The Measures of Cyber Team Performance

Geoffrey B. Dobson

Carnegie Mellon University, Pittsburgh PA 15213, USA
gdobson@andrew.cmu.edu

Abstract. This paper defines a set of cyber team performance measures that can lead to a comprehensive understanding of how cyber teams perform. As cyber teams have matured in recent years, researchers are beginning to study how cyber teams perform. A virtual experiment based method to computationally model cyber team performance is demonstrated. In this work, the Cyber-FIT modeling and simulation framework is updated to support virtual experiments that can assist in developing more cyber team performance measures.

Keywords: agent-based modeling, simulation, cyber, military.

1 Introduction

On May 3, 2019, the white house published [1] Executive Order number 13870 titled "America's Cybersecurity Workforce". The executive order requires, among other things, various federal agencies to provide a plan on how hold a large cybersecurity competition that will challenge and reward the best individuals and teams of the nation. The executive order also directs the agencies to define "the parameters for the competition, including the development of multiple individual and team events that test cybersecurity skills". Furthermore, the order requires that the competition incorporate the skills already identified in the National Institute for Standards and Technology's National Initiative for Cybersecurity Education (NICE) Framework [2]. The formalization of events and clearly defined measures that prove which cyber teams and individuals are best, will be no small task. The NICE Framework approaches this issue by defining cyber work roles that have associated tasks, knowledge, skills, and abilities. For example, the cyber defense analyst work role has thirty-four tasks, seventy required knowledge items, fifteen skills, and six abilities. Creating a scoreboard, for the aforementioned cybersecurity competition, will take significant effort.

In this working paper, I will use an existing agent-based modeling and simulation framework to define twenty-one cyber team performance measures. Each of the measures can be simulated and observed programmatically. One virtual experiment is run as a proof of concept that simulates a cyber team operation and collects several performance measures.

2 Background

Assessing the performance of teams is difficult, in almost any endeavor. Even in sporting events, experts are regularly surprised when seemingly great teams are upset by lesser teams. As cyber operations mature, the need to delineate performance becomes more pressing. This problem is not unique to the military, it is present across industry in general. In order for an organization to recruit talent, it must know what types of talent it needs. To retain talent, organizations must reward performance, which has yet to be universally defined in cybersecurity. To design an organization efficiently, cyber work roles and performance outcomes should be clearly delineated.

In order to make an initial attempt at defining cyber team performance measures, the focus of this working paper will be on military cyber team operations. That way, the operations can be based on already established doctrine and the scope of measures will be limited. Militaries are especially concerned with the notion of team (or unit, squad, platoon, etc.) assessment. Most military units are required to formally assess their performance, sometimes referred to as readiness, against their assigned tasks. Usually, to do that, the team will engage in a military exercise. In October of 2016, United States Cyber Command announced [3] that all 133 cyber protection teams achieved "initial operating capability". This means that the teams were assessed, to some level of satisfaction, and their performance was deemed capable. What is not clear is how well the teams performed on said tasks, or which teams performed better than others. Military planners understand well the need to have better definitions of cyber team performance. There is existing work in this field. Henshel et al designed [4] an assessment strategy at a large military cyber exercise that is meant to evolve over time as systems are developed to track performance data automatically. They defined four performance measures that were collected throughout the exercise manually by observers and concluded that systems should be designed that support performance data collection for future events. By pre-defining all possible metrics of cyber teams would allow exercise planners to properly instrument the range for collection purposes. In similar work, Magdalena and Andersson collected [5] both subject and objective data from a large NATO cyber exercise. Some of the objective data included metrics that are similar to the Hershel et al paper. They were able to create one automated scoring criteria (service availability), but still relied on a number of manual scoring measures such as correct incident reports and number of vulnerabilities removed. Correct incident reports would be very difficult to automate, but vulnerability removal is something that can be programmatically implemented. This working paper is attempting to assist in the automation of such performance scores.

3 The Measures of Cyber Team Performance

The measures of cyber team performance could be grouped in many ways. In this working paper I will group by: defending team measures, defending individual measures, and attacking team measures. In the current state of the art, cyber defensive operations are much more clearly defined than the attacking operations, which leads to

attacking team measures being included in this work but not attacking individual measures. Table 1 lists all measures and provides a definition of each measure.

Table 1. Proposed Cyber Team and Individual Performance Measures

Defending Team Measures	
Measure	Explanation
Time to react	Time to observe and log new vulnerability, indicator of compromise, or exploit
Time to restore	Time to restore unavailable system
Time to plan	Time to complete mission planning phase
Time to survey	Time to complete survey phase
Time to secure	Time to complete secure phase
Cyber situational awareness	Total cyber situational awareness of team as it relates to terrain status, prioritizations of activities, and awareness of what teammates are working on
Operational balance	Temporal based view showing percentage of time all team members are operating on various tasks
Communication balance	Temporal based view showing types of communication occurring amongst team members
Planning efficiency	The ratio of planned actions versus unplanned actions
Terrain vulnerability level	Total vulnerability of all assigned cyber terrain
Terrain vulnerability change	Change in vulnerability since beginning operations
Terrain compromises	Total number of compromised terrain
Terrain compromise change	Change in compromised terrain since beginning operations
Terrain compromise time	Total time assigned terrain is in compromised state
Cyber mission capability rate	Ratio of mission supported terrain system availability versus unavailability
Defending Individual Measures	
Measure	Explanation
Operational success rate	Ratio of operations taken to remove vulnerabilities, update systems, restore compromises, etc. that are successful versus unsuccessful
Operational efficiency	Ratio of operational activity that is correct according to the plan, indicator of

	compromises, or terrain status versus activity that is wasteful
Communication efficiency	Ratio of communicating actions taken that further the mission versus those that are miscellaneous or detrimental to mission operations
Cyber situational awareness	Awareness of cyber terrain system vulnerability level and availability
Attacking Team Measures	
Measure	Explanation
Time to breach	Time to access cyber terrain that is unauthorized
Time to deliver	Time to deliver attack or malware payload to system
Time to exploit	Time for exploit to succeed once triggered
Exploit success rate	Ratio of successful versus unsuccessful exploits
Stealth rate	Ratio of time unnoticed versus total time operating within unauthorized cyber terrain

4 Cyber-FIT Simulation Framework

The Cyber-FIT simulation framework [6] is an agent-based modeling tool that provides a mechanism to run virtual experiments that test assumptions about the deployment of cyber forces. The current version was built using Repast Symphony in the Java programming language. The most recent version of the framework includes attacker (adversary) agent tier levels based on the United States Department of Defense Scientific Advisory Board report. The report defines [7] six tiers of adversary, ranging from least to most sophisticated. The Cyber-FIT code base implements the tier levels by altering the success rate of exploitation attempts as well as the difficult of the attacks that the adversary has available.

5 Virtual Experiments

One virtual experiment was run to test the model and collect simulated cyber team performance data. Table two describes the virtual experiment.

Table 2. Virtual Experiment Design

Independent Variables		
IV	Variants	Values
Defender Agents	1	10
Terrain Agents	1	211
Attacker Agents	1	2
Attacker Agent Tier	2	[2,5]
Dependent Variables		
DV	Type	
Terrain Compromise Time	Integer	
This experiment will be 2X10 runs = 20 replications		

The virtual experiment was run using Repast Symphony and Figure one shows the results of the experiment.

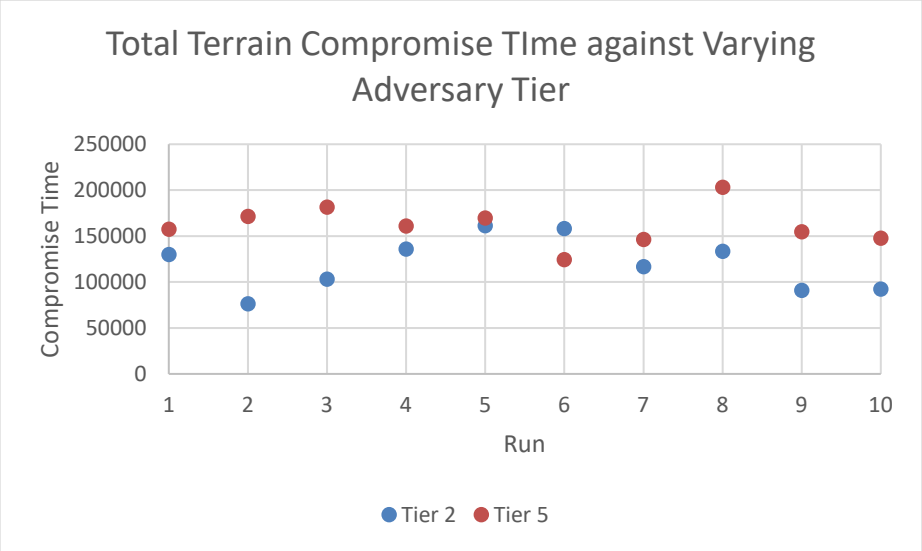

Figure 1. Results of virtual experiment

As shown in the scatter plot, a single cyber team of ten members should expect to see more system compromises, when facing a tier five adversary (high sophistication), over the course of an operation than a tier two adversary (low sophistication). Over all runs of the virtual experiment, the cyber team saw average terrain compromise time of 119,936 minutes against a tier two adversary and 161,902 minutes against a tier five adversary. The total simulated time was a five day operation (7,200 minutes). With 211 systems to defend, there are 1,519,200 total terrain minutes. Therefore, the tier two adversary realized a .08 terrain compromise rate, while the tier five adversary realized a .11 terrain compromise rate. These results show that the model is working, in some ways, as expected. A tier five adversary should outperform a tier two adversary, on average, against a defending cyber team. However, should the tier five adversary only

outperform a tier two adversary by 37.5 percent? This is an open question with many complex design decision that must be addressed.

6 Conclusion

This working paper described an initial set of measures that can be used to assess cyber team performance. The goal of this work is to move closer to a comprehensive list of all possible cyber team performance measures. As measures are developed, opportunities to collect empirical data will be sought out through cyber exercises and, where available, operational logs.

References

1. The White House, "Executive Order on America's Cybesecurity Workforce," 2 May 2019. [Online]. Available: https://www.whitehouse.gov/presidential-actions/executive-order-americas-cybersecurity-workforce/. [Accessed 3 May 2019]
2. W. Newhosue, S. Keith, B. Scribner and G. Witte, "National Initiative for Cybersecurity Education (NICE) Framework," National Institute for Standards and Technology, Washington D.C., 2017.
3. U.S. Department of Defense, "All Cyber Mission Force Teams Achieve Initial Operating Capability," 24 October 2016. [Online]. Available: https://dod.defense.gov/News/Article/Article/984663/all-cyber-mission-force-teams-achieve-initial-operating-capability/. [Accessed 14 May 2019].
4. D. S. Henshel, G. M. Deckard, B. Lufkin, N. Buchler, B. Hoffman, P. Rajivan and S. Collman, "Predicting proficiency in cyber defense team exercises," in *IEEE Military Communications Conference*, Baltimore, 2016.
5. M. Granåsen and D. Andersson, "Measuring team effectiveness in cyber-defense exercises: a cross-disciplinary case study," *Cognition, Technology & Work ,* vol. 18, no. 1, pp. 121-143, 2016.
6. G. B. Dobson and K. M. Carley, "Cyber-FIT: an agent-based modelling approach to simulating cyber warfare," in *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, Washington D.C., 2017.
7. U.S. Department of Defense Defense Science Board, "Resilient Military Systems and the Advanced Cyber Threat," 2013.