알려지지 않은 신종 웜 탐지 기술

연재순서

1. 행위기반 봇넷 탐지기술
2. 알려지지 않은 스파이웨어 탐지기술
3. VoIP공격 탐지기술
4. **알려지지 않은 신종 인터넷 웜 탐지기술**
5. 악성코드 사전 탐지기술
6. IP 스푸핑 탐지기술

컴퓨터와 인터넷의 발달과 함께 인터넷 웜의 공격 방법은 더욱 지능화 되어가고 있으며, 경쟁이라도 하는 듯이 인터넷을 통한 웜의 전파속도는 급속도로 빨라지고 있으며, 그로인한 피해 규모는 날이 갈수록 증폭되고 있다. 이러한 인터넷 웜에 효과적으로 대응하기 위하여 본서에서는 알려지지 않은 신종 웜 탐지 기술에 대하여 알아본다.

글 · 박현도 고려대 컴퓨터보안연구실 연구원, 이희조 고려대 교수

"사장님, 컴퓨터에 바이러스가 감염되었습니다."라고 엔지니어가 사장님에게 보고를 하자 사장님은 "그러 길래 진작 소독약 뿌리랬잖아!"라고 호통을 쳤다는 유머는 인터넷이 지금처럼 발달해 보급되기 이전에 컴퓨터 바이러스에 대한 일반적인 무지를 풍자하는 유머로 유명하다.

그러나 이제는 초등학생들도 바이러스, 인터넷 웜, 백도어 등의 악성 프로그램이나 행위들에 대한 용어에 대해 잘 알고 있으며 그 중 일부는 인터넷에 유포되어있는 웜이나 바이러스 제작 도구를 사용할 줄 알고 있으며 백도어를 어떻게 설치해야 하는지까지 알고 있을 정도로 누구나 아는 일반적인 상식이 됐다. 이는 인터넷에서의 범죄는 끊이지 않고 그 심각성 또한 날이 갈수록 높아가고 빈번하게 발생되는 것에 기인한다. 얼마 전 뉴질랜드 10대 소년이 지구촌 사이버 범죄조직의 일명 '큰 형님'이며 FBI에서 조사를 받고 있다는 신문 기사는 실로 혀를 내두르게 하는 반면에 인터넷에서의 사이버 범죄의 현 주소를 여실히 보여주고 있는 사례라 할 수 있다.

이에 이번 연재에서는 자기 스스로 인터넷에 취약한 호스트에 대해 조사 · 공격 · 감염을 통해 인터넷에 확산하는 자동화 프로그램인 인터넷 웜에 더욱 효과적으로 대응하기 위한 알려지지 않은 웜의 조기 탐지 기술에 대해서 기술한다.

인터넷 웜, 인터넷의 공공의 적

최초의 인터넷 웜으로 불리는 Morris웜이 1988년 11월에 발생된 이래로 인터넷 웜은 해가 거듭될수록 더욱 악성화 되며 지능화 되어 우리의 인터넷을 호시탐탐 공격하고 있다. 그 피해 규모도 날이 갈수록 증가하고 있다. 초기 웜들은 확산 속도도 느리고 확산되는 네트워크 규모도 소규모였던 것에 반해 근래에 출현하는 웜들은 확산을 위한 네트워크 규모가 인터넷 전체를 대상으로 함을 우리는 여러 경험으로 잘 알고 있다. 2001년 CodeRed의 출현으로 인터넷에 연결되어 있는 수많은 컴퓨터들이 이에 공격을 받아 엄청난 피해를 입었으며 2003년 1월에는 우리가 1.25 인터넷 대란으로 잘 알고 있는 SQL Slammer 웜의 경우 단 10분 만에 인터넷의 7만5,000여대 이상의 호스트를 감염시키고 인터넷 전체를 마비시키는 결과를 초래했다.

이러한 웜이 새로이 출현하게 되면 웜 제작자는 보안 회사를 상대로 게릴라전을 방불케 할 정도로 수많은 변종들을 쏟아낸다. 그 결과 이와 관련한 변종들이 지속적으로 발견되게 된다. 또한 끊이지 않고 지속적으로 새로운 웜이 출현하고 있는 실정이다. 이렇게 인터넷 웜은 더욱 고속화 지능화를 반복하며 진화하고 있으며 이제 인터넷에서 보안업체 뿐만 아니라 모든 국가와 모든 인터넷 사용자들에게 이러한 치명적인 웜에 대비를 하지 않으면 안되는 명실상부한 공공의적으로 자리매김했다.

이러한 웜의 심각성은 웜이 다른 컴퓨터를 감염시킨 이후에 이 감염된 컴퓨터를 악의적인 목적에 사용한다는 것이다. 웜에 감염된 호스트들은 특정 호스트에 DDoS(분산 서비스 거부 공격, Denial-of-Service)공격을 감행하거나 중요한 정보를 유출하

표 1. KISA 해킹 바이러스 통계

구분		2004년	2005년	2006년
신종	A사	28	63	200
	B사	24	9	20
변종	A사	132	527	2,383
	B사	116	38	72

기 위한 백도어를 설치하는 등의 공격 메커니즘을 내포하고 있어 대단히 심각한 악성 코드이다. 근래에는 이메일이나 메신저 등을 이용해 전파되기 때문에 친구로부터 온 메일이나 메신저를 통해 전송된 프로그램을 여과 없이 다운받아 그 심각성이 더 커지고 있다.

그리고 인터넷 인프라가 발전하고 또한 휴대폰이나 PDA와 같은 소형 단말기들의 성능이 날로 발전함에 따라서 웜에 의한 피해 가능성은 날로 증가하고 있는 실정이다. 실제로 2004년 6월 최초의 휴대폰 웜인 Cabir웜이 유포되어 경기장에서 경기를 관람하던 관중들 수십명의 휴대폰에 블루투스를 이용해 전염되는 사례가 보고됐다. 또 모바일 게임 회사가 불법 게임을 퇴치하기 위해 불법게임에 대한 보고를 단문자 서비스(SMS)를 통해 하도록 프로그래밍했다가 엄청난 SMS의 폭주를 경험하기도 했다. 이와 관련해 아직까지는 큰 피해나 손실이 있지는 않았지만 향후 휴대폰을 통한 웜의 피해를 예상할 수 있는 사례라 할 수 있다.

또한 이전에는 취약점이 인터넷에 보고되고 이 취약점이 보완되어 패치가 발표된 후에 수일 또는 수개월 이후에 그와 관련한 웜이 등장했지만 그 주기가 짧아져서 패치가 발표되기 이전에 웜이 발생하는 제로데이(zero-day)웜이 발견되고 있는 실정이다. 취약점에 대한 패치가 공표되기 이전에 퍼지는 제로데이 웜에 대해서 효과적이고 능동적인 대처를 하지 않는다면 속수무책으로 당할 수밖에 없다.

이러한 웜에 대처하기 위해서는 조기에 이러한 웜을 탐지하고 피해가 커지기 전에 이를 막기 위한 연구가 절실하다. 이에 이 글에서는 앞으로 알려지지 않은 인터넷 웜을 조기에 탐지할 수 있는 탐지 기법들에 대해 소개한다.

인터넷 웜, 조기 탐지로 피해 최소화

인터넷 웜을 탐지하기 위한 방법으로는 크게 패턴매칭 기법과 행위인식 기반으로 구분할 수 있다. 패턴 매칭 기법은 현재 폭넓게 사용되고 있으며 이전에 발견된 웜의 패턴 정보와 비교해 봄으로써 향후 동일한 패턴의 웜을 탐지하기 위한 웜 탐지 기술이다.

예를 들어 CodeRed웜은 특정 코드 값을 가지고 있다. 차후에 다른 패킷들에게서 이 특정 값과 동일한 코드값이 존재하면 CodeRed웜으로 판별하는 기법을 말한다. 이러한 방법은 정확하게 웜을 탐지할 수 있는 장점을 가지고 있다. 그러나 인터넷 웜은 진화하고 있으며 새로운 기법들을 이용해 더욱 고도화 지능화 되어 수많은 신종 웜이 출현하고 있는 현실이다. 정확도에서는 패턴매칭을 통한 탐지 기법을 따라갈 탐지 기술은 아직까지는 없다. 그러나 향후 발생할 수 있는 웜을 예측하여 대응하기에는 패턴매칭 알고리즘으로는 역부족이며 이전에 발견되지 않은 새로운 웜의 출현에는 속수무책으로 당할 수밖에 없다.

이를 보완할 수 있는 기술이 바로 알려지지 않은 웜을 탐지할 수 있는 비정상 행위 기반 기술이다. 이 기술은 웜의 코드 내의 특정 값이나 운영체제에 프로그램을 등록시키거나 침입하는 특정 행위 패턴에 기반 하여 웜을 탐지하는 것이 아니라 웜의 행동 양상을 탐지하여 일반적인 네트워크 상황에서의 행위와 비교하여 비정상 행위를 탐지함으로서 이전에 발견되지 않은 웜이라도 탐지할 수 있는 기법이다. 이러한 웜 탐지 기술은 알려지지 않은 새로운 웜을 조기에 탐지하여 이에 대한 대처를 할 수 있게 해줌으로써 상당히 매력적인 이점을 가지고 있다. 그러나 이러한 방법들은 오탐율이 알려진 웜의 패턴매칭을 통한 탐지기술에 비해 상대적으로 높아서 아직 정확한 탐지가 어렵고 그 결과 구현이 쉽지 않다는 단점이 있다.

그러나 알려진 웜을 탐지하는 기술만으로는 웜에 대처할 수 없기 때문에 이러한 비정상 행위 기반의 탐지 기술은 현재 많은 대학과 연구기관에서 연구가 진행되고 있으며 앞으로도 더욱더 많은 발전이 있을 것이라 본다.

알려지지 않은 신종 웜, 비정상 행위 탐지로 대응

연결 실패 횟수를 통한 웜 탐지 기법

미국의 Dartmouth대학의 Vincent Berk가 제안한 방법으로 인터넷에서 연결 실패 횟수를 기반으로 웜을 탐지하는 기술이다. 웜은 인터넷에 전파를 목적으로 다른 호스트를 감염시키고자 한다. 다른 호스트를 감염시키기 위해 어떠한 취약한 호스트가 인터넷에 분포되어 있는지 알아보기 위해 웜에 감염된 호스트는 공격 가능한 취약한 호스트들의 정보를 얻기 위해 인터넷에 많은 ICMP패킷을 전송한다.

웜은 이렇게 전송한 ICMP패킷들의 응답을 토대로 감염시킬 호스트를 결정하게 된다. 그러므로 일반적인 네트워크에서 연결 실패가 일어나는 빈도보다 훨씬 많은 연결 실패가 웜에 의해 발생되고 이 메커니즘은 이러한 연결 실패 횟수를 기반으로 웜을 탐지하게 된다.

일반적으로 연결 실패와 관련한 패킷으로는 TCP_RST패킷, ICMP 목적지 도달 불가능(Destination Unreachable) 패킷, TCP 유휴제한 시간 초과(Timeout)패킷들이 포함된다. 이 중 본 메커니즘은 ICMP 목적지 도달 불가능 패킷의 횟수를 기반으로 웜을 탐지하게 된다. 만약 연결 실패 횟수가 임계치를 넘

게 되면 웜이라 간주하는 심플한 모델이다. 그러나 만약 IP 스푸핑 기술을 통해 자신의 IP를 속인다면 ICMP패킷은 다시 웜이 감염된 호스트나 네트워크로 유입되지 않을 것이다.
이러한 경우 웜을 탐지하지 못하게 되는 결과를 초래한다. 또한 웜이 스캐닝을 수반하지 않는다면 다시 말해서 무작위로 감염 시도만을 수행한다면, ICMP 프로토콜을 사용하지 않기 때문에 이러한 웜에 대해서는 탐지할 수 없는 결과를 초래한다.

DEWP

DEWP(Detecting Early Worm Propagation through packet matching)는 ISI(Information Science Institute)에서 제안한 모델이다. 일반적으로 네트워크에서 트래픽의 성격은 갑자기 변화되지는 않는다. 네트워크가 제공하는 서비스가 갑자기 변화하는 경우가 아니라면 이전과 크게 다르지 않은 트래픽들이 네트워크에 존재하게 된다. 이에 반해 웜은 특정 네트워크의 서비스 포트로 유입되는 트래픽이 갑자기 폭주하게 된다. 이 메커니즘은 이러한 현상에 기반 하여 웜을 탐지하는 기술이다.
이 메커니즘은 특정 네트워크로 유입되는 트래픽의 IP주소와 포트를 체크하여 평균적인 값을 분석한 후에 갑자기 특정 포트로 평균값 이상의 값이 유입되면 이를 웜이라 간주한다.

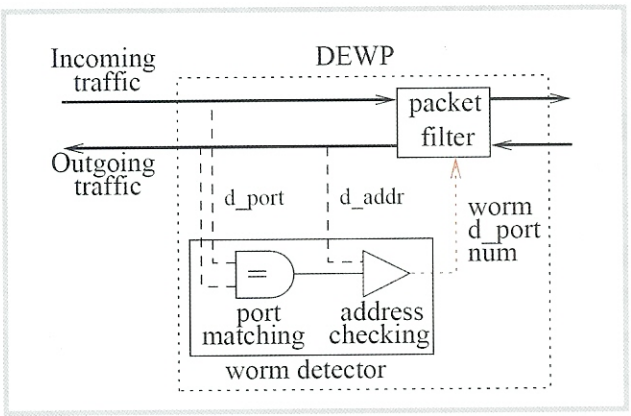

그림 1. DEWP 구조도

이 메커니즘의 경우 구조적인 면에서나 계산 복잡도가 상당히 낮은 반면에 Flash Crowd와 같은 일반적인 사용자들이 정상적인 요청에 의한 연결(월드컵 기간에 FIFA 홈페이지가 트래픽이 폭주하는 경우와 같은)이 폭주하는 경우에는 이를 웜으로 오인하게 되고 이러한 경우에 오탐율이 높다는 단점이 있다.

TRW

TRW(Threshold Random Walk)는 미국 MIT에서 연구한 비정상 행위 탐지 기술로서 네트워크 트래픽 패턴을 분석하여 웜을 탐지하는 조기 탐지기술이다.
이 기술은 TCP프로토콜을 이용하는 웜을 탐지하기 위한 기술이며 이 기술은 TCP/IP프로토콜이 Three-way handshake를 통해 호스트 간에 연결을 이루는 것에 기반해 TCP 프로토콜에서 이를 따르지 않는 연결 시도가 얼마나 있는지를 조사하여 웜을 탐지하는 기술이다. 일반적으로 웜은 감염을 시도하기 위해 많은 스캐닝 트래픽을 유발하며 이에 반해 연결이 성공되어 감염을 시도하는 비율이 낮음에 착안한 알고리즘이다.
TCP/IP 프로토콜에서 호스트사이에 연결을 맺고 통신을 하고자 하면 클라이언트는 SYN패킷을 전송하고 이를 받은 서버는 SYN/ACK 패킷을 클라이언트에 전송하며 클라이언트는 ACK 패킷을 전송함으로써 통신을 위한 연결을 맺게 된다. TRW는 SYN이 네트워크에서 최초 발견되면 그 연결 시도에 대해 연결이 성공적으로 이루어지는지를 모두 체크한다.
TRW는 TCP 프로토콜에서 SYN 패킷이 도달하게 되면 값을 증가시키고 이 연결 시도가 성공적으로 이루어지면 값은 감소하게 된다. 일반적인 네트워크의 경우 연결이 실패하는 경우가 빈번히 발생하지 않기 때문에 이 값은 증가와 감소를 동일하게 하게 되며 값은 계속 유지된다.
그러나 일반적으로 웜은 많은 스캐닝 트래픽을 발생시키며 이 양에 비해 연결 되어 감염이 성공하는 비율이 상대적으로 작은 값을 갖는다. 그러므로 어떤 호스트가 계속적으로 연결 실패가 일어나게 되면 그 값은 결국 임계치를 넘어서게 되고 TRW는 웜을 탐지하게 된다.

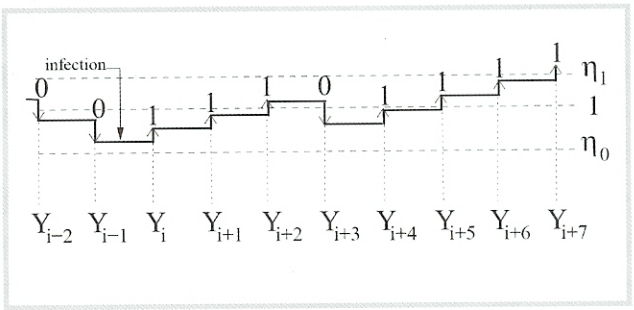

그림 2. TRW에 의한 웜 탐지

TRW는 수학적 기반을 이용하기 때문에 수식은 복잡하나 구현 복잡도는 상당히 적으며 빠르게 조기에 TCP웜을 탐지할 수 있는 장점이 있다. 그러나 UDP를 이용하는 웜은 탐지할 수 없는 단점을 가지고 있다.

DNS 트래픽 기반 웜 탐지 기법

캐나다의 Carleton 대학의 David Whyte에 의해 제안된 방법으로서 일반적인 트래픽은 DNS 서버에 주소 요청 쿼리를 통하여 통신을 수행한다. 특별한 경우가 아닌 다음에야 IP주소를 이

용하여 직접적인 통신을 하는 경우는 극히 드물다. 이에 착안하여 고안된 기술이다.

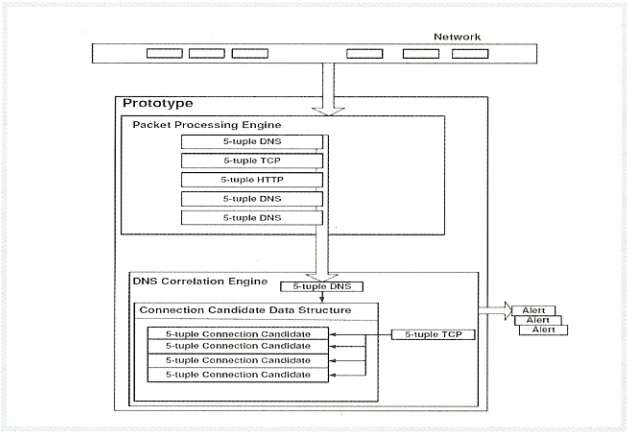

그림 3. DNS 기반 웜 탐지 기술 프로토타입

일반적인 네트워크에서는 엔터프라이즈 네트워크로부터 인터넷으로 유출되는 트래픽의 경우 DNS 쿼리를 우선으로 해당 목적지 IP 주소를 획득 한 이후에 통신을 시작한다. 그러나 웜은 일반적으로 랜덤 생성기를 통해 목적지 IP 주소를 생성하기 때문에 DNS에 목적지 주소 확인 요청을 하지 않게 된다. 이러한 차이를 통하여 웜을 탐지할 수 있게 된다.

현재 인터넷에서 많은 비중을 차지하고 있는 트래픽 중 한 가지가 P2P에 의한 트래픽이다. P2P의 경우 프로그램을 통해 직접적으로 IP를 이용해서 서버 또는 파일을 공유할 해당 호스트들과 통신을 하게 된다. 그러므로 P2P트래픽을 웜으로 오탐할 수 있는 여지를 가지고 있다.

과다 트래픽 탐지 기법(ETRI)

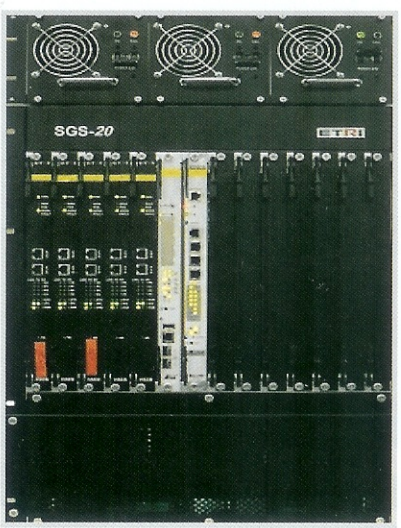

그림 4. SGS-20 침입방지 시스템

SGS-20은 ETRI에서 개발한 침입 방지 시스템 제작 프로젝트명이다. 이 침입 방지 시스템은 웜이나 DDoS공격과 같은 침입에 대한 패턴을 저장하고 있으며 이를 통해 침입을 방지하고 비정상 과다 트래픽 탐지 시스템을 포함하고 있다. 비정상 과다 트래픽 탐지 시스템은 웜과 DDoS공격을 탐지할 수 있도록 설계되어 있다.

이 SGS-20 시스템의 비정상 과다 트래픽 탐지 시스템은 트래픽을 플로우(flow)단위로 정리한 후 이들의 BPS(Bandwidth Per Second)와 PPS(Packet Per Second)값을 정리하여 비교하도록 설계되어 있다. 이 탐지 시스템은 얼마나 현재 BPS나 PPS가 이전과 비교하여 급증하는지를 체크하게 된다. 또한 이에 추가하여 TCP 연결 성공에 비해 연결 실패가 어느 정도 빈번하게 일어나는지 그 비율 값을 추가적으로 사용하게 된다. 만약 TCP연결 실패 비율이 높으며 상대적으로 이전에 비해 BPS와 PPS의 값이 급증하였다면, 이를 웜으로 간주하여 보고하도록 설계되어 있다.

ADUR(Korea University)

ADUR(Anomaly Detection Using Randomness check)는 고려대학교 컴퓨터보안 연구실에서 제안한 알려지지 않은 신종 웜 탐지기술이다. 일반적으로 웜은 인터넷에 유포되기 위해 다른 호스트에 감염을 시도한다. 이 때에 감염을 시도하고자 하는 대상 컴퓨터를 웜은 랜덤 생성기를 통해 IP 주소를 생성하여 대상을 물색한다. 이렇게 될 경우 랜덤 생성기를 통해 생성된 트래픽이 네트워크에 유입되게 되며 이렇게 유입된 네트워크 트래픽은 네트워크 트래픽의 랜덤성을 증가시키게 된다. ADUR는 이에 착안하여 네트워크 트래픽의 새로운 연결들의 랜덤성을 측정하여 이를 기반으로 웜을 탐지하는 메커니즘이다.

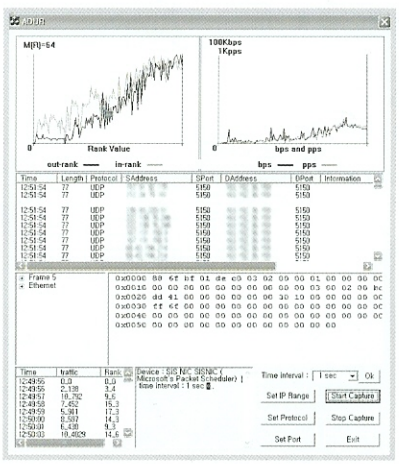

그림 5. ADUR의 실제 구현 · 동작 화면

일반적인 네트워크 트래픽은 서로 정보를 주고받기 위해 일정기간 호스트간의 연결을 유지한다. 그러나 웜에 의해 발생하는 트래픽은 호스트 스캐닝 트래픽이 많고 또한 감염을 시키기 위해 전송하는 트래픽의 양도 상당히 작다. 그러므로 많은 랜덤한 호스트에게 연결을 시도하고 짧은 기간 동안만 연결을 유지한다. 이러한 웜의 특징에 착안하여 제작된 ADUR는 트래픽을 행렬에 표현하여 짧은 시간을 주기적으로 반복하여 트래픽을 저장하며 바로 이전의 행렬에 표현된 트래픽과 현재 행렬에 표현된 트래픽을 행렬 XOR연산을 통해 지속적으로

연결이 유지되는 트래픽은 소거시킨 후에 새로운 연결 시도 트래픽만의 랜덤성을 체크하게 된다. 행렬의 랜덤성을 측정하기 위해 ADUR는 Gaussian 소거법을 통해 행렬의 rank 값을 도출하게 된다.

10×10크기 이상의 정방행렬의 특징 중에 흥미로운 점은 이 rank값이 임계치를 넘어가면 행렬의 원소들의 분포가 랜덤하다는 것을 측정할 수 있다는 것이다. rank값이 임계치를 넘어가게 되면 ADUR는 웜을 탐지하게 된다. 그 결과 다양한 실험을 통해 웜을 잘 탐지하고 있음을 보여주고 있다.

인터넷 웜, 비정상 행위의 조기 탐지와 대처로 피해 확산 최소화

이제까지 우리는 비정상 행위 탐지를 통해 조기에 웜을 탐지할 수 있는 국내외 메커니즘들에 대해서 알아보았다. 웜이 인터넷에 유포되는 초기에는 감염된 호스트의 숫자가 인터넷에 일정 이상 퍼지기 이전에는 상대적으로 적을 수밖에 없다.

그러나 일단 어느 정도 감염된 호스트들의 숫자가 불어나게 되면 그 이후부터는 감염된 호스트들 모두가 또다시 여러 호스트를 상대로 감염을 시도하기 때문에 기하급수적으로 늘어나고 이때가 되면 웜에 대응하기는 거의 불가능하게 될 수 있다.

이는 전염병이 마을과 도시를 휩쓰는 것과 동일하게 볼 수 있다. 처음에는 한두 명이 감염 증세를 보이다가 이후 마을 전체로 삽시간에 퍼져나가는 것과 동일하다. 그렇기 때문에 웜에 효과적으로 대응하기 위해서는 초기에 웜을 탐지하여 손을 쓸 수 없을 정도까지 확산이 되기 전에 웜에 대응해야만 인터넷을 안전하게 유지할 수 있다.