256에 근접하고 있다. 즉, $N=10^6$, $\beta=10^3$, $L=2^{16}$인 환경에서 $\alpha=0.03$일 때에 252를 초과하는 랭크값이 얻어진다. 또한, 스캔율이 600스캔/초와 300스캔/초의 경우 또한 인터넷 전체에서의 5%와 10%의 취약 호스트가 감염된 경우에도, 랭크값은 이미 252를 넘어서서 최대값인 256에 근접하고 있다. 즉, $N=10^6$, $L=2^{16}$인 환경에서 $\beta=6\times10^2$, 3×10^2일 때에 $\alpha=0.05$, 0.1일 때에 252를 초과하는 랭크값이 얻어진다.

[0101] 도 7의 시뮬레이션에서, 랭크값이 급격하게 증가하는 시점은 웜에 감염된 호스트의 수가 급격하게 증가하는 시점보다 약 30초 선행하고 있음을 알 수 있다.

[0102] 오늘날의 모든 자동화된 공격에 대하여는 인력의 개입을 통한 대응조치는 너무 늦어서 전파를 막을 수가 없다. 예컨대, 슬래머 웜이 네트워크 내의 호스트에 최대로 전파되려면 10분만으로 충분하다. 그러나, 본 발명의 실시예에 따른 웜 탐지 방법에 의하면, 랭크값의 급격한 변화를 인지하는 것만으로도 웜의 침입을 수십 초 가량 먼저 예측할 수 있다.

[0103] 본 발명에 의하여 얻어지는 이러한 수십 초의 여유 시간에 힘입어, 웜 전파의 예방 또는 전파속도 지연을 위하여 필요한 조치를 행할 수 있게 된다.

[0104] 도 8은 256×256 트래픽 행렬에서 2종의 웜의 랭크값을 스캔율β의 함수로서 나타낸 그래프이다.

[0105] 또한, 도 9는 64×64 트래픽 행렬에서 2종의 웜의 랭크값을 스캔율β의 함수로서 나타낸 그래프이다.

[0106] 도 8 및 도 9에서 나타낸 바와 같이, 2종의 웜은 각각 랜덤 스캔 타입의 웜과 순차적 스캔 타입의 웜을 사용하였다.

[0107] 도 8 및 도 9의 시뮬레이션은 대학교 캠퍼스 구내의 /16 네트워크 내에 인터넷 웜(랜덤 스캔 웜 및 순차적 스캔 웜)을 실제로 투입한 후 네트워크의 게이트웨이에서 이를 추적하는 방식으로 수행되었다. 또한, 이 때의 랭크값은 10,000대의 호스트가 감염된 경우($\alpha=0.01$)에 대하여 측정되었다.

[0108] 도 8 및 도 9에서 공통적으로 알 수 있는 바와 같이, 랜덤 스캔 웜의 경우, 랭크값은 스캔율이 증가할수록 비약적으로 증가한다.

[0109] 반면, 순차적 스캔 웜의 경우, 랭크값은 $\beta=1,000$ 이후에 돌연 하락한다.

[0110] 이러한 특성을 통하여 비균일 스캔 웜(non-uniform scan worm), 예컨대 블라스터와 같은 것은 IP 주소 공간의 단일 블록을 모니터링함으로써 탐지할 수 있게 된다. 반면, 이전의 웜 모니터링 방법은 모니터링 대상 주소 공간이 광범위하게 흩어져있을 때에만 효과적일 뿐이다.

[0111] 즉, 웜의 활동성이 증가함에 따라 그 행렬 랭크값은 극단적으로 높아지거나 낮아짐으로써 웜의 침입을 예보하는 분명한 표지가 된다. 이와 달리, 만약 네트워크의 상태가 정상적이어서 웜의 활동성이 증가하지 않는 경우에는 랭크값이 시간에 따라 변화하더라도 극단적으로 증가하거나 감소하지는 않는다.

[0112] 도 10은 순차적 스캔 웜의 개수가 증가하더라도 스캐닝 웜이 한 개라도 섞여있다면, 랭크값은 극단적으로 0에 근접한 값을 가짐을 보여주고 있다.

[0113] 이러한 경우는 크게 두가지 경우가 있을 수 있다. 한가지는 여러가지 복합적인 스캐닝 기법을 가지고 있는 여러 웜이 인터넷에 동시에 전파되는 경우가 있을 수 있으며, 다른 한가지는 랜덤 스캐닝 행위의 특징을 속이기 위해 중간 중간 순차적 스캐닝 트래픽을 섞어서 스캐닝을 하는 것과 같이 회피방법을 사용하더라도 결국에는 순차적 스캐닝 기법을 함께 사용한다면 결과는 순차적 스캐닝 기법에 의한 탐지로서 웜의 전파 상황을 탐지할 수 있음을 보여주고 있다.

[0114] 이는 랭크값을 계산하는 수학적 공식인 가우시안 소거법(Gaussian Elimination)에 의한 결과이다.

[0115] 가우시안 소거법을 통해 랭크값을 계산하기 위해서는 행렬의 행들끼리 XOR 연산을 수행하게 된다. 이는 본 기술

이 랭크값을 계산하는 행렬이 이진 행렬이기 때문이다.

[0116] 이에 순차적 스캐닝 윔이 전파되면 행렬의 여러 행이 모두 1의 엔트리로 되어 있는 행들이 중복적으로 그리고 순차적으로 존재하게 되며, 이는 가우시안 소거법에 의해 모두 0으로 가는 것에 기인한다.

[0117] 위의 결과는 수학식 8과 같은 결과를 도출한다.

[0118] 수학식 8은 가우시안 소거법에 의하여 모든 엔트리들이 1인 행렬들이 0으로 변화됨을 쉽게 확인하기 위해 4×4 행렬을 예로 들었다.

[0119] 좌측의 첫번째 행부터 세번째 행까지 순차적 스캐닝 윔에 의해 모든 엔트리가 순차적으로 1의 값을 갖게 된다면, 랭크값을 구하기 위한 가우시안 소거법의 과정에 의해 우측 행렬과 같이 하나의 행만이 남고 나머지는 XOR연산에 의해 모두 0의 값을 갖는 행으로 변한다.

[0120] 이러한 가우시안 소거법의 효과에 의해 순차적 스캐닝 윔이 전파될 경우 랭크값은 0에 근접한 값으로 급격하게 감소하게 된다.

수학식 8

[0121]
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

[0122] **VI. 스캐닝하는 패킷 수의 결정**

[0123] 252 등의 문턱값을 초과하는 높은 값으로 랭크값이 증가할 때, 윔 전파가 개시되었음을 긴급 경고하기 위해서 트래픽 행렬에 수집하여야 하는 스캐닝 패킷의 수는 아래와 같은 계산을 통하여 산출할 수 있다.

[0124] γ를 랜덤 지수(random factor)라 하고, 극단적으로 높은 랭크값을 가지는 "랜덤" 바이너리 행렬을 γm^2개의 1이 채우도록 한다.

[0125] 이 때, 랜덤 스캐닝 윔이 트래픽 행렬 내에서 0으로 기록되어 있는 성분들을충분히 많은 개수만큼 1로 랜덤하게 변경하기 위해서는, γm^2개의 스캐닝 패킷이 필요하다고 볼 수도 있다.

[0126] 그러나, 사실상 필요한 랜덤 스캐닝 패킷의 개수는 그보다 적다.

[0127] 왜냐하면, 앞서 살펴본 바와 같이, 트래픽 행렬 내에서 정상 트래픽을 소거하는 동작(예컨대 $M(t) \oplus M(t-1)$)을 수행한 후에는, $M'(t)$ 내의 0이 아닌 성분이 거의 2배로 증가하는 효과가 나타나기 때문이다.

[0128] 따라서, γm^2의 절반만으로도 랜덤 m ×m 행렬을 구축할 수 있다.

[0129] 따라서, 윔 전파의 개시를 경고하기 위하여 수집하여야 하는 스캐닝 패킷의 수는 아래 수학식과 같이 결정할 수 있게 된다.

수학식 9

[0130]
$$\alpha N \cdot \beta \cdot L/2^{32} \geq \gamma \cdot m^2/2$$

[0131] 한편, N, β, α 및 L은 결과 행렬 $M'(t)$가 랜덤인지 여부와 무관하게 결정된다. 즉, 파라미터 N 및 L은 호스트와 네트워크의 설정값에 의하여 결정되며, β는 윔의 특성에 따라 결정되고, m은 이상유무 탐지 모듈의 설정값에 따라 결정된다. 또한, 그 밖의 파라미터 γ는 행렬의 랜덤성의 척도로서, m에 의해서만 결정된다.

[0132] 표 1은 랜덤 지수 트래픽 행렬의 관계를 나타낸 표이다.

표 1

[0133]

m	32	64	128	256	512	1024
γ	0.063	0.041	0.025	0.014	0.008	0.005

[0134] 평균적인 랜덤 지수(γ)는 랜덤 행렬 구축의 반복을 통하여 실험적으로 측정할 수 있다. 이를 표 1에 나타내었다.

[0135] 랜덤 스캔 패킷에 의하여 단지 트래픽 행렬 영역의 작은 일부만이 활성화되더라도, 본 발명의 실시에에 의한 웜 탐지 방법에 따라 정상 트래픽 소거 행렬의 랭크값을 통하여 랜덤성 탐지가 가능함을 알 수 있다.

[0136] 또한, 표 1은 랜덤성 탐지에 사용되는 경우의 랭크 메트릭의 민감도(sensitivity)의 관점에서도 이를 살펴볼 수 있다.

[0137] 한편, 수학식 9에 의하여 파라미터 간의 상호작용을 더욱 잘 이해할 수 있다.

[0138] 즉, 감염률(α)이 낮은 상태에서도 웜 스캔율(β)을 증가시킴으로써 웜의 전파를 조기에 탐지할 수 있게 되는 것이다.

[0139] 예컨대 도 11에 나타난 시뮬레이션에서 감염률(α)이 0.03일 때에 웜 스캔율(β)은 1,000가 되는데, 만약 동일한 조건에서 웜 스캔율(β)을 3,000으로 증가시키는 경우에는 감염률(α)은 0.01로 감소한다. 즉, 감염률이 더욱 낮은 상태(웜이 네트워크 내에 덜 전파된 상태)에서도 웜을 탐지할 수 있음을 알 수 있다.

[0140] 수학식 9에서 나타낸 바와 같은 α와 β 간의 이러한 반비례 관계에 의하여, 본 발명의 실시예에 의한 웜 탐지 방법에서는 전파속도가 빠른 웜에 대해서도 스캔율을 증가시킴으로써 더욱 빠른 탐지를 실현할 수 있다.

[0141] 도 11은 웜의 전파가 정상 트래픽 소거 행렬의 랭크값 변화를 통하여 탐지되기 위하여 필요한 감염률과 웜 스캔율 및 호스트 집단의 크기의 관계를 나타낸 그래프이다.

[0142] 예컨대, 웜에 취약한 호스트 집단의 크기(N)가 10^5 내지 10^6이고, 웜 스캔율(β)이 5,000 내지 10,000인 경우에는, 탐지된 웜에 의한 네트워크 내 호스트의 감염률(α)은 고작 0.3% 내지 1.2%에 불과한 것을 도 11의 그래프를 통하여 확인할 수 있다.

[0143] 즉, 전역적인 웜 전파로 전개되는 파라미터 범위에 대해서도 본 발명에 의한웜 탐지 방법을 사용하면 웜 전파에 대응하기 위하여 필요한 시간을 충분히 확보할 수 있을 정도로 이른 시점에 네트워크에 침입한 인터넷 웜을 탐지할 수 있게 되는 것이다.

[0144] 이상에서 살펴본 바와 같이, 행렬은 네트워크를 통하여 침입하는 웜을 조기에 탐지하기 위하여 필요한 다양한 연산을 적용함에 있어 많은 잘 정의된 좋은 동작을 위한 편리한 데이터 구조이다. 본 발명에 따르면, 행렬 기반 침입 탐지 메커니즘으로서 네트워크 트래픽으로부터 트래픽 행렬을 구축하고, 정상적인 트래픽에 대한 필터링 동작 후의 행렬의 랭크값을 측정함으로써 랭크값의 크기에 따라 간단하게 네트워크 상에 웜이 침입했는지 여부를 판단할 수 있다.

[0145] 더욱이, 본 발명에 의한 행렬적 접근은 인터넷 웜의 탐지에만 적용될 수 있는 것이 아니며, 네트워크 트래픽 상에서 랜덤성을 증가시키는 다른 타입의 공격에 대해서도 마찬가지로 적용될 수 있다.

[0146] 또한, 본 발명의 실시예는 이상에서 설명한 장치 및/또는 방법을 통해서만 구현이 되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하기 위한 프로그램, 그 프로그램이 기록된 기록 매체 등을 통해 구현될 수도 있으며, 이러한 구현은 앞서 설명한 실시예의 기재로부터 본 발명이 속하는 기술분야의 전문가라면 쉽게 구현할 수 있는 것이다.

[0147] 또한, 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면의 간단한 설명

[0148] 도 1은 본 발명의 제1 실시예에 따른 웜 탐지 장치를 도시한 블록도,

[0149] 도 2는 균일 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면,

[0150] 도 3은 서브넷 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면,

[0151] 도 4는 순차적 스캔이 적용된 경우의 IP 주소의 형태를 나타낸 도면,

[0152] 도 5는 본 발명의 실시예에 의한 웜 탐지 방법에서 사용되는 필터링 및 랭크값 측정의 메커니즘을 나타낸 도면,

[0153] 도 6은 본 발명의 실시예에 따른 필터링의 원리와 트래픽 행렬의 관계를 나타낸 벤다이어그램,

[0154] 도 7은 시뮬레이션의 수행 결과를 감염된 호스트 수와 랭크값의 관계로 나타낸 그래프,

[0155] 도 8은 256×256 트래픽 행렬에서 2종의 웜의 랭크값을 스캔율β의 함수로서 나타낸 그래프,

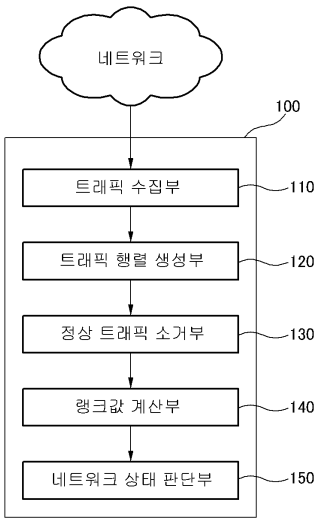
[0156] 도 9는 64×64 트래픽 행렬에서 2종의 웜의 랭크값을 스캔율β의 함수로서 나타낸 그래프,

[0157] 도 10은 순차적 스캐닝 웜 1개가 인터넷에 전파되는 상황을 기반으로 랜덤 스캐닝 웜이 0개, 1개, 2개, 3개로 차례로 증가시켜도 결국 랭크값은 0에 근사한 값으로 급감함을 나타낸 그래프,

[0158] 도 11은 웜의 전파가 정상 트래픽 소거 행렬의 랭크값 변화를 통하여 탐지되기 위하여 필요한 감염률과 웜 스캔율 및 호스트 집단의 크기의 관계를 나타낸 그래프이다.

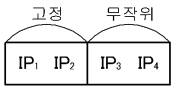
도면

도면1

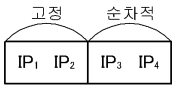


도면2

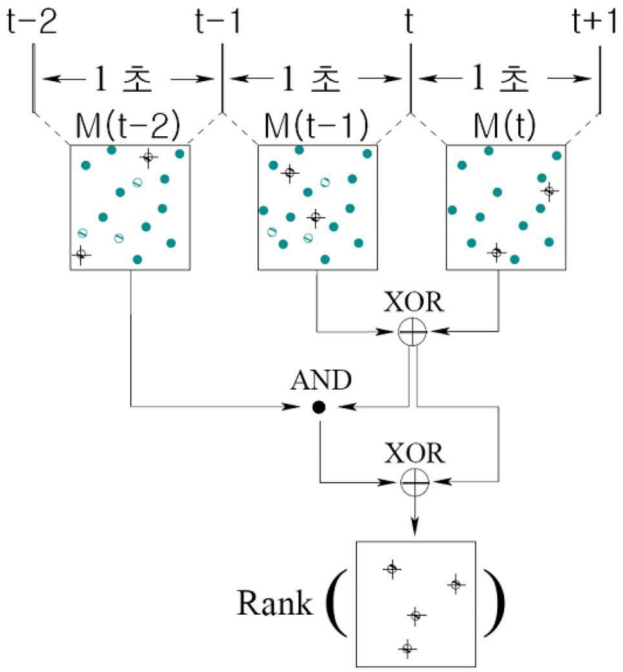
도면3


고정
무작위
IP1 IP2 IP3 IP4

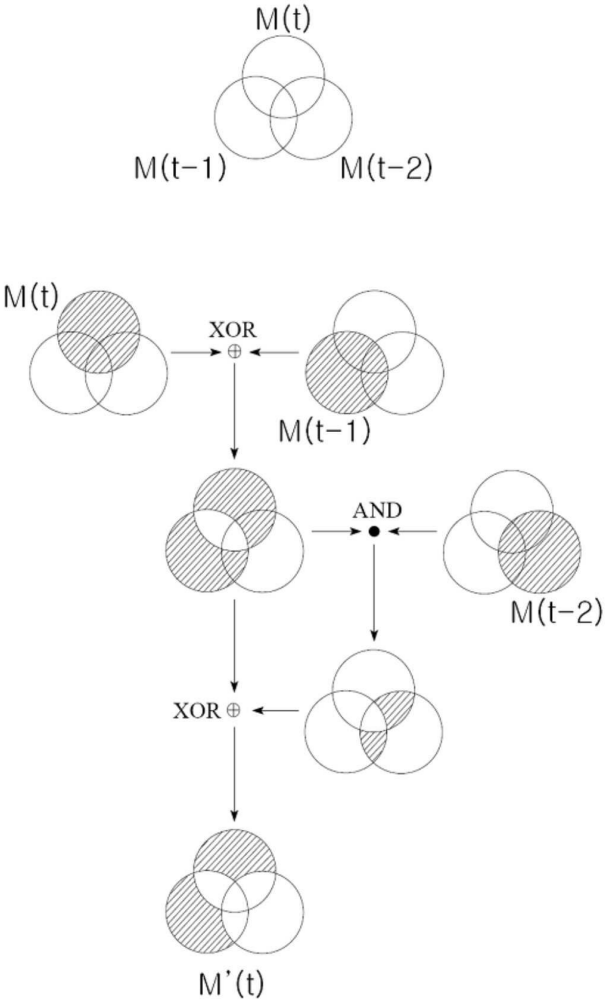
도면4


고정
순차적
IP1 IP2 IP3 IP4

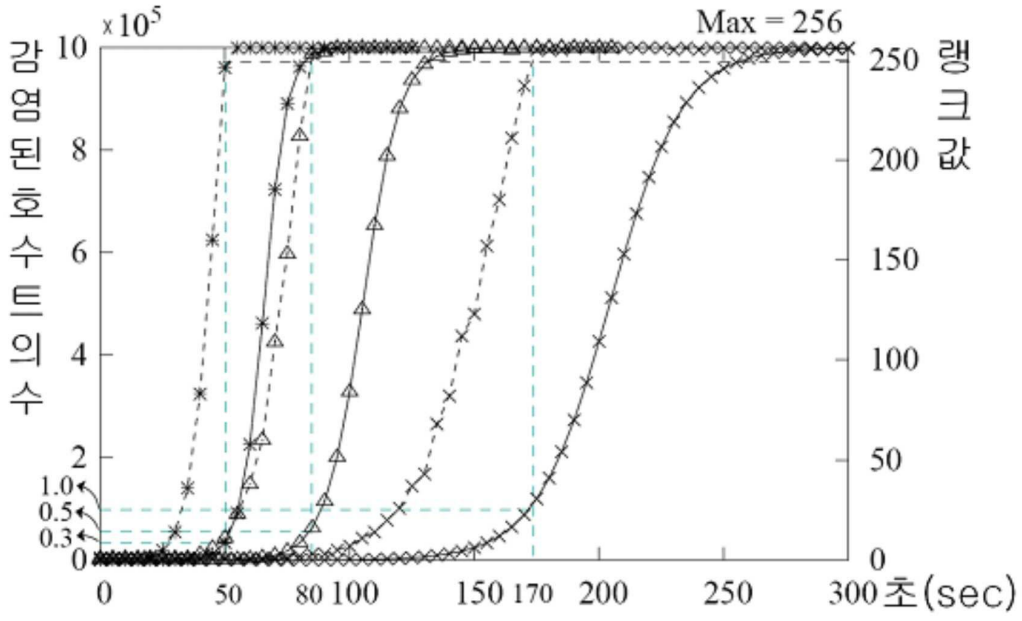
도면5


t-2
t-1
t
t+1
1 초
1 초
1 초
M(t-2)
M(t-1)
M(t)
XOR
AND
XOR
Rank

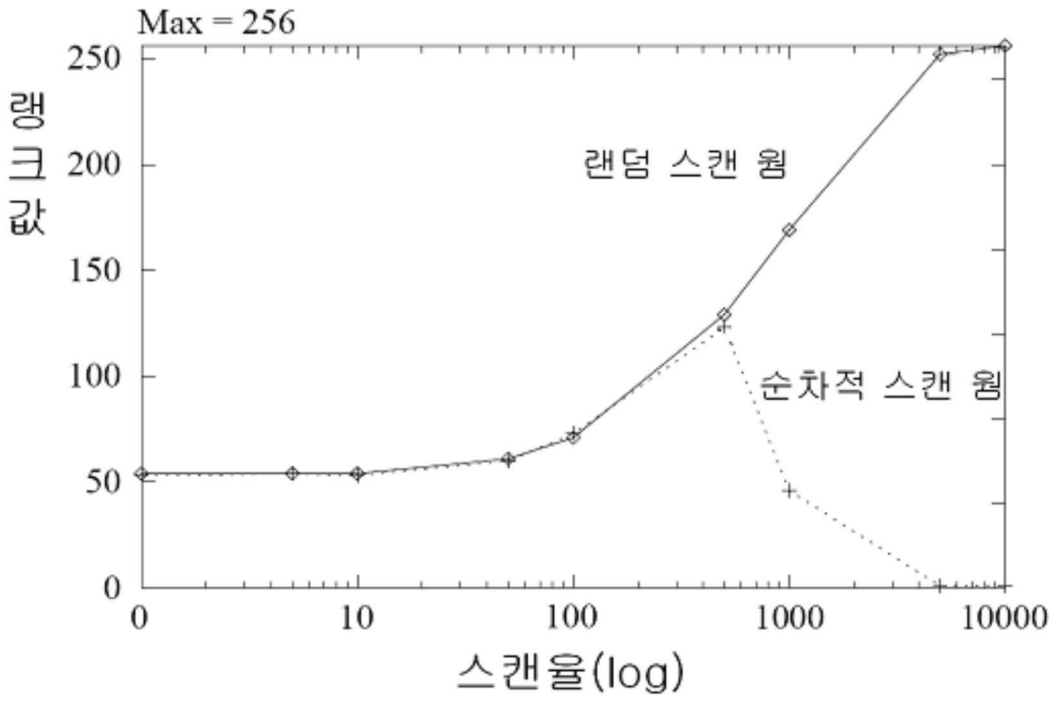
도면6


M(t)
M(t-1)
M(t-2)
M(t)
XOR
M(t-1)
AND
M(t-2)
XOR
M'(t)

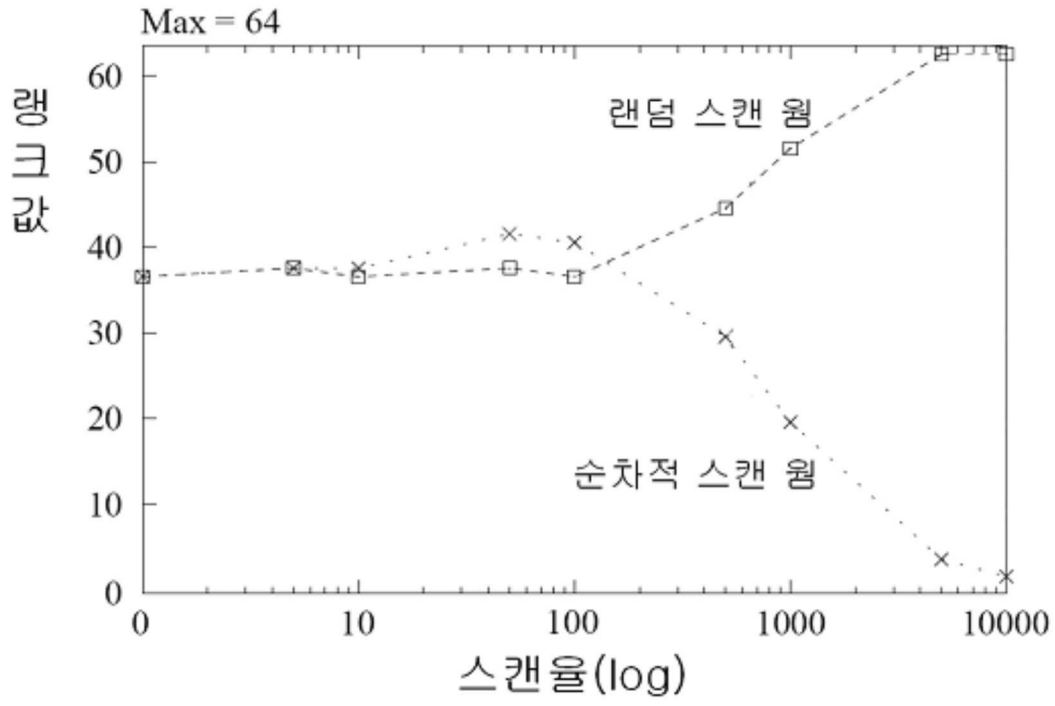
도면7


x 10^5
Max = 256
감염된 호스트의 수
랭크 값
10
8
6
4
2
1.0
0.5
0.3
0
250
200
150
100
50
0
0
50
80
100
150
170
200
250
300
초(sec)

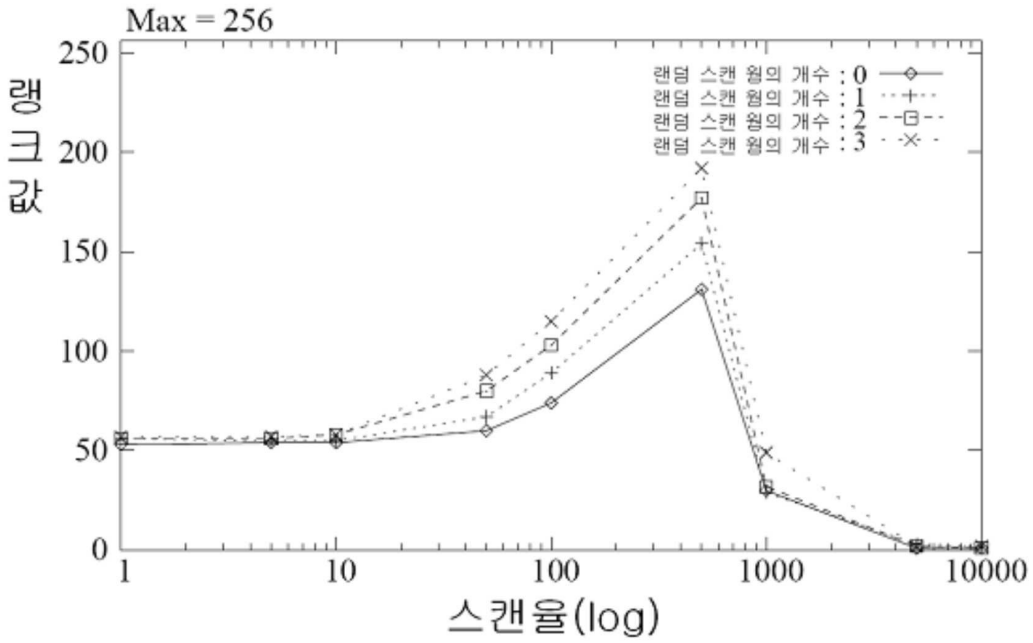
도면8


Max = 256
랭크값
250
200
150
100
50
0
랜덤 스캔 웜
순차적 스캔 웜
0
10
100
1000
10000
스캔율(log)

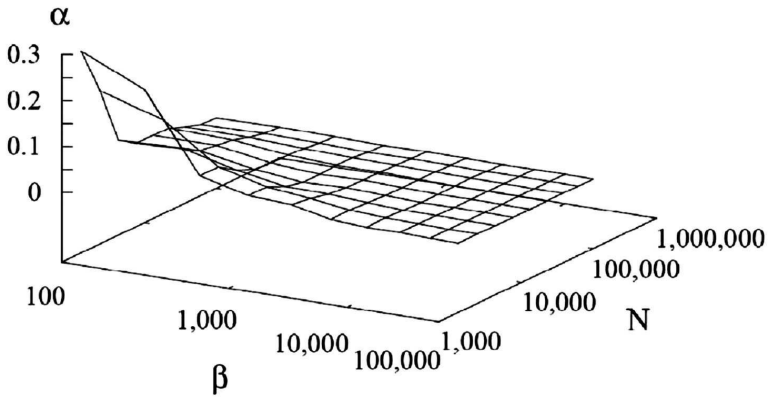
도면9


Max = 64
랭크값
60
50
40
30
20
10
0
랜덤 스캔 웜
순차적 스캔 웜
0
10
100
1000
10000
스캔율(log)

도면10


Max = 256
랭크값
250
200
150
100
50
0
1
10
100
1000
10000
스캔율(log)
랜덤 스캔 원의 개수 : 0
랜덤 스캔 원의 개수 : 1
랜덤 스캔 원의 개수 : 2
랜덤 스캔 원의 개수 : 3

도면11


α
0.3
0.2
0.1
0
100
1,000
10,000
100,000
β
1,000
10,000
100,000
1,000,000
N