

# Il worm Code Red

Simone Bassani  
sbassani92@gmail.com

Simone Falvo  
smvfal@gmail.com



UNIVERSITA' degli STUDI di ROMA  
TOR VERGATA

<i>INDICE</i>	1
---------------	---

## **Indice**

<b>1</b>	<b>Introduzione</b>	<b>2</b>
<b>2</b>	<b>Dettagli incidente</b>	<b>2</b>
<b>3</b>	<b>Diffusione e sistemi coinvolti</b>	<b>2</b>
<b>4</b>	<b>Conseguenze e impatto economico</b>	<b>5</b>
<b>5</b>	<b>Esperimento: se Code Red non fosse stato scoperto</b>	<b>5</b>
<b>6</b>	<b>Contromisure</b>	<b>8</b>
<b>7</b>	<b>Conclusioni</b>	<b>8</b>

## 1 Introduzione

## 2 Dettagli incidente

## 3 Diffusione e sistemi coinvolti

Non esistono molti dati riguardo la diffusione e l'impatto di Code Red, ma sicuramente l'analisi svolta da Moore et al. [1] è la più completa e significativa che è stata effettuata.

La loro analisi si è svolta analizzando due set di dati relativi al monitoraggio di pacchetti TCP SYN indesiderati che giungevano rispettivamente nella rete /8 di ricerca dell'università della California a San Diego e in altre due reti /16 del Lawrence Berkeley Laboratory.

Analizzando gli indirizzi IP di provenienza sono riusciti a determinare l'estensione della diffusione del worm e contando il numero di diversi indirizzi IP che effettuavano le scansioni ripetute è stato possibile effettuare una stima sul numero di host infettati.

I risultati hanno mostrato che tra la mezzanotte del 19 Luglio a quella del 20 Luglio sono stati infettati intorno ai 359000 distinti indirizzi IP provenienti da ogni parte del mondo, la figura 1 mostra la distribuzione geografica delle macchine infette. Inoltre poiché i dati raccolti costituiscono soltanto un campione delle richieste di connessione, il numero di host rilevati fornisce un lower bound per il numero totale di host compromessi.

Le figure 2 e 3 danno un'idea del forte impatto che ha avuto la versione di Code Red a seme dinamico, infatti si vede che a partire dalla mattina del 19 Luglio c'è stato un improvviso incremento del tasso di infezione che ha raggiunto un valore di 2000 host al minuto. È interessante notare anche la decrescita esponenziale di tale tasso, dovuta probabilmente al conseguente stato di indisponibilità dei server, all'adozione di contromisure e ai gravi problemi causati alla rete globale che hanno portato ad un inevitabile rallentamento del traffico.

La figura 4 mostra il numero di host che hanno smesso di sondare la rete al variare del tempo e, a conferma di quanto detto sopra, tale numero era già pari a circa 200000 unità (oltre il 50% delle infezioni totali) prima che il worm cessasse definitivamente l'attività di diffusione per procedere alla fase di attacco DDoS.

Per comprendere la composizione demografica dell'utenza coinvolta, i ricercatori del CAIDA [1] hanno esaminato i vari livelli di dominio e la locazione geografica degli host infetti.

La figura 5 riassume i risultati di tale studio: per quanto riguarda i domini di primo livello la proporzione rispetta la allora attuale situazione dei web server, mentre è curioso notare che il 10% delle macchine compromesse sono state localizzate in Korea; i principali nomi di dominio sono costituiti da server provider per infrastrutture casalinghe e piccole imprese, da qui si vede che anche queste piccole realtà hanno un ruolo rilevante riguardo la salute globale di internet.

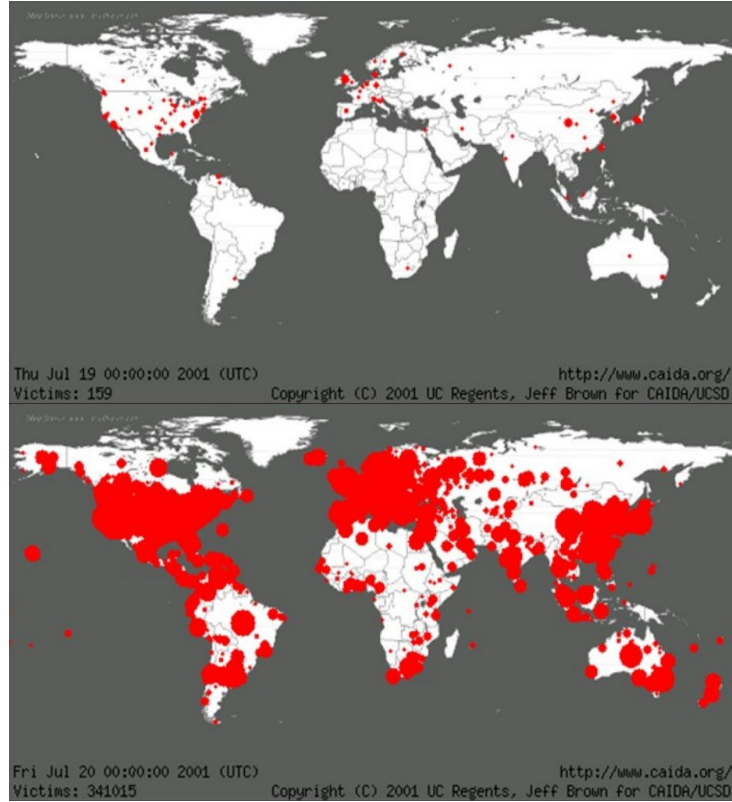


Figura 1: diffusione Code Red

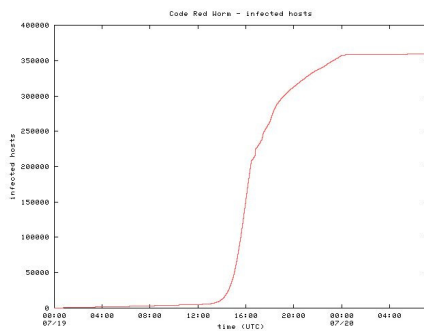


Figura 2: totale host infettati

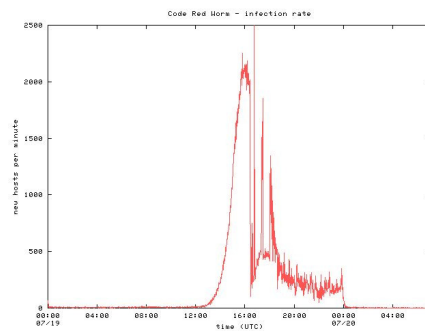


Figura 3: tasso di infezione

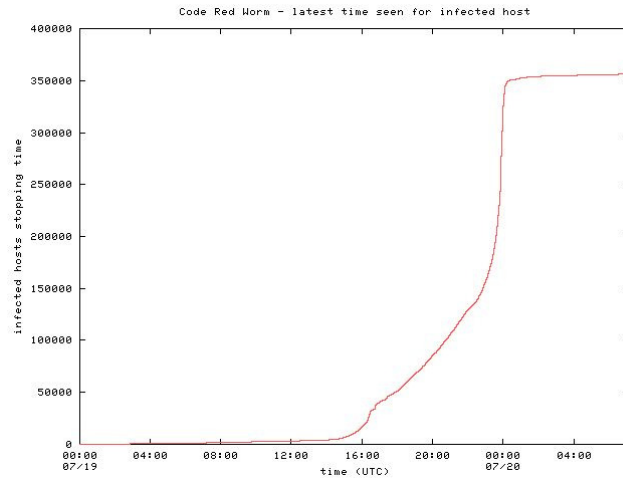


Figura 4: host "disattivati"

Top 10 Countries		
Country	hosts	hosts(%)
United States	157694	43.91
Korea	37948	10.57
China	18141	5.05
Taiwan	15124	4.21
Canada	12469	3.47
United Kingdom	11918	3.32
Germany	11762	3.28
Australia	8587	2.39
Japan	8282	2.31
Netherlands	7771	2.16

TABLE I

TOP TEN COUNTRIES WITH CODE-RED INFECTED HOSTS ON TOP TEN TOP-LEVEL DOMAINS WITH CODE-RED INFECTED HOSTS ON JULY 19.

Top 10 Top-Level Domains		
TLD	hosts	hosts(%)
Unknown	169584	47.22
net	67486	18.79
com	51740	14.41
edu	8495	2.37
tw	7150	1.99
jp	4770	1.33
ca	4003	1.11
it	3076	0.86
fr	2677	0.75
nl	2633	0.73

TABLE II

Top 10 Domains		
Domains	hosts	hosts(%)
Unknown	169584	47.22
home.com	10610	2.95
rr.com	5862	1.63
t-dialin.net	5514	1.54
pacbell.net	3937	1.10
uu.net	3653	1.02
aol.com	3595	1.00
hinet.net	3491	0.97
net.tw	3401	0.95
edu.tw	2942	0.82

TABLE III

TOP TEN DOMAINS WITH CODE-RED INFECTED HOSTS ON JULY 19.

Figura 5: risultati analisi demografica

## 4 Conseguenze e impatto economico

I principali effetti del worm Code Red furono il degradamento delle prestazioni e la perdita di stabilità dei sistemi coinvolti. Il costo globale complessivo stimato fu di 2.6 miliardi di dollari [3, 4], di cui 1.1 miliardi impiegati nell'ispezione ed il recupero dei server ed i restanti 1.5 miliardi riguardarono le perdite di produttività a seguito dell'indisponibilità dei sistemi.

Quest'ultimi comprendevano non soltanto le macchine server degli utenti finali, ma anche vaste porzioni dell'infrastruttura di rete che furono completamente disabilitate, molte compagnie provider di dispositivi di rete sperimentarono un'indisponibilità media di ben 36 ore [4].

Il processo di propagazione del worm ha generato un'enorme quantità di pacchetti. Sebbene il volume di questi pacchetti era relativamente piccolo rispetto al normale traffico di rete, l'ingente quantità ha causato congestionamento e gravi problemi ad alcuni router, specialmente quelli con risorse limitate. Per esempio, a causa della generazione randomica degli indirizzi IP, molti pacchetti non venivano inoltrati poiché la destinazione risultava sconosciuta, finendo così per riempire le cache ARP, esaurire la memoria e provocare il riavvio dei dispositivi [5].

La figura 6 mostra il costo complessivo di Code Red in relazione agli incidenti più rilevanti del periodo [3].

Analysis by Incident			
Year	Code Name	Worldwide Economic Impact (\$ U.S.)	Cyber Attack Index
2001	Nimda	\$635 Million	0.73
2001	Code Red(s)	2.62 Billion	2.99
2001	SirCam	1.15 Billion	1.31
2000	Love Bug	8.75 Billion	10.00
1999	Melissa	1.10 Billion	1.26
1999	Explorer	1.02 Billion	1.17

Figura 6: confronto danno economico

Il sito web della Casa Bianca riuscì ad evitare conseguenze sostanzialmente “disattivando” l'indirizzo IP bersaglio dell'attacco DDoS, ovvero reindirizzando tutte le richieste non malevole verso altri indirizzi associati, questo è stato possibile perché il worm è stato progettato per inviare traffico verso un unico indirizzo IP, invece dell'intero blocco di indirizzi relativi al dominio della Casa Bianca.

## 5 Esperimento: se Code Red non fosse stato scoperto

Come appena visto Code Red ha costituito una vera e propria piaga, che si è propagata in ogni angolo del pianeta e ha causato gravi danni nonostante sia stato scoperto diversi giorni prima che la seconda versione facesse la sua comparsa ed iniziasse a diffondersi in modo efficace.

A fronte di ciò, abbiamo effettuato una simulazione per vedere cosa sarebbe accaduto se il worm non fosse stato scoperto, e quindi avere una misura dell'entità

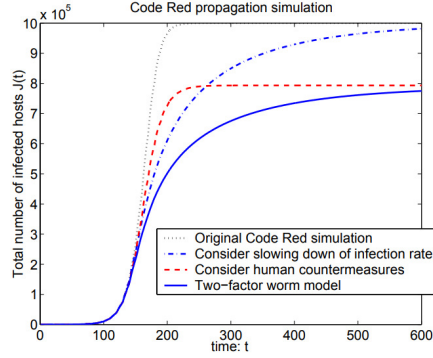


Figura 7: infezioni al variare del tempo

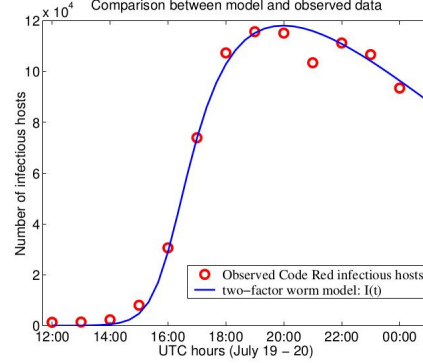


Figura 8: confronto dati-modello

del danno in termini di numero di host compromessi, così poi da confrontare i risultati con il caso reale in cui la consapevolezza della presenza di tale minaccia ha spinto gli utenti ad adottare contromisure come patch, firewall, antivirus e simili.

Per fare questo, innanzitutto è stato necessario l'ausilio di un modello epidemico che simulasse fedelmente il comportamento ed in particolare la diffusione del worm.

I classici modelli per lo studio dello sviluppo epidemico, però, non sono adatti a replicare il comportamento di Code Red, perché non tengono conto delle contromisure umane intraprese durante il processo di diffusione e inoltre considerano il tasso di infezione costante nel tempo. Questi due fattori hanno portato Zou et al. [2] alla derivazione del “Two Factor Worm Model” con cui hanno dimostrato tramite simulazione che i risultati approssimano bene l'andamento dei dati osservati, in particolare sono riusciti a giustificare il rallentamento delle scansioni che si è verificato appena prima il cessamento dell'attività di diffusione del worm. Tale evento, infatti, è la conseguenza della violenta propagazione su larga scala che ha provocato congestionamento e danni alla rete.

Le figure 7 e 8 mostrano i risultati ottenuti da Zou et al., in particolare la figura 8 mostra il confronto con i dati reali osservati da Goldsmith and Eichman [7, 6].

Di seguito è riportato il set di equazioni differenziali che descrivono il sistema, caratterizzanti sono il fattore  $\beta(t)$  che rappresenta la frequenza di infezione variabile nel tempo, ed i fattori  $R(t)$  e  $Q(t)$  che costituiscono il numero di host rimossi dalla popolazione infetta a seguito di contromisure a valle del contagio, ed il numero di host rimossi dalla popolazione suscettibile (vulnerabile) a seguito di contromisure a monte, le altre variabili sono riassunte nella tabella di figura 9.

$$\left\{ \begin{array}{l} dS(t)/dt = -\beta(t)S(t)I(t) - dQ(t)/dt, \\ dR(t)/dt = \gamma I(t), \\ dQ(t)/dt = \mu S(t)J(t), \\ \beta(t) = \beta_0[1 - I(t)/N]^\eta, \\ N = S(t) + I(t) + R(t) + Q(t), \\ I(0) = I_0 \ll N; S(0) = N - I_0; R(0) = Q(0) = 0; \end{array} \right. \quad (1)$$

La simulazione è stata eseguita con lo stato iniziale specificato dai seguenti parametri:  $N = 1000000$ ,  $I_0 = 1$ ,  $\eta = 3$ ,  $\gamma = 0.05$ ,  $\mu = 0.06/N$ , e

$\beta_0 = 0.8/N$ .

$I_0$  e  $\beta_0$  rappresentano rispettivamente numero di host infetti e tasso di infezione iniziali,  $\gamma$  e  $\mu$  i tassi di rimozione, infine  $\eta$  è un parametro di sensibilità che regola  $\beta(t)$  in funzione di  $I(t)$ , sostanzialmente è un fattore che indica quanto il tasso di infezione dipende dal numero di host infetti in un certo istante di tempo, ad esempio  $\eta = 0$  implica un tasso di infezione costante.

Il nostro studio è consistito nel semplificare il “Two Factor Model” eliminando

Notation	Explanation
$S(t)$	Number of susceptible hosts at time $t$
$I(t)$	Number of infectious hosts at time $t$
$R(t)$	Number of removed hosts from the infectious population at time $t$
$Q(t)$	Number of removed hosts from the susceptible population at time $t$
$N$	Total number of hosts under consideration, $N = I(t) + R(t) + Q(t) + S(t)$
$J(t)$	Number of infected hosts at time $t$ , i.e., $J(t) = I(t) + R(t)$
$C(t)$	Total number of removed hosts at time $t$ , i.e., $C(t) = R(t) + Q(t)$
$\beta(t)$	Infection rate at time $t$
$D(t)$	Infection delay time in simulation, representing the time for a Code Red worm to find an IIS server

Figura 9: notazioni variabili del modello

il fattore relativo alle contromisure intraprese dagli utenti, proprio come se il worm avesse potuto agire indisturbato senza che nessuno si accorgesse della sua presenza, pertanto le variabili  $R(t)$  e  $Q(t)$  relative agli host in stato di “rimosso” sono state annullate, ottenendo così il seguente modello semplificato:

$$\begin{cases} dS(t)/dt &= -\beta(t)S(t)I(t), \\ \beta(t) &= \beta_0[1 - I(t)/N]^\eta, \\ N &= S(t) + I(t), \\ I(0) &= I_0 \ll N; S(0) = N - I_0; \end{cases} \quad (2)$$

La simulazione è stata realizzata tramite Simulink, a partire dalle stesse condizioni iniziali poste da Zou et al. e per un tempo di simulazione equivalente a circa 24h.

I risultati ottenuti (Figura 10) mostrano che in 12 ore (metà tempo di simulazione) circa il 70% degli host suscettibili sono stati infettati, approssimativamente la stessa proporzione della simulazione effettuata da Zou et al., quindi in definitiva il fatto che Code Red sia stato scoperto prima della sua violenta diffusione non ha alleviato significativamente le conseguenze in termini di macchine infettate.

Questo risultato mette in evidenza che di fronte ad un worm così aggressivo come Code Red, con un tasso di propagazione che ha raggiunto un picco di 2000 infezioni al minuto, le contromisure che vengono adottate a seguito dell’inizio dell’epidemia non contribuiscono a porre rimedio in maniera efficace da un punto di vista globale, ma ciò che può fare la differenza sono le contromisure preventive che vengono attuate prima degli attacchi, ovvero quelle che tendono a ridurre il numero di host potenzialmente a rischio. Quindi il vero fattore che ha favorito gli effetti devastanti del worm è stata la mancata applicazione della patch che andava ad eliminare la vulnerabilità sfruttata, nonostante fosse stata resa disponibile non poche settimane prima della comparsa della minaccia.



## 6 Contromisure

## 7 Conclusioni

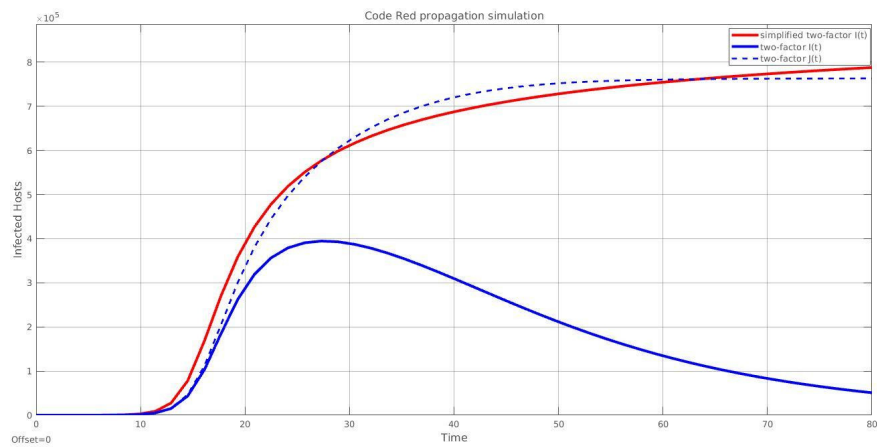


Figura 10: modello semplificato vs. two-factor

## Riferimenti bibliografici

- [1] Moore, David & Shannon, Colleen & Brown, Jeffery. *"Code-Red: a case study on the spread and victims of an Internet worm"* CAIDA.  
<http://www.caida.org/publications/papers/2002/codered/codered.pdf>
- [2] Cliff Changchun Zou, Weibo Gong, Don Towsley. *"Code Red Worm Propagation Modeling and Analysis."* Dept. Electrical & Computer Engineering Univ. Massachusetts Amherst, MA  
<http://www.cs.ucf.edu/~czou/research/codered.pdf>
- [3] *"Malicious Code Attacks Had \$13.2 Billion Economic Impact in 2001."* Computer Economics. September, 2002  
<https://www.computereconomics.com/article.cfm?id=133>
- [4] Schauer, Renee C.. *"The Mechanisms and Effects of the Code Red Worm."* Sans Institute. 2001  
<https://www.sans.org/reading-room/whitepapers/dlp/the-mechanisms-and-effects-of-the-code-red-worm-87>
- [5] *"Code Red Worm - Customer Impact."* Cisco. July 20, 2001  
<https://tools.cisco.com/security/center/content/CiscoSecurityAdvisory/cisco-sa-20010720-code-red-worm>
- [6] K. Eichman. *"Possible CodeRed Connection Attempts."* July 20, 2001  
<http://lists.jammed.com/incidents/2001/07/0159.html>
- [7] D. Goldsmith. *"Possible CodeRed Connection Attempts."* July 20, 2001  
<http://lists.jammed.com/incidents/2001/07/0149.html>