ReSCIND Performer HSR Dataset Cover Sheet
Guarding Against Malicious Biased Threats (GAMBiT) HSR2

ReSCIND

Dataset Details

Dataset Title:	Guarding Against Malicious Biased Threats (GAMBiT) HSR2		
Dataset Citation:	GAMBiT HSR2 Dataset Citation TBD		
Data Format:	Available on S3 bucket, zip files	Data Size:	2.1 TB
Dates & Duration:	November 9, 2024 – January 29, 2025 Two 8-hour days per participant	Time Zone:	EST/EDT
How to access dataset:	Rachelle Thomas rthomas@bullsrungroup.com		
Point of contact for data questions:	Peggy Wu Peggy.Wu@rtx.com		

Description of Scenario

Objectives

This experiment observed how skilled attackers behaved during a cyber-attack scenario. The goal was to collect detailed behavioral data to help researchers develop new ways to classify attacker actions and decision-making patterns.

Experiment Description

Over two days, 20 red team participants were given access to a simulated enterprise network (a "cyber range") and instructed to conduct self-paced cyberattacks. In the beginning, participants received operational instructions and credentials for initial network access. From there, they pursued realistic objectives—such as identifying sensitive systems and exfiltrating valuable data—at their own pace. Participants also completed periodic surveys and maintained written notes to document their reasoning and tactical choices throughout the exercise.

Experimental Results

Analysis of cyber data, skills tests, self-reports, and operational notes found that higher-skilled individuals made more progress in cyber-attacks.

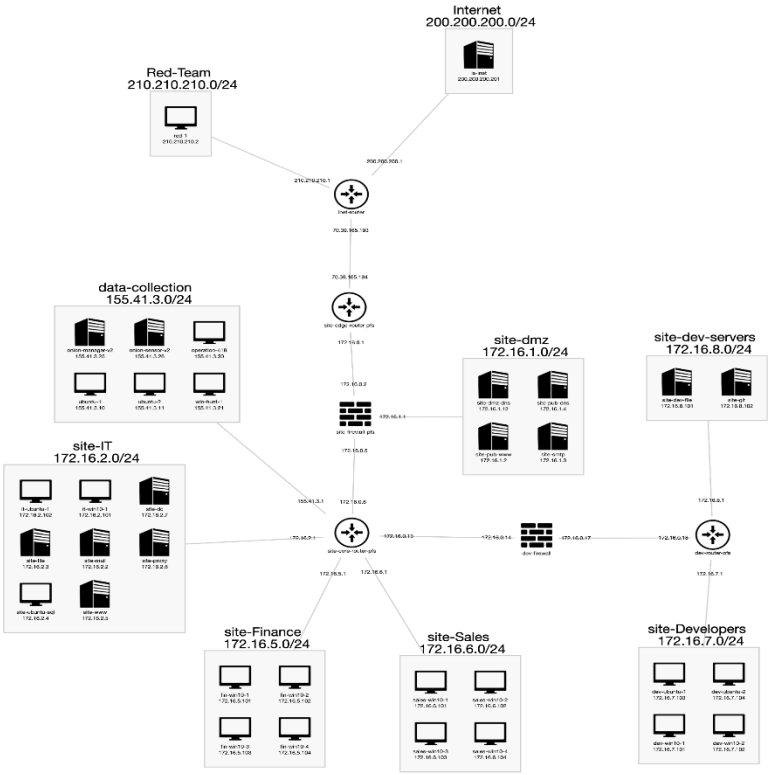
Cyber Environment

Experiment 2 used the SimSpace Cyber Force Platform to design and implement the GAMBiT cyber range, which simulated an enterprise business information system. This cyber environment comprised approximately 40 virtual devices organized into subnetworks, incorporating routers, switches, and user traffic commonly found in operational networks. Figure 1 illustrates the most recent network topology of the GAMBiT cyber range. Each participant operated within a designated cyber range and initiated all challenges using a virtual machine running Kali Linux.

The starting box was 10.10.0.5. Each participant had identical IP addresses within their assigned range, ensuring consistency across individual environments. The network included restricted subnets designated for managing the environment during the engagement, which were classified as no-strike targets. The experiment environment consisted of one range with seven triggers.

Image 1: Network topology for GAMBiT cyber range for Experiment 2, for reference only. Detailed network diagram and details about each host is included with the full dataset.

The following subnets were strictly off-limits and not to be scanned or accessed:



- 10.10.0.0/16
- 155.41.3.0/24
- 192.168.0.0/21
- 172.16.100.0/22
- 3.136.223.108

DATA

DATA SOURCES

Primary Data Sources

These data were collected directly from the cyber range experiment environment.

Category	Data Source	Examples of Select Data Features
Self Reports/Background Data *	Screening Questionnaire	Years of experience, type of cyber experience, team size, preferred OS, length of campaign.
	Demographics Questionnaire	Age, gender, native language, education level
Self Reports/Psychometric Data *	Cognitive Reflection Test (CRT)	3 multiple-choice items to assess the ability to override an intuitive but incorrect response and engage in more deliberate, analytical thinking. (Frederick, S. 2005)
	Big Five Inventory extra-short form (BFI-2)	15 items to indicate personality (Soto & John, 2017)
	General Risk Propensity Scale (GRiPS)	8 items on tendency to risky behaviors (Zhang et al., 2019)
	Adult Decision-Making Competence Scale (A-DMC)	Resistance to Framing Positive (7 items) & Negative (7 items) and Resistance to Sunk Cost (10 items) (Bruine de Bruin et al., 2007)
Self Reports/ Questionnaires	Applied Techniques	Hourly Stage reports (X.1-X.3): intended/applied MITRE ATT&CK techniques
	Reasoning and Affect Changes	Hourly Stage reports (X.1-X.3): 5-point Likert scale items on reasoning (6 items) and mood (5 items) changes
	OPNOTES	CherryTree file with Operation Notes from participant.

Network Data	PCAP	Timestamps, source & destination packets & protocols, payloads
Network Data Kali Host Data	NIDS - Suricata	Monitors network traffic for suspicious activities based on predefined rules
	Keylog	Keylogger where each line records an individual keystroke. Particularly useful as it collects text that participants copy to their clipboard.
	Terminal histories	Bash and zsh histories, timestamps, order of commands

**Provide a citation for each psychometric assessment in the References section below.*

Derivative Data Sets

These datasets were created from aggregating, analyzing, curating, and labeling the source data.

Category	Data Source	Examples of Select Data Features
	Clean Log	Keylogger where each line records an individual keystroke. Particularly useful as it collects text that participants copy to their clipboard
		Result of running a post-processing script on Admin VM to remove certain keystrokes for better readability.

RESEARCH

Hypotheses

The GAMBiT HSR2 dataset was used to answer the following hypotheses:

[H1] Expert participants will perform better than open-division cyber attackers. Experts will stay on the attack path longer than open-division attackers, and will make more progress along the attack path.

Publications

1. Shuo Huang, Frederick Jones, Nikolos Gurney, David Pynadath, Kunal Srivastava, Stoney Trent, Peggy Wu, and Quanyan Zhu (2024). PsybORG+: Modeling and Simulation for Detecting Cognitive Biases in Advanced Persistent Threats. In proceedings of 2024 IEEE Military Communications Conference (MILCOM). Washington, DC. Oct 28 to Nov 1, 2024.

References

1. Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. Journal of personality and social psychology, 92(5), 938.
2. Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic perspectives, 19(4), 25-42.
3. Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. Journal of Research in Personality, 68, 69-81.
4. Zhang, D. C., Highhouse, S., & Nye, C. D. (2019). Development and validation of the general risk propensity scale (GRiPS). Journal of Behavioral Decision Making, 32(2), 152-167