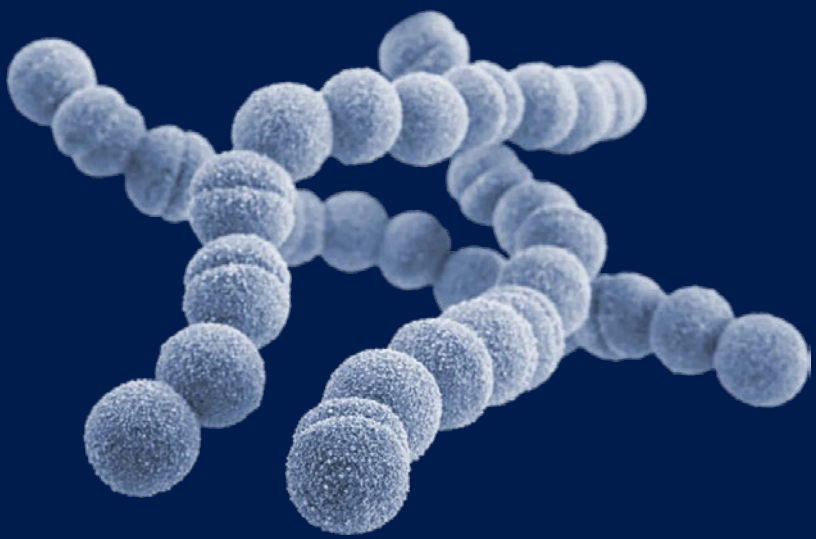
White Paper

The Anatomy of SCADA Risk: Leveraging Lifesaving Epidemic Models for a Novel Evaluation of SCADA/ICS Risks

Yehonatan Kfir
CTO, Radiflow



radiflow

ABSTRACT

- How does a virus propagate within a real network?
- What is the single best node to immunize?
- Which connections are best removed from the network?

While these questions seem to have been taken from a computer network domain, they are in fact questions that have been researched for several decades for the sole purpose of eliminating biological viruses.

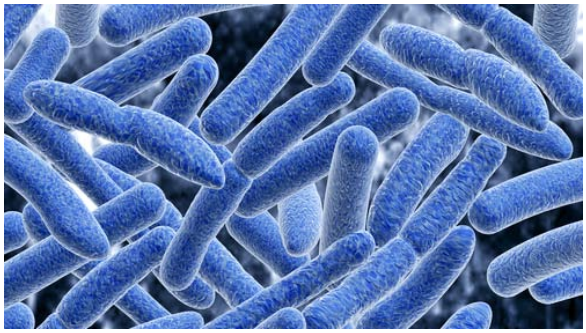
The well-researched biological epidemic models demonstrate astounding results in the prediction of disease and planning of immunization programs.

WE SHOW HOW THIS EPIDEMIC MODEL CAN BE USED TO PRIORITIZE SECURITY MITIGATIONS WITHIN A SCADA/ICS NETWORK.

In this whitepaper, we investigate these types of models and reveal how ideas derived from biological epidemic models can be replicated in a SCADA/ICS cybersecurity environment. We present an epidemic-based mathematical definition for SCADA/ICS *network vulnerability* and we show how this epidemic model can be used to *prioritize* security mitigations within a SCADA/ICS network.

INTRODUCTION

Reducing the vulnerability of a SCADA/ICS network can be an exhaustive task. Devices in such systems are rarely patched or hardened and therefore large numbers of vulnerabilities can easily be exposed. Additionally, the network architecture of these systems has not always been developed with security in mind.



In recent years, awareness of SCADA/ICS security has risen. Numerous tools, guidelines and regulations have been developed to support the transition of these types of systems to be more secure and robust. Security recommendations, such as best practices and mitigations, are frequently published. The detailed information shared on each recommendation provides security managers with the information they need to implement the various recommendations within their network.

Implementing security recommendations in SCADA/ICS networks can be time consuming and complicated, thus heightening the potential impact on the processes controlled by the network. Therefore, the implementation rate of security recommendations can be hindered. Furthermore, due to the increasing rate of reported vulnerabilities, there is a further increase in the backlog of

security recommendations. So much so that security managers are constantly required to prioritize a growing list of recommendations and choosing the mitigation measures that would have the highest impact is becoming a must-have best-practice.

In this whitepaper, we present the foundations for a risk-oriented system that prioritizes security recommendations according to their contribution to reducing the vulnerability of the network. For this analysis, we assume that the risk-oriented system is aware of the network topology and the characteristics of its connected assets.

"A RISK-ORIENTED SYSTEM THAT PRIORITIZES SECURITY RECOMMENDATIONS ACCORDING TO THEIR CONTRIBUTION TO REDUCING THE VULNERABILITY OF THE NETWORK"

In addition, we assume that the risk-oriented system is aware of the vulnerabilities and security recommendations. All assumptions are based on capabilities that Radiflow provides in its visibility and threat detection system - iSID.

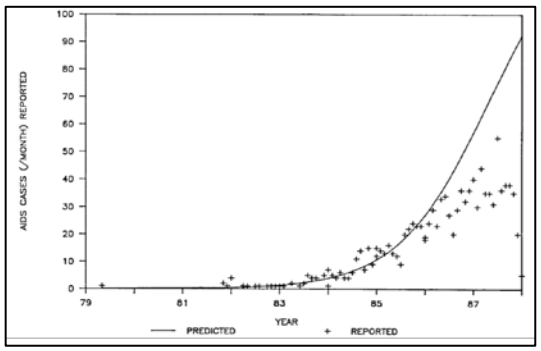
To solve the prioritization problem, we first need to define "network vulnerability" and quantify this value. To achieve this, models used in epidemic research to measure a population's vulnerability to a biological virus must be adopted. The biological model and its basic interpretation to SCADA/ICS networks are explained in the following section.

After introducing this model, we explain various challenges in the basic adoption of the biological model and various methods to overcome those challenges. Finally, we demonstrate how the enhanced real-world example compares with the systems prioritization output and compares that against human risk analysis.

BASIC EPIDEMIC MODEL FOR SCADA/ICS NETWORKS

THE MATHEMATICAL EPIDEMIC MODEL

Given a social network where links represent who has the potential to infect whom – what are we able to predict regarding the potential of this population being infected by a given virus? In other words, what is the infection and recovery time? Can a small infection infect the entire population? In comparison then, what would change if nodes had permanent, temporary or no immunity?

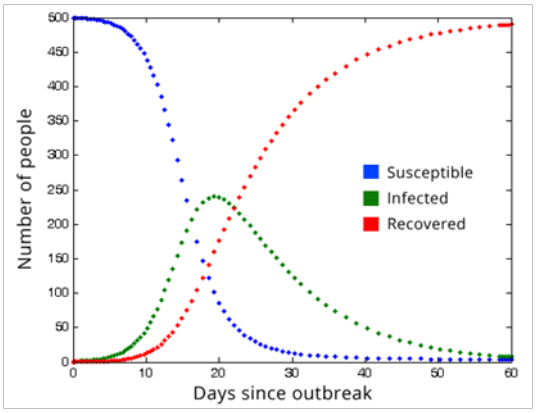


Number of AIDS cases predicted vs. reported. Recent mathematical models were improved and generalized to support more complicated infection and immunization types. Source: Anderson, R.M., 1988, The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. *Journal of Acquired Immune Deficiency Syndromes*, *1*(3), pp.241-256

The issue of viral propagation has a deep research tradition. In early 1985, Anderson et al. presented a model for the infection and spread of the AIDS virus. Although, as an initial model at the time, it was able to provide a basic prediction rate for the

transmission of the AIDS virus in the population (see next figure). During that year more advanced models were developed and in the last decade extensive progress in the area of epidemic modeling has been made.

- Among the many proposed models for viral propagation, two have garnered wide acceptance. The first, the SIS model, considers individuals as being either susceptible (S) or infective (I); a susceptible individual can become infective on contact with another infective individual, then heal herself with some probability to become susceptible again. The second, the SIR model, is similar the only difference being that once healed, an individual is considered removed (R) from the population and immune to further infection. Intuitively, SIS models the flu, while SIR models mumps.



The SIR model assumes recovery over time.

Will a virus be able to infect the population or die out?

In order for this to be the case, the biological models must be defined via an *epidemic threshold*. The epidemic threshold is defined as the minimum level of virulence to prevent a viral contagion from dying out quickly. Research carried out by Prakash et. al.[1] discovered that for any given network there is a mathematical method to determine the epidemic threshold for all virus propagation models (more than 25 models, including HIV.). We will now briefly explain the mathematical method.

Any given undirected graph can be represented by its *adjacency matrix*. The adjacency matrix A of an undirected graph G is a matrix from size N x N, where N is the number of devices in the network. The value of every cell every a_{ij} depends on whether there is a link between device *i* and device *j*. If there is a link between device *i* and device *j*, then $a_{ij} = 1$. If there is no link, then $a_{ij} = 0$.

Mathematical matrixes have a lot of interesting properties that we will not discuss in this whitepaper. However, one interesting property is the matrix '*Eigenvalues*'.

The eigenvalues of a matrix are very important in physics and engineering and their exact value provides useful insight into various common applications, such as stability analysis of systems, oscillations of vibrating systems and much more.

In the example to which we refer, Prakash et. al. discovered that the largest eigenvalue of this matrix, denoted by λ , is the only parameter that determines the epidemic threshold, namely, the epidemic threshold is derived from λ. As this λ increases the network is more vulnerable to viruses.

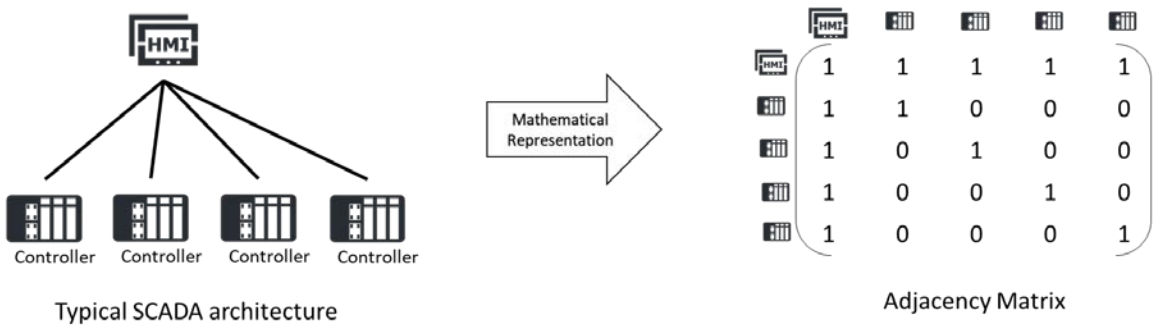
[1] Prakash, B.A., Chakrabarti, D., Valler, N.C., Faloutsos, M. and Faloutsos, C., 2012. Threshold conditions for arbitrary cascade models on arbitrary networks. *Knowledge and information systems*, *33*(3), pp.549-575.

SCADA/ICS NETWORK VULNERABILITY

Unlike software vulnerabilities where a metric exists in the form of the CVSS, there is no metric for network vulnerability or network risk. Although *network vulnerability* and *vulnerable network* are commonly used terms, there is no widely-accepted metric to measure network vulnerability.

In this section, we provide some thought as to why the leading eigenvalue λ is a good representation for measuring network vulnerability.

Given a computer's network topology, we model each host as a node in the graph G and we model all the communication links between devices as edges in G. We calculate the adjacency matrix of G, as well as λ, as we described in the previous paragraphs. See next figure for an example of SCADA/ICS topology, and the corresponding adjacency matrix.

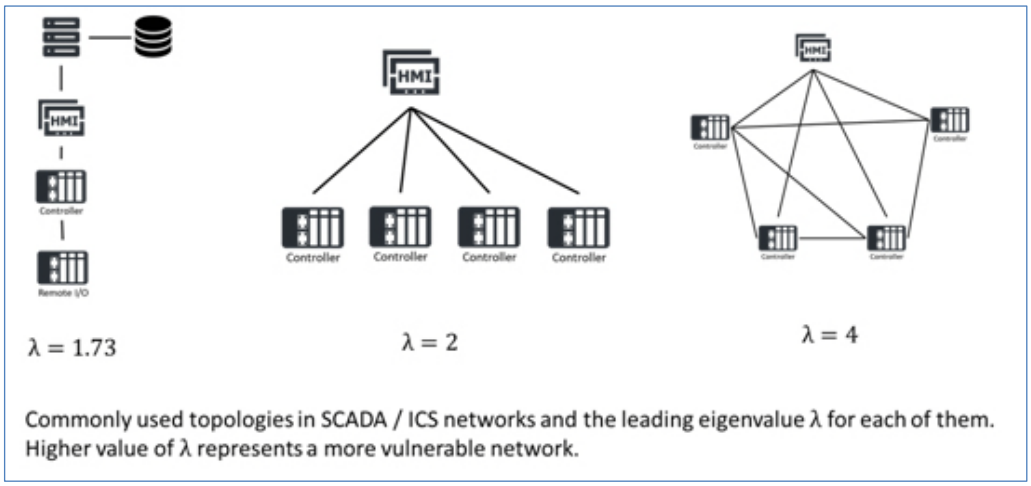


Typical SCADA architecture, and its mathematical representation.

How is λ related to SCADA/ICS networks vulnerability? To explore the interpretation of λ we will use a few examples of typical SCADA/ICS topologies.

Consider the following three networks all with five nodes. Intuitively, the fully-connected graph is the most vulnerable – every node can infect all others. Contradictorily, the chain architecture is the least vulnerable to virus propagation.

This intuition is reflected clearly in the corresponding leading eigenvalue λ of each of the networks. The chain network has a lower λ value than the fully connected network.



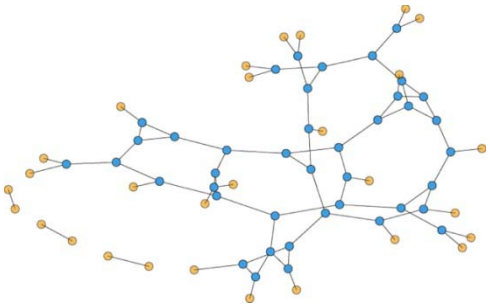
Commonly used topologies in SCADA / ICS networks and the leading eigenvalue λ for each of them. Higher value of λ represents a more vulnerable network.

ENAHNCED MODEL FOR SCADA/ICS SECURITY

As shown above, the epidemic model where λ provides a good alternative for defining a score for network vulnerability. However, there are several challenges in this simplified network model. In the next subsections, we disclose some of the challenges we faced in developing Radiflow's risk evaluator algorithm. That algorithm is focused on risk; however, in this whitepaper, we will mainly describe estimating the vulnerability section.

EDGES MODELING

Computer viruses can spread through existing communication links or create new ones. Even if there is no live communication between two hosts, but the possibility for communication exists, the virus will be able to spread from one host to another. This is contradictory to the epidemic model were edges represent possible infections paths.



Edge-based compartmental modeling for infectious disease spread (Miller, Slim and Volz, 2011, Journal of the Royal Society Interface)

One possible extension is to model edges in the network that include all *possible* communication paths between hosts. In other words, edges will represent the possible routing in the network. By adopting this model, two devices in the same subnetwork that are not communicating, but have the applicable routing to communicate, will have an edge between them.

COMPUTER VIRUSES MAY SPREAD BY EXPLOITING TARGET DEVICE VULNERABILITIES. THEREFORE, SOME OF THE POSSIBLE ROUTES ARE CREATED BY VULNERABILITIES INSIDE DEVICES.

Another extension is to consider device vulnerabilities and not only protocols. Computer viruses may spread by exploiting target device vulnerabilities. Therefore, some of the possible routes are created by vulnerabilities inside devices. Those routes should also be considered.

BENCHMARKING

The epidemic model provides a numerical value. This value can significantly differ between different network topologies. Additionally, different network sizes will have a different range for this value.

To overcome this challenge, we utilize a benchmarking methodology. For any given network, we compare the epidemic threshold value to the values received from similar networks. In our analysis, we compare any given network to more than 200 similar networks.

Using this benchmarking methodology, we provide a value that represents the comparison between the client's epidemic threshold value and the group benchmarked networks. Using the benchmark, we normalized the value to vary from 0 to 100. The lowest score reflects the client's network and has a lower risk compared to the benchmark.

DEVICE VULNERABILITY

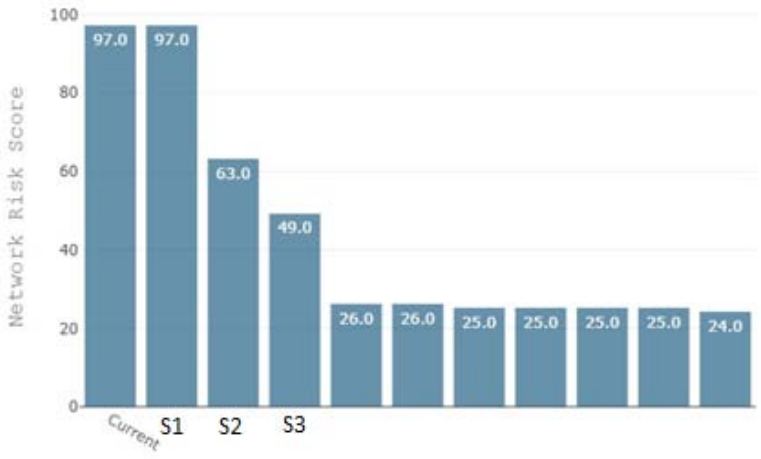
Device vulnerability is another term that does not have a widely accepted definition. In our model, we define device vulnerability as the *contribution of this device to the overall risk of the network*. To evaluate this value, we compare the network risk of two cases: (1) case where the device is fully patched and hardened and (2) the current network risk with the current hardening/patching level of the device. The difference in the network risk is defined as the device risk.

Based on this definition, we can now apply the metric to all networked devices and prioritize the devices according to their contribution to the overall risk. Devices that contribute more are those that should receive greater attention and more frequent patches.

Moreover, using this device vulnerability metric we can now automatically build a hardening and mitigation plan. The algorithm can now look for a group of devices that will mostly reduce the network risk. We called this analysis a *hardening plan*.

REAL-WORLD CASE STUDY

This case study describes an assessment carried out within a Building Management System (BMS). The monitored network included more than 100 assets and around 1000 communication links. In the assessment, we leveraged our risk algorithm to prioritize the mitigation activities and efficiently reduce the network risks to a minimum.



The calculated risk on the network was very high – 97.

The assessment report also described the main cause for this high risk: many assets contained end-of-life software, limited or no segmentation, insecure protocols and more. Fixing all of these issues would have been lengthy in terms of time and effort. From a human analyst perspective, all of these issues were critical and the client almost got into a long and costly upgrade process.

By executing Radiflow's algorithm, we were able to determine that the main risks on the network were caused by three devices. Those three devices were three SCADA servers. Each of the three connected to multiple PLCs, while there were very few connections between the different SCADA servers. Our algorithm was able to detect those sensitive devices automatically. The automatic algorithm reached the same conclusion as the extensive manual analysis performed by an expert analyst.

During the next step, the algorithm continued to prioritize the recommendations and build a mitigation plan for the CISO.

This graph shows the optimal order of actions that should be taken, starting with the current risk score of 97. The algorithm suggested that the starting block lies with the hardening and segregation of SCADA Server S1. According to the algorithm, this action will not make any significant change to the network risk, but the next steps show why it is essential; this is the optimal path for reducing the risk the lowest level.

When we analyzed the result of the algorithm, we discovered that S1 SCADA server for process A was a sperate physical process than the one controlled by S2 and S3 – process B. There was an unwanted connection, due to low segmentation, between those two business processes. As a result. even after hardening S1, the network risk did not change. In addition, S2 and S3 were redundant SCADA servers, while S1 remained as the only SCADA server controlling process A. The algorithm was able to detect this redundancy and to recommend hardening S1 first.

THIS RESULT ALLOWS THE CISO TO FOCUS ON FIVE RECOMMENDATIONS RELATED TO THOSE DEVICES FROM A TOTAL OF MORE THAN TWENTY RECOMMENDATIONS RELATED TO THE ENTIRE NETWORK.

Additional insight from the results showed that S1, S2 and S3 were the main contributors to network risk. After performing the relevant actions on those three devices, the network reached a significantly lower risk score of around 26. Reducing the risk even further below this low score requires significant effort, but would not necessarily result in a decrease in the overall network risk. This result allows the CISO to focus on five recommendations related to those devices from a total of more than twenty recommendations related to the entire network.

CONCLUSION

In this whitepaper, we presented several biological models that evaluated network vulnerabilities, all of which can be compared to a biological virus.

In conclusion, we believe that a risk-orientated system with the ability to automatically prioritize risk mitigation measures should be standard business practice. The increase of vulnerabilities found, and the various security practices developed, proved that it should be mandatory to have a single system that aggregates all of this information. A single system that ultimately supports the decisions of those security officers challenged with making those critical decisions and limits the impact on business operations.