Computers & Security (2005) **24**, 280–286

ELSEVIER

Computers & Security

www.elsevier.com/locate/cose

Infection dynamics on the Internet

David B. Chang[a,*], **Carl S. Young**[b]

[a] *Consultant*
[b] *Goldman Sachs & Co., Office of Global Security and Office of Information Security, 85 Broad Street, New York, NY 10004, USA*

Accepted 15 March 2005

KEYWORDS
Network security;
Virus;
Scale-free

Abstract In previous works, the connectivity of nodes in social networks such as the Internet has been shown to follow a scale-free distribution in which there is a larger probability of nodes with lower connectivity and a smaller probability of nodes with higher connectivity. This network structure facilitates communication but also aids in the propagation of viruses. In this work, solutions have been obtained for a dynamical mean-field equation that characterizes virus infections and growth in scale-free networks. In contrast to previous findings, a threshold condition has been found for the persistence of computer infections. The effect of connectivity-dependent growth and recovery rates is also reported. It has been found that it is possible to reduce the deleterious effects of viruses by preferentially discouraging growth and enhancing recovery in high-connectivity nodes. Significantly, a security ''figure-of-merit'' has been derived that will allow network administrators to sample their environment in real time and measure the risk relative to E-mail-borne threats.

Introduction and background

Simplified mathematical descriptions of the dynamic behavior of viruses in biological and computer systems involve the well-known logistic equation. This is a first order non-linear differential equation of the form

$$da/dt = \alpha a(1-a) \tag{1}$$

where a represents the fraction of infected nodes, t is time, and α is the rate at which nodes become infected. Re-infection of disinfected and therefore susceptible nodes is not considered in this simplified model of behavior. A solution to Eq. (1) is given by

$$a = \frac{e^{\alpha t}}{1+e^{\alpha t}} \tag{2}$$

for the case where $a = 1/2$ at $t = 0$.

A plot of Eq. (2) yields the familiar sigmoid where the initial fraction of infected nodes is small. Some time later, the fraction of infected

* Corresponding author.
E-mail addresses: dbcsfc@aol.com (D.B. Chang), carl.young@gs.com (C.S. Young).

0167-4048/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.cose.2005.03.004

nodes rises precipitously. For large time t, a approaches unity, as all nodes have either been infected and died or have developed an immunity from infection. For biological systems the logistic equation describes a population where a fraction of the community has either died or developed antibodies to the infection. The analogue of developing antibodies in a computer network is characterized by the remediation and patching of nodes. It is clear from Eq. (3) that the dynamic behavior of an infection is solely dependent on the infection rate α.

However, this model assumes equal probabilities for node linking and a constant network size. In other words, assumptions inherent in Eq. (1) are that the probability of infecting a particular node is independent of the particular node itself, and that the network adds no new nodes with time. E-mail-type networks fall into a category known as social networks that exhibit both growth and preferential attachment (Barabasi and Albert, 1999).

With respect to growth, standard network models often assume there are a fixed number of nodes that are either randomly connected (Erdos and Renyi) or exhibit small world behavior and clustering (Watts and Strogatz), but where the total number of nodes never changes. Networks such as the Internet are continuing to add nodes, thereby increasing the number of vertices with time.

Some networks also display preferential attachment, where the probability of connecting to a new node is greater for nodes that already exhibit a higher number of connections. This characteristic is an important feature of the Internet, and accounts for many of the important behavioral phenomena associated with the propagation of viruses. Moreover, the combination of preferential attachment and the continuous addition of vertices leads to a model of network growth that is scale invariant (Barabasi and Albert, 1999).

In contrast with other network models, the topology of social networks such as the Internet can be characterized by a scale-free distribution of network nodes. In these types of networks, the probability of connectivity $P(k)$ for any node of connectivity k, scales as a power law:

$$P(k) = k^{-\gamma} \tag{3}$$

for $m < k < k_{\max}$.

Eq. (3) suggests that for scale-free networks, a large number of its nodes have a small number of links to other nodes, and a small number is highly-linked. Moreover, this inverse power law distribution is thought to have important security implications, where the highly-connected nodes play a critical role in facilitating virus propagation (Ebel et al., 2002). Therefore, smaller values for γ imply a greater number of highly-connected nodes in the network. Typical values have been calculated to be in the 2–4 range, and one study revealed a measured value of 1.81 (Ebel et al., 2002).

In an important work published in 2001, an analysis of the propagation of computer viruses was performed using a ''mean field'' analysis (Pastor-Satorras and Vespignani, 2001). In this paper, data on viral infections on the Internet was analyzed, and a mean field equation depicting the time evolution of the probability of viral infection as a function of the node's connectivity was introduced. Mean field approximations represent a form of averaging over many elements of a system, and are often used in physics and phase transition-type calculations.

Pastor-Satorras and Vespignani (2001) used numerical simulation to study the time behavior and steady state of virus propagation, as well as to obtain analytic expressions for the steady state virus-spreading condition. The time rate of change of the probability ρ_k of a node with connectivity k infected with a virus was found to equal the decay in the probability of infection resulting from applying network remediation (e.g., patching infected nodes) plus a term proportional to the probability of linking to an already-infected node. In the steady state, $\partial\rho_k/\partial t = 0$.

The authors also relied on a widely-cited result by Albert et al. (2000) that specified a value for the exponent γ in Eq. (3). Importantly, a narrow range of nodes relative to their connectivity was examined in this work. The values of connectivity examined ranged from nodes of low connectivity where the virus decay rate exceeded the growth rate and included nodes of higher connectivity where the virus growth rate exceeded the decay rate. This analysis yielded an expression for the steady-state probability Θ^{ss} that a given node in a scale-free network pointed to an infected node. This important expression was given by

$$\Theta^{ss} = \frac{\exp^{-1/\lambda m}}{\lambda m} \tag{4}$$

- δ denotes the remediation rate of infected nodes (i.e., the rate of nodes being restored following infection).
- υ is the infection rate of an uninfected node if it is connected to an infected node.
- k is the number of connections or links of a node.

- m is the minimum number of nodes available for connection.
- $\lambda = \upsilon/\delta$.

Eq. (4) implies that zero values of Θ^{ss} are not permitted for any finite λ. This suggests that a computer virus can pervade a network with finite prevalence in sufficiently large networks; once established, viruses will grow or decay but not remain static under steady-state conditions. The authors also concluded that these results implied scale-free networks of sufficient size required no threshold for epidemic spreading. These results dramatically departed from previously held notions on infections since it was believed that viruses died out (i.e., the prevalence is zero) below some threshold infection rate. The explanation given for this departure was the increased statistical likelihood of encountering nodes with higher connectivity in scale-free networks.

In the data analysis portion of Pastor-Satorras and Vespignani (2001), the surviving probabilities of 814 different viruses in the 50-month-period between February 1996 and March 2000 were examined. It was found that file viruses (i.e., those that infect a computer when it runs an infected application) exhibited an exponential decay in time with a characteristic time constant of seven months. Boot viruses (i.e., those that spread by infected applications but copy themselves on to the boot sector of the hard drive) and macro viruses (i.e., those that infect data files and are therefore platform-independent), also exhibit exponential decays but with a characteristic time constant of 14 months. Some of the data examined also suggested that there might be a low level persistence in the viral infection. These findings tended to support the analytical conclusions as expressed in Eq. (4).

Network viruses in steady-state conditions

Further examination of the steady state condition yields interesting properties of virus propagation in scale-free networks. Applying Eq. (4) to the aforementioned range of steady-state values of connectivity yields the condition

$$1 \ll \exp(1/\lambda m) \ll (k_{max}/m) \tag{5}$$

$k_{max} = N-1 =$ the maximum number of nodes that a single node can connect to, and $m =$ the minimum number of nodes available for connection.

Eq. (5) sets an upper limit on the magnitude of the remediation-to-infection rate (i.e., $\delta/\upsilon = 1/\lambda$). In fact, Eq. (5) defines the condition that separates a persistent infectious state from a non-persistent one. We also see that the larger the k_{max}, the easier it is to satisfy Eq. (5).

By applying Eq. (5) to the mean field equation for the steady-state condition (i.e., when $\partial\rho_k/\partial t = 0$) and evaluating this expression under various network connectivity conditions, we can further characterize the probability that a node will be infected in the steady state, ρ_k^{ss}.

Such an analysis reveals that when there is low node connectivity, i.e., $(k/m)\exp(-1/\lambda m) \ll 1$,

$$\rho_k^{ss} = (k/m)\exp(-1/\lambda m) \tag{6}$$

When there is high connectivity, i.e. $(k/m)\exp(-1/\lambda m) \gg 1$,

$$\rho_k^{ss} \sim 1 \tag{7}$$

Therefore, when the steady-state condition applies, the probability that a node with small connectivity is infected can be much less than 1 (increasing linearly with connectivity k), and the probability that a node with large connectivity becomes infected is almost 1.

As noted previously, Pastor-Satorras and Vespignani (2001) assumed values of node connectivity such that the decay rate exceeds the growth rate for nodes of low connectivity and where the growth rate exceeds the decay rate for nodes of high connectivity. However, there are two other important ranges of network connectivity conditions to consider.

The first case is when the infection growth rate υ greatly exceeds the decay rate δ *for all node connectivity values* k. This situation exists in a network that has little anti-viral prevention and little remediation software. By applying a steady state condition to the mean field equation it becomes apparent that a persistent infection state is possible in which the probability that a given link points to an infected node is close to unity for all connectivity values k.

In the second case, the infection decay (i.e., remediation) rate greatly exceeds the growth rate *for all* k. This can occur in networks for which attention is paid to maintaining viral prevention software, rapid incorporation of patches, and diligence in implementing remediation measures. In this case we find no non-zero steady state exists when $\exp(1/\lambda m) > k_{max}/m$.

To recap, we have shown that a network infection condition can exist under two conditions in the steady state:

1. When infection growth is larger than decay for large connectivity k and infection growth is smaller than infection decay for small connectivity.
2. When infection growth is larger than infection decay for all k.

No steady state condition is possible when the decay rate is greater than the infection rate for all connectivity k.

This shows that a threshold condition does indeed exist for network infection persistence in the steady state, even for a scale-free network. This condition depends on the size of the network through the maximum number of nodes available for connection k_{max}, and is given by

$$\lambda = \upsilon/\delta > \lambda_{threshold} \tag{8}$$

where $\lambda_{threshold} = [m \ln(k_{max}/m)]^{-1}$.

Eq. (8) implies that the larger the network, the lower the threshold condition for infection persistence and hence a greater vulnerability to infection. As indicated above, k_{max} can be set to $N-1$, where N is the number of nodes in the network (a node cannot connect to itself, hence the $N-1$ term). We also note that the logarithmic condition for the threshold condition only applies for the scaling exponent $\gamma = 3$ (considered to be a typical value for Internet/E-mail networks).

Fig. 1 below shows the variation of $\lambda_{threshold}$ with network size for $\gamma = 3$ (the value used in Barabasi and Albert, 1999 and Pastor-Satorras and Vespignani, 2001). Infection persistence occurs when $\lambda > \lambda_{threshold}$. If $\lambda < \lambda_{threshold}$, the decay rate exceeds the growth rate and the infection dies out.

The three possible behaviors of viral infections in the network are shown in Fig. 2 below. In this figure, the $\lambda_{threshold}$ curve of Fig. 1 is reproduced, along with the curve for condition 1 ($\lambda = 1/m = 1/3$) discussed previously, where the virus growth rate exceeds the decay rate for all nodes. Above the $\lambda = 1/3$ curve the probabilities of persistence approach unity. Between the two curves in Fig. 2 the virus persists but with smaller probability. The persistence probabilities decrease as the ordinate position is decreased, and they become vanishingly small as the lower curve is approached. Below the lower curve, there is no persistence in the virus infection.

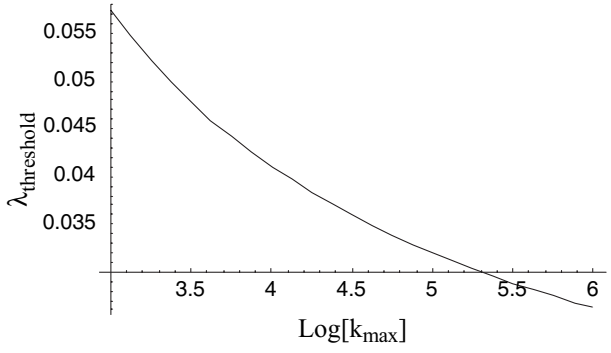

Figure 1 Variation of threshold with network size.

General network infection conditions

In the previous section we examined virus infections only for the steady-state network condition. As we noted earlier, in the steady-state the time rate of change of probability of linking to an infected node is zero. A more general situation relative to virus propagation can be obtained by considering the time evolution of infections leading up to a steady-state condition. We wish to explore the general time-dependence of infection spreading, and what happens near the threshold of viral persistence.

We assume that a virus is introduced into the network at nodes that do not have a specified connectivity value. Based on the discussion in Sections Introduction and background and Network viruses in steady-state conditions, it might be expected that the most damage would occur if the virus is introduced into the network via high connectivity nodes. However, we address the more general case in which the connectivity of the initially infected nodes is arbitrary.

At time $t = 0$ for a small group of initially infected nodes, the mean field equation for the time rate of change of the probability of linking to an infected node simplifies to

$$\partial\rho_k/\partial t \sim -\delta\rho_k \tag{9}$$

Direct integration of Eq. (9) yields

$$\rho_k(t) = \exp(-\delta t) \tag{10}$$

Therefore the infection probability from the small group of initially infected nodes drops off rapidly in time, with a time scale determined by the recovery rate δ. The mean field equation describing the probability of linking to an infected node now derives from two parts: (1) nodes not initially infected and (2) nodes initially infected.

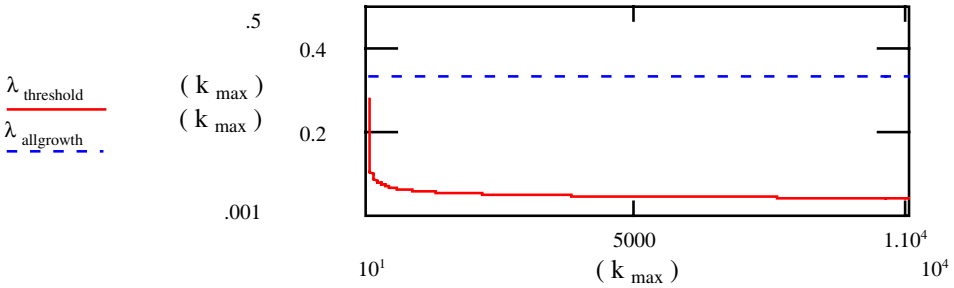

Figure 2 Boundaries between three regions of virus behavior.

The probability of linking to an infected node via a node not initially infected can be obtained by taking the first moment of the modified mean field equation. This yields an expression in terms of the first and second moments of the scale-free distribution $P(k)$.

Recall that the first and second moments of $P(k)$ are defined as $\int kP(k)dk = D_1$ and $\int k^2P(k)dk = D_2$, respectively.

Using this method, a condition for the growth of network infections to a persistent state has been found to exist when $(\upsilon/\delta)(D_2/D_1) > 1$. Conversely, the condition for non-persistence of infection can be shown to be $(\upsilon/\delta)\ (D_2/D_1) < 1$.

The probability of infection by nodes that were initially infected continues to grow until it reaches the steady state or persistent value as specified in the previous section. Specifically, when the scaling exponent γ is 3, the condition for persistent infectious growth becomes $(k_{max}/m) > \exp(1/\lambda m)$ as before.

For nodes not initially infected, a similar analysis reveals that when $(\upsilon/\delta)(D_2/D_1) < 1$, the probability of infection grows to a maximum value and then decays to zero. The probability of linking to an infected node that was not initially infected is directly proportional to its connectivity k.

We can also estimate the time required to achieve a persistent viral state by setting the general probability of linking to an infected node equal to the probability in the steady-state. It has been found that for $\gamma = 3$, the time to achieve non-zero persistence can be made quite long if the values for k_{max}, λ and m are kept small.

Summary of results

It has been found that a threshold exists for the persistence of an infection in scale-free networks such as the Internet. Figs. 1 and 2 plot threshold conditions of k_{max} (network size) versus $\lambda = \upsilon/\delta$, the ratio of intrinsic growth to intrinsic decay rates for a scaling exponent $\gamma = 3$.

In particular, Fig. 2 shows three regions separated by two curves: below the lowest curve no persistent infection exists. Between the two curves, infections persist, but at a low level when near the lower curve. Above the upper curve, the infection probability of each node is close to unity.

When nodes in a narrow range of connectivity are initially infected, there will be no persistent viral infection in the network if $1 > \lambda m \ln(k_{max}/m)$ for $\gamma = 3$. Since $\lambda = \upsilon/\delta$, this suggests increasing the intrinsic decay rate and decreasing the intrinsic growth rate. In addition, the no-persistence condition will be easier to satisfy with smaller networks, since k_{max} in the logarithm term is given by $N-1$, where N is the number of nodes in the network.

The time for the infection probability to reach a maximum in those nodes not initially infected is inversely proportional to $\ln(k_{max}/m)$. This suggests that a larger network will also result in a shorter incubation time for a virus. Once infected, the decay time can become very long as $\lambda m \ln(k_{max}/m)$ approaches unity from below. This again implies increasing the intrinsic decay rate and decreasing the growth rate of a node. Smaller networks have shorter decay times.

Our results show that despite the fact that the probability of a link being connected to an infected node that was initially uninfected increases with network size, the individual node infection probabilities decrease with larger networks. This implies that an increase in network size is favorable relative to the chances of infecting any specific node.

We therefore see that increased network size has competing effects on security. On one hand, the no-persistence condition is easier to satisfy with smaller networks, as well as producing shorter infection decay times. On the other hand, the probability of a particular node being infected increases with network size.

The infection probabilities are proportional to the node connections that have been previously infected. This would suggest that the most damage

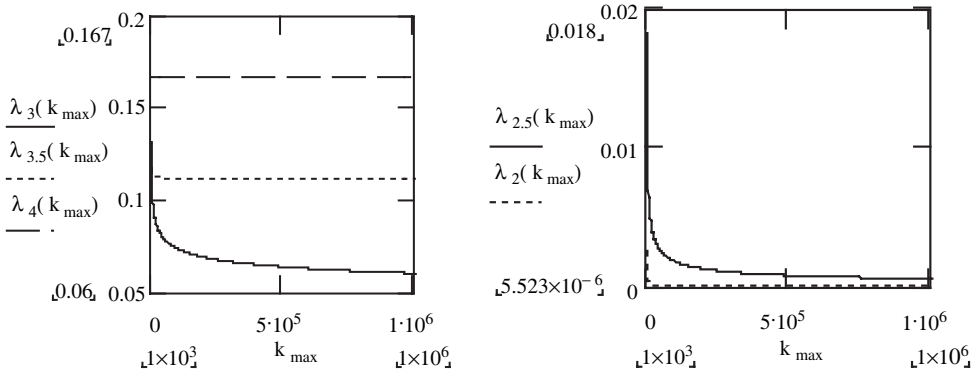

Figure 3 Threshold λ_γ vs k_{max} for various $\gamma = 3$, 3.5 and 4 [plot on left] and $\gamma = 2$ and 2.5 [plot on right]. The subscript of λ indicates the corresponding network scaling exponent γ.

is achieved by infecting high-connectivity nodes, in agreement with intuition. However, if a steady state condition applies (i.e., $\lambda m \ln(k_{max}/m) > 1$), the infection probability is independent of the connectivity of originally targeted nodes. We also confirmed that adjusting the infection growth and decay rates induces the probability of node infection to change maximally for the highest connectivity nodes.

In Fig. 3 below, the threshold λ is plotted against the connectivity k_{max} of the network for various scaling exponents γ. It is interesting to note that:

- the larger the network, the lower the threshold value of λ,
- the larger the exponent γ for a distribution $P(k)$, the higher the allowable value of λ.

Network security implications

The results herein suggest alternative approaches to network organization and surveillance in order to enhance security. Networks have traditionally been organized into subnets based on differences in functionality or user groups, rather than according to topological features. However, the confirmation of the existence of a threshold for infection persistence has significant implications, since any actions that contribute to remaining below that threshold decrease the vulnerability to infection spreading.

First, it is clear that priority quarantining and patching of high-connectivity nodes is mandated. These results as well as the results of others (Albert et al., 2000; Ebel et al., 2002; Pastor-Satorras and Vespignani, 2001) argue strongly for preferentially monitoring these specific nodes for infections. This strategy is consistent with published recommendations for defending against self-propagating code such as Code-Red (Moore et al., 2003, 2002). Quick intervention and remediation of high-connectivity nodes will increase the virus incubation time by decreasing the value of λ, m, and k_{max} which appear in the denominator of the expression for the time-to-persistence.

There has been considerable documentation of modes of infection spreading. These typically involve variations on a similar theme, where viruses self-replicate and then distribute themselves to address book entries, MAPI mailboxes or some other means of E-mail-based distribution. Examples of such viruses include Nimda, SoBig-A, and variants of Melissa (Information World Review, 2001, 2003, 1999). Furthermore, in at least one case it has been explicitly demonstrated by direct measurement that a seemingly typical E-mail network obeyed a scale-free distribution with $\gamma = 1.81$ (Ebel et al., 2002). The continued exploitation of E-mail as a means of virus transmission coupled with the prevalence of contact and/or address lists creates a ready means of directed attacks.

We are not aware of an automated method of examining server logs in order to determine the changing hierarchy of node connectivity, and thereby monitor the risk of infection in a targeted fashion. In lieu of this capability, the number of entries in network users' contact lists might be considered to identify the high-connectivity nodes. It is not unreasonable to assume that the number of entries in contact lists follows a scale-free distribution across the network community and might mirror the distribution data containing in the server logs. Future security products might include those that identify and monitor high-connectivity network nodes in real time.

In view of the direct dependence on the number of available nodes for connection k_{max} by the persistence threshold value, segmenting the

network into a hierarchy according to the number of nodes would appear to be advantageous. In that vein, one might envision a pyramid-shaped network topology, such that the segment with the lowest population has a single node. In some sense this implies a re-examination of the fundamental notion of a node, where each segment consisting of a varying number of nodes might be considered a node unto itself.

Finally, and for what is believed to be the first time, a true security metric can be explicitly communicated based on these results. This metric will enable network administrators to sample their environment and actually measure the exposure to risk relative to E-mail-borne viruses in real time. Specifically, these results suggest the creation of a security figure-of-merit

$$S = D_1/(D_2\lambda) \tag{11}$$

where D_i represents the ith moment of the connectivity probability distribution $P(k)$, and as before $\lambda = \upsilon/\delta$ is the ratio of infection growth-to-decay rates.

Larger values of S imply an enhanced defense relative to the susceptibility to computer virus infection. In particular, a value of $S = 1$ represents the threshold condition for viral persistence once the virus has been introduced into the network. Such a metric may offer opportunities for the development of security software designed to measure, report, and alert on the value of S as the network connectivity evolves with time.

References

Albert R, Jeong H, Barabasi AL. Nature 27 July 2000;406: 378–381.

Barabasi AL, Albert R. Science 1999;286:509–11.

Ebel H, Mielsch LI, Bornholdt S. Physical Review E 2002;66. 035103(R).

Erdos P, Renyi A. Publ. Math. Inst. Hung. Acad. Sci. 1960;5:17.

Moore D, Shannon C, Claffy K. Code-red: a case study on the spread and victims of an internet worm. Internet measurement workshop. In: Proceedings of the second SIGCOMM workshop on internet measurement; 2002. p. 273–84.

Moore D, Shannon C, Voelker GM, Savage S. Internet quarantine: requirements for containing self-propagating code; April 2003. Infocomm 2003, San Francisco, Ca.

New melissa virus variant on the loose. Information World Review October 19, 1999.

Nimda worm most virulent virus ever. Information World Review September 21, 2001.

Pastor-Satorras R, Vespignani A. Physical Review Letters 2001; 86:3200.

SoBig virus infections on the rise. Information World Review January 15, 2003.

Watts DJ, Strogatz SH. Nature 1998:292–440.

Dr Chang has 45 years of experience in industry, government, and academia. He's served in a number of technology director, chief scientist, and senior technical management and research positions at Hughes Electronics, Occidental Research, Boeing, the U.S. Department of Commerce, and USC. He has held adjunct professorships at UCI, UW, USC, and CSULA, and currently consults for a variety of organizations. Dr Chang has over 200 publications and patents in several areas of basic and applied physics.

Carl Young is an applied physicist with a specific focus on quantifying risk and solving complex security problems. Mr. Young spent 15 years in the US government designing, developing, and deploying security technology. In 1997 he was awarded the James R. Killian medal by the White House for individual contributions to national security. He is currently the Director of Research and Analysis for the Office of Global Security at Goldman Sachs & Co., and lectures on science and technology applied to security as an adjunct professor at Polytechnic University in New York City. He has authored a wide variety of papers on technical security and risk-related problems, and is the author of *The Science of Security…And the Fundamentals of Risk Mitigation* (to be published). Mr. Young holds bachelors and masters degrees in applied mathematics and physics from the Massachusetts Institute of Technology, Cambridge, Massachusetts.