Code Red 웜 전파 패턴 분석을 위한 시뮬레이션

Simulation for the Propagation Pattern Analysis of Code Red Worm

강구홍
서원대학교 컴퓨터정보통신공학부

Koo-Hong Kang(khkang@seowon.ac.kr)

요약

Code Red와 같은 인터넷 웜이 얼마나 심각하게 우리들 일상생활에 영향을 미쳤는지 잘 알려져 있다. 오늘날 인터넷의 고속화와 함께 이러한 웜의 피해는 단기간 내에 발생할 것이 자명하다. 따라서 이러한 웜에 대항하기 위해서 웜의 전파 특성을 분석하는 것이 무엇보다 중요하다. 본 논문에서는 컴퓨터 시뮬레이션을 통해 Code Red 웜의 전파 특성을 분석한다. 특히 Code Red 웜 감염 호스트 수에 관한 기존 시뮬레이션 연구 결과가 관측된 자료와 매칭되지 않음을 보이고 이를 해결하기 위해 개선된 시뮬레이션 환경을 제시하였다. 또한 웜에 대한 초기 대응과 감염에 따른 대응이 웜의 전파 속도 변화에 어떠한 영향을 미치는지 그 결과를 보였다.

■ 중심어 : | 인터넷 웜 | 인터넷 보안 | Code Red |

Abstract

It was well known that how much seriously the Internet worm such as the Code Red had an effect on our daily activities. Recently the rapid growth of the Internet speed will produce more swift damage us in a short term period. In order to defend against future worm, we need to understand the propagation pattern during the lifetime of worms. In this paper, we analyze the propagation pattern of the Code Red worm by a computer simulation. In particular, we show that an existing simulation result about the number of infectious hosts does not match the observed data, and then we introduce a factor of revised human countermeasures into the simulation. We also show the simulation results presenting the importance of patching and pre-patching of the Internet worm.

■ keyword : | Internet Worm | Internet Security | Code Red |

I. 서 론

오늘날 산업, 정부, 그리고 심지어 일반 개인 생활이 인터넷에 점점 더 의존해감에 따라 네트워크를 통한 정보 흐름에 관한 중요성이 더욱 강조되고 있다. 그러나 해커들에 의한 네트워크 혹은 주요 서버 컴퓨터에 대한 침입이나 공격은 네트워크를 마비시키거나 서비스거부공격(DoS: Denial of Service)으로 인해 전자 상거래 서비스 중단과 같은 심각한 경제적 손실뿐만 아니라 인터넷 서비스 중단에 따른 극심한 사회적 혼란을 초래하고 있

접수번호 : #061012-002
접수일자 : 2006년 10월 12일
심사완료일 : 2006년 11월 20일
교신저자 : 강구홍, e-mail : khkang@seowon.ac.kr

다. 특히 2001년 말에 발생한 Cod Red와 CodeRedII 웜으로 인한 26억 달러 상당의 경제적 손실[1][2]을 감안할 때 인터넷 보안은 우려의 수준을 넘어서 반드시 극복해야할 당면 과제라고 볼 수 있다.

2001년 7월 19일 인터넷에 연결된 359,000개 이상의 마이크로소프트 IIS (Internet Information Services) 서버 컴퓨터가 불과 14시간 만에 Code Red 웜에 감염되었다[1][2]. 이러한 인터넷 대란은 국내에서는 더욱 심각하여 미국에 이어 Code Red 웜에 감염된 컴퓨터의 수가 세계 2위 수준인 37,94B에 달했다. 이러한 수치는 전 세계적으로 감염된 총 컴퓨터의 10.57%에 이른다[2]. 외국의 경우, Code Red 웜 감염에 따른 실측 자료[8][9] 제공과 eEye Digital Security 사와 CAIDA에 의해 Code Red 웜에 대한 자세한 동작 분석[1][2] 결과가 제공되고 있다. 그러나 국내의 경우, Code Red 웜의 이러한 엄청난 위력에도 불구하고 분석결과는 차치하고 변변한 자료 하나 구하기가 쉽지 않은 것이 현 실정이다.

기존 소프트웨어의 문제점들을 이용해 패치하지 않은 컴퓨터를 대상으로 공격하는 대부분의 인터넷 웜과 마찬가지로 Code Red 웜 역시 .ida 문제점[5][6]을 패치하지 않은 마이크로소프트 IIS 웹 서버 컴퓨터를 감염시킨다. 특히 오늘날 이러한 소프트웨어의 문제점들이 발견되고 일반 대중들이 이들 문제점을 알기까지 시간 차이가 거의 나지 않는다는 zero-day 공격으로 인해 웜의 전파 특성을 분석하는 것은 인터넷 보안 분야의 선행 연구과제로 볼 수 있다. 즉 인터넷 웜의 전파 패턴을 분석함으로서 이들 웜에 대처할 보다 효율적인 방법을 찾을 수 있을 뿐만 아니라 웜에 대한 특성을 잘 이해할 수 있다.

본 논문에서는 Code Red 웜을 대상으로 웜의 전파 패턴을 시뮬레이션을 통해 분석한다. Code Red 웜의 전파 특성은 수집된 자료를 통한 분석[8][9], 전염성 모델(epidemic model)에 의한 분석[2-4], 그리고 시뮬레이션에 의한 분석[3]이 있었다. 앞에서 설명한 바와 같이 Code Red 웜은 14 시간 만에 전 세계의 인터넷을 마비시킨 치명적인 사건이었다. 그럼에도 불구하고 웜의 전파 특성을 분석한 자료는 의외로 찾아보기 힘이 든다. 특히 시뮬레이션을 통한 Code Red 웜의 전파 패턴을 분석한 자료는 참고문헌[3]이 유일하다. 그러나 참고문헌[3]의 시뮬레이션에 의한 Code Red 감염 호스트(infectious hosts) 수의 패턴은 실측 자료[8][9]와 비교하면 상당한 차이를 보인다. 이러한 문제점은 감염가능 호스트(susceptible host)가 패치(patch)되는 현상과 감염상태 호스트(infectious host)가 패치되는 현상을 하나의 수식으로 통합 처리함으로서 발생된다. 따라서 본 논문에서는 이러한 문제점을 개선하기 위해 감염된 호스트로부터 패치되는 현상을 Kermack-McKendrick 모델[7]을 이산시간 모델로 변경하여 사용하고, 감염가능 호스트로부터 패치되는 현상은 일반 전염성 모델로부터 면역되는 모델을 사용하였다. 개선된 시뮬레이션 환경은 실측자료와 비교하면 거의 일치된 전파 패턴을 나타낸다. 한편, 소프트웨어의 보안상 문제점을 초기에 패치하는 초기 패치율과 웜에 감염된 호스트의 패치율에 따른 웜의 전파 패턴 분석 자료를 제시함으로서 오늘날 인터넷 웜의 심각한 전파 특성을 논한다.

본 논문은 서론에 이어 제2장에서는 Code Red 웜에 대해 간략히 언급하고, 제3장에서는 기존 시뮬레이션에 의한 결과분석 및 문제점을 설명한다. 제4장에서는 본 논문에서 제안하는 새로운 시뮬레이션 환경을 설명하고 시뮬레이션 결과를 기준으로 인터넷 웜의 전파 패턴을 분석한다. 마지막으로 제6장에서 결론을 맺는다.

II. Code Red 웜

본 장에서는 시뮬레이션 환경을 이해하기 위해 필요한 Code Red 웜에 대한 최소 설명만 한다. 윈도우즈 IIS 문제점을 이용한 최초의 공격 Code Red 웜은 2001년 6월 13일 발생하였다. 그러나 랜덤 수(random number) 발생에 문제점이 있어 웜 전파가 잘 이루어지지 못했으며 2001년 7월 19일 10시 경 부터 랜덤 수 발생을 수정한 Code Red 버전 2 (이하, 본 논문에서는 Code Red라 칭함) 웜이 본격적으로 활동하였다[1].

앞서 설명한 바와 같이 Code Red 웜은 .ida 버퍼 넘침 공격(buffer overflow attack)을 이용해 인터넷 상의 마이크로소프트 IIS 웹 서버들을 통해 전파된다. Code Red는 웜을 전파하기 위해 감염된 웹 서버 상에서 다음과 같

이 동작한다[1].

(i) 감염된 시스템 상에 초기 웜 환경을 설정한다.

(ii) 100개의 쓰래드(thread)를 생성한다.

(iii) 99개 쓰래드는 다른 웹 서버(TCP 포트 80)를 공격하여 웜을 전파하고 나머지 1개 쓰래드는 현재 시스템이 영문판 윈도우즈 NT/2000 시스템인지 확인하여 로컬 웹 서버의 웹 페이지를 손상시킨다. 만약 이러한 영문판 버전이 아니면 다른 쓰래드와 마찬가지로 다른 웹 서버를 감염시키는 역할을 하게 된다.

III. 기존 시뮬레이션 연구

앞서 설명한 바와 같이 Code Red의 전파 특성을 분석한 시뮬레이션 자료는 참고문헌[3]이 유일하다. 본 장에서는 이들 시뮬레이션 환경 및 결과에 대해 간략하게 언급하고 시뮬레이션 결과에 대한 문제점을 지적 한다.

1. 시뮬레이션 환경

시뮬레이션 내 개의 호스트가 직접 연결되어 있으며 이들 호스트들은 [그림 1]과 같이 세 가지 상태 - (i) S (susceptible): 감염 가능 상태, (ii) I (infectious): 감염 상태, 그리고 (iii) R(R*) (removed): 패치된 상태 - 중 하나의 상태에 있게 된다.

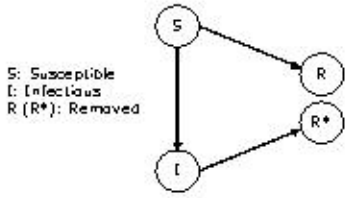


그림 1. 호스트 상태도

[그림 1]에서 보듯이 호스트는 감염 가능 상태에서 패치된 상태(S->R)로 천이되거나, 감염 가능 상태에서 감염 상태로 그리고 패치된 상태(S->I->R*)로 천이된다. 참고문헌[3]에서 시뮬레이션에 사용된 기호들은 다음과 같다.

- : 시간 에서 감염 가능 상태의 호스트 수.
- : 시간 에서 감염상태의 호스트 수
- : 시간 에서 감염 상태에서 패치된 상태로 천이한 호스트 수 (I -> R*)
- : 시간 에서 감염 가능 상태에서 패치된 상태로 천이한 호스트 수 (S -> R)
- : 전체 호스트 수
  ( ).
- : 시점 에서 감염 호스트 수
  ( )
- : 시점 에서 패치된 호스트의 전체 수
  ( )

감염 상태의 호스트는 전체 호스트를 대상으로 무작위로 선택해 감염을 시도하며 연속된 두 감염 시도 사이 시간을 감염지연시간(: infection delay time)이라 정의한다. 따라서 감염 상태에 있는 호스트는 이후 하나의 윈도우즈 IIS 서버를 발견하게 된다. 이때 선택된 호스트가 감염 가능 상태에 있다면 감염상태로 바뀌게 되나 이미 감염 상태에 있거나 패치된 상태라면 상태 변화가 일어나지 않게 된다. 는 다음과 같이 표현된다.

(1)

여기서, 는 기본 감염지연시간이고 는 다음과 같이 구할 수 있다.

□ □ (2)

한편, 는 다음과 같이 정규분포(normal distribution)으로 구할 수 있다.

~ (3)

여기서, 는 평균값이 이고 편차가 인 정규분포를 나타내며 은 모델 파리미터이다. 또한

. 한편 는 다음과 같이 결정된다.

. (4)

여기서, ≤ .

2. 시뮬레이션 결과

시뮬레이션은 다음과 같이 네 가지 경우에 대해 각각 이루어졌다.

(i) case1:
(ii) case2:
(iii) case3:
(iv) case4:

Case1은 단순 전염성 모델[4]을 가정한 것이며, case2는 감염지연시간 변화를 고려한 것이며, case3는 패치율을 고려한 것이며, 마지막으로 case4는 감염지연시간 변화와 패치율 모두를 고려한 것이다.

시뮬레이션 초기 상태 즉 초기 감염된 상태의 호스트 수가 10이고, 시뮬레이션에 사용된 파라미터는

로 각각 주어졌다. [그림 2]는 참고문헌[3]에서 제시된 시뮬레이션 결과 그래프이다.

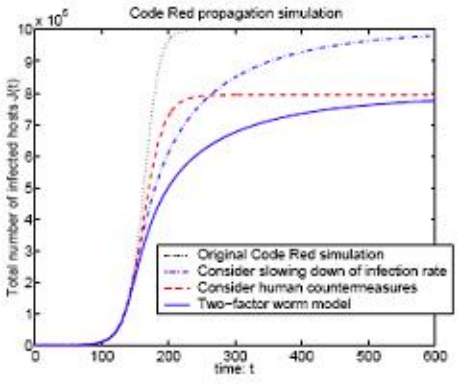


그림 2. 관련 시뮬레이션 결과[3]

3. 자체 시뮬레이션 및 토의

본 논문에서는 참고문헌[3]에서 수행한 시뮬레이션을 C-언어로 구현해 이산시간(discrete time)시뮬레이션을 수행하였다. 이때 참고문헌[3]과 동일한 환경을 제공하기 위해 동일한 시뮬레이션 파라미터 값을 사용했으며 시뮬레이션 시간을 600으로 고정하고 총 100회 시뮬레이션 한 후 각 시뮬레이션 타임유니트에서 평균값을 결과 그래프에 찍었다. [그림 3]과 [그림 4]는 이러한 시뮬레이션을 통해 얻은 와 관련 결과 그래프이다. [그림 2]와 [그림 3]을 비교해 보면 동일한 시뮬레이션 결과 를 얻을 수 있었다. 한편, 참고문헌[3]에서는 관련 시뮬레이션 결과 그래프를 보여주지 않고 있어 비교가 불가능하였다. 그러나 [그림 3]과 [그림 4]는 하나의 시뮬레이션 수행을 통해 얻어진 결과 값들이기 때문에 본 논문을 통해 제작된 시뮬레이션 환경이 참고문헌[3]과 정확하게 동일하다고 볼 수 있다.

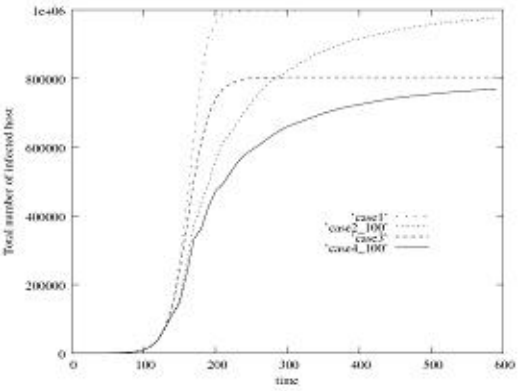


그림 3. 관련 시뮬레이션 결과

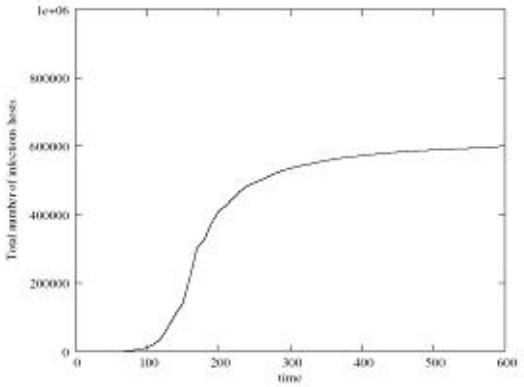


그림 4. 관련 시뮬레이션 결과

[그림 5]는 code red 웜 발생 당시 측정된 의 변화를 보여준다. 참고문헌[3]에서 언급한 바와 같이 감염지연시간과 호스트들의 패치를 고려한 시뮬레이션 모델(case4, two-factor worm model[3])이 관찰값을 가장 정확하게 근사함을 확인할 수 있다. [그림 3]의 case4 그래프와 [그림 5]의 상위값과 서로 다른 것은 code red 웜이 자정(00:00)을 기준으로 감염행위를 중단함에 따른 것이다. 이러한 code red의 동작은 [그림 6]의 자료를 통해 더욱 분명해진다. 즉 [그림 6]에서 보듯이 시간 00:00을 기준으로 code red 웜 감염 호스트로부터 더 이상의 감염행위가 발생하지 않는다. 특히 [그림 6]에서 보듯이 19:00를 기준으로 감염 상태 호스트들의 패치행위로 인해 감염 상태 호스트 수가 줄어들고 있음을 확인할 수 있다. 그러나 [그림 4]의 시뮬레이션 결과는 이러한 패치 결과가 제대로 반영되지 않고 있음을 보여준다. 따라서 참고문헌[3]에서 사용하는 시뮬레이션 환경이 의 변화는 잘 반영하고 있으나 의 변화는 실제 관찰된 데이터와 매우 상이함을 보여준다.

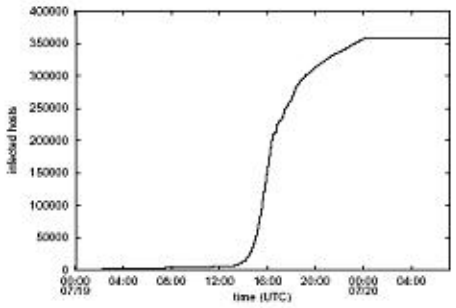


그림 5. 관찰된 [2]

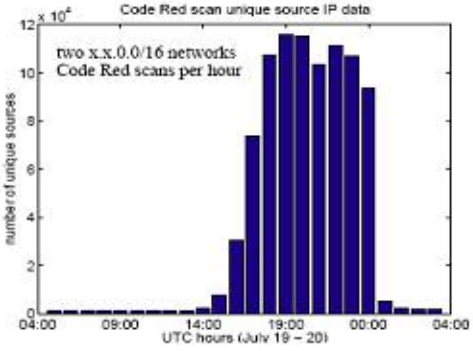


그림 6. 관찰된 [3][8][9]

IV. 새로운 시뮬레이션 환경

1. 시뮬레이션 파라미터 수정

기존의 시뮬레이션은 감염 가능상태(susceptible)와 감염상태(infectious)에서 패치되는 호스트를 하나의 수식으로 표현하였다(식4 참조). 그러나 감염 가능 상태의 호스트가 패치되는 현상과 이미 감염된 호스트가 패치되는 현상은 매우 다르다. 본 시뮬레이션에서는 감염된 호스트로부터 패치되는 현상을 Kermack-McKendrick 모델[7]을 이산시간 모델로 변경하여 다음과 같이 사용한다.

(5)

한편, 감염 가능상태에서 패치되는 현상은 식(4)를 변형하여 다음과 같이 사용한다. 즉 감염 호스트 수 가 증가하면 (혹은 감염 가능 호스트 수 가 감소하면) 감염 가능 상태에서 패치된 호스트 수 는 증가(혹은 감소) 하게 된다.

(6)

2. 결과 분석

시뮬레이션에 사용된 관련 파라미터 ,

로 각각 설정하였다. [그림 7]은 본 논문에서 제안한 시뮬레이션 환경과 제3장에서 설명한 기존 시뮬레이션 결과를 비교하기 위해 식5와 식6을 이용해 구한 와 를 각각 [그림 3]과 [그림 4]에서 구한 그래프와 함께 나타내었다.

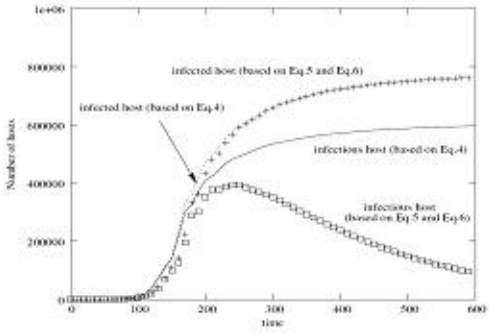


그림 7. 와 시뮬레이션 결과 비교

[그림 7]에서 보듯이 는 기존 시뮬레이션 결과와 비교해 거의 차이가 없음을 확인할 수 있으나 는 기존의 시뮬레이션 결과와 확연히 차이가 남을 볼 수 있다. 수정된 시뮬레이션 환경에 의한 [그림 7]의 즉 감염 상태의 호스트 수는 시간이 지남에 따라 패치 상태로 급격히 천이하여 그 수가 감소함을 보여준다. 이러한 현상은 실측자료인 [그림 6]과 동일한 현상이다. 따라서 본 논문에서 제시한 새로운 시뮬레이션 환경이 Code Red 웜의 전파패턴을 보다 정확하게 반영함을 확인할 수 있다.

[그림 8]은 실측자료([그림 5]와 [그림 6])와 시뮬레이션 결과를 비교하기 위해 시뮬레이션 시간 영역을 설정해 보여준다. [그림 8]의 시뮬레이션 시간 300까지 영역을 구별해 실측자료와 비교해보면 본 논문에서 실시한 시뮬레이션 결과가 기존 시뮬레이션에 비해 실측 자료를 매우 잘 나타내고 있음을 확인할 수 있다. 즉 [그림 8]의 는 지수 함수적으로 증가하다가 지점에서부터 증가 속도가 둔화된다. 이러한 현상은 [그림 5]의 실측 자료에서도 확인할 수 있다. 뿐만 아니라 [그림 8]의 는 증가하다가 가 전체 호스트의 60% 지점에서 감소하기 시작한다. 이러한 현상은 그림6의 실측 자료에서도 확인할 수 있다.

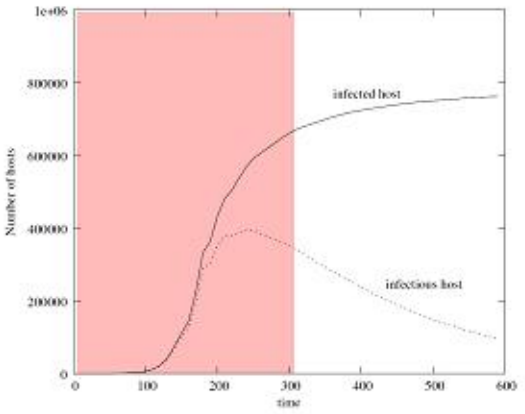
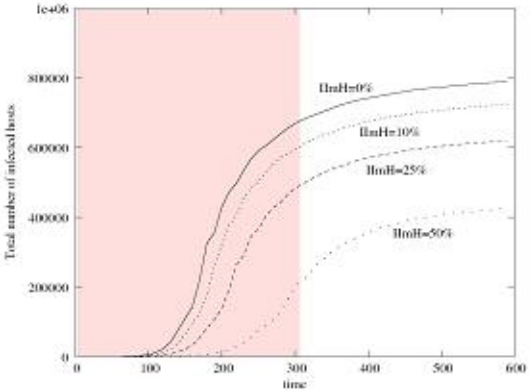


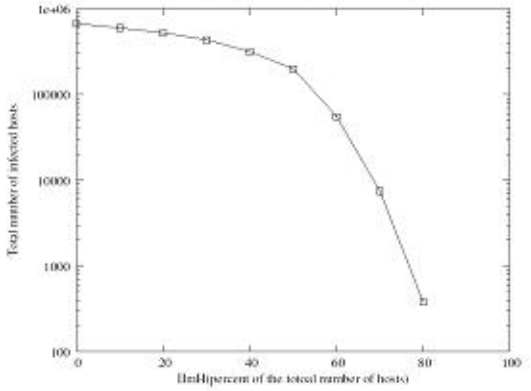
그림 8. 실측 자료와 비교를 위한 시뮬레이션 시간 영역

3. 토의

[그림 9]는 Code Red 웜이 발생하기 전에 패치된 호스트(IImH: Initially immune Host) 수에 따른 의 변화를 보여준다. [그림 9](a)에서 보듯이 IImH가 전체 호스트 수의 50%일 경우 시뮬레이션 시간 300을 기준으로 가 20% 수준에 불과함을 볼 수 있다. 이러한 사실은 Code Red 웜이 사용한 IIS 취약점을 한 달 전에 이미 발표했고 이를 막기 위해 패치를 전체 감염가능 호스트의 절반 정도 수준으로 했다면 그날과 같은 심각한 인터넷 대란을 사전에 어느 정도 수준에서 막을 수 있었다는 결론을 내릴 수 있다. 그러나 IImH가 전체 호스트 수의 25%정도 수준에서는 그 효과가 극히 부족해 보인다. 이러한 현상은 Code Red 웜 자체가 워낙 전염성이 강하기 때문이다. 따라서 오늘날 인터넷 보안은 zero-day 공격의 중요성이 더욱 강조될 것으로 예상된다. 한편 [그림 9](b)에서 보듯이 패치의 효과가 뚜렷해지는 지점은 IImH가 전체 호스트 수의 60% 정도로 추정된다.



(a) 시간에 따른 변화



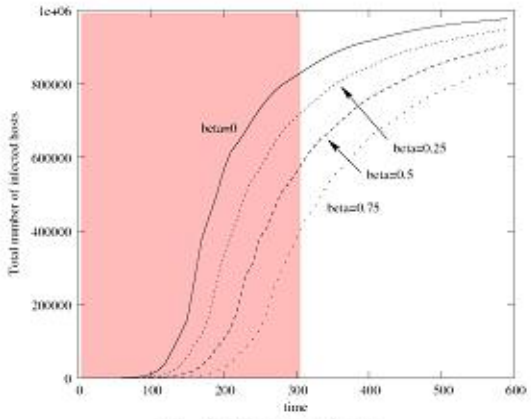
(b)시뮬레이션 시간 300에서 측정값

그림 9. IImH(초기 패치된 호스트) 값에 따른 의 변화

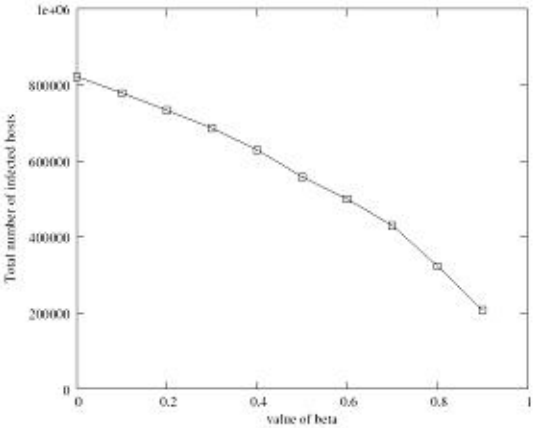
2001년 7월 19일 Code Red 웜에 감염된 호스트를 대상으로 2001년 8월 14일 패치된 호스트 분포를 조사한 결과를 살펴보면 영국 65.65%, 미국 59.59%, 그리고 한국 20.20% 수준이다[2]. 이제 Kermack-Mckendrick 모델[7]를 이용해 감염된 호스트가 패치되는 정도에 따른 웜의 확산효과를 알아보자. Kermack-Mckendrick 모델은 다음과 같은 패치모델을 가정한다. 즉 감염 가능 상태에서 패치되는 호스트는 무시하고 감염상태의 호스트가 얼마나 적극적으로 패치 과정에 참여하는지를 파라미터 를 이용해 나타낸다(식7 참조).

(7)

[그림 10]은 값에 따른 의 변화를 보여준다. 그림에서 보듯이 에서 를 기준으로 가 50% 수준으로 떨어짐을 볼 수 있다. 그러나 전체 감염상태 호스트 수의 75%가 패치된다는 가정에도 불구하고 그 효과는 의외로 적음을 알 수 있다. 이러한 사실은 Code Red 웜의 전파 능력이 탁월하다는 것을 의미한다. 따라서 그림9에서 보듯이 초기 패치율이 오늘날 인터넷 환경에서 무엇보다 중요하다고 결론 내릴 수 있다.



(a) 시간에 따른 변화



(b) 시뮬레이션 시간 300에서 측정값

그림 10. 값에 따른 의 변화

V. 결 론

본 논문에서는 컴퓨터 시뮬레이션을 통해 Code Red 웜의 전파 특성을 분석하였다. 특히 Code Red 웜 감염 호스트 수에 관한 기존 시뮬레이션 결과가 관측된 자료와 매칭되지 않음을 보이고 이를 해결하기 위해 개선된 시뮬레이션 환경을 제시하였다. 또한 웜에 대한 초기 대응과 감염에 따른 대응이 웜의 전파 속도 변화에 어떠한 영향을 미치는지 그 결과를 보였다.

참 고 문 헌

[1] http://research.eeye.com/html/advisories/published/AL20010717.html

[2] http://www.caida.org/publications/papers/2002/codered/codered.pdf

[3] C. Zou, W. Gong, and D. Towsley, "Code Red Worm Propagation Modeling and Analysis," CCS'02, 2002.

[4] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," 11th Usenix Security Symposium, 2002.

[5] http://research.eeye.com/html/advisories/published/AD20010618.html

[6] C. Cowan, P. Wagle, and C. Pu, "Buffer Overflows: Attacks and Defenses for the Vulnerability of the Decade," SANA 2000, 2000.
[7] N. T. Bailey, *The Mathematical Theory of Infectious Diseases and its Application*, Hafner Press, New York, 1975.
[8] http://lists. jammed.com/incidents/2001/07/0159.html
[9] http://lists. jammed.com/incidents/2001/07/0149.html

저 자 소 개

강 구 홍(Koo-Hong Kang) 정회원

- 1985년 8월 : 경북대학교 전자공학과 (공학사)
- 1990년 2월 : 충남대학교 전자공학과 (공학석사)
- 1998년 2월 : 포항공과대학교 전자계산학과 (공학박사)
- 2000년 9월 ~ 현재 : 서원대학교 컴퓨터정보통신공학부 부교수

<관심분야> : 컴퓨터 네트워크, 네트워크 프로그래밍