This document outlines a probabilistic framework for modeling and fitting parameters associated with types of attack steps in a discrete-choice scenario, as might be used in cybersecurity or decision-making simulations.

1. Problem Setup

We start with a set of *attack steps*:

$$S = \{s_1, \ldots, s_n\}.$$

Each attack step $s \in S$ is associated with a *type* drawn from a finite set:

$$T = \{1, \ldots, m\}.$$

A function $t : S \to T$ maps each attack step to its type.

Example

If S represents specific cyberattack techniques (e.g., phishing, malware injection), the set T might represent broader categories of attacks (e.g., social engineering, software exploitation).

2. Parameterization

Each type $j \in T$ has a scalar parameter $\theta_j \in R$, collected into a parameter set:

$$\Theta = \{\theta_1, \ldots, \theta_m\}.$$

These parameters control the likelihood of selecting an attack step based on its type.

To determine probabilities, we use an *exponential weight function*:

$$p_s(\Theta) = \exp\big(\theta_{t(s)}\big).$$

This ensures positivity (important for probabilities) and scales naturally with θ_j.

3. Selection Probabilities in a Horizon

An *attack horizon* $H \subseteq S$ represents a subset of attack steps under consideration at a given time. The probability of selecting an attack step $s \in H$ is:

$$P_\Theta(s \mid H) = \frac{p_s(\Theta)}{\sum_{u \in H} p_u(\Theta)}.$$

This uses a softmax function over the weights to normalize the probabilities within the horizon.

Intuition

If $t(s)$ has a high $\theta_{t(s)}$, the probability of selecting s increases relative to other steps in H. Conversely, a low $\theta_{t(s)}$ reduces its likelihood.

4. Learning the Parameters

Given a sample of attack horizons:

$$\mathcal{H} = \{H_1, \ldots, H_K\},$$

and observed selections $\sigma(H_k) \in H_k$ for each horizon H_k, we aim to fit the parameters Θ so that the model's predicted probabilities match these observations. This is achieved by maximizing the likelihood:

$$\max_\Theta \prod_{k=1}^{K} P_\Theta(\sigma(H_k) \mid H_k).$$

Equivalently, we maximize the log-likelihood for computational convenience:

$$\max_\Theta \sum_{k=1}^{K} \ln P_\Theta(\sigma(H_k) \mid H_k).$$

5. Why Use Exponentials (Softmax)?

The exponential function:

- Ensures all probabilities are positive.
- Automatically normalizes probabilities within the horizon.
- Is computationally efficient and widely used in machine learning (e.g., softmax layers in neural networks).

6. Applications

This framework can be applied to problems where:

- A large sample of choices (attack steps) is grouped into subsets (horizons).
- Probabilities depend on intrinsic properties (types) of each item.
- Observed selections inform the estimation of these intrinsic properties.

7. Numerical Example

We illustrate the framework with a simple case of two attack-step types:

$$T = \{1, 2\}, \quad S = \{s_1, s_2\},$$

where $t(s_1) = 1$ and $t(s_2) = 2$. Thus,

$$\Theta = \{\theta_1, \theta_2\}, \quad p_{s_1}(\Theta) = \exp(\theta_1), \quad p_{s_2}(\Theta) = \exp(\theta_2).$$

7.1 Observed Horizons and Selections

Suppose we observe five horizons:

$$H_1 = H_2 = H_3 = H_4 = \{s_1, s_2\}, \quad H_5 = \{s_1\}.$$

The selections are:

$$\sigma(H_1) = s_1, \quad \sigma(H_2) = s_1, \quad \sigma(H_3) = s_1, \quad \sigma(H_4) = s_2, \quad \sigma(H_5) = s_1.$$

Hence, in the first four horizons (each containing both s_1 and s_2), we see s_1 selected three times and s_2 once. In the fifth horizon, H_5, only s_1 is available, so it is selected with certainty.

7.2 Probability Computation

For any horizon containing both $\{s_1, s_2\}$, the softmax probabilities are

$$P_\Theta(s_1 \mid \{s_1, s_2\}) = \frac{\exp(\theta_1)}{\exp(\theta_1) + \exp(\theta_2)}, \quad P_\Theta(s_2 \mid \{s_1, s_2\}) = \frac{\exp(\theta_2)}{\exp(\theta_1) + \exp(\theta_2)}.$$

For the horizon $H_5 = \{s_1\}$, the probability of selecting s_1 is trivially

$$P_\Theta(s_1 \mid \{s_1\}) = 1.$$

7.3 Likelihood and Parameter Fitting

The overall likelihood is the product of the probabilities for each selection:

$$L(\Theta) = P_\Theta(s_1 \mid \{s_1, s_2\})^3 \times P_\Theta(s_2 \mid \{s_1, s_2\}) \times P_\Theta(s_1 \mid \{s_1\}).$$

Because $P_\Theta(s_1 \mid \{s_1\}) = 1$, it does not affect parameter estimation. So the key part is

$$L(\Theta) = \left(\frac{\exp(\theta_1)}{\exp(\theta_1) + \exp(\theta_2)}\right)^3 \times \left(\frac{\exp(\theta_2)}{\exp(\theta_1) + \exp(\theta_2)}\right).$$

Taking logs,

$$\ln L(\Theta) = 3\theta_1 + \theta_2 - 4\,\ln\big(\exp(\theta_1) + \exp(\theta_2)\big).$$

Maximizing this with respect to θ_1 and θ_2 amounts to matching the ratio $\exp(\theta_1) : \exp(\theta_2)$ to the observed frequency (3:1) from the horizons that offered both s_1 and s_2.

7.4 Maximum Likelihood Estimate

Let $\alpha = \exp(\theta_1)$ and $\beta = \exp(\theta_2)$. Because s_1 is selected 3/4 of the time when both options are present, we have

$$\frac{\alpha}{\alpha + \beta} \approx 0.75,$$

implying $\alpha : \beta = 3 : 1$. One straightforward solution is

$$\alpha = 3, \quad \beta = 1 \implies \theta_1 = \ln(3), \quad \theta_2 = 0.$$

This yields

$$P_\Theta(s_1 \mid \{s_1, s_2\}) = 0.75, \quad P_\Theta(s_2 \mid \{s_1, s_2\}) = 0.25, \quad P_\Theta(s_1 \mid \{s_1\}) = 1.$$

Any constant shift added to both θ_1 and θ_2 leaves these probabilities unchanged, so the essential feature is $\exp(\theta_1)$ being three times $\exp(\theta_2)$.

Interpretation: Despite the extra horizon containing only s_1, the likelihood for horizons that offer both s_1 and s_2 drives the ratio of $\exp(\theta_1)$ to $\exp(\theta_2)$. The single-element horizon H_5 contributes a probability of 1 to the selection of s_1, so it does not affect the balance between θ_1 and θ_2.

Concise Formalism

1. **Attack Steps:**
$$S = \{s_1, \ldots, s_n\}.$$
2. **Attack Step Types:**
$$T = \{1, \ldots, m\}, \quad t : S \to T.$$
3. **Parameters:** Each type $j \in T$ has a scalar parameter θ_j. Collect these in
$$\Theta = \{\theta_1, \ldots, \theta_m\}.$$
4. **Probability Weight Function:**
$$p_s(\Theta) \;=\; \exp\big(\theta_{t(s)}\big).$$
5. **Selection Probability in a Horizon $H \subseteq S$:**
$$P_\Theta(s \mid H) \;=\; \frac{p_s(\Theta)}{\sum_{u \in H} p_u(\Theta)} \quad \textit{for } s \in H.$$
6. **Sample of Attack Horizons:**
$$\mathcal{H} = \{H_1, \ldots, H_K\},$$
with observed selections $\sigma(H_k) \in H_k$.
7. **Likelihood and Fitting:**
$$\max_\Theta \prod_{k=1}^{K} P_\Theta\big(\sigma(H_k) \mid H_k\big) \iff \max_\Theta \sum_{k=1}^{K} \ln P_\Theta\big(\sigma(H_k) \mid H_k\big).$$