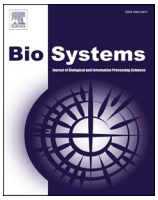
BioSystems 229 (2023) 104918

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems



Short Communication

Why the power of diversity does not always produce better groups and societies

Yoshiyasu Takefuji

Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo, 135-8181, Japan

ARTICLE INFO

Handling Editor: Dr. A.U. Igamberdiev

Keywords:
Diversity prediction theory
Real number
Complex number
Individual ability
Nature-inspired intelligence
Collective intelligence

ABSTRACT

Diversity is supposed to create better groups and societies but sometimes fails. It is explained why the power of diversity may not create better groups in the current diversity prediction theory. Diversity may hurt civic life and introduce distrust. This is because the current diversity prediction theory is based on real numbers that ignore individual abilities. Its diversity prediction theory maximizes performance with infinite population size. Contrary to this, collective intelligence or swarm intelligence is not maximized by infinite population size, but by population size. The extended diversity prediction theory using the complex number allows us to express individual abilities or qualities. The diversity of complex numbers always produces better groups and societies. The wisdom of crowds, collective intelligence, swarm intelligence or nature-inspired intelligence is implemented in the current machine learning or artificial intelligence, called Random Forest. The problem of the current diversity prediction theory is detailed in this paper.

1. Introduction

The importance of diversity was discovered by Francis Galton in 1907 in decision-making (Galton, 1907). The wisdom of crowds, collective intelligence, or nature-inspired intelligence has suggested that, under the right conditions, groups of people with diversity can make better decisions than individuals (Galton, 1907). Or, a diverse group of problem-solvers made a better collective guess than that produced by the group of best-performing solvers. Galton's diversity theory suggests that diverse minds do better in decision-making. Heiko et al. formulated diversity of the wisdom of crowds (Heiko et al., 2011).

The wisdom of crowds is a touting catchphrase like crowdsourcing, big data and predictive analytics (Grep, 2017). The current Diversity Prediction Theory (DPT) uses the real number (Page, 2007). Diversity is supposed to create better groups and societies but sometimes fails. The wisdom of crowds is called collective intelligence, swarm intelligence, or nature-inspired intelligence in different study areas.

Robert Putnam suggests that diversity hurts civic life and that differences can actually translate into distrust (Putnam, 2007). There are pros and cons against Putnam's finding (McKenna et al., 2018; Grewal, 2016; Robin et al., 2020). Diverse societies are less cohesive (McKenna et al., 2018). Diversity creates distrust (Grewal, 2016).

Algorithms exist that construct higher dimensional number systems from lower dimensional number systems (Hilbert, 2020). The number system starts from the real numbers, complex numbers, quaternions, octonions, and higher dimensional numbers (Hilbert, 2020). In the number system, the higher dimensional numbers, the better versatility we can obtain (Hilbert, 2020; Baez, 2012).

This paper shows why the current DPT sometimes fails in creating better groups. This is because the current DPT is expressed in real numbers (Page, 2007). The current DPT lies in that the goal is to minimize the squared difference between the crowd's prediction and the truth. The squared difference always produces positive numbers regardless of individual abilities or qualities as long as real numbers are used in diversity. In real society, a negative individual always degrades the quality of the group, but the squared difference (positive or negative) will always generate positive numbers in the current DPT. This means that infinite population can achieve the maximum performance.

The extended DPT is proposed where individuals are expressed by the complex number with the better versatility. The squared difference can be distinguished by individual abilities. The squared difference by negative individuals can generate negative numbers in the extended DPT because of the power of complex numbers.

The wisdom of crowds or swam intelligence is actually implemented in ensemble machine learning of artificial intelligence which is called Random Forest (Stephan et al., 2015; Buckley et al., 2021). The number of trees in Random Forest is equivalent to swam population size.

Constructing optimal binary decision trees is NP-complete (Hyafil

E-mail address: takefuji@keio.jp.

https://doi.org/10.1016/j.biosystems.2023.104918
Received 10 April 2023; Received in revised form 6 May 2023; Accepted 7 May 2023
Available online 15 May 2023

and Rivest, 1976). In other words, swarm intelligence with optimum population size is unknown. In other words, the optimum number of trees in the Random Forest algorithm is also unknown. However, the latest swarm intelligence study shows that too small population or too large population does not obtain the good performance (Piotrowski et al., 2020). Besides, the large number of trees in Random Forest does not mean the good performance (Oshiro et al., 2012).

Most systems based on artificial intelligence or machine learning are using the real number. When we can use the higher numbers, the better versatility can be added to the future AI systems. This paper shows how to use the higher numbers for creating smart human groups or systems from the viewpoint of collective intelligence, swam intelligence or nature-inspired intelligence.

The proposed idea using the complex number is derived from the energy maximization of impedance matching. In impedance matching, we need to nullify the imaginary part to maximize the transmission energy. In other words, the impedance matching problem is equivalent to the diversity prediction problem. The goal of DPT is to minimize the squared difference between the crowd's prediction and the truth while that of impedance matching using the complex number is to maximize the transmission energy.

In order to express individual abilities in diversity prediction theory, constrained complex numbers are introduced where they can be expressed in real or complex numbers only in the imaginary part for simplicity.

If individual ability x_i is expressed by $x_i = a_i$ (positive individuals) or $x_i = b_i$ j (negative individuals) in the proposed diversity prediction model, ${x_i}^2$ can be represented by ${a_i}^2$ or -${b_i}^2$ where ${a_i}^2$ and -${b_i}^2$ are both real numbers. As with impedance matching, the imaginary parts are all nullified in the proposed diversity prediction model. Note that a_i and b_i are real numbers and j is an imaginary number ($j = \sqrt{-1}$).

If individual ability x_i is expressed by $x_i = a_i + b_i$ j, ${x_i}^2$ can be represented by ${x_i}^2 = {a_i}^2$-${b_i}^2$+2 $a_i b_i$ j where the imaginary term 2 $a_i b_i$ j still remains. As with impedance matching, we need to nullify the imaginary parts, but we cannot. Therefore, $x_i = a_i + b_i$ j expression as individual ability cannot be used in the proposed diversity prediction model.

The complex number allows us to express two classes such as positive and negative individual abilities in the proposed DPT. In other words, the conventional DPT based on the real number cannot express individual abilities while the proposed DPT based on the complex number can express individual abilities or qualities. Individual abilities or qualities will be detailed in this paper.

2. Rationale

Diversity is the quality of being diverse or different; difference or unlikeness. The DPT is introduced by Page where it shows how the power of diversity creates better groups, firms, schools, and societies (Page, 2007). The term, "collective intelligence", swam intelligence or "wisdom of crowd" is used for explaining intelligence of sociological diversity.

The goal of DPT is to minimize the squared difference between the crowd's prediction (C) and the truth (X). In the DPT, the number system is based on the real numbers (Page, 2007):

$$(C-X)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - X)^2 - \frac{1}{n}\sum_{i=1}^{n}(x_i - C)^2$$

where C is the crowd's prediction, X is the truth, x_i is individual i's prediction, and n is the number of individuals respectively.

The DPT can be summarized as follows:

[crowd's square error] = [mean square error] – [diversity of crowds].

The last term of [diversity of crowds] can reduce the value of [crowd's square error] as long as the diversity of crowds is greater than zero.

If the number system is based the real number, then it is true that the term of [diversity of crowds] is always greater than zero:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - C)^2 \geq 0$$

In other words, the power of diversity always creates better groups by the diversity term.

However, the problem of the DPT lies in ignoring individual abilities. Can you imagine that negative diverse individuals can create better groups? The current DPT is missing an important indicator to express individual abilities.

Besides, based on the DPT model, the higher the number of individuals, the better the diversity intelligence. Based on the DPT model, the maximum diversity intelligence can be obtained by the infinite number of individuals. However, the latest study of swarm intelligence shows that the best performance was obtained with swarms composed of 70–500 particles (Piotrowski et al., 2020). This result shows that too small population or too large population does not show the good performance in optimization. Similarly, in Random Forest algorithms, increasing the number of trees does not mean the performance improvement (Oshiro, 2021).

In the proposed DPT using the complex number, we can express positive and negative individuals for organizing better groups or societies while the conventional DPT only allows us to express positive individuals.

3. Complex number

In order to express the versatility of individuals or individual abilities, assume the individual i can be given by $x_i = a_i$ *or* $x_i = jb_i$ where a_i and b_i are real numbers and j is an imaginary number with $j = \sqrt{-1}$. And assume C = 0 for simplicity.

Assuming "$x_i = a_i$ *or* $x_i = jb_i$" means that we can have two types of individuals: positive and negative person. Versatility of the complex number allows us to express positive and negative individuals. In the current DPT, remember that all persons are expressed by the real number.

If the number system is based on the complex number with the above conditions

since $j^2 = \sqrt{-1}^2 = -1$, then

the value of [diversity of crowds] expressed by the complex number can be converted to the real number fortunately:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i)^2 \geq 0 \; or \; \frac{1}{n}\sum_{i=1}^{n}(x_i)^2 \leq 0$$

As long as individuals are expressed by $x_i = a_i$ *or* $x_i = jb_i$, the square of individuals can be converted to the real number (positive or negative). Remember that the square of individuals using the real number is always greater than zero. In other words, there is no negative value of [diversity of crowds] in the conventional DPT.

If a team of individuals composed of diverse negative persons, then diversity of negative crowds can be calculated as the negative real number:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i)^2 < 0$$

which is smaller than zero. In other words, the power of diversity using the complex number allows us to express negative groups. This is theoretically correct because the group is made up of diverse negative individuals.

4. Discussion

When expressing individuals with the complex number, we are allowed to express positive and negative individuals for creating better groups while the current DPT using real numbers only allows us to express positive individuals.

In the current DPT, because of the diversity term, the larger number (the number of members in the group), the better group can be created regardless of individual qualities. The current DPT using real numbers misses an important indicator of individual abilities such as positive and negative ability. Based on the real number, the current DPT shows such unpleasant behaviors in diversity.

In the extended DPT, the diversity term can be expressed by the complex number so that the better group can be created by selecting individuals with their qualities or abilities (positive and negative), not by the number of members.

In the extended DPT, the proposed method in the complex number allows us to express positive and negative individuals to organize better groups and societies while the conventional DPT in real numbers only allows us to express positive individuals or qualities. Diversity using the complex number allows us to select a team of members (positive or negative individuals) and plays a key role in always creating better groups or societies.

As mentioned in INTRODUCTION Section, all traditional machine learning algorithms are based on real numbers. Complex numbers are not yet used in current machine learning. By introducing complex numbers to machine learning, we may be able to turn the algorithms into a new dimension. Because, a complex number consists of the ordered pair: the real component and the imaginary component. Complex numbers can be expressed in terms of their magnitude and phase angle. Complex numbers play a key role in machine learning (Bassey et al., 2021; Jesper et al., 2021; Zhang et al., 2021).

5. Conclusion

The current DPT in real numbers always suggests that the power of diversity creates better groups, firms, schools, and societies where individual abilities are ignored. In the current DPT, the best performance can be obtained by the infinite number of individuals. However, the latest swarm intelligence study reveals that too small population or too large population does not obtain the good performance. The current DPT contradicts with the latest result of swarm intelligence.

The complex number allows us to express individual abilities by positive and negative individuals in the extended DPT.

Diversity using the complex number suggests that selecting a team of members (positive or negative individuals) plays a key role in always creating better groups or societies.

The diversity prediction problem is equivalent to the impedance matching problem. The goal of diversity prediction problems is to minimize the squared difference between the crowd's prediction and the truth while that of impedance matching is to maximize the transmission energy. The diversity problem is a reinvention of the impedance matching problem.

There is a problem of how to build a good team in order to achieve a given goal with the best performance. This paper concludes that diversity should not be simply introduced into an organization, but rather diversity should be introduced by taking into account the capabilities of individuals and the number of individuals in a team or group.

In collective intelligence or wisdom of crowds, the size of the group with its individual capabilities infer intelligence.

This study proposed methodological improvements to existing research methods. However, as Future Work, an empirical study or Case Study is needed to demonstrate the effectiveness of the new methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Baez, J.C., 2012. Division Algebras and Quantum theory. Found. Phys. 42, 819–855. https://doi.org/10.1007/s10701-011-9566-z, 2012.

Bassey, J., et al., 2021. A Survey of Complex-Valued Neural Networks arXiv:2101.12249 [stat.ML].

Buckley, S.J., et al., 2021. Application of the random forest algorithm to *Streptococcus pyogenes* response regulator allele variation: from machine learning to evolutionary models. Sci. Rep. 11, 12687 https://doi.org/10.1038/s41598-021-91941-6, 2021.

Galton, F., 1907. Vox populi (The wisdom of crowds). Nature 75, 450–451.

Grep, P., 2017. Wisdom of the Crowd: from TV Show to Retail Business Model. https://www.forbes.com/sites/gregpetro/2017/11/21/wisdom-of-the-crowd-from-tv-show-to-retail-business-model/?sh=68262d5a38ca.

Grewal, D., 2016. Does Diversity Create Distrust? Doubts about a Harvard Professor's Landmark Finding. https://www.scientificamerican.com/article/does-diversity-create-distrust/.

Heiko, R., et al., 2011. The wisdom of crowds in one mind: how individuals can simulate the knowledge of diverse societies to reach better decisions. J. Math. Psychol. 55 (2), 191–197, 2011.

Hilbert, D., 2020. Hilbert Book Model Project/Quaternions. https://en.wikiversity.org/wiki/Hilbert_Book_Model_Project/Quaternions.

Hyafil, L., Rivest, R.L., 1976. Constructing optimal binary decision trees is NP-complete, 1976 Inf. Process. Lett. 5 (Issue 1), 15–17. https://doi.org/10.1016/0020-0190(76)90095-8. ISSN 0020-0190.

Jesper, S.D., et al., 2021. Complex-valued neural networks for machine learning on non-stationary physical data. Comput. Geosci. 146 (2021), 104643 https://doi.org/10.1016/j.cageo.2020.104643. ISSN 0098-3004.

McKenna, S., et al., 2018. Are diverse societies less cohesive? Testing contact and mediated contact theories, 2018 PLoS One 13 (3), e0193337. https://doi.org/10.1371/journal.pone.0193337, 2018 Mar 29.

Oshiro, T.M., Perez, P.S., Baranauskas, J.A., 2012. How many trees in a random forest?. In: Perner, P. (Ed.), Machine Learning and Data Mining in Pattern Recognition. MLDM 2012. Lecture Notes in Computer Science, vol. 7376. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31537-4_13.

Page, S.E., 2007. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press.

Piotrowski, A.P., et al., 2020. Population size in particle swarm optimization. Swarm Evol. Comput. 58 (2020), 100718 https://doi.org/10.1016/j.swevo.2020.100718. ISSN 2210-6502.

Putnam, R.D., 2007. E Pluribus Unum: diversity and community in the twenty-first century the 2006 Johan Skytte prize lecture. Scand. Polit. Stud. 30, 137–174. https://doi.org/10.1111/j.1467-9477.2007.00176.x.

Robin, J.E., et al., 2020. Getting Serious about Diversity: Enough Already with the Business Case. https://hbr.org/2020/11/getting-serious-about-diversity-enough-already-with-the-business-case#.

Stephan, J., et al., 2015. A random forest approach to capture genetic effects in the presence of population structure. Nat. Commun. 6, 7432. https://doi.org/10.1038/ncomms8432, 2015.

Zhang, H., Gu, M., Jiang, X.D., et al., 2021. An optical neural chip for implementing complex-valued neural network. Nat. Commun. 12, 457. https://doi.org/10.1038/s41467-020-20719-7.