Pathogen Dose based Natural Killer Cell Algorithm for Classification

Dongmei Wang
Wuhan University
Wuhan, China
dmwang@whu.edu.cn

Yiwen Liang
Wuhan University
Wuhan, China
ywliang@whu.edu.cn

Chengyu Tan
Wuhan University
Wuhan, China
nadinetan@163.com

Hongbin Dong
Wuhan University
Wuhan, China
hbdong@whu.edu.cn

Xinmin Yang
Wuhan University
Wuhan, China
1565010158@qq.com

ABSTRACT

Negative selection algorithm(NSA), which is the most representative classification algorithm among immune heuristic algorithms, has been successfully applied to solve many classification problems. However, there are two obvious problems with NSA. First, NSA establishes specific antibodies for each antigen based on the mechanism of specific antigen matching antibody, which leads to too many and redundant detectors. Second, the detection stage needs to calculate the matching degree of the antigen with all the detectors, and the detection efficiency is low. This paper proposes a new non-specific natural killer cell algorithm (NKA) based on pathogen dose. NKA draws on the mechanism of NK cells constructing phenotype detectors based on pathogen dose. First, NKA defines dose and phenotype detector, and optimizes it based on the memory evolution mechanism of phenotype; then establishes k-d tree for the optimized phenotype, and pathogens only need to match the dose with the nearest phenotype detector. Experimental results show that the method proposed in this paper, NKA, can not only achieve a better performance through fewer detectors, but also has a higher efficiency in the training and detection phase, compared with three NSA-based algorithms.

CCS CONCEPTS

• **Theory of computation** → *Models of learning*;

KEYWORDS

Innate Immune Memory, NK Cell Model, Artificial Immune System, classification, Negative Selection Algorithm

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
GECCO '21 Companion, July 10–14, 2021, Lille, France
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8351-6/21/07…$15.00
https://doi.org/10.1145/3449726.3459434

ACM Reference Format:
Dongmei Wang, Yiwen Liang, Chengyu Tan, Hongbin Dong, and Xinmin Yang. 2021. Pathogen Dose based Natural Killer Cell Algorithm for Classification. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459434

1 MOTIVATION

Artificial Immune System (AIS) is a type of computational intelligence inspired by the Biology Immune System (BIS), and more precisely, the Human Immune System (HIS). Negative selection algorithm (NSA) was proposed by Forrest [2], and it is one of the most representative algorithms in AIS. NSA classifies pathogens through self (normal) and nonself (abnormal) discrimination mechanisms. It has been successfully applied to various fields, especially classification problems, such as computer intrusion detection, anomaly detection, and fault detection. At the same time, many people have made various improvements to the algorithm to improve the efficiency of the algorithm and expand the application of the algorithm [7].

However, there are two problems with NSA in nature that limit the improvements of NSA :

(1) The algorithm establishes specific antibodies for each antigen based on the mechanism of specific antigen matching antibody, which leads to an excessive number of detectors and numerous redundant detectors;
(2) The detection phase needs to calculate the antigen and matching degree of all the detectors, the time complexity is high, and the detection efficiency is low.

2 PROPOSED METHOD

First, we abstract the dose-driven innate immune memory establishment process of NK cells , as shown in figure 1.

There are three states of NK cells: resting NK cells, effector NK cells and memory NK cells [4]. The phenotype, which is used by NK cells to identify and study functional subsets, has served as an accessible means for profiling NK cells in different states [5, 6]. When NK cells are in a resting state, the pathogen dose is very small, and their surface phenotypes are mainly $KLRG1^{lo\ and\ hi}$, $CD62L^{lo\ and\ int}$, etc. At this time, no immune response is produced. When resting NK cells are stimulated by a certain dose of pathogens, they will transform into effector NK cells. The phenotypes on the

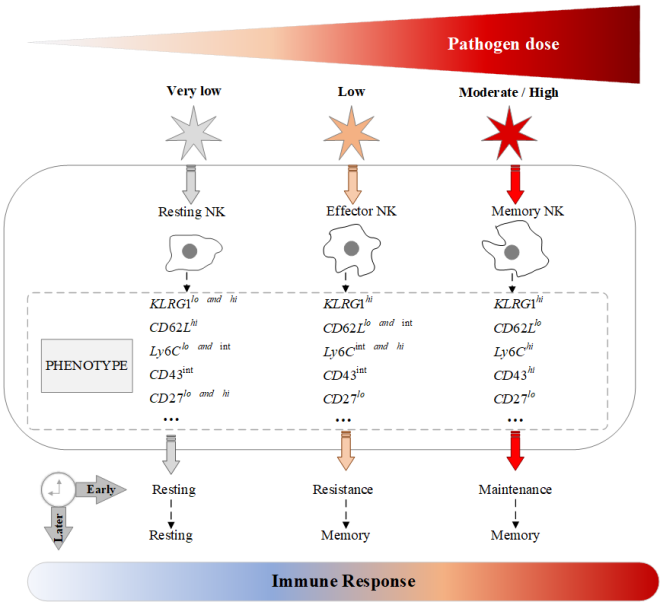

Figure 1: The innate immune memory establishment process of NK cells

surface of effector NK cells are mainly $KLRG1^{hi}$, $CD62L^{int\ and\ hi}$, and the immune system produces a mild immune response. When the pathogen dose further increases, the effector NK cells will differentiate into memory NK cells and produce the function of innate immune memory. The main phenotypes on the surface of memory NK cells are $KLRG1^{hi}$, $CD62L^{lo}$, etc. Memory NK cells with the phenotype against this pathogen invasion have a longer survival time than other NK cells, so when non-specific pathogens invade for a second time, only a lower dose of pathogens can activate memory NK with these phenotypes cells, the immune system will recognize pathogens with a faster and stronger immune response and protect the body.

The pathogen dose based NKA has two notable features which avoid the two problems of NSA:

(1) The establishment of the phenotype detector of natural killer cells is not specific to a certain pathogen, but non-specifically to the dose of all pathogens with certain characteristics in a certain dangerous zone. The dangerous zone is inspired by the Danger Theory which was proposed by Polly Matzinger [3]. And NKA establishes different NK cell phenotypes based on the characteristic dose of pathogens;
(2) When classifying the pathogen, NKA only needs to compare with the core pathogen in the nearest dangerous zone.

NKA mainly consists of four stages: definition of the pathogen feature space, generation of the phenotype detectors, evolution of phenotype detectors and detection phase. The definition of the pathogen feature space stage maps the input data to n-dimensional real-valued vectors to characterize the pathogens feature space. The generation of the phenotype detectors stage defines the pathogen dose and clusters the abnormal pathogen dose as the phenotype detector. The evolution of phenotype detectors stage achieves the purpose of optimizing the detector through the cloning, tolerance, and extinction operations of the phenotype detector, and then the k-d tree is established for the optimized phenotype detectors (see [1] for details). The detection stage uses dose matching with the nearest phenotype detector to check whether testing sample is normal or abnormal. The time complexity is reduced from $O(n)$ to $O(log\ n + f)$, where n Is the total number of phenotype detectors, and f is the number of leaf nodes.

3 EXPERIMENTS

We use two data sets in the UCI data set as experimental data, and compare the proposed model with the traditional negative selection algorithm RNS based on the self-nonself discrimination mechanism and its classic improved version V-detector and NSA-PSO. The experimental results show that the NK cell model based on pathogen dose proposed in this paper can not only achieve a better effect to the negative selection algorithm through fewer detectors, but also has a higher efficiency in the training and detection phase.

4 CONCLUSIONS

Aiming at the problem that distance-based NSA cannot establish an effective detector in normal and abnormal mixed areas, a natural killer algorithm (NKA) based on pathogen dose is proposed in this paper. A theoretical analysis and experimental results show that NKA has a better accuracy, detector generation quality, training and testing complexity compared to three versions of distance-based NSA. Hence, it is suitable to use NKA for classification tasks in both low-dimensional and high-dimensional space. Future work will mainly focus on more in-depth theoretical analysis and applying NKA to more different situations to explore more possibilities of this method, such as intrusion detection, anomaly detection, fault diagnosis.

ACKNOWLEDGMENTS

The authors want to thank NSFC- http://www. nsfc.gov.cn/ for the support through Grants Number 61877045, and Military Commission Innovation Special Zone for the support through Grants Number 17-H863-01-ZT-002-016-02.

REFERENCES

[1] Jon Louis Bentley and Jerome H Friedman. 1979. Data structures for range searching. *ACM Computing Surveys (CSUR)* 11, 4 (1979), 397–409.
[2] Stephanie Forrest, Alan S Perelson, Lawrence Allen, and Rajesh Cherukuri. 1994. Self-nonself discrimination in a computer. In *Proceedings of 1994 IEEE computer society symposium on research in security and privacy*. Ieee, 202–212.
[3] Polly Matzinger. 2012. The evolution of the danger theory. *Expert review of clinical immunology* 8, 4 (2012), 311–317.
[4] Mihai G Netea, Leo AB Joosten, Eicke Latz, Kingston HG Mills, Gioacchino Natoli, Hendrik G Stunnenberg, Luke AJ ONeill, and Ramnik J Xavier. 2016. Trained immunity: a program of innate immune memory in health and disease. *Science* 352, 6284 (2016).
[5] Gaitan Fabrice Njiomegnie, Scott A Read, Nicole Fewings, Jacob George, Fiona McKay, and Golo Ahlenstiel. 2020. Immunomodulation of the natural killer cell phenotype and response during HCV infection. *Journal of clinical medicine* 9, 4 (2020), 1030.
[6] Sophie M Poznanski and Ali A Ashkar. 2019. What defines NK cell functional fate: phenotype or metabolism? *Frontiers in immunology* 10 (2019), 1414.
[7] Chikh Ramdane and Salim Chikhi. 2017. Negative selection algorithm: recent improvements and its application in intrusion detection system. *Int. J. Comput. Acad. Res.(IJCAR)* 6, 2 (2017), 20–30.