OBESITY HEURISTIC, NEW WAY ON ARTIFICIAL IMMUNE SYSTEMS

M. A. El-Dosuky[1], M. Z. Rashad[1], T. T. Hamza[1], and A.H. EL-Bassiouny[2]

[1] Department of Computer Sciences, Faculty of Computers and Information sciences, Mansoura University, Egypt

mouh_sal_010@mans.edu.eg
magdi_12003@yahoo.com
Taher_Hamza@yahoo.com

[2] Department of Mathematics, Faculty of Sciences, Mansoura University, Egypt

el_bassiouny@mans.edu.eg

ABSTRACT

There is a need for new metaphors from immunology to flourish the application areas of Artificial Immune Systems. A metaheuristic called Obesity Heuristic derived from advances in obesity treatment is proposed. The main forces of the algorithm are the generation omega-6 and omega-3 fatty acids. The algorithm works with Just-In-Time philosophy; by starting only when desired. A case study of data cleaning is provided. With experiments conducted on standard tables, results show that Obesity Heuristic outperforms other algorithms, with 100% recall. This is a great improvement over other algorithms.

KEYWORDS

Artificial Immune Systems, Obesity Heuristic, Data Cleaning

1. INTRODUCTION

Metaheuristics are global heuristic optimizers, usually inspired by nature [1], such as Artificial Immune Systems [2], genetic algorithms [3], ant colony optimization [4], particle swarm optimization [5] and Artificial Bee Colony [6]. All metaheuristics apply two strategies of **intensification** and **diversification** in effective exploration of search space [7]. Intensification eliminates the search space by examining neighbors of elite solutions, while diversification is a stochastic component that widens the search space by examining unvisited regions [8].

The diverse research into Artificial Immune Systems is surveyed by listing the application areas [9]. The major application areas are: Clustering/Classification, Anomaly Detection, Computer Security, Numeric Function Optimization, Combinatoric Optimization, and Learning. Minor application areas are: Bio-informatics, Image Processing, Control, Robotics, Virus Detection, and Web Mining.

Now, let us specify the problem statement as follows. There is an emphasis on the need for a support of a long-lasting extraction of new metaphors from immunology to flourish the application areas of Artificial Immune Systems [9].

We develop a new heuristic called Obesity Heuristic that is inspired from advances in medicine, especially in anti-Inflammatory treatment of obesity. **Section 2** reviews main Artificial Immune Systems notions and algorithm. **Section 3** introduces the underpinning principles for Obesity Heuristic. **Section 4** is the proposed algorithm. **Section 5** provides a case study of data cleaning in a typical data warehouse.

2. ARTIFICIAL IMMUNE SYSTEMS

Artificial Immune Systems algorithms are inspired by the studies of human Immune System [10]. Different types of proteins in human body can be classified into "self" and "non-self". T-cells are called "detector" cells, derived by the thymus, opposed to B-cells derived by bone marrow. It was noted that the immune system has a "memory" to remember illness-causing proteins (antigens) [10]. Ideas from human immune systems are applied in network security, especially intrusion detection system ([11], [12])

Clonal selection algorithm ([13], [14]) is described as follows:

1. Generate initial antibodies, where each antibody is a solution
2. Compute the fitness of each antibody.
3. Select highest fitness antibodies from population to generate new antibodies.
4. For each antibody, generate clones and mutate each clone.
5. Lower fitness antibodies are deleted from the population
6. New generated antibodies are added to the population
7. Repeat steps from 2- 6 until stop criterion is met. The stop criterion can be The number of iterations

2. OBESITY HEURISTIC

The direct relation among nutrition, infection and immunity seems to be intuitive. It took time to find the proof of the relation between inflammation and metabolic consequences of obesity, by clarifying the molecular mechanisms of inflammatory processes at that take place at the genetic level ([15], [16], [17]).

To facilitate describing the proposed heuristic, let us first review the obesity, processes involved in anti-Inflammatory treatment of obesity.

3.1 Obesity and the Perfect Nutritional Storm

The term "Perfect Nutritional Storm" [18]   refers to the major changes in dietary. First, the increased consumption of refined carbohydrates notably increases the glycemic load of the diet, defined as the amount of consumed carbohydrate multiplied by the glycemic index [19]. Second, the increased consumption of refined vegetable oils rich in *omega-6* fatty acids is due to hormonal factors influencing desaturation by enzymes delta-6 and delta-5, strongly activated by insulin [20]. The third major change in dietary is the decreased consumption of long-chain *omega-3* fatty acids [21].

Obesity is the accumulation of excess body fat safely stored in *adipose tissue* in the form of triglycerides for long-term storage [22]. However, if the excess fat continues to increase and become deposited in any organ other than adipose tissue, this is known as *lipotoxicity*, and this induces chronic diseases [23].

Insulin signaling is a main player in fat circulation, to provide the sufficient glucose levels into the fat cell that can be converted to glycerol [24].

3.2 Innate Immune System

Burns, injuries, microbial invasion are factors that can trigger inflammation through the innate immune response, now we know that and obese diet can do, as the innate immune system is

highly evolved to become very sensitive to nutrients [25]. Advances in molecular biology shed light on inherent mechanisms of a typical innate immune system [26].

Inflammatory process is a function of two supposedly-balanced phases of pro- and anti-inflammatory phases. *Pro-inflammatory* mechanisms are indicators of cellular destruction, such as pain, swelling, redness and heat. *Anti-inflammatory* mechanisms are responsible for cellular repair and regeneration [27].

We may link pro- and anti-inflammatory mechanisms with intensification and diversification strategies mentioned in section 1.

Now, we need to show how our awareness of Perfect Nutritional Storm principle mentioned in section 3.1 can affect the behavior immune system.

A common *omega-6* fatty acid is Arachidonic acid (AA) and can induce inflammatory responses, so it must be reduced [28].

The explanation of the role of AA is as follows:

- Necrosis or cell death of a fat cell can occur, when the accumulation of AA in that fat cell increases over a threshold [29].

- This causes a migration of macrophages into the adipose tissue [30]

- These macrophages causes inflammatory mediators, such as IL-1, IL-6 and TNFα ([31] , [32])

Omega-3 fatty acids, which is said to have an impact on brain development and intelligence [33], can activate anti-inflammatory gene transcription factors, so it must be increased enough [34].

4. PROPOSED OBESITY HEURISTIC

The proposed metaphor depends on obesity and immune system. The main forces of the algorithm are the generation *omega-6* and *omega-3* fatty acids*.* Within a large data set, we may think of raw data as simple amino acids and generated information and mined rules as complex fatty acids. Both omega-6 and omega-3 are generated and stored. Roughly speaking, *omega-6* can be a metaphor for undesired generated information and *omega-3* to be a metaphor for intelligent desired generated information. If the amount of *omega-6* is above a certain threshold, this is considered inflammation sign, and requires the start of immune system mechanisms to handle.

Adipose tissue storing triglycerides can be thought of as a metaphor for long-term data storage. If the excess fat continues to increase and become deposited in any location other than adipose tissue, this is known as *lipotoxicity*, and this induces malicious behaviors such as denial-of-service [35].

The hybrid method is described as follow:-

1. *Initialize Adipose tissues, the long-term data storage locations, to be empty*

2. Get the sequence of simple amino acids, the raw data from input streams

3. Generate complex fatty acids

4. *Store generated fatty acids*

5. If the storage location is not in one adipose tissue

   a. *If the size of this storage location is above threshold (**lipotoxicity**)*

   ➔ **start immune system mechanisms** to handle

6. Calculate the amount of generated ***omega-6*** and ***omega-3*** fatty acids

7. If the amount of ***omega-6*** is above a certain threshold,

   ➔ **start immune system mechanisms** to handle

8. *Repeat step 2- 7.*

9. *The performance measure of the system is* amount of generated ***omega-3***

The main feature of the proposed algorithm is that it operates with Just-In-Time philosophy, by starting with full power only when desired as in 5th and 7th steps. We can utilize any AIS algorithm, but we prefer Clonal Selection Algorithm. In step 3, to generate complex fatty acids, we can apply any mining algorithm. However, we can assume that complex fatty acids are the amino acids, i.e., the output is the input for simplicity. This leads us for the following case study of data cleaning.

5. DATA CLEANING, A CASE STUDY

With Data Warehouse is an **integrated, subject-oriented, time-variant and non-volatile** database [36]. Integrating data from two or more heterogeneous sources shall cause some inevitable "*dirt*". In order to improve data consistency by reducing data redundancy, data cleaning is required [37]. Data cleaning is the automated detection and correction of missing and incorrect values [38].

Many methods are proposed for eliminating duplicate records in large data, such as sorting [39], sorting and clustering using a scanning window [40], basic field matching algorithm [41], pre-processing by tokenizing fields [42], and declarative language for data cleaning [43]. Simple performance measures can be:

$$\text{Recall} = \frac{|\text{correctly identified duplicates}|}{|\text{identified duplicates}|}$$

(1)

$$\text{False positive error} = \frac{|\text{wrongly identified duplicates}|}{|\text{identified duplicates}|}$$

(2)

Experiments are conducted using data warehouse with tables listed in [44].

Results show that Obesity Heuristic outperforms other algorithms, with a recall close to 100%, or false positive error close to zero. This is a great improvement over the algorithms described in [45].

6. CONCLUSIONS

We have introduced Obesity Heuristic a new metaphor derived from anti-Inflammatory treatment of obesity to the meet the need for a support of a long-lasting extraction of new metaphors from immunology to flourish the application areas Artificial Immune Systems. We believe that may extend the meaning and the application areas of Artificial Immune Systems. We review main Artificial Immune Systems notions and algorithm. Then we introduce the underpinning principles for Obesity Heuristic. We see that the proposed algorithm is promising especially in long-term huge data storage. Finally, the main feature of the proposed algorithm is that it links artificial immune system to huge storage such as data warehouses and makes it operates with Just-In-Time philosophy, by starting with full power only when desired as in 5th and 7th steps.

Future work may be the application of the proposed algorithm in other areas such as network security.

ACKNOWLEDGEMENTS

REFERENCES

[1] Yang, X. S., (2008) "Nature-Inspired Metaheuristic Algorithms", Luniver Press

[2] Farmer, J.D.; Packard, N.; Perelson, A., (1986) "The immune system, adaptation and machine learning". Physica D 22 (1-3): 187–204.

[3] Goldberg, D. E., (1989) "Genetic Algorithms in Search", Optimisation and Machine Learning, Reading, Mass., Addison Wesley.

[4] Dorigo, M,. (1992) "Optimization, Learning and Natural Algorithms" (Phd Thesis). Politecnico di Milano, Italy.

[5] Kennedy, J.; Eberhart, R. (1995) "Particle Swarm Optimization". Proceedings of IEEE International Conference on Neural Networks. IV. pp. 1942–1948

[6] Karaboga, D., (2005) "An Idea Based On Honey Bee Swarm For Numerical Numerical Optimization". Technical Report-TR06 (Erciyes University, Engineering Faculty, Computer Engineering Department).

[7] Blum, C. and Roli, A., (2003) "Metaheuristics in combinatorial optimization: Overview and conceptural comparision", ACM Comput. Surv., vol. 35, pp. 268-308.

[8] Glover, F. and M. Laguna. (1997) "Tabu Search". Kluwer Academic Publishers.

[9] Hart, E. and J. Timmis. (2005) "Application areas of AIS: The Past, Present and Future". In C. Jacob, M. Pilat, P. Bentley, and J. Timmis, editors, Proceedings of the 4th International Conference on Artificial Immune Systems, volume 3627 of LNCS, pages 483–499.

[10] Janeway, C. A. and P. Travers, (1996) "Immunobiology: The Immune System in Health and Disease", Current Biology Ltd., 2nd edition.

[11] Forrest, S., S. A. Hofmeyr, and A. Somayaji, (1997) "Computer Immunology," Communications of the ACM, vol. 40, no.10, pp. 88-96.

[12] Somayaji, A., S. A. Hofmeyr, and S. Forrest, (1998) "Principles of a computer immune system," in New Security Paradigms Workshop, pp. 75-82.

[13] De Castro L.N.&VonZuben,F.J. (2000) "The Clonal Selection Algorithm with Engineering Applications", GECCO'00 – Workshop Proceedings, pp. 36-37

[14] De Castro L. and Timmis, J. (2002) "Artificial Immune Systems: A New Computational Approach", Springer-Verlag New York, Inc.

[15] Olefsky, J. M., (2009) "IKKε: a bridge between obesity and inflammation," Cell, vol. 138, no. 5, pp. 834–836.

[16] Vachharajani, V. and D. N. Granger, (2009) "Adipose tissue: a motor for the inflammation associated with obesity," IUBMB life, vol. 61, no. 4, pp. 424–430.

[17] Wellen, K. E. and G. S. Hotamisligil, (2003) "Obesity-induced inflammatory changes in adipose tissue," Journal of Clinical Investigation, vol. 112, no. 12, pp. 1785–1788.

[18] Sears, B., (2008) "Toxic Fat", Regan Books, New York, NY, USA.

[19] Ludwig, D. S., (2002) "The glycemic index: physiological mechanisms relating to obesity, diabetes, and cardiovascular disease," Journal of the American Medical Association, vol. 287, no. 18, pp. 2414–2423

[20] Brenner, R. R., (1982) "Nutritional and hormonal factors influencing desaturation of essential fatty acids," Progress in Lipid Research, vol. 20, no. 1, pp. 41–48

[21] Serhan C. N., M. Arita, S. Hong, and K. Gotlinger, (2004) "Resolvins, docosatrienes, and neuroprotectins, novel omega-3-derived mediators, and their endogenous aspirin- triggered epimers," Lipids, vol. 39, no. 11, pp. 1125–1132.

[22] Karelis,. A. D., M. Faraj, J.-P. Bastard et al., (2005) "The metabolically healthy but obese individual presents a favorable inflammation profile," Journal of Clinical Endocrinology and Metabolism, vol. 90, no. 7, pp. 4145–4150.

[23] Unger R. H., (2002) "Lipotoxic diseases," Annual Review of Medicine, vol. 53, pp. 319–336.

[24] Massiera, F., P. Saint-Marc, J. Seydoux et al., (2003) "Arachidonic acid and prostacyclin signaling promote adipose tissue development: a human health concern?" Journal of Lipid Research, vol. 44, no. 2, pp. 271–279.

[25] Hotamisligil, G. S., and E. Erbay, (2008) "Nutrient sensing and inflammation in metabolic diseases," Nature Reviews Immunology, vol. 8, no. 12, pp. 923–934.

[26] H. Hölschermann, D. Schuster, B. Parviz, W. Haberbosch, H. Tillmanns, and H. Muth "Statins prevent NF-κB transactivation independently of the IKK-pathway in human endothelial cells," Atherosclerosis, vol. 185, no. 2, pp. 240–245, 2006.

[27] Serhan, C. N., S. D. Brain, C. D. Buckley et al., (2007) "Resolution of inflammation: state of the art, definitions and terms," FASEB Journal, vol. 21, no. 2, pp. 325–332.

[28] Heller, A., et al., (1998) "Lipid mediators in inflammatory disorders", Drugs, vol. 55, pp. 487-496.

[29] C. Pompeia, T. Lima, and R. Curi, (2003) "Arachidonic acid cytotoxicity: can arachidonic acid be a physiological mediator of cell death?" Cell Biochemistry and Function, vol. 21, no. 2, pp. 97–104.

[30] Weisberg, S. P., D. McCann, M. Desai, M. Rosenbaum, R. L. Leibel, and A. W. Ferrante Jr., (2003) "Obesity is associated with macrophage accumulation in adipose tissue," Journal of Clinical Investigation, vol. 112, no. 12, pp. 1796–1808.

[31] Festa, A., R. D'Agostino Jr., G. Howard, L. Mykkänen, R. P. Tracy, and S. M. Haffner, (2000) "Chronic subclinical inflammation as part of the insulin resistance syndrome: the insulin resistance atherosclerosis study (IRAS)," Circulation, vol. 102, no. 1, pp. 42.

[32] Ruan, H. and H. F. Lodish, (2003) "Insulin resistance in adipose tissue: direct and indirect effects of tumor necrosis factor-α," Cytokine and Growth Factor Reviews, vol. 14, no. 5, pp. 447–455.

[33] Helland IB, Smith L, Saarem K, Saugstad OD, Drevon CA. (2003) "Maternal supplementation with very-long-chain n-3 fatty acids during pregnancy and lactation augments children's IQ at 4 years of age", Pediatrics. Jan; vol. 111, no. 1, pp. 39-44.

[34] Sears, B., Camillo Ricordi, (2011) "Anti-Inflammatory Nutrition as a Pharmacological Approach to Treat Obesity", Journal of Obesity, no. 431985

[35] Bicakci, K., Tavli, B., (2009) "Denial-of-Service attacks and countermeasures in IEEE 802.11 wireless networks", Computer Standards & Interfaces, vol. 31 , pp 931–941

[36] W.H Inmon. (1996) "Building the Data Warehouse", second edition, John Wiley.

[37] B. Delvin. (1997) "Data warehouse from architecture to implementation", Addison- Wesley.

[38] E. Simoudis and B. Livezey and R. Kerber. (1995) "Using Recon for Data Cleaning". In Proceedings of KDD 1995, pp 282-287.

[39] D. Bitton and D.J. Dewitt. (1983) "Duplicate Record Elimination in Large Data Files", ACM Transactions on Database Systems, Vol. 8, No. 2, PP 255 - 265.

[40] M.A Hernandez and S.J Stolfo. (1998) "Real-World Data is Dirty: Data Cleansing and the Merge/Purge Problem", Data Mining and Knowledge Discovery, vol. 2, pp. 9 - 37.

[41] A.E Monge and C.P Elkan. The Field Matching Problems: Algorithms and Applications. Proceedings of the 2nd Int'l Conference on Knowledge and Data Mining pp 267 - 270, 1996.

[42] M.L. Lee and L. Hongjun and W.L Tok and T.K Yee. (1999) "Cleansing Data for Mining and Warehousing", In Proceedings of the 10th Int'l Conference on Database and Expert Systems Applications (DEXA 99), Florence, Italy.

[43] H. Galharda and D. Florescu and D. Shasha and E. Simon and C. Saita. (2001) "Declarative Data Cleaning: Language, Model and Algorithms", Proceedings of the 27th VLDB conference, Roma, Italy.

[44] Timothy E. Ohanekwu and C.I. Ezeife, (2003) "A Token-Based Data Cleaning Technique for Data Warehouse Systems", proceedings of the International Workshop on Data Quality in Cooperative Information Systems, held in conjunction with the 9th International Conference on Database Theory (ICDT 2003), Siena, Italy, pp. 21-26.

[45] M.A Hernandez and S.J Stolfo. (1995) "The Merge/Purge Problem for Large Databases", In Proceedings of the ACM SIGMOD Int'l Conference on Management of Data, pp 127 - 138.

Authors

Mohammed El-Dosuky is a PhD student at Computer sciences Department, Faculty of Computers and Information, Mansoura University, Egypt. He received his Master in Computer Sciences about quantum computing from the Mansoura University in 2008. His current research interests are Multi-Agent Systems, intelligent communication, and evolutionary computing.

Ahmed EL-Bassiouny is currently a Professor at Mansoura University, Egypt. Currently, he is the dean of the Faculty of Sciences, Mansoura University.

Taher Hamza is currently an associate Professor at Mansoura University, Egypt. He is the vice dean For Higher Studies, Faculty of Computer and Information, Mansoura University.

Magdy Rashad is currently an associate Professor at Mansoura University, Egypt. He is the head of Computer Sciences department and the vice dean for Community Service and Environmental Development, Faculty of Computer and Information, Mansoura University.