

## **The Weapon-Target Assignment Problem [Kline et al., 2019]**

Lorem ipsum 1 ...

## **Applying reinforcement learning to the weapon assignment problem in air defence [Mouton et al., 2011]**

Lorem ipsum 2 ...

## **Optimization of Weapon-Target Pairings Based on Kill Probabilities [Bogdanowicz et al., 2013]**

Lorem ipsum 3 ...

## **A New Approach to Weapon-Target Assignment in Co-operative Air Combat [Chang et al., 2017]**

Lorem ipsum 4 ...

## **A Coordinated Air Defense Learning System Based on Immunized Classifier Systems [Nantogma et al., 2021]**

Multi-agent systems can be applied to defense and security applications, specifically to solve threat evaluation and weapon assignment problems. These applications require complex systems, due to both the complexity of the application domain and incompleteness and uncertainty of perceived information. In addition, many battlefield decisions are time sensitive. All of the above make it difficult to design an effective model. This paper combines an artificial immune system and a learning classifier system to tackle common threat evaluation and weapon assignment problems.

An artificial immune system is a multi-agent, decentralized information processing system that is capable of learning and remembering. It uses computational techniques derived from defense mechanisms native to biological immune systems to accomplish its goals. In an immune system, an antigen is a harmful foreign substance, like a bacterium. B-cells, or lymphocytes, are immune cells that produce antibodies: markers that mark an antigen to be eaten by a T-cell.

- **CLONALG**, one such technique based on clonal selection, uses the idea of immunological memory to implement the learning and remembering component of the artificial immune system. Biological immune systems implement memory by cloning two types of B-cell. One type of B-cell immediately acts to combat infection, while the other type is retained by the immune system and grants immunity to future infection.
- **Danger theory** is another technique based on the fact that immune systems only respond to harmful foreign substances. Similarly, in an immunized air defense system, it's important to determine which threat to prioritize. When making a weapon target assignment, it's important to take the threat level of the targets into account.

Whenever a target, or hostile entity is detected, a danger zone is constructed, based on the trajectory of the detected target. The value of the combat units inside this danger zone signals how much priority should be given in intercepting the target.

- **Network theory**: In an immune system, the antibodies of many different B-cells, of the same and different type, are able to recognize and interact with each other, even without the presence of foreign substances. When one antibody recognizes the other, B-cells will produce different antibodies meant to suppress the recognized antibody, while the antibody that did the recognizing will get a boost in production. The resulting network of antibody interactions serves to broaden the diversity of the existing antibody population.

A learning classifier system is a decision making system that makes use of simple if-then rules, or classifiers. As it executes, the classifiers are slowly improved. In operation, it consists of three phases: performance, reinforcement, and discovery, in that order.

- **Performance**: In the performance phase, actions are selected and executed using an existing batch of classifiers. The classifiers that have that specific action are taken and stored in an action set. The executed action has an associated reward.
- **Reinforcement**: In the reinforcement phase, the parameters in the classifiers are updated using the reward from the performance phase.

The update equations for this phase for a classifier  $cl$ , are as follows:

– **Prediction**  $cl.p$ :

$$cl.p = cl.p + \beta(R - cl.p) \quad (1)$$

– **Prediction Error**  $cl.\epsilon$ :

$$cl.\epsilon = cl.\epsilon + \beta(|R - cl.p| - cl.\epsilon) \quad (2)$$

– **Fitness**  $cl.F$ :

$$cl.F = cl.F + \beta(\hat{\lambda}(cl) - cl.F) \quad (3)$$

– **Accuracy**  $\lambda(cl)$ :

$$\lambda(cl) = \begin{cases} 1 & \text{if } cl.\epsilon < \epsilon_0 \\ \alpha(\frac{cl.\epsilon}{\epsilon_0})^{-v} & \text{if } cl.\epsilon \geq \epsilon_0 \end{cases} \quad (4)$$

Where  $\epsilon_0$  is an accuracy criterion constant. A classifier is accurate if  $cl.\epsilon$  is smaller than  $\epsilon_0$ .  $\alpha$  and  $v$  are hyper-parameters used to control the rate at which the accuracy reduces.

– **Relative accuracy  $\hat{\lambda}(cl)$ :**

$$\hat{\lambda}(cl) = \frac{cl.n \times \lambda(cl)}{\sigma_{cl_b \in [A]} \lambda(b) \times b.n} \quad (5)$$

- **Discovery:** In the discovery phase, the algorithm applies a genetic algorithm to generate a new batch of classifiers. Two parents with high fitness are selected from the action set, and offspring are produced using a crossover algorithm and a random bit-flip mutation. Classifiers with low fitness are deleted if the number of total classifiers in the batch get too large, and classifiers whose conditionals are already included in more accurate, experienced classifiers can be deleted.

In a generic air defense mission, the two problems to solve are action strategy selection and cooperation. Thus, we use two agents, one in charge of generating strategies, and one in charge of coordinating defenses. The strategy generation agent uses a learning classifier system to learn the right weapon and amount of ammunition to use against given targets. These classifiers are passed to the strategy coordination agent. Then, the strategy coordination agent uses immune network dynamics to decide what to do, given the classifiers that the strategy generation agent created.

The strategy generation agent treats each weapon type as a B-cell, or decision unit, that each have their own classifiers, or antibodies. Each classifier will have control decisions as their consequent (the THEN component of the IF-THEN statement). Enemy forces are expressed in terms of their intent, capability, and opportunity. When individual B-cells activate on detection of antigens, or hostile units, the strategy generation agent generates a set of classifiers, from which the strategy generation agent selects the most appropriate action. The strategy generation agent interfaces with the environment through an intermediate module, which also provides local detectors for the B-cells.

The strategy coordination agent receives classifiers from the strategy generation agent, establishes connections between the classifiers, forming a network. Then, it applies immune network dynamics. Classifier  $i$  is said to be connected to classifier  $j$  if they produce antibodies for the same target. The concentrations for each classifier is calculated, and the one with the highest concentration ultimately is chosen as the action. The output classifiers from this immune network dynamics process are the ones that undergo genetic operations and updates.

### Specific Equations:

- **Objective Function:** At time  $T$ , for targets detected  $H$  and  $C$  categories of weapons for each agent, with each category  $C_p$  having  $j$  ammunition, the objective function is

$$O_b = \begin{cases} \min & \sum_{j=1}^{H'(t)} \sum_{i=1}^C \beta_i q_{ij} \delta_{ij} \\ \max & \sum_{j=1}^{H'(t)} v_j (1 - \prod_{i=1}^C (1 - p_{ij})^{\delta_{ij}(s)}) \end{cases} \quad (6)$$

Where  $H'$  is the remaining number of unassigned detected targets/enemy units,  $\delta_{ij}$  is the minimum delay before weapon  $i$  can be deployed against target  $j$  based on ready

time and current allocation of that weapon,  $q_{ij}$  is the quantity of ammunition of weapon of type  $i$  allocated to target  $j$ ,  $b_i$  is the unit cost of the ammunition of weapon  $i$ ,  $v_j$  is the threat value of target  $j$ , and  $p_{ij}$  is the weapon kill probability.

- **Expected Action Payoff  $P(a_i)$ :** Given a set of classifiers matching the attributes of a specific B-cell  $[M]$ , the expected action payoff for a specific action  $a_i$  is

$$P(a_i) = \frac{\sum_{cl_k \in [M]|a_i} cl_k.p \times cl_k.F * \omega}{\sum_{cl_k \in [M]|a_i} cl_k.F * \omega} \quad (7)$$

Where  $\omega$  is the affinity between the B-cells' classifiers and antigens. This is the reward that the agent receives for performing a specific action.

- **Classifier-Antigen Affinity  $\omega$ :**

$$\omega = [1 - \prod_{i=1}^q (1 - d_g * p_g * w_e * r_t)] * v_c \quad (8)$$

Where  $q$  is the quantity of ammo suggested by the classifier,  $d_g$  is the normalized distance between target  $c$  and the friendly unit (B-cell),  $p_g$  is the speed advantage of firing the weapon against the target,  $w_e$  is the kill probability of the weapon suggested,  $r_t$  is the ready time of the weapon if it were to be deployed, and  $v_c$  is an estimate for the value of the target.

This value tells you how good a weapon (described in a classifier) is, against a specific antigen.

- **Classifier Concentration  $a_i(t+1)$ :**

$$\frac{da_i(t+1)}{dt} = \left( \alpha \sum_{j=1}^N m_{ji} a_j(t) - \beta \sum_{j=1}^N m_{ik} a_k(t) + \gamma m_i - k \right) a_i(t) \quad (9)$$

Where  $N$  is the number of classifiers that deal with the target antigen,  $m_i$  is the affinity between classifier  $i$  and the target antigen,  $m_{ji}$  is the mutual stimulus coefficient of antibody  $j$  on classifier  $i$ ,  $m_{ki}$  is the inhibitory effect of classifier  $k$  on classifier  $i$ ,  $k$  is the natural death rate of classifier  $i$ ,  $a_i(t)$ ,  $a_j(t)$ , and  $a_k(t)$  are the bounded concentrations that are imposed on the classifiers, and coefficients  $\alpha$ ,  $\beta$ , and  $\gamma$  are weight factors that determine the significance of each term.

This ordinary differential equation embodies the immune network dynamics that this algorithm leverages. Classifiers compete with each other via this concentration value. Since it is an ordinary differential equation, you have to calculate the next step's value using previous values, like any other algorithmic differential equation solver.

# The state-of-the-art review on resource allocation problem using artificial intelligence methods on various computing paradigms [Joloudari et al., 2022]

This is a comprehensive literature study on ML and deep learning methods for the resource allocation problem in different computing environments (distributed, IoT, etc.)

Interesting Resource Allocation articles and their main ideas:

- **Othman and Nayan (2019):**

Review of solutions based on reinforcement and heuristic learning using a dynamic and adaptive allocation resource. Deep learning methods had faster, more accurate convergence.

A. Othman and N. A. Nayan, "Efficient admission control and resource allocation mechanisms for public safety communications over 5G network slice," *Telecommunication Systems*, vol. 72, no. 4, pp. 595-607, 2019.

- **Yousefzai et. al. (2017):**

Investigation of the resource allocation problem in cloud computing. Examination of different schemes based on cloud computing resources using effective features like optimization goals, optimization methods, design approaches, and useful functions.

A. Yousafzai et al., "Cloud resource allocation schemes: review, taxonomy, and opportunities," *Knowledge and Information Systems*, vol. 50, no. 2, pp. 347-381, 2017.

- **Han et. al. (2018):**

Describes an evolutionary reinforcement learning / genetic algorithm that works 90% better than the naive strategy in long term use of a network.

B. Han, J. Lianghai, and H. D. Schotten, "Slice as an evolutionary service: Genetic optimization for inter-slice resource management in 5G networks," *IEEE Access*, vol. 6, pp. 33137-33147, 2018.

- **Bega et. al. (2017):**

Describes

D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, and X. Costa-Perez, "Optimising 5G infrastructure markets: The business of network slicing," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, 2017: IEEE, pp. 1-9.

communications over 5G network slice," *Telecommunication Systems*, vol. 72, no. 4, pp. 595-607, 2019.

## References

- [Bogdanowicz et al., 2013] Bogdanowicz, Z. R., Tolano, A., Patel, K., and Coleman, N. P. (2013). Optimization of weapon-target pairings based on kill probabilities. *IEEE Transactions on Cybernetics*, 43(6):1835–1844.
- [Chang et al., 2017] Chang, Y.-z., wu Li, Z., xin Kou, Y., peng Sun, Q., yan Yang, H., and yan Zhao, Z. (2017). A new approach to weapon-target assignment in cooperative air combat. *Mathematical Problems in Engineering*, 2017.
- [Joloudari et al., 2022] Joloudari, J. H., Mojrian, S., Saadatfar, H., Nodehi, I., Fazl, F., shirkharkolaie, S. K., Alizadehsani, R., Kabir, H. M. D., Tan, R.-S., and Acharya, U. R. (2022). The state-of-the-art review on resource allocation problem using artificial intelligence methods on various computing paradigms.
- [Kline et al., 2019] Kline, A., Ahner, D., and Hill, R. (2019). The weapon-target assignment problem. *Computers & Operations Research*, 105:226–236.
- [Mouton et al., 2011] Mouton, H., Roodt, J., and le Roux, H. (2011). Applying reinforcement learning to the weapon assignment problem in air defence. *Scientia Militaria: South African Journal of Military Studies*, 39:99–116.
- [Nantogma et al., 2021] Nantogma, S., Xu, Y., and Ran, W. (2021). A coordinated air defense learning system based on immunized classifier systems. *Symmetry*, 13(2):271.