

Simulating Network Robustness for Critical Infrastructure Networks

Anthony H. Dekker

Defence Science and Technology Organisation
Department of Defence, Canberra ACT 2600

dekker@ACM.org

Abstract

We examine the robustness of critical infrastructure networks in the face of terrorist attack, using a simulation experiment that incorporates link *capacity*; and an extension of data farming which we call *network farming*. Our results show that symmetrical designed networks generally outperform randomly generated networks, although ring-like structures are very vulnerable. Under targeted attacks, most networks begin to fail when the number of attacks is equal to the *node connectivity*. Examining the distribution of real-world terrorist attacks, we show that these can be modelled by a Poisson statistical distribution, leading to recommendations for the node connectivity required at different threat levels.

Keywords: network robustness, graph connectivity, simulation, terrorism.

1 Introduction

Modern technological civilisation is dependent on its *critical infrastructure networks*: communication, electrical power, rail, and fuel distribution networks. Failure of any of these critical infrastructure networks can bring the ordinary activities of work and recreation to a standstill.

This dependence has led to the frequent selection of critical infrastructure networks as military targets in times of war. In the US Civil War, the rail junction of Chattanooga became a key military objective, and telegraph networks were also attacked (Dickson 2001). In the Second World War, Allied bombers targeted rail, fuel, and electrical power networks in the German Ruhr. More recently, in the former Yugoslavia, the US Air Force temporarily disabled electrical power stations by dropping conductive fibres (Jones & Geppert 2002).

The same vulnerability to attack that makes critical infrastructure networks military targets also makes them targets for terrorist attacks. In particular, critical infrastructure networks are often targeted by terrorists wishing to destroy society as a whole, or to replace it by some “ideal” society of their own. For example, the Anarchists and Syndicalists in Spain in the early 20th Century engaged in terrorism in order to “annihilate the

state in one great revolution and initiate a life for all in agrarian communes and free municipalities” (Sinclair 2003). Terrorist attacks on electrical power networks, rail networks, and oil pipelines have occurred in Colombia, India, Pakistan, Turkey, Algeria, and Spain (ICT 2004).

An important aspect of critical infrastructure networks is their *interdependence*. Attacks on the electrical power and communications networks in particular have a “force multiplier” effect on other services. For example, the terror attacks which destroyed the Twin Towers in New York City on 11 September 2001 had the side-effect of severely damaging Verizon’s central telephone switch, and destroying many communications antennas. Police and other emergency services lost both communications and electrical power (Van Ooyen *et al* 2002). This underlined the vulnerability of network nodes to terrorist attack and the potential “force multiplier” effect in terms of the impact on emergency services. It also demonstrated the ability of terrorist organizations to launch multiple synergistic attacks.

The critical infrastructure networks we have discussed, like all other networks, consist of *nodes* (railway stations, pumps, transformers, switches, etc.) and *links* (tracks, pipes, cables, etc.). Trains, oil or gas, electrical power, and messages flow through the networks, and importantly each link has a fixed *capacity*. A single track can carry only so many railway carriages per hour, a communication line can carry only so many bits per second, etc. Our simulation therefore incorporates link capacities.

Redundancy in the network comes from the presence of *alternate paths* along which traffic (trains, fuel, electrical power, or messages) can travel. However, if the usual (shorter) path on which traffic travels is unavailable, and traffic is re-routed along a longer path, the total load on the network is increased. This in turn can result in further overloaded links and more re-routing. If this process “snowballs” out of control, the result is a *cascading failure* of the network. This is more likely to occur if the overloaded condition causes the links or the nodes to fail completely (Motter & Lai 2002). Cascading failure has occurred several times in the North American electrical power grid, with major blackouts occurring in March 1989, August 1996, and August 2003. These failures could have been avoided if some load had been shed instead of being re-routed (Amin 2001).

In this paper, we examine the robustness of critical infrastructure networks, specifically aspects which relate to the *network topology*, rather than other network characteristics, such as management and control, or physical security. We describe a simulation experiment which identifies a number of characteristics that make

networks robust in the face of terrorist attack. Specifically, networks begin to fail under targeted attacks when the number of attacks is equal to the *node connectivity*, with the severity of failure dependent on the *average degree* and the amount of network *symmetry*.

Our simulations utilise a tool suite for analysing, visualising, and simulating networks called CAVALIER (Dekker 2001, Dekker 2003), which we have developed. Included within CAVALIER is a network simulator, which we use to study the effect of terrorist attacks on networks. CAVALIER also includes facilities for calculating network metrics, and the visualisation capabilities of CAVALIER were used to produce some of the diagrams in this paper (Figures 3, 4, 6 & 9).

The process we use for analysing our simulations is an extension of data farming (Horne 1997, Brandstein & Horne 1998), which we call *network farming* (Dekker 2004b).

We also survey *historical data* on terrorist attacks by seven terrorist groups (ICT 2004), and show that the distribution of these attacks can be modelled by a Poisson statistical distribution. This in turn can be used to estimate how large a node connectivity is “enough” to withstand anticipated numbers of synergistic terrorist attacks.

2 Network Farming

In order to study network-based processes, we use *network farming*, a technique we have developed (Dekker 2004b) as an extension of *data farming*. Data farming was developed by Gary Horne for the US Marine Corps (Horne 1997, Brandstein & Horne 1998).

Data farming studies a complex parameterised process $P(x_1, \dots, x_n)$. Since data farming was developed for the US Marine Corps, the parameters x_1, \dots, x_n originally referred to such technical and social variables as vehicle speed or morale. The process $P(x_1, \dots, x_n)$ is simulated by a computer model (often agent-based, and with a random component), and experiments are conducted by running the model $P(x_1, \dots, x_n)$ with some parameters fixed ($x_i = a_i$) and some parameters varying across a range ($b_j \leq x_j \leq c_j$). As these experiments are repeated, an overall perspective of process behaviour can be obtained by *zooming out*, i.e. varying more parameters, over wider ranges $b_j \dots c_j$, but for fewer values in those ranges. A more detailed understanding of particular parameter ranges of interest can be obtained by *zooming in*, i.e. varying fewer parameters, over narrower ranges $b_j \dots c_j$, with more values in those ranges. Zooming in also requires more iterations of the simulation in order to average out any random factors and better understand the landscape of possibilities. The understanding obtained by zooming in can then be discussed with domain experts and applied to other experimental or analytical strategies. This will in turn suggest other parameter ranges of interest which might be explored.

A key aspect of data farming is visualisation and analysis of the data produced. In the case where we are varying only two parameters and measuring only one output

(performance) variable, a 3-dimensional *fitness landscape* provides a useful way to visualise average behaviour, as shown in Figure 1.

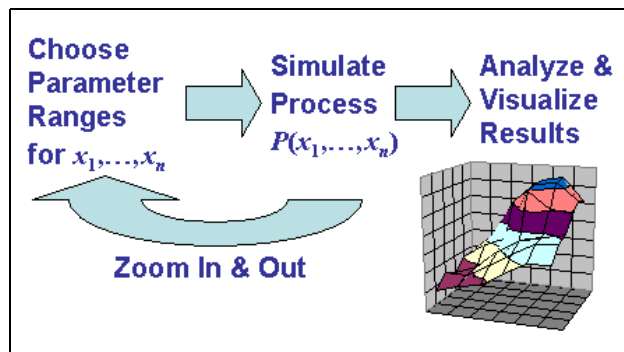


Figure 1: Data Farming Process

In order to study network-based processes, we have developed an extension of data farming which we call *network farming*. In network farming, the process $P(N, x_1, \dots, x_n)$ which we are studying is parameterised not only on a list of numbers x_1, \dots, x_n , but also on a particular network topology N . The performance of a network in the face of terrorist attacks is one such parameterised process.

In order to study the process $P(N, x_1, \dots, x_n)$, we must generate a large list of different networks N , as well as choosing values or ranges for the other parameters x_1, \dots, x_n . This generation process may involve a pre-existing library of different networks and/or the use of various random graph generation approaches (Bollobás 2001). We then simulate the process using the different networks and parameter values. Table 1 illustrates a simple example, examining the performance of different network topologies under terrorist attack, and drawn from the study described later in this chapter (in fact, it is a subset of Table 3).


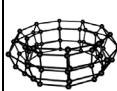


		Network N			
					
No. of Nodes Lost (x)	1	100%	100%	100%	100%
	2	100%	100%	96.6%	100%
	3	96.5%	100%	93.2%	100%
	4	96.4%	98.6%	89.5%	100%
	5	92.0%	98.5%	89.2%	100%
	6	94.8%	98.5%	85.6%	100%

Table 1: Network Farming Example — Performance of Networks under Terrorist Attack

Unfortunately, although data in the form of Table 1 is easy to understand, it is difficult to analyse mathematically. For the purpose of analysis, we therefore derive various numerical *metrics* $M_1(N), \dots, M_m(N)$ for each network N , such as node connectivity and average degree (described below). These metrics provide a numerical summary of the structure of the networks N .

We can then use the same visualisation and analysis techniques that data farming uses, in order to see how performance varies with the parameters x_1, \dots, x_n and the metrics $M_1(N), \dots, M_m(N)$. However, it is important to verify that the distribution of values for the metrics $M_i(N)$ is not so irregular as to give deceptive results. In particular, many statistical techniques require that the metrics be normally distributed. Table 2 below shows some example metrics for the networks in Table 1, while Figure 2 illustrates the network farming process.

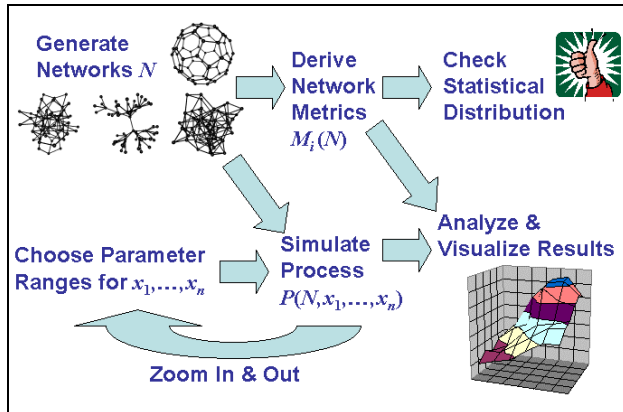


Figure 2: Network Farming Process

3 Graph Theory

The mathematical discipline of *graph theory* (Biggs 1993, Gibbons 1985) is a natural way to study network topologies. Graph theory models the topology of a critical infrastructure network as an undirected *graph*. Graph theory also provides the metrics necessary for network farming.

If the number of incoming links of a node is d , then we say that the node has *degree* d . For the network as a whole, we can then calculate the minimum (d_{\min}), maximum (d_{\max}), and average degrees (d_{ave}), and these are all potentially useful metrics. In previous work (Dekker & Colbert 2004a, Dekker 2004a), we identified *node connectivity* κ as the best robustness metric for a network. The node connectivity of a network is the smallest number of nodes whose removal disconnects the network, or equivalently, the smallest number of node-distinct paths between any pair of nodes. The CAVALIER tool, which we have developed, includes a facility for calculating node connectivity and other metrics.

The *diameter* D of a network is the longest of all the shortest paths between pairs of nodes. It provides an upper bound on the *average distance* D_{ave} , which is the average length of all shortest paths between pairs of nodes. Networks with low diameter and average distance are desirable, because traffic uses fewer links to get to its destination, and therefore places less load on the network as a whole (Dekker & Colbert 2004a).

We have also found it useful to measure the *symmetry* of the network. After some investigation, we have identified the *symmetry ratio* r as a more useful metric than other options, such as the size of the automorphism group (Dekker & Colbert 2005). The symmetry ratio is

calculated by finding the different eigenvalues of a network (Biggs 1993). The number of different eigenvalues will be at least one more than the diameter (Biggs 1993, p 10), and dividing by this lower bound yields the symmetry ratio. For highly symmetrical networks, the symmetry ratio will be low (in the range 1...2), while for unsymmetrical networks the symmetry ratio will be high (approaching one third of the number of nodes). Table 2 shows these metrics for three example 60-node networks.

	Network N		
Node Connectivity	3	4	1
Average Degree	3	4	3.83
Maximum Degree	3	4	16
Symmetry Ratio	1.5	2.11	10
Diameter	9	8	5

Table 2: Some Example Network Metrics

The concept of *scale-free networks* was introduced by Barabási & Albert (1999), and scale-free networks have attracted a great deal of interest (Barabási 2002, Dekker & Colbert 2004b). Scale-free networks grow by a process of *random preferential attachment*. In particular, a k -linked scale-free graph grows by incrementally adding nodes, and randomly connecting each new node by k links to existing nodes. These new links are preferentially connected to existing highly connected “hub” nodes. In particular, the nodes that the new links go to are chosen randomly with probability proportional to their degree (it is possible for some or all of these links to go to the same node).

Bollobás (2001) has shown that such a randomly generated network is almost certainly connected (if $k \geq 2$). Scale-free networks have small diameter, and are resistant to random attacks, but are not particularly robust against deliberate attacks (Albert & Barabási 2002, Barabási & Bonabeau 2003, Bollobás & Riordan 2003). These studies (which do not take link capacity into account) are confirmed by our simulation results (using link capacity).

4 The CAVALIER Terrorist Attack Simulator

The CAVALIER tool, which we have developed, includes a simulator for studying the effect of terrorist attacks on networks. The terrorist attack simulator studies the effect of destroying nodes in a network which sends “packets” of traffic back and forth along links, i.e. a network such as a rail or communications network. The simulator assumes that the same number of packets are

sent between every pair of nodes. The simulator also assumes that each link in the network has just enough capacity to handle the load that would exist in a ring-shaped network topology.

The simulator then simulates the behaviour of the network with between 1 and 6 nodes destroyed by terrorist attack, and measures the percentage of packets which successfully reached their destination (over 10 choices of node removal, and 10,000 packets sent). Table 1 shows some results from the simulator. Clearly, this is a network farming process, and the simulator is simulating a terrorist attack process $P(N,x)$, where x is the number of attacks. For the purpose of analysis, described below, we derive various numerical metrics $M_1(N), \dots, M_m(N)$ for each network N , such as the node connectivity κ , average degree d_{ave} , diameter D , and symmetry ratio r .

The network farming process requires methods for generating lists of different networks N . For the study described in this chapter, we generated a variety of networks with 60 nodes, using a combination of three techniques:

- Two scale-free networks, one 2-linked and the other 4-linked, as discussed above.
- Fifty randomly generated networks, with average degree ranging from 2 to 10, created by taking a randomly generated tree network (to ensure connectedness) and adding additional links with the identical probability for a link being added between any pair of nodes (i.e. the Erdős-Rényi model).
- A fixed list of nine symmetrical networks, shown in Figure 3. These are all *Cayley graphs*, and can be generated using *group theory* (Biggs 1993, Dekker & Colbert 2004a). Six of these networks (all except 3(e), 3(f), and 3(g)) are also *planar graphs*, which are of particular interest for infrastructure networks.

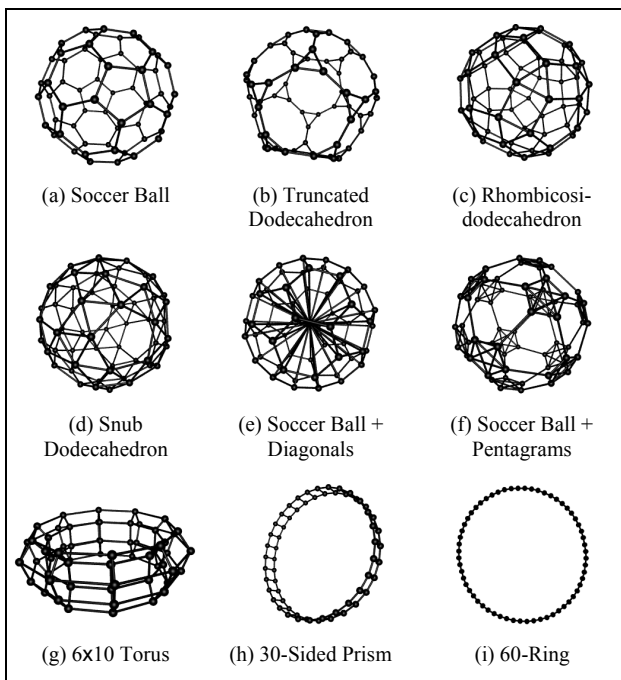


Figure 3: Nine Sixty-Node Networks

Our simulator uses a shortest-path routing algorithm, i.e. packets are sent along the shortest path or paths. The simulator balances traffic between all the shortest paths if there is more than one. However, the simulator assumes that, if all the shortest paths are loaded to maximum capacity, traffic is lost rather than re-routed on longer paths. This simulates strategies for preventing cascading failure, which can result from such re-routing.

Two terrorist attack strategies were used:

- Centralised attack: where terrorists select the most central, or most important, node (or one of the most central nodes, if there are several choices).
- Random attack: where terrorists choose a node to be destroyed purely at random.

The concept of centrality used is that in Dekker (2002), i.e. the most central node is the one with the largest average closeness to all the other nodes (where closeness is the inverse of distance). Centralised attack is the more realistic model of terrorist operations, since terrorists can be expected to target precisely those nodes whose destruction has the most impact.

5 Simulation Results: Centralised Attack

For the case of centralised attacks, the results of our simulation are shown in Table 3. The measure of performance we use is the average percentage of packets which successfully reached their destination. Table 3 shows the averages for different numbers of attacks α .

Table 3 shows that most networks begin to fail when the number of attacks α is equal to the node connectivity κ . There are exceptions however: some networks are inherently robust, e.g. the Rhombicosidodecahedron network in Figure 3(c) and the Snub Dodecahedron network in Figure 3(d) do not fail with up to six terrorist attacks. These two planar networks are characterized by:

- A reasonably high node connectivity κ (4 or 5).
- A high degree of symmetry, i.e. a low symmetry ratio $r < 2$.
- An even distribution of links across the network, as opposed to the clustering in the “Soccer Ball + Pentagrams” network of Figure 5(f), which performs quite poorly, given in its high degree.
- An absence of large ring structures.

The simulator was designed so that each link in the network had just enough capacity to handle the load that would exist in a ring network. Unsurprisingly, the ring network begins to fail with just a single attack, since the links at the centre of the remaining linear network are unable to carry the load. A similar phenomenon is responsible for the poor performance of the 30-Sided Prism network in Figure 3(h), which consists of a double ring.

The 2-linked scale-free network begins to fail on the second attack, but if one node is removed, the remaining network has $\kappa = 2$, and so the 2-linked scale-free network still fits the pattern of failing when $\alpha = \kappa$.

Network	Messages Received for Numbers of Attacks					
	1	2	3	4	5	6
Soccer Ball	100%	100%	96.5%	96.4%	92%	94.8%
Truncated Dodecahedron	100%	100%	94.9%	97.9%	95.6%	93.7%
Rhombicosidodecahedron	100%	100%	100%	100%	100%	100%
Snub Dodecahedron	100%	100%	100%	100%	100%	100%
Soccer Ball + Diagonals	100%	100%	100%	96.4%	96.4%	96%
Soccer Ball + Pentagrams	100%	100%	100%	100%	87.1%	86%
30-Sided Prism	100%	100%	97.2%	75.1%	77.4%	74.8%
6x10 Torus	100%	100%	100%	98.6%	98.5%	98.5%
Ring	71.5%	58.4%	44.6%	36.4%	29.5%	22.2%
2-linked scale-free	100%	96.6%	93.2%	89.5%	89.2%	85.6%
4-linked scale-free	100%	100%	96.6%	96.5%	92.9%	92.6%
Random nets (averages)	84.4%	79.2%	76.6%	75.2%	73.2%	71.2%

Table 3: Simulation Results — Centralised Attack

To understand the performance of the randomly generated networks, we use network farming with the metrics in Table 4. These metrics are:

- The average degree d_{ave} .
- The *node connectivity* κ .
- The *diameter* D .
- The *symmetry ratio* r .

Statistical analysis of the variation in performance showed that:

- The average degree d_{ave} alone explained 68% of the variation in performance, since even random networks with low node connectivity benefited from adding more and more links.
- The difference between the node connectivity κ and the number of attacks α explained an additional 12% of the variation in performance (this number is fairly low, since the difference between, for example, 2 and 3 attacks is much less than the difference between a good network and a bad network).
- The symmetry ratio r explained an additional 6% of the variation in performance (with the more symmetrical networks, i.e. those with lower symmetry ratios, performing better).
- Unknown factors explained 16% of the variation in performance.

The three effects listed were all statistically highly significant (to ensure a valid statistical distribution, logarithms were taken in each case). Figure 4 shows the relationship between network performance and these

effects. Our analysis showed that network performance P (i.e. the percent of messages successfully received) could be estimated by the equation:

$$\log(101 - P) = -4.64 \log \log d_{ave} + 0.77 \log r + 0.61 \log(\max(1, 2 + \alpha - \kappa)) + 1.35$$

or:

$$P = 101 - 3.86 (\log d_{ave})^{-4.64} r^{0.77} (\max(1, 2 + \alpha - \kappa))^{0.61}$$

The points marking the performance of the designed networks in Figure 4 are each marked with the corresponding value $\alpha - \kappa$. Randomly generated networks in Figure 4 are not marked. The statistical correlation between this estimate and actual performance was 0.91, with some of the randomly generated networks performing even worse than estimated. The 30-Sided Prism network in Figure 3(h), and the ‘‘Soccer Ball’’ variations in Figure 3(e) and 3(f) performed poorly in spite of being symmetrical.

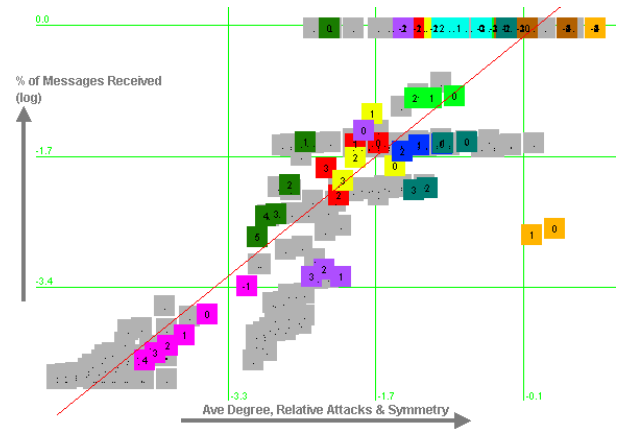


Figure 4: Regression Equation for Centralised Attack

For an alternative perspective, we can combine the effect of average degree and symmetry ratio by forming the adjusted average degree d_{adj} , which is the average degree divided by some function of the symmetry ratio r . For 60-node networks, the best results are obtained by using the fourth root of the symmetry ratio (which ranges from 1 to 2):

$$d_{adj} = d_{ave} / \sqrt[4]{r}$$

The adjusted average degree provides a way of characterising several important aspects of network structure using a single number, and allows us to plot the average percentage of messages received as a function of two variables: $\alpha - \kappa$ and d_{adj} , as shown in Figure 5. These two variables predict 82% of the variation in performance. Figure 5 suggests that the most robust networks (in the sense of being able to absorb multiple attacks) are those where the adjusted average degree d_{adj} is at least 3, i.e.:

- The Rhombicosidodecahedron and the Snub Dodecahedron (which did not fail with up to six terrorist attacks).
- The two ‘‘Soccer Ball’’ variations (although these had disappointing performance).

- The 6x10 Torus (with at least 98.5% of messages received even at six terrorist attacks).
- The 4-linked scale-free network (although this also had disappointing performance).
- Some random networks (random networks with $d_{adj} \geq 3$ had at least 95% of messages successfully received).

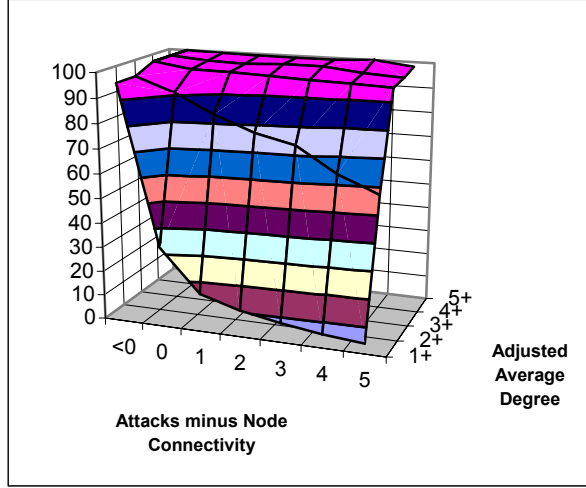


Figure 5: Percent of Messages Successfully Received as a Function of $\alpha - \kappa$ and d_{adj} — Centralised Attack

6 Simulation Results: Random Attack

For the case of random attacks, the results of our simulation are shown in Table 5 and Figure 6. Random attacks are less realistic as a model of terrorist behaviour, and correspond more closely to accidental failures. For the random attack case, the only networks which failed were the ring, the 30-Sided Prism, and the randomly generated networks. Very slight failures also occurred in the 2-linked scale-free network.

Network	κ	d_{ave}	D	r	d_{adj}
Soccer Ball	3	3	9	1.5	2.71
Truncated Dodecahedron	3	3	10	1.18	2.88
Rhombicosidodecahedron	4	4	8	1.44	3.65
Snub Dodecahedron	5	5	7	1.88	4.27
Soccer Ball + Diagonals	4	4	5	2.5	3.18
Soccer Ball + Pentagrams	5	5	6	1.29	4.70
30-Sided Prism	3	3	16	1.82	2.58
6x10 Torus	4	4	8	2.11	3.32
Ring	2	2	30	1	2
2-linked scale-free	1	3.833	5	10	2.16
4-linked scale-free	3	6.7	4	12	3.60
Random nets (averages)	1.9	6	5.02	10.88	3.26

Table 4: Network Metrics for Network Farming Simulation Results

Network	Messages Received for Numbers of Attacks					
	1	2	3	4	5	6
Soccer Ball	100%	100%	100%	100%	100%	100%
Truncated Dodecahedron	100%	100%	100%	100%	100%	100%
Rhombicosidodecahedron	100%	100%	100%	100%	100%	100%
Snub Dodecahedron	100%	100%	100%	100%	100%	100%
Soccer Ball + Diagonals	100%	100%	100%	100%	100%	100%
Soccer Ball + Pentagrams	100%	100%	100%	100%	100%	100%
30-Sided Prism	100%	100%	100%	94.4%	93.7%	92.9%
6x10 Torus	100%	100%	100%	100%	100%	100%
Ring	71.3%	58.5%	52.2%	41.8%	28.6%	31%
2-linked scale-free	99.7%	99.6%	100%	99.0%	99.6%	98.6%
4-linked scale-free	100%	100%	100%	100%	100%	100%
Random nets (averages)	96%	95.2%	94.6%	94.2%	92.6%	92.3%

Table 5: Simulation Results — Random Attack

The impact of node connectivity κ was less significant in this case. The previous section described how coordinated terrorist attacks began to have an impact when the number of attacks reached the node connectivity, but for uncoordinated random attacks, this did not occur (largely because most of the designed networks did not fail). Network farming showed that:

- The average degree d_{ave} explained 73% of the variation in performance.
- The number of terrorist attacks α in combination with the average degree d_{ave} explained an additional 3% of the variation in performance (with the impact of α reduced for high d_{ave}).
- The symmetry ratio r explained an additional 4% of the variation in performance (with the more symmetrical networks, i.e. those with lower symmetry ratios, performing better).
- Unknown factors explained 20% of the variation in performance.

These effects were all statistically highly significant. The equation estimating network performance P , as shown in Figure 6, was:

$$\log(101 - P) = 0.76 + 0.088r + 0.150 \alpha / \log d_{ave} - 3.21 \log \log d_{ave}$$

or:

$$P = 101 - 2.14 (\log d_{ave})^{-3.21} 1.09^r 1.16^{(\alpha / \log d_{ave})}$$

The statistical correlation between estimate and performance was 0.89, with the ring and 30-Sided Prism performing even less well than estimated, as in the case of centralised attack.

Figure 7 shows the average percentage of messages received as a function of the number of attacks α and the adjusted average degree d_{adj} . These two variables predict 76% of the variation in performance.

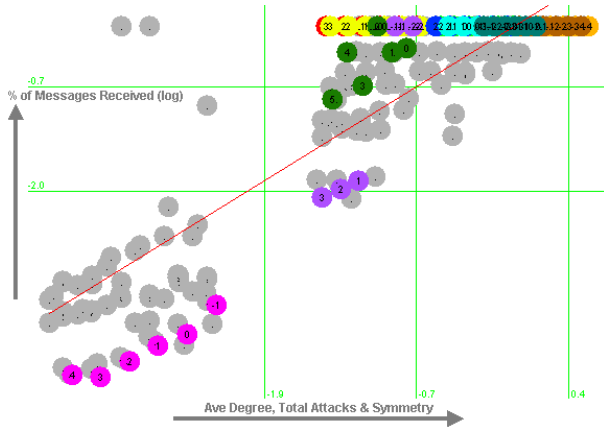


Figure 6: Regression Equation for Random Attack

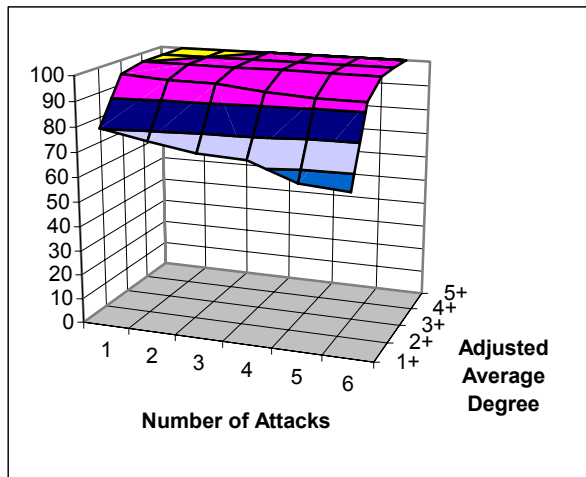


Figure 7: Percent of Messages Successfully Received as a Function of α and d_{adj} — Random Attack

7 The Distribution of Terrorist Attacks

In order to relate our simulation results to real-world terrorist attacks, we extracted data on seven terrorist groups from ICT's *International Terrorism Database* (ICT 2004). We examined the number of terrorist attacks within each year (or quarter, where appropriate).

The number of terrorist attacks per period ranged from 0 to 13. For example, there were 13 ETA attacks in Spain from July 12 2000 to September 21 2000, significantly more than the average of 2.7 attacks per quarter. The terrorist attacks were not evenly distributed in time, but were clustered, as shown in Figure 8 for Al Qa'ida.

Table 6 summarises data on the seven terrorist groups (a is the average number of attacks in a year or quarter).

The use of χ -squared statistical tests showed that the distribution of terrorist attacks in general did not fit a uniform distribution (i.e. with the same number of attacks per period). The clustering of terrorist attacks continued on smaller scales, with multiple attacks sometimes

occurring in the same month, or even on the same day (as on 11 September 2001). We therefore considered whether the attack distribution would be scale-free, and indeed, log-log plots of number of attacks against rank (adjusted for differences in attack rate) were approximately linear, (Figure 9). Yet χ -squared tests showed that a scale-free (power-law) distribution did not always fit either. In the case of the Tamil Tigers, there was a less than 0.001% chance that the attacks fitted a scale-free distribution, largely because quarters with exactly one attack were far more common than quarters with zero or two attacks. However, χ -squared tests showed that, in all seven cases, the distribution of terrorist attacks fitted a Poisson distribution (Saaty 1961), as illustrated for Al Qa'ida attacks in Figure 8.

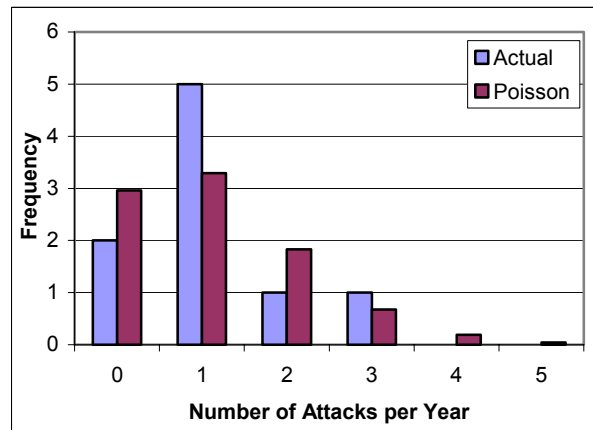


Figure 8: Al-Qa'ida Attacks vs Poisson Distribution

Terrorist Group	Period	a	Distribution		
			Uniform	Scale-Free	Poisson
Fuerzas Armadas Revolucionarias de Columbia (FARC)	years: 1992 – 2002	3.4	No $p < 0.01\%$	Yes	Yes
Euskadi Ta Askatasuna (ETA) – Spain	quarters: 2000 – 2002	2.7	No $p < 0.0001\%$	Yes	Yes
Al-Qa'ida (AQ) – Worldwide	quarters: 2000 – 2002	1.1	Yes	No $p < 1\%$	Yes
Irish Republican Army (IRA) – Northern Ireland	years: 1989 – 2001	1.6	No $p < 0.01\%$	Yes	Yes
Liberation Tigers of Tamil Eelam (LTTE) – Sri Lanka	quarters: 1996 – 2001	1.1	Yes	No $p < 0.001\%$	Yes
New People's Army (NPA) – Philippines	quarters: 1989 – 1990	1.2	Yes	No $p < 2\%$	Yes
Sendero Luminoso (SL) – Peru	years: 1989 – 2002	1.4	No $p < 5\%$	No $p < 5\%$	Yes

Table 6: Statistical Distribution of Terrorist Attacks

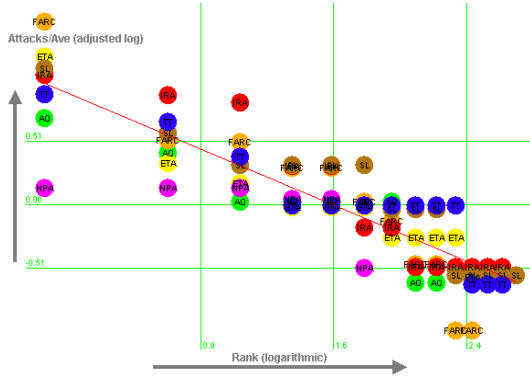


Figure 9: Test for Scale-Free Distribution of Attacks

The Poisson statistical distribution says that if the average number of attacks in a year (or quarter) is a , then the probability of n or more attacks during the year (or quarter) is:

$$1 - e^{-a} \left(1 + a + \frac{a^2}{2} + \frac{a^3}{6} + \dots + \frac{a^{n-1}}{(n-1)!} \right)$$

Consequently, although the most likely number of attacks is about a , there is an exponentially decaying possibility of many more attacks. Since χ -squared tests showed that this distribution fits, we can use it for planning purposes. In particular, we can use it to answer the question: how much node connectivity is enough? We have already established that the network will typically begin to fail when the number of nodes destroyed is equal to the node connectivity.

To answer this question, we assume that the network has node connectivity at least 2, and that a terrorist attack has already destroyed one node. We then consider the possibility of terrorist attacks during the period in which the damaged node is being repaired. In the case of Verizon in New York, this repair period was about one week (CTIA 2001). If we know the typical repair period, and we can estimate the average number of attacks per year, then we can easily calculate the average number of attacks during the repair period (i.e. a in the formula above). For example, if the network has node connectivity 3, it can survive the initial attack, and one additional attack during the repair period.

Average Number of Attacks during Repair Period	Node Connectivity Required for ...	
	1 in 1,000 Failure	1 in 1,000,000 Failure
3	12	16
1	7	11
0.3	5	7
0.1	4	6
0.03	3	5
0.01	3	4
0.003	3	4
0.001	2	3

Table 7: Connectivities Required for Specified Network Failure Probabilities

Table 7 shows the node connectivity required to reduce the chance of network failure to less than 1 in 1,000 or less than 1 in 1,000,000, for different values of a .

For example, with an expected number of one attack per year, and a repair period of a little over one month, the expected number of attacks during the repair period would be 0.1, and a node connectivity of 6 would be required to reduce the chance of network failure to less than 1 in 1,000,000. On the other hand, with an expected number of one attack per decade, and a ten-day repair period, the expected number of attacks during the repair period would be 0.003, and a node connectivity of 3 would be required to reduce the chance of network failure to less than 1 in 1,000.

Therefore, even on fairly optimistic estimates of attack frequency and repair time, critical infrastructure networks in high-threat environments should have a node connectivity of at least 3, in order to absorb two synergistic targeted attacks without failing.

8 Conclusions

In this paper, we have discussed a simulation experiment analysing the robustness of the critical infrastructure networks on which civilization depends, and which are now under the threat of terrorist attack. Our simulation takes into account *link capacity* and the possibility of failure due to link overload. We have used an extension of data farming, which we call *network farming*, as a framework for conducting the experiment. This process involved generating a test set of network topologies, and using metrics derived from *graph theory* to analyse the results. The network topologies we used included the symmetrical networks in Figure 3, two *scale-free networks*, and 50 randomly generated networks.

For each network, we simulated the effect on performance of targeted and random terrorist attacks. With targeted attacks, most networks began to fail when the number of attacks was equal to the *node connectivity*, either because of disconnection, or because of overloaded links. However, some highly symmetrical and well-connected networks did not fail, even with up to six terrorist attacks. On the other hand, randomly generated networks, and networks containing large rings, were not robust. Scale-free networks were robust against random attacks, but not against targeted attacks. These are general observations, however, and the behaviour of a specific real-world critical infrastructure network will depend on sector-specific factors. For this reason, we expect to see more sophisticated sector-specific models of terrorist attack emerge over the next few years.

Examining the distribution of real-world terrorist attacks, we showed that these could be modelled by a Poisson statistical distribution. This allows us to determine the probability of a given number of terrorist attacks within a certain time period, and hence the node connectivity required in order to reduce the chance of network failure to some specified small probability, such as 1 in 1,000 or 1 in 1,000,000. Our simulation results also show the benefits of symmetrical network design, where this is possible.

9 Acknowledgements

The author is greatly indebted to Bernard Colbert for extensive discussions on graph theory, scale-free networks, network robustness, and the symmetry ratio. Figure 3 was produced by interfacing CAVALIER to the Persistence of Vision™ ray-tracing package.

10 References

- Albert, R. & Barabási, A.-L. (2002), "Statistical mechanics of complex networks," *Reviews of Modern Physics*, **74**, 47–97. Available electronically at <http://www.nd.edu/~networks/Papers/review.pdf>
- Amin, M. (2001), "Towards Self-Healing Energy Infrastructure Systems," *IEEE Computer Applications in Power*, January 20–28.
- Barabási, A.-L. & Albert, R. (1999), "Emergence of scaling in random networks," *Science*, **286**, 509–512. <http://www.nd.edu/~networks/Papers/science.pdf>
- Barabási, A.-L. (2002), *Linked: The New Science of Networks*, Perseus Publishing.
- Barabási, A.-L. & Bonabeau, E. (2003), "Scale-Free Networks" *Scientific American*, **288**, 50–59. [http://www.nd.edu/~networks/PDF/Scale-Free Sci Amer May03.pdf](http://www.nd.edu/~networks/PDF/Scale-Free_Sci_Amer_May03.pdf)
- Biggs, N. (1993), *Algebraic Graph Theory*, 2nd edition, Cambridge University Press.
- Bollobás, B. (2001), *Random Graphs*, 2nd edition, Cambridge University Press.
- Bollobás, B. & Riordan, O. (2003), "Robustness and Vulnerability of Scale-Free Random Graphs," *Internet Mathematics* **1** (1), 1–35.
- Brandstein, A. G. & Horne, G. E. (1998), "Data Farming: A Meta-technique for Research in the 21st Century," *Maneuver Warfare Science 1998*, US Marine Corps Combat Development Command Publication, http://www.mcwl.quantico.usmc.mil/divisions/albert/research/documents/data_farming.rtf
- CTIA (2001), "Report to NRIC: Network Impact and Recovery Efforts – September 11 2001," http://files.ctia.org/pdf/CTIA_NRIC_0911.pdf
- Dekker, A. H. (2001), "Visualisation of Social Networks using CAVALIER," *Proc. Australian Symposium on Information Visualisation*, P. Eades & T. Pattison (Eds.), Sydney, Australia. *Conferences in Research and Practice in Information Technology*, **9**, 49–55. <http://crpit.com/confpapers/CRPITV9Dekker.pdf>
- Dekker, A. H. (2002), "C4ISR Architectures, Social Network Analysis and the FINC Methodology: an Experiment in Military Organisational Structure," DSTO Report DSTO-GD-0313, January. <http://www.dsto.defence.gov.au/corporate/reports/DSTO-GD-0313.pdf>
- Dekker, A. H. (2003), "Using Agent-Based Modelling to Study Organisational Performance and Cultural Differences," *Proc. MODSIM 2003 International Congress on Modelling and Simulation*, Townsville, Queensland, 1793–1798. Available electronically at http://mssanz.org.au/modsim03/Media/Articles/Vol_4_Articles/1793-1798.pdf
- Dekker, A. H. (2004a), "Simulating Network Robustness: Two Perspectives on Reality," *Proceedings of SimTecT 2004 Simulation Conference*, Canberra, 126–131.
- Dekker, A. H. (2004b), "Network Farming for the Analysis of Complex Network Systems," Presentation to the Complex Adaptive Systems in Defence Workshop, University of Adelaide, July.
- Dekker, A. H. & Colbert, B. (2004a), "Network Robustness and Graph Topology," *Proc. 27th Australasian Computer Science Conference*, V. Estivill-Castro (Ed.), Dunedin, New Zealand. *Conferences in Research and Practice in Information Technology*, **26**, 359–368. Available electronically at <http://crpit.com/confpapers/CRPITV26Dekker.pdf>
- Dekker, A. H. & Colbert, B. (2004b), "Scale-Free Networks and Robustness of Critical Infrastructure Networks," to appear in *Proceedings of 7th Asia-Pacific Conference on Complex Systems*, Cairns, Australia, 6–10 December.
- Dekker, A. H. & Colbert, B. (2005), "The Symmetry Ratio of a Network," to appear in *Proceedings of Computing: The Australasian Theory Symposium*, Newcastle, Australia, *Conferences in Research and Practice in Information Technology*, **41**.
- Dickson, K. D. (2001), *The Civil War for Dummies*, IDG Books, Indianapolis.
- Gibbons, A. (1985), *Algorithmic Graph Theory*, Cambridge University Press.
- Horne, G. E. (1997), "Data Farming: A Meta-technique for Research on 21st Century Questions," briefing presented at the US Naval War College, Newport, Rhode Island, November.
- ICT (2004), *International Terrorism Database*, <http://www.ict.org.il/>, accessed 27 April.
- Jones, W. D. & Geppert, L. (2002), "9/11: One Year Later," *IEEE Spectrum*, September, 35–36.
- Motter, A. E. & Lai, Y.-C. (2002), "Cascade-based attacks on complex networks," *Physical Review E*, **66**, 065102. Available at http://arxiv.org/PS_cache/cond-mat/pdf/0301/0301086.pdf
- Saaty, T. L. (1961), *Elements of Queueing Theory with Applications*, McGraw-Hill.
- Sinclair, A. (2003), *An Anatomy of Terror: A History of Terrorism*, Pan Macmillan, London.
- Van Ooyen, M., Noe, N. & Lynn, J. (2002), Technology Lessons Learned From New York City's Response To 9/11, Council of the City of New York Report, August. http://www.nycouncil.info/pdf_files/reports/9_11techreport.pdf