Date: 28 September 2018

To: Romcholo Macatula

From: Cyber Warriors

Subject: Technical Memo 4: Evaluating Competing Methods to Identify an Optimal Solution

1 Summary

In this technical memo, we shall select our solution strategies through quantitative criteria ranking and scoring matrices. For quantitative criteria ranking we assign weights to each of our criterion for both data collection and mathematical modeling. For our data collection component we ranked *Ease* with the highest weight (0.25) *Data Format* with the lowest weight (0.05). For our mathematical modeling component we ranked *Accuracy, Interpretation,* and *Ease* with the highest weight (0.20) and *Language* with the lowest weight (0.05).

After completion of the scoring matrices with the criteria from the quantitative ranking, we found the optimal solution strategy for our data collection component as synthetic data. We found the best solution strategy for our mathematical modeling component as Markov Chains.

2 Quantitative Criteria Ranking and Scoring Matrices

In this section, we shall assign weights to each of the criterion for each of our project components. We shall utilize the criterion weights in our scoring matrices to determine the optimal solution strategy for each component.

2.1 Quantitative Ranking of Criteria

For each criterion, we assign a weight, w_j for the jth criterion, with $w_j \in [0, 1]$. All of the weights sum to one: for n criteria [1],

$$\sum_{j=1}^{n} w_j = 1.$$

Data Collection:

criterion	weight
Number of data sources	.10
Data Format	.05
Access	.20
Accuracy	.20
Integration	.20
Ease	.25

Mathematical Modeling:

criterion	weight
Language	.05
Speed	.10
Accuracy	.20
Factors	.10
Ease	.20
Interpretation	.20
Visualization	.10
Scope	.15

2.2 Scoring Matrices

Next, we set up a table with our various solution strategies as labels of the columns, and the criteria to label the rows. Each solution strategy yields two sub-columns, as depicted below.

For each criterion, we select one solution strategy that satisfies the criterion in an "average" way, and place a "3" in the respective column.

Now, we rank the rest of the solutions on a scale of 1–5 (1: poorly satisfies the criterion; 5: excellently satisfies the criterion).

To complete the scoring matrix, we multiply each score by the weight for each criterion, and sum across the row [1].

Data Collection:

		Synthetic data		Free online sources	
Number of data sources	.10	4	0.40	2	0.20
Data Format	.05	4	0.20	2	0.10
Access	.20	4	0.80	2	0.40
Accuracy	.20	2	0.40	3	0.60
Integration	.20	4	0.40	3	0.60
Ease	.25	4	1.00	2	0.50
	score		3.20		2.40

We create another table for our second component.

Mathematical Modeling:

		Network of Servers		Markov Chain		Network of Users		Regression Analysis	
Language	.05	3	0.15	3	0.15	3	0.15	3	.015
Speed	.10	3	0.30	3	0.30	3	0.30	3	0.30
Accuracy	.20	3	0.60	4	0.80	3	0.60	2	0.40
Factors	.10	2	0.20	5	0.50	2	0.20	2	0.20
Ease	.20	1	0.20	3	0.60	1	0.20	4	0.80
Interpretation	.20	3	0.60	3	0.60	3	0.60	3	0.60
Visualization	.10	3	0.30	3	0.30	3	0.30	3	0.30
Scope	.15	3	0.45	3	0.45	3	0.45	3	0.45
score			2.80		3.70		2.80		3.20

3 Solution Strategy

The team shall utilize synthetic data for our data collection component. By web scrapping through online job applications the team aims to collect useful data pertaining to specific software a company uses. From this data the team will infer, from readily available online vulnerabilities, what type of software a company uses that has existing vulnerabilities. The team will also create virtual networks and simulate, by hand, how an attacker would gain access to higher level users. The team will research most commonly used attacks to infiltrate a network and use those as factors in the teams mathematical model.

The team shall utilize Markov Chains for our mathematical model. Contingent on the data collected the team will generate probabilistic weights that justify how an attacker navigates from node to node. Each node within the Markov Chain shall represent a designated user. Each user shall have respective factors that relate to their level of security. For instance, a user might have admin access with an IDS and firewall to prevent against attacks. That same user might have a strong password and only use a specific set of software. All these factors will way in to the rate at which an attacker can move from node to node.

References

[1] M. Embre, *Technical memo 4: Evaluating competing methods to identify an optimal solution*, tech. rep., Virginia Tech, 2018.