NetAttack: Co-Evolution of Network and Attacker

Holly Arnold*[1], David Masad†[2], Giuliano Andrea Pagani‡[3], Johannes Schmidt§[4], and Elena Stepanova¶[5]

[1] University of Oregon, Eugene, Oregon, USA,
[2] George Mason University, Fairfax, Virginia, USA,
[3] University of Groningen, Groningen, The Netherlands,
[4]University of Natural Resources and Life Sciences, Vienna, Austria,
[5]Santa Anna School of Advanced Studies, Pisa, Italy

September 16, 2013

Abstract

Interactions between individuals or organizations, and the changes and evolutions that result, are a central theme of complexity research. NetAttack aims at modeling a network environment where an attacker and a defender compete to disrupt a network or keep it connected. The choices of how to attack and defend the network are governed by a Genetic Algorithms (GA) which is used to choose among a set of available strategies. Our analysis shows that the choice of strategy is particularly important if the resources available to attacker and defender are similar. In such a situation, the defender and attacker genomes co-evolve and find an equilibrium. The best strategies found through GAs by the attackers and defenders are based on betweenness centrality. Our results agree with previous literature assessing strategies for network attack and defense in a static context. However, our paper is the first to show how a GA approach can be applied in a dynamic game on a network. This research provides a starting-point to further explore strategies and to optimize network disruption and reconstruction. Many applications for our kind of analysis may be found in the field of security and safety dealing with social (criminal networks) and technological (computer networks) contexts.

*email: arnold3@uoregon.edu
†email: dmasad@gmu.edu
‡email: g.a.pagani@rug.nl
§email: johannes.schmidt@boku.ac.at
¶email: e.stepanova@sssup.it

1 Introduction

Network theory may be used to elegantly model systems with many interacting elements, and is applied in many different disciplines [30, 26], including but not limited to biology [17], chemistry [8], linguistics and social sciences [31], computer networks and the web [9, 2], epidemics [18, 6], infrastructures [21, 13, 28], and banking [4]. In general, empirical studies of networks can be classified into biological, social, technological, or information networks [26]. By modeling a system as a network, one can ask questions about how the interactions between the elements within it affect the overall network behavior. A network element is represented by a node, and an edge between two nodes represents an interaction or relationship between the two elements [27].

A central question in network science is to understand the robustness of a network if nodes or edges fail or come under attack [16, 3]. The study of network robustness has many different applications, depending on the system which the network represents [16]. For example, social networks of interest are covert networks such as criminal or terrorist organizations [20]. In this article, we aim at examining the interactions between attackers and defenders on a network. In particular, we are interested in which strategies may be applied by the two groups to efficiently reach their goal of either destroying the network coherence or restoring it, respectively.

Practical applications include the removal of individuals from criminal organizations by police forces and attempts to restore the full communication capability of the network by the attacked group. Analogously, the maintainers of computer networks might attempt to identify the best strategy to defend against cyber attacks or random failures.

Several factors determine how effective an attack is, and how the network is able to defend itself. For example, the network's topology can have a large affect on its robustness. Albert et al. [3] showed that scale-free networks are very robust to random failure but vulnerable to targeted attacks. Holme et al. [15] consider attacks on edges as opposed to nodes, and suggest edge centrality as an effective target of an attacker. Domingo-Ferrer et al. [7] showed that the attacker's knowledge of the network is also an important factor in the effectiveness of an attack. Random disruptions and targeted attacks on networks have been considered with particular attention in the context of infrastructures represented by network. Vulnerability of power grids [29, 1], subway networks [21], and airline transportations [14, 13] have been analyzed to understand the key points of their safety or vulnerability. Several researchers have addressed the issue of network robustness using iterative attack and defense games on networks, however they all use static attack and defense strategies [24, 15, 7].

We extend their approach by allowing the attacker and defender to operate with a set of strategies in each time step and to make decisions based on mixing strategies. Both players choose a given strategy based on their genome, which represents a weight on all available strategies. Changes in the choice of strategies and optimization of strategies over time is a more realistic assumption for modeling covert social networks than relying on static approaches. We apply a

genetic algorithm (GA) to allow for dynamic development of strategies.

GAs apply the principles of natural selection in order to computationally optimize solutions to a particular problem [23, 10, 11]. The idea is to encode a problem of interest to be solved or optimized in a large population of individuals. As in natural populations of individuals, each individual in a GA population is characterized by a set of genes (i.e., the genome) that encode parameters the individual uses to solve the problem. These parameters are then evaluated by a fitness function; individuals with higher fitness functions are more likely to reproduce. Reproduction is repeated for several generations. In that way, a certain fitness function can be optimized over generations, however, solutions do not guarantee strict optimality.

This paper is structured as follows. Section 2 provides a description of our model. In section 3, we introduce the details of the GA, the genomes of attackers and defenders, the reproduction process, and the fitness function. In section 4 we outline the scenarios that we have run for this article. Section 5 reports on simulation results, discusses its main implications, and also shows the results of our sensitivity analysis. Section 6 introduces to related work, and Section 7 concludes the paper. An appendix section contains simple numerical examples to better illustrate the computation of attacker and defender genomes and the fitness of individuals.

2 The Model

In our model we have three fundamental entities that we deal with:

1. A **network** composed by a set of n nodes and m edges.
2. An **attacker** attempting to disrupt the network.
3. A **defender** attempting to repair the network after an attack to guarantee its continuing functionality.

An attacker disrupts the network by percolation, or the process of removing a node and all its associated edges. The defender, on the other hand, may reintroduce a node previously disconnected to the network and add (or rewire) some edges within the network. The attacker and defender each have an assigned set of resources that they can use in their attack or defense process. The resources for the attacker corresponds to the number of nodes that he can remove, whereas defender resources correspond to the number of edges that can be added to the network following an attack. We assume that attackers and defenders have complete knowledge of the network topology.

A particular simulation starts by generating an initial (first generation) population of an equal number of attackers and defenders. Their genes are initialized randomly, and attackers and defenders are randomly paired up. Each attacker-defender pair is assigned a network of n vertices and no edges. Based on the rules defined by their genomes (which are explained in detail in section 3), each defender adds new edges to the network, up to a total number of m edges.

After the network is initially built, the attacker removes k nodes, k being the amount of resources assigned to the attacker, in the network. The choice of the nodes to remove depends on the attacker genome. Once the attack phase is completed, the defender is allowed to add a total of w edges, w being the amount of resources assigned to the defender, to the network. First, the nodes removed by the attacker in this round are re-connected to the network. The nodes to which they will be connected depends on the defender genome. If defender resources allow additional edges to be inserted into the network, those edges are added to the network by the following rule: the starting point for the edges is a random node from the list of nodes which lost edges in the previous attack. The end point is determined by the weighting algorithm. If even more resources are available, random nodes in the network are picked as starting points while again, the ending points are determined by the weighting algorithm which is driven by the GA.

This process of attack and defense on the network is repeated for r rounds. In summary, a round is an execution of the game with iterative attacks each based on the k resources for the attacker and a (re-)wiring process consisting of w resources for the defender. In our simulations r is equal to 20, i.e. a total of 20 attack-defense rounds is played in each generation of the genetic algorithm. Figure 1 shows a scheme of the overall process.

After each round, the fitness (see Section 3 for the thorough fitness description) of the attackers and defenders[1] is calculated and a final average fitness after r rounds is computed for each individual in the population (see Appendix 3 for an example of average fitness computation). Recombination of individuals and mutations which are necessary to generate a new generation of attackers and defenders are discussed in the next section.

[1]The attacker's fitness is the opposite of the defender's fitness and is discussed in Section 3.

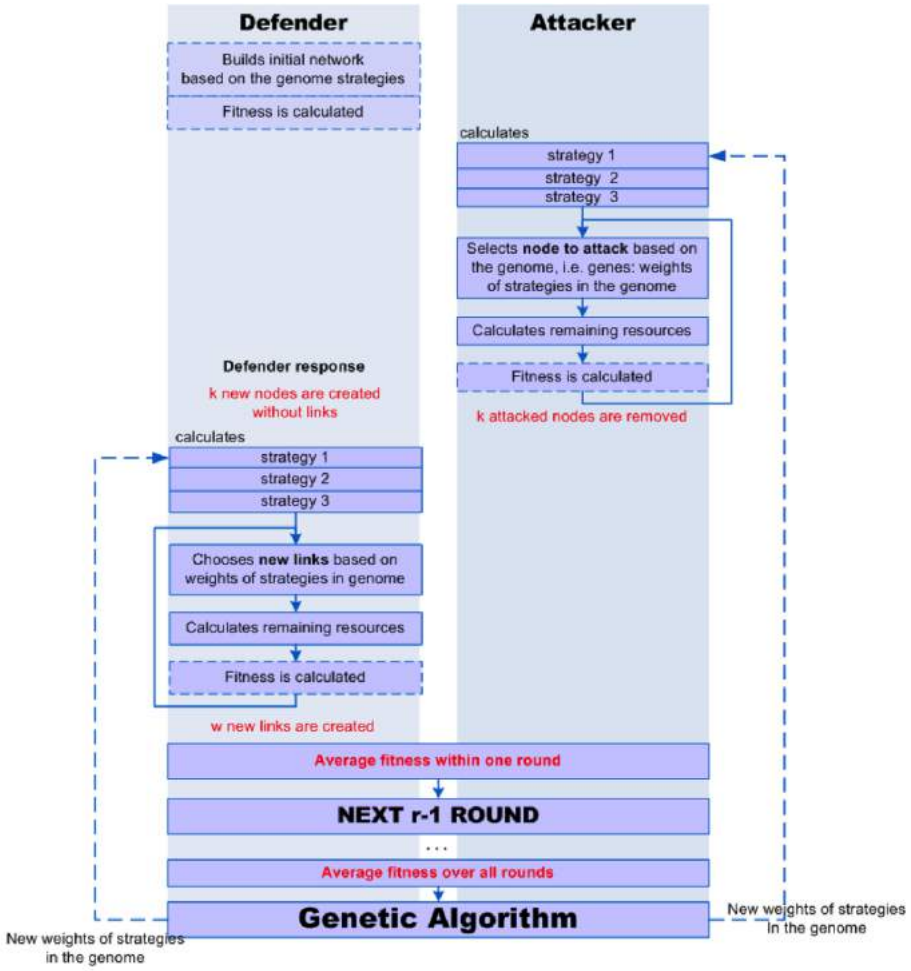

Figure 1. Sequence of actions of one attacker/defender individual in one generation.

3 Genetic Algorithm

The GA is used to evolve the strategies applied by the attackers and defenders and thus, allows for a dynamic development of the strategies that are applied by the two groups. First, we define the fitness function, then we discuss the genomes of attackers and defenders, and finally we present recombination and mutation strategies.

3.1 The fitness function

We define fitness to be the size (i.e., number of nodes) of the Largest Connected Component (LCC). The size of the LCC is a good proxy of the resilience of the

network, its ability to keep its structure connected and thus allow interaction between the nodes. The same metric has been used in previous studies [19, 25], allowing our results to be compared to previously-published ones. However, depending on the application of our model, different fitness functions may be appropriate. In section 6 we discuss this aspect in more detail.

3.2 Attacker genome

A set of strategies[2] is available to the attacker indexed by $j = \{1, 2, 3\}$ – these strategies have been developed previously in the literature [19, 25, 7]:

1. High-degree removal: nodes are prioritized for removal in decreasing order with respect to their degree.
2. High-centrality removal: nodes are prioritized for removal in decreasing order with respect to their betweenness centrality, which is known to be more related to connectivity than other centrality measures.
3. Random removal: nodes are prioritized randomly.

Each gene G_j corresponds to a weight on one of the strategies, and its value varies from 0 to 100. A strategy assigns to each node i in the network a value N_{ij} in the interval $[0, 1]$. For each node in the network, the attacker's genome assigns a number $TotalN_i = \sum_j G_j N_{ij}$ which is a linear combination of all available strategies weighted by the attacker genome. The probability of a node to be attacked Pr_i is this number $TotalN_i$ divided by sum of these numbers for all network nodes, $Pr_i = \frac{TotalN_i}{\sum_i TotalN_i}$. In each round a node with all its edges is removed with probability Pr_i (see Appendix A for a numerical example).

3.3 Defender Genome

The strategies of the defender are similar to the attacker strategies as they are based on the same weighting algorithm. The starting point of an edge that is added to the network is not determined by this weighting algorithm, but by a sequence of rules as outlined in the previous section. Only the endpoint of the new edge is determined by the defender's genome.

The following strategies are available to the defender indexed by $j = \{1, 2, 3\}$ - these strategies have been developed previously in the literature [19, 25, 7]:

1. Preferential replenishment: nodes are ranked in decreasing order with respect to their degree.
2. Balanced replenishment: nodes are ranked in increasing order with respect to their betweenness centrality.
3. Random replenishment: nodes are ranked randomly.

[2] A strategy is a mechanism for both the attacker and the defender to decide which node to attack or edge to create/rewire based on some rules, measures or indicators on the network.

Weighting of strategy is performed similar to the attacker, i.e. the genome determines how the value of a certain metric for the nodes is weighted. See the description of the attacker genome above for details. Appendix A presents a numerical example for the defender case.

3.4 Genome reproduction process

The ordered set of genes G_j, $j = \{1, 2, 3\}$ representing the attacker and the defender genome are initially randomly sampled from a uniform distribution in the range $[0, 100]$. Reproduction consists of gene recombination: two attackers or defenders from the current population are randomly chosen from the current generation. The probability of being picked is not uniform, but is proportional to the fitness of the agent. A random position in the genome is chosen for crossover. At this position, the two individuals will exchange their genetic material, taking the first part from the first parent and the second part from the second parent (as we have only 3 genes in the genome, there are only two possibilities: the offspring will inherit the first gene from his first parent and second and third genes from his second parent, or he will inherit the two first genes from the first parent and the third gene from the second parent).

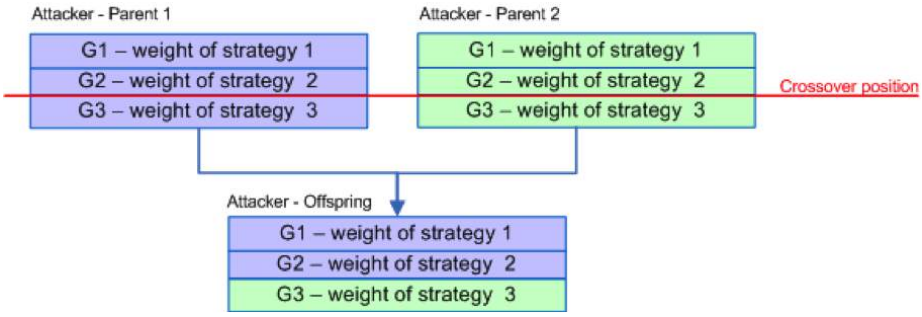

Figure 2. Example gene crossover

A mutation process occurs with a fixed 5% probability. The mutation in a gene is obtained by sampling a value from a Gaussian distribution with the mean equal to the current value of the gene and a standard deviation of 5.

4 Scenarios

We are interested in the following research problems: first, how does an attacker applying a genetic algorithm perform against a static defender, i.e. a defender with only one, fixed defense strategy. We next look at the inverted scenario, i.e. how a static attacker performs against an evolving defender. Finally, we allow both the attacker and defender to co-evolve against each other. For the purpose of comparison, we also run each static attacker strategy against each

static defender strategy. Both defender and attacker have 3 different strategies each. This implies that there are 16 different scenarios to assess in total.

In the base run, we start with a population of 200 attackers and defenders, operating on a network of 100 nodes and 150 edges, and run the GA for 500 generations. Attackers are allowed to remove 3 nodes while defenders rewire 5 edges. In a sensitivity analysis we test different defender budgets of 3,7, or 9 edges.

5 Results

5.1 Scenarios Results

5.1.1 Static Defenders

Figure 3 shows that the dynamic attacker quickly approaches the fitness of the single best attacker strategy against a static random defender. The genes evolve accordingly, prioritizing high weights for the betweenness strategy and much lower weights for the other two strategies. It can also be observed that the standard deviation in the genes decreases over time, indicating that the individuals in the population converge. Playing against the other two static defender strategies show similar results (middle and bottom in Figure 3). The worst static defense strategy is preferential attachment which can be derived from the fact that the attacker fitness is highest in that case (middle in Figure 3). The best possible static defense strategy is balanced replenishment as indicated by the low attacker fitness (bottom in Figure 3). In all cases, the betweenness attack strategy is selected by the attacker's GA.

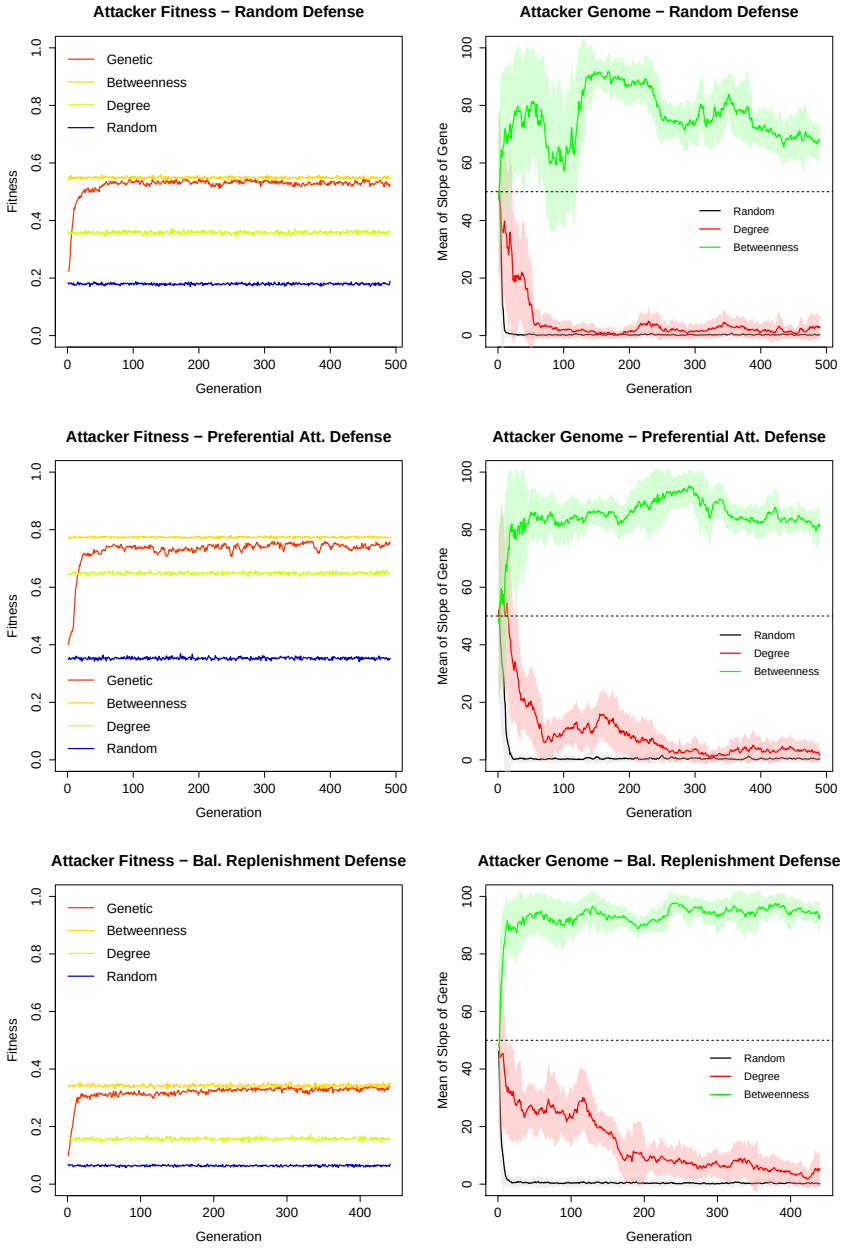

Figure 3. Left: evolution of the mean of fitness in the attacker population when attackers use the genetic algorithm against 3 static strategies. Right: Evolution of the mean of attacker weights for different strategies in the genetic case. The transparent areas indicate the standard deviation. Top: Attacker vs. Random defender. Middle: Attacker vs. Preferential defender. Bottom: Attacker vs. Balanced Replenishment.

5.1.2 Static Attackers

Also the defender has a preferred strategy, independent of the static attacker strategy. It is balanced replenishment. However, the GA takes more time to find the dominating strategy in comparison to the attacker's GA in some cases. Defending against a random attacker (top in Figure 4) shows that the defender's fitness approaches the fitness of the best possible solution only after 400 generations - even though the balanced replenishment strategy is selected earlier as can be observed by the top graph on the right of Figure 4. However, as long as the random strategy has a rather high weight, the fitness of the defender is not significantly increased. Only after ruling out the random defense, the fitness increases rapidly. That indicates that even a small amount of mixing of strategies may cause a rather bad performance of the defender. This is not the case for the second and third comparison in Figure 4 - if the attacker applies the degree attack and betweenness strategy respectively, the defender evolves rapidly in using the balanced replenishment strategy only. The fitness, accordingly, increases quickly in both cases. The defender can deal best with the random attack strategy, as indicated by the high overall fitness in Figure 4, while the best strategy for the attacker seems to be betweenness attacks, as also confirmed by the results in the previous section.

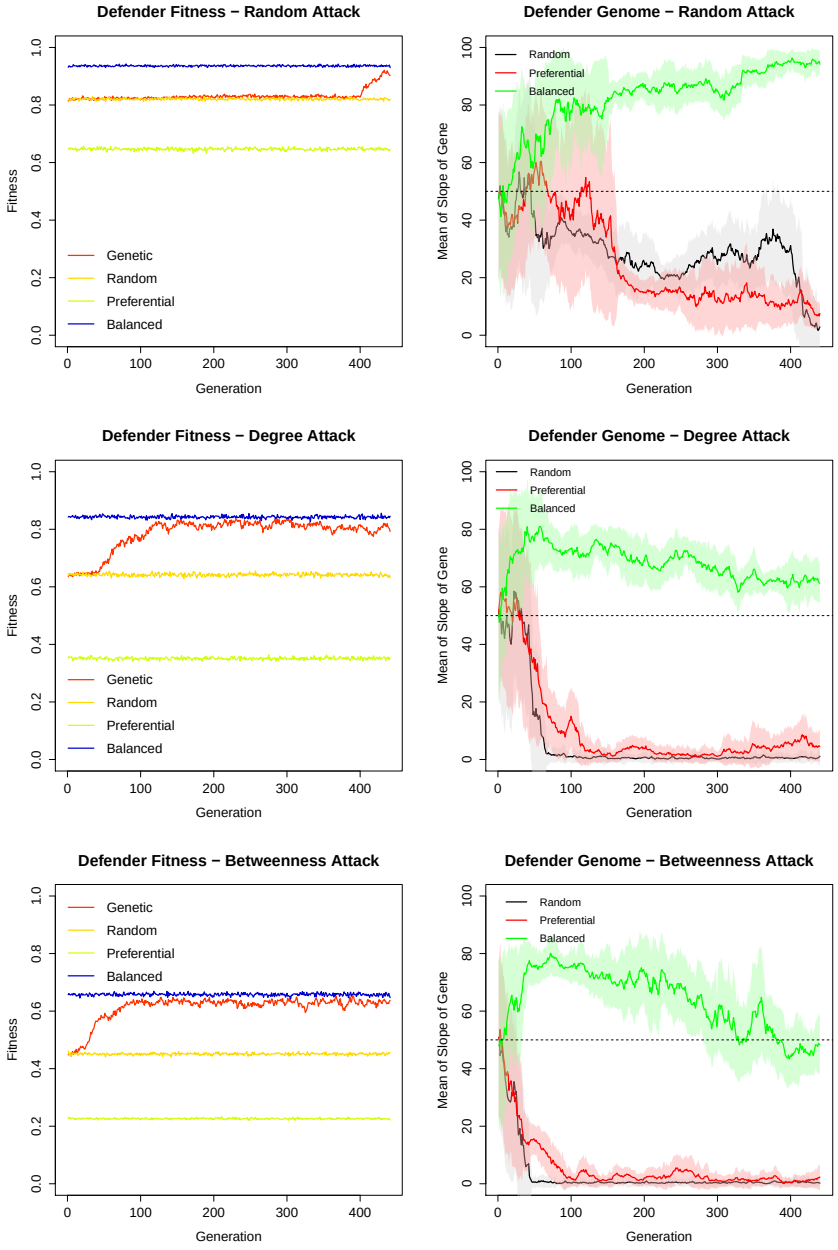

Figure 4. Results of simulation runs: Defenders applying the genetic algorithm against 3 static attack strategies. Left: Mean of fitness of attacker. Right: Mean of attacker genes. The transparent areas indicate the standard deviation. Top to bottom: Random attack vs. Defender, Degree attack vs. Defender, Betweenness attack vs. Defender.

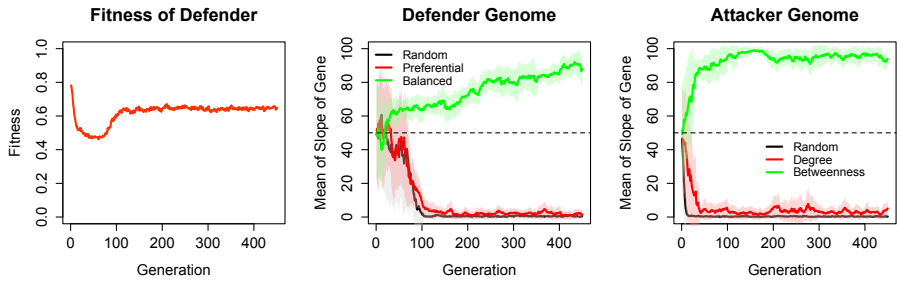

Figure 5. Results of simulation runs. Left: evolution of the mean of fitness in the defender population in the co-evolution case. Middle: Evolution of the mean of defender weights for different strategies in the GA case. The transparent areas indicate the standard deviation. Right: Evolution of the mean of attacker weights for different strategies in the GA case. The transparent areas indicate the standard deviation.

5.1.3 Co-Evolution

In the case of co-evolution, i.e. both, defenders and attackers employ a genetic algorithm to select their strategy, attackers evolve quicker towards the more efficient strategy, causing a decline in the fitness of the defender (see Figure 5). However, after about 50 generations, there is a turn-around and the defender starts selecting the best defense strategy, causing an increase in the defender's fitness. After defender and attacker have evolved into applying the balanced replenishment and betweenness attack strategies respectively, the fitness function stabilizes and no further major fluctuations are observed – an equilibrium is reached.

5.2 Sensitivity analysis

The sensitivity analysis shows that a decrease in the amount of edges increases the attacker fitness at the beginning and at the end of the simulation and vice versa. In addition, a high number of edges and an efficient defense strategy (i.e. balanced replenishment) almost completely reduces the possibility of the attacker to increase her fitness - see Table 1, row GA vs. Balanced Replenishment and a budget of 9. On the other hand, a low budget decreases the fitness improvements over time for the defender - see Table 1, budget of 3. This indicates that a meaningful game can only be played if the available resources are in a certain, rather small corridor - too many resources for one of the two sides will make any response strategy inefficient. In the co-evolution case, the defender shows a decreasing fitness over time if the budget is smaller or equal to 5 edges, while it is the other way round for a budget above that level.

Defender Budget	**3**		**5**		**7**		**9**	
Attacker vs.	FAS	FAE	FAS	FAE	FAS	FAE	FAS	FAE
Random Defense	0.38	0.63	0.22	0.52	0.12	0.37	0.08	0.18
Preferential Defense	0.48	0.76	0.40	0.76	0.36	0.64	0.32	0.62
Balanced Replenishment	0.37	0.54	0.10	0.34	0.01	0.02	0.01	0.01
Defender vs.	FDS	FDE	FDS	FDE	FDS	FDE	FDS	FDE
Random Attack	0.67	0.68	0.81	0.92	0.90	0.97	0.94	0.98
Degree Attack	0.49	0.54	0.63	0.81	0.81	0.98	0.90	0.98
Betweenness Attack	0.36	0.42	0.45	0.63	0.61	0.95	0.82	0.97
Co-Evolution	FDS	FDE	FDS	FDE	FDS	FDE	FDS	FDE
GA vs. GA	0.62	0.38	0.78	0.66	0.88	0.95	0.92	0.98

Table 1. Fitness of attackers and defenders with varying budget (number of edges to be added). FAS and FAE indicate the average fitness of the attacker at the start and the end of the simulation (i.e. generation 1 and generation 500), respectively. FDS and FDE indicate the average fitness of the defender at the start and at the end of the simulation, respectively.

6 Related Work

In the network literature several authors have considered the topic of network robustness in case of attacks on nodes or edges. Here we look more in detail to studies where the concepts of evolution of a network, in terms of its topology, is tied to the behavior of an attacker of the network. In a seminal paper by Albert et al. [3], the authors demonstrate that scale-free networks are vulnerable to targeted attacks of nodes of high degree, while fairly robust to random attacks. Holme et al. [15] consider attacks on edges as opposed to nodes, and suggest edge centrality as an effective target of an attacker. We also look at studies where the co-evolution term is taken into consideration.

As already mentioned in Section 1, the work of Nagaraja and Anderson [25] is relevant to our paper since it considers an evolutionary game theory approach that takes place on a network. In a way similar to our interpretation of the evolutionary game, their game is organized in rounds and each round consists of an attack followed by a recovery. The attack consists of targeting a number of nodes to be removed, depending on the attacker budget. However, the recovery is different than the one we propose in this paper, and consists in two stages, namely replenishment and adaptation. The first stage deals with inserting new nodes into the network and establishing new connections based on the defender's budget, while the second deals with rewiring existing links. The objective for the attacker is to split the network in separate components. The authors examine attacks on scale free network, based on node degree; the defense strategies tested are ring replenishment, clique replenishment, and random. The first defense consists in substituting a node with high degree (therefore likely to be attacked) with a set of n nodes. The high degree node's edges are equally split between the members of the ring. The second strategy has a similar flavor, but the high degree node is replaced with a clique of n nodes. In this first set of experiments random replenishment is basically ineffective and the network is almost immediately disrupted. Also the ring replenishment provided limited benefits, while the clique substitution is the most effective one giving the network

a robust connectivity. The authors also consider betweenness as a type of attack and the effects are more disrupting against all types of defense. Another defense strategy that is proposed in the paper is delegation: rewiring part of a node's connections to a chosen neighbor that becomes the 'deputy' of the node. This last strategy in combination with cliques proves quite efficient in maintaining a component at least half the size of the original one. The paper is a good inspiration for NetAttack, however, our approach is more flexible giving the possibility to the attacker and defender to adapt or change their strategies (i.e., type of attack/defense) during the game, while in [25] the strategies are chosen and kept fixed through the game. Our model allows to identify the strategies for attackers and defenders that provide the maximum fitness out of a potentially broad set of strategies. In [25] the test performed takes into account scale free networks as initial topologies, whereas our approach starts with an initial topology that is already optimized by the defender under the assumption that the defender initially generates the network (e.g., criminal social networks). One aspect that NetAttack proves through the evolution of the genome is the superiority in attack of the balanced replenishment strategy that is highlighted also in [25]. Nagaraja and Anderson's work is not without limitations, however. The cost of implementing an edge is essentially zero since the network is allowed to rewire with an arbitrary amount of newly added edges.

Kim and Anderson [19] expand upon the work of Nagaraja and Anderson. Kim and Anderson give each attacker and defender a fixed budget, or cost to add nodes and edges after an attack, and analyze the effect of attacks on a variety of different network topologies. They find a strategy of connecting low centrality nodes is the best defense strategy. However, as the edge to node ratio increases, the network becomes more robust, and even adding edges randomly is effective against targeted attacks. They find that there is a threshold value for the proportion of edges to nodes at which point the effectiveness of attacks decreases drastically.

The work of Domingo-Ferrer and Gonzalez-Nicolas [7] is based on the ideas and findings of previous work by Nagaraja and Ross [25] and Kim and Ross [19] and adds further properties to the networks and the experiment set. In the paper the authors analyze the evolution of the *order* and average path length of scale-free networks under attack and defense. In particular, the tests consider networks that are undirected and directed, and unweighted and weighted. In addition, the only strategy of attack considers betweenness centrality as the measure to identify the most critical node; whereas defense is achieved following two types of strategies: delegation and node replenishment. The first works by reducing the degree of the nodes with highest degree and distributing part of the edges between the neighbors of the node. The second is node replenishment and consists of replacing a node with a clique of n nodes. One of the novelty of the paper consists in considering costs for attacks and defenses for which the attacker and defender have a budget each. In addition, the attacks and defenses do not take place synchronously (an attack followed immediately by a defense), but a defense is performed after a number of attacks (node removals) take place. The experiments are conducted once again on scale-free networks

(weighted/unweighted and directed/undirected) and a sample of the Internet (more than 20000 nodes). The results show basically that an important factor is the visibility that an attacker has of the network (attackers with 20%, 40%, 60%, 80% and 100% visibility are considered), while there is basically no difference in the disruption behavior of weighted and unweighted networks (although different measures in the path centrality computation are used). A remarkable aspect is the better resistance to attacks that directed network have compared to undirected ones, partially due to the higher density of the former. The article brings interesting points such as introducing explicit cost functions and testing on different type of networks (un/weighted, un/directed), however always scale-free. Our approach is more flexible considering the possibilities of different strategies of attack and defense and networks that are not fixed a priori, but built by the defender that is usually the organization that has to defend from the attacks.

Another work that considers the concept of attack and defense in a network is the paper by Chi et al. [5]. The paper focuses on considering only two types of well-known network topologies such as random graphs and Barabasi-Albert scale-free networks. The idea of the authors is to evaluate the appearance of a stability condition after a long series of attacks and repair. The type of attack considered targets all the links of the node with highest degree, while the repair process consists in defining a probability that enables the node that has lost the connections to connect to any other node. The simulation shows that the network tends to have a convergence of the invulnerability already from the initial evolution and then it gets even more stable. Both the random graphs and the scale free network reach a steady state condition for the instability index. The main summary of the findings are that the overall statistics of the network do not change considerably (e.g., node degree distribution), while the local properties undergo an interesting evolution that leads to networks with more local clusters with similar nodes. Comparing this paper to NetAttack, the network topologies investigated are fixed as well as the strategies of attack that do not change over time. We consider the ability of the defender and the attacker to adapt and change the strategies according to a GA and topologies that do not follow standard models, two important innovations that bring the model closer to real-life network games.

In NetAttack the term co-evolution has a well defined meaning: the network changes its topology due to the combined evolution of an attacker and a defender of the network that use GA to choose their strategy. Other studies on networks use the term co-evolution in a different context: they aim at describing the combined evolution of the topology of a network and the state of nodes in the network. In this context the term is used in [12]. The paper is a review on the topic of adaptive networks. Adaptive networks are network thats exhibit a relationship between the evolution of the topology and the state/condition of a node, therefore creating a sort of feedback loop between the state and topology; these networks are also referred as coevolutionary networks. This type of coevolution is different from what we consider coevolution. However, the similarity relies in the dynamic condition that characterizes the network

and changing topological properties (adding/removing nodes and links). The main idea of the paper is to present scenarios where adaptive networks appear, main concepts, and previous work on the issue. Adaptive networks appear in several fields from biology (e.g., food webs) to social sciences (e.g., opinion spread) to physiology (e.g., blood vessels) and health (e.g., epidemics). Several examples of networks are provided and one of the main characteristics is the self organization that sustains the evolution of the networks and the formation of network structures.

Another work on the coevolution of network as evolution of node state (or behavior) and network topology is the paper by Zhang et al. [32]. The co-evolution considers a network of individuals that has local interactions with their neighbors and can collaborate or defect in sharing goods, therefore influencing their payoff and the payoff of nodes directly attached. Considering the evolution of the topology, which is most relevant for our article, Zhang et al. show that the average node degree stabilizes to a value a little higher than 4. During the process of evolution in a collaborator's neighborhood the number of collaborators is generally stable, while the defectors almost disappear; a defector's neighborhood is almost only composed by defectors with the number of collaborators slowly disappearing. From a node degree distribution perspective, there is an evolution towards a long tailed distribution with the majority of the nodes that have a very low degree and few nodes (that are cooperators) with an high amount of connections. The term co-evolution here once again is considered for the behavior/state of the node. The network does not disrupt, but the changes in the topology are due to the evolution in the state of the nodes (especially the payoff of the neighborhood of a node) and in the probability parameters that rule the change in strategy or connections, thus an important underlying difference compared to our work.

The work of Louzada et al. [22] does not focus on the attack, but more on the mechanism to make a network more robust through rewiring edges. The approach proposed enables a better robustness than a random rewiring approach. The mechanism proposed, called smart rewiring, consists in picking a random node and selecting the lowest and highest degree neighbors of the node. From these two selected nodes, one random neighbor of each one is selected and the connecting edges are removed. The two edges are instead placed to connect the two pairs of selected nodes (the neighbors of the initial node and the neighbors of the neighbors). In such a way the rewiring procedure connects parts of the network that would have been disconnected upon the failure of an hub. Compared to random rewiring, the smart one proposed is more efficient: to achieve an increase of 30% robustness it only needs a rewiring of 9% of the links, whereas the random one requires 15% of rewiring steps. The smart rewiring increases modularity in the network by creating triangles and it creates connections between hubs and leaves. The solution proposed for rewiring is interesting and the result are beneficial for robustness, however the strategies of attack and defense of the network are static and no changes are allowed. The attack always targets the nodes with highest degree and the strategy to improve the network is always a smart rewiring. Our solution gives more flexibility in both the type of attack

and defense.

7 Conclusions and future work

We have shown that our approach to model interactions between attackers and defenders can be successfully modeled using genetic algorithms. Our results confirm what has been found in previous papers which compared various static strategies. In addition, our work shows that successful strategies can also be applied to generate networks from scratch (in contrast to other papers, which only used them to rewire networks after they have been attacked)[3]. Obviously, the success of a defense and attack depends on the available resources. The choice of the strategy matters primarily when the defender's resources are slightly larger than the attacker's resources. In any other case, the results of the game are going to be biased towards the side with the resource advantage. However, if resources are distributed so that a small advantage for the defender arises and if the defender's goal is to maintain or increase the LCC and the attacker aims for the opposite, the balanced replenishment and betweenness attack strategy, respectively, can be considered to be the most efficient strategies among the ones tested in this work – independent of which strategy is applied by the opponent. An equilibrium situation arises if the two opponents apply these strategies, although the defender appears to evolve slower than the attacker.

This result may be applied to social networks, computer networks, or any other kind of network. From an empirical perspective, it would be interesting if similar strategies are observed in real networks (i.e. where they have evolved 'naturally'). From a normative point of view, the results of this paper and related work can be used to design strategies to defend against attacks or to target attacks against certain nodes in networks.

Future work will include the development and testing of new defender and attacker strategies - currently, only three strategies are included. A larger number of strategies may make the game dynamics more complex than the current version, which allows for a stable equilibrium in the co-evolution case. Additionally, the current fitness function emphasizes connectedness of the network, but does not assess the efficiency of the network in providing transportation or communication services. Different fitness functions which may include a combination of the largest connected component with some measure of efficiency as, for example, the diameter or effective diameter of the network, therefore might be considered interesting options for future research.

Acknowledgements

This paper is the outcome of a group project started at the Complex Systems Summer School of the Santa Fe Institute in June, 2013. We are very grateful

[3]However, this difference is somehow minor if we consider that many attack-defense rounds applying the same defense strategies will cause the network topology to resemble a network that was built from scratch using the very same defense strategy.

to the organizers of the summer school, and to all the lecturers and presenters. To all of our fellow summer school students: thanks for the good times – and the greenhouse sessions. Thanks to Mauricio Cantor and Bruno Pace for very valuable discussion on the subject of the paper. Very special thanks go to Tom Carter for all of his efforts and in particular for his enlightening performance of the "power-law blues". The simulations for this paper were partly executed on the Vienna Scientific Cluster (VSC) – thanks a lot to the VSC team for their support during the whole process.

Appendix

Definition of the genome - Example Attacker

An attacker has 2 strategies j=2 (so he has 2 genes in the genome representing weights of these strategies - see table 2). The two strategies are attacking nodes with high degree centrality and random attack.

G1	G2
32.4	62.3

Table 2. Attacker genome

We consider a network as Example network 1 shown in Figure 6.

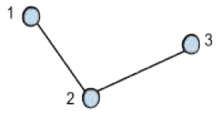

Figure 6. Example network 1

So that

node i=1 has degree centrality of 0.5 and the random number assigned to it is 0.37.

node i=2 has degree centrality of 1 and the random number assigned to it is 0.73.

node i=3 has degree centrality of 0.5 and the random number assigned to it is 0.16.

And consequently N_{ij} are the following

$N_{11} = 0.5$, $N_{12} = 0.37$

$N_{21} = 1$, $N_{22} = 0.73$

$N_{31} = 0.5$, $N_{32} = 0.16$

For each node the attacker's genome delivers the following number ($TotalN_i = \sum_j G_j N_{ij}$):

$TotalN_1 = 32.4 * 0.5 + 62.3 * 0.37 = 39.25$

$TotalN_2 = 32.4 * 1 + 62.3 * 0.73 = 77.88$

$TotalN_3 = 32.4 * 0.5 + 62.3 * 0.16 = 26.17$

The probability of each node to be attacked is thus the following (Pr_i)

$Pr_1 = 39.25/(39.25 + 77.88 + 26.17) = 0.27$

$Pr_2 = 77.88/(39.25 + 77.88 + 26.17) = 0.54$

$Pr_3 = 26.17/(39.25 + 77.88 + 26.17) = 0.19$

Definition of the genome - Example Defender

A defender has 2 strategies j=2 (so he has 2 genes in the genome representing weights of these strategies - see table 3). The strategies are preferential attachment to nodes with high degree centrality and random defense.

G1	G2
32.4	62.3

Table 3. Defender genome

Let's assume that node 1 has been removed during the last attack and that the node is chosen to be rewired. We consider a network as Example network 2 shown in Figure 7.

Figure 7. Example network 2

As calculated by node 1:

node i=2 has degree centrality of 1 and random number assigned to it is 0.37.

node i=3 has degree centrality of 1 and random number assigned to it is 0.73.

And consequently N_{ij} are the followings

$N_{21} = 1$, $N_{22} = 0.37$

$N_{31} = 1$, $N_{32} = 0.73$

Round	Action	Fitness - Defender	Fitness - Attacker
0	Initial network setup	0.6	0.4
1	Attack	0.4	0.6
1	Rewire	0.5	0.5
1	Rewire	0.6	0.4
1	Rewire	0.6	0.4
2	Attack	0.45	0.55
2	Rewire	0.52	0.48
2	Rewire	0.58	0.42
2	Rewire	0.6	0.4
	Final Fitness	0.54	0.46

Table 4. Development of fitness of defender and attacker over two rounds of attack. Final fitness is calculated, assuming that there are only two rounds of attacks.

For each node defenders genome delivers the following number ($TotalN_i = \sum_j G_j N_{ij}$):

$$TotalN_2 = 32.4 * 1 + 62.3 * 0.37 = 55.45$$

$$TotalN_3 = 32.4 * 1 + 62.3 * 0.73 = 77.88$$

Probabilities of node 2 and 3 to be connected with node 1 are thus the followings (Pr_i)

$$Pr_1 = 55.45/(55.45 + 77.88) = 0.42$$

$$Pr_2 = 77.88/(39.25 + 77.88) = 0.58$$

Example of the computation of the fitness function for attackers and defenders

Assume that the attacker has one resource to use in its attack and the defender has three resources for its defense. There are two attacks and two defenses. Table 4 shows the calculation of the fitness function in that case.

Flow diagram

A flow diagram for the model described in Section 2 for the generation of a network and an attack and defense is shown in Figure 8. In the left side of the figure the operations of the defender are depicted, while those involving the attacker are represented as a flow on the right side of the figure. At the center of the figure the run manager is the entity (object at code level) that has the duty of operating on the network (adding/removing nodes and links); the run manager has also the task of computing the fitness at the end of each round and at the end of r rounds for an individual attacker and defender.

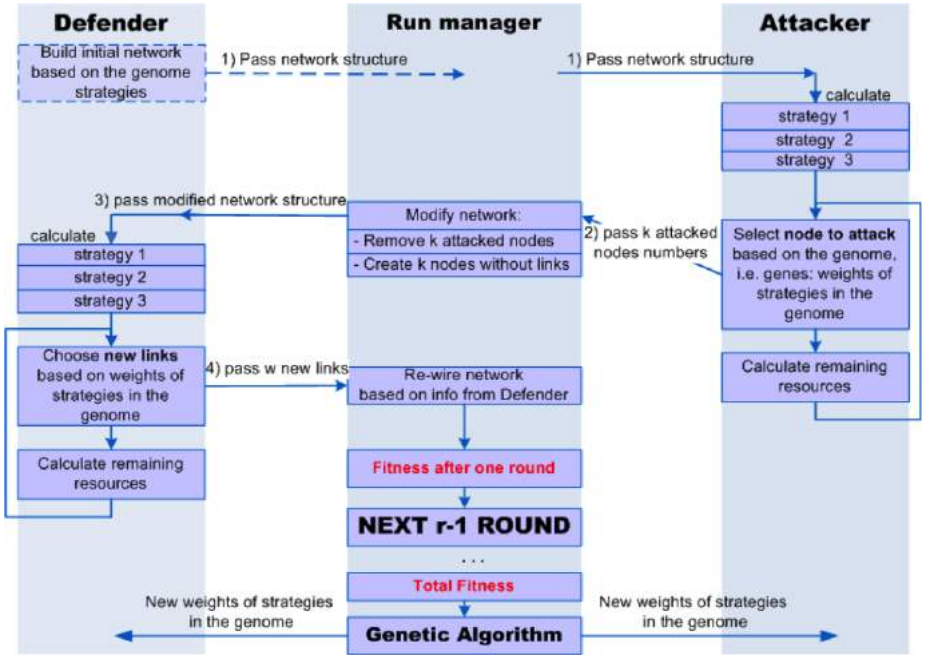

Figure 8. Flowchart of the interactions in an attack/defense round between the logical entities involved.

References

[1] R. Albert, I. Albert, and G. Nakarado, *Structural vulnerability of the north american power grid*, Physical Review E, 69 (2004).

[2] R. Albert, H. Jeong, and A.-L. Barabási, *Internet: Diameter of the World-Wide Web*, Nature, 401 (1999), pp. 130–131.

[3] R. Albert, H. Jeong, and A. L. Barabási, *Error and attack tolerance of complex networks*, Nature, 406 (2000), pp. 378–382.

[4] M. Boss, H. Elsinger, M. Summer, and S. Thurner, *The network topology of the interbank market*, Quantitative Finance, 4 (2004), pp. 677–684.

[5] L. Chi, C. Yang, and X. Cai, *Stability of complex networks under the evolution of attack and repair*, arXiv preprint cond-mat/0505197, (2005).

[6] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, *Predictability and epidemic pathways in global outbreaks of infectious diseases: the sars case study.*, BMC Med, 5 (2007), p. 34.

[7] J. Domingo-Ferrer and rsula Gonzlez-Nicols, *Decapitation of networks with and without weights and direction: The economics of iterated attack an d defense*, Computer Networks, 55 (2011), pp. 119 – 130.

[8] J. Doye, *Network Topology of a Potential Energy Landscape: A Static Scale-Free Network*, Physical Review Letters, 88 (2002), pp. 1–4.

[9] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *On power-law relationships of the internet topology*, in Proceedings of the conference on Applications, technologies, archit ectures, and protocols for computer communication, ACM, 1999, p. 262.

[10] S. Forrest, *Genetic algorithms: principles of natural selection applied to computation*, Science, 261 (1993), pp. 872–8.

[11] D. E. Goldberg and J. H. Holland, *Genetic algorithms and machine learning*, Machine learning, 3 (1988), pp. 95–99.

[12] T. Gross and B. Blasius, *Adaptive Coevolutionary Networks: A Review*, Journal of The Royal Society Interface, 5 (2007), pp. 259–271.

[13] R. Guimerà and L. a. N. Amaral, *Modeling the world-wide airport network*, The European Physical Journal B - Condensed Matter, 38 (2004), pp. 381–385.

[14] R. Guimerà, S. Mossa, a. Turtschi, and L. a. N. Amaral, *The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles.*, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005), pp. 7794–9.

[15] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, *Attack vulnerability of complex networks*, Physical Review E, 65 (2002), p. 056109.

[16] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang, *Attack robustness and centrality of complex networks*, PloS one, 8 (2013), p. e59613.

[17] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, *The large-scale organization of metabolic networks.*, Nature, 407 (2000), pp. 651–4.

[18] J. Kephart and S. White, *Directed-graph epidemiological models of computer viruses*, in Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on, May 1991, pp. 343 –359.

[19] H. Kim and R. Anderson, *An experimental evaluation of robustness of networks*, Systems Journal, IEEE, 7 (2013), pp. 179–188.

[20] V. E. Krebs, *Mapping networks of terrorist cells*, Connections, 24 (2002), pp. 43–52.

[21] V. Latora and M. Marchiori, *Is the boston subway a small-world network?*, Physica A: Statistical Mechanics and its Applications, 314 (2002), pp. 109 – 113.

[22] V. H. Louzada, F. Daolio, H. J. Herrmann, and M. Tomassini, *Smart rewiring for network robustness*, arXiv preprint arXiv:1303.5269, (2013).

[23] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, USA, 1998.

[24] S. Nagaraja, *Topology of covert conflict*, in Security Protocols, Springer, 2007, pp. 329–332.

[25] S. Nagaraja and R. Anderson, *The topology of covert conflict*, Tech. Rep. UCAM-CL-TR-637, University of Cambridge, Computer Laboratory, July 2005.

[26] M. Newman, *The structure and function of complex networks*, SIAM review, 45 (2003), pp. 167–256.

[27] ——, *Networks: an introduction*, Oxford University Press, 2009.

[28] G. A. Pagani and M. Aiello, *The power grid as a complex network: A survey*, Physica A: Statistical Mechanics and its Applications, 392 (2013), pp. 2688 – 2700.

[29] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, *Robustness of the European power grids under intentional attack*, Physical Review E, 77 (2008), pp. 1–7.

[30] S. H. Strogatz, *Exploring complex networks*, Nature, 410 (2001), pp. 268–276.

[31] J. Travers and S. Milgram, *An experimental study of the small world problem*, Sociometry, 32 (1969), pp. 425–443.

[32] C. Zhang, J. Zhang, G. Xie, and L. Wang, *Coevolution of strategy and structure on social networks*, in Decision and Control (CDC), 2010 49th IEEE Conference on, IEEE, 2010, pp. 1804–1809.