Information Sciences 430–431 (2018) 487–503

Contents lists available at ScienceDirect

Information Sciences

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

An influence analysis of diversity and collective cardinality on collective performance

Van Du Nguyen[a,b], Ngoc Thanh Nguyen[a,*]

[a] Faculty of Computer Science and Management, Wroclaw University of Science and Technology, Wyb. St. Wyspianskiego 27, 50-370 Wroclaw, Poland

[b] Faculty of Information Technology, Nong Lam University, Ho Chi Minh City, Vietnam

ARTICLE INFO

Article history:
Received 28 April 2017
Revised 21 November 2017
Accepted 26 November 2017
Available online 5 December 2017

Keywords:
Collective intelligence
Collective performance
Integration computing

ABSTRACT

This paper presents a general framework to demonstrate the prominent role of diversity in the effectiveness of collective performance. There appears to be ample evidence that diversity is one of the essential criteria of which a collective to be intelligent. Intuitively, a collective involving diverse individuals may add new information, new perspectives, and so forth on the problem that needs to be solved. Moreover, the diversity of individual solutions to the given problem has been proven helpful in eliminating the phenomenon of correlated errors. The objective of the paper is to investigate the influence of the latter kind of diversity on the collective performance by taking into account the collective cardinality. Our findings qualify the positive impact of diversity on collective performance. Particularly, collectives with higher diversity levels will lead to better collective performances. Subsequently, expanding the collective cardinality that causes an increase in its diversity will also be positively associated with the collective performance. With some restrictions, the hypothesis *"the more diverse the collective, the higher the collective performance"* is formally proved. Furthermore, the conditions under which increasing the cardinality of a collective will cause its diversity to be increased (or decreased) are worked out.

1. Introduction

Several publications have appeared in recent years documenting the effectiveness of a collective of individuals (even uninformed individuals) in solving some difficult problems such as judging and predicting problems [4,9,37,42]. Even though, in [2,23] the predictions based on such collectives could outperform those produced by the traditional forecasting methods. It is due to the fact that a collective may have additional knowledge that single individuals do not possess, and such knowledge may be relevant to the problem that needs to be solved [44]. For example, if member *A* knows that "{P}" and member *B* knows that "{P→Q}" then collectively they not only know that "{P, P→Q}" but also know that "{Q}". From this fact, it is reasonable to conclude that collective knowledge is often not a *"normal sum"* of individual knowledge [18,29]. In general, the superiority of a collective has correspondingly been termed *Collective Intelligence* that often considered as the intelligence emerged from the collaboration and competition of many individuals [26]. It has existed for a very long time in forms of a collective of individuals doing things at least sometimes intelligent [43].

In [42], the author has opened up the underlying criteria of intelligent collectives. They are *diversity, independence, decentralization,* and *aggregation. Diversity* is understood as the variety of individual backgrounds or individual solutions to a

* Corresponding author.
E-mail addresses: van.du.nguyen@pwr.edu.pl (V.D. Nguyen), ngoc-thanh.nguyen@pwr.edu.pl (N.T. Nguyen).

https://doi.org/10.1016/j.ins.2017.11.053
0020-0255/

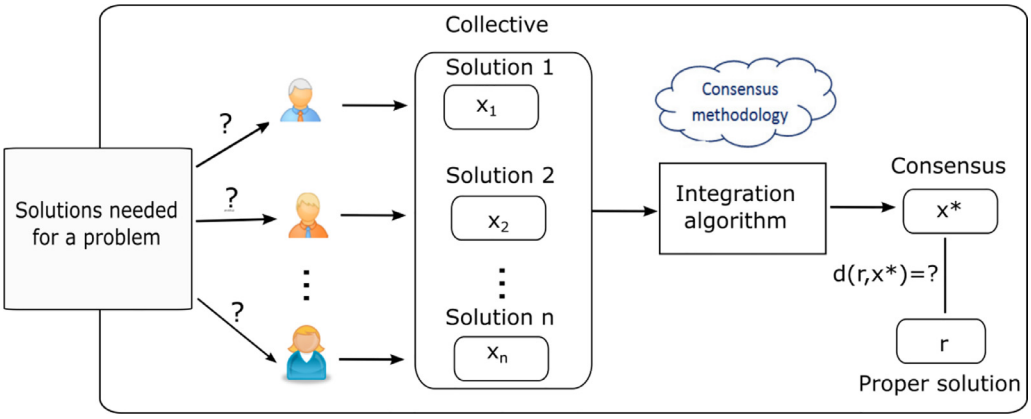

Fig. 1. The process of consensus determination.

given problem in the real world. Intuitively speaking, a collective seems not to be intelligent if all members have the same background or the same viewpoint on a given problem. In [4] the authors have indicated that using heterogeneous collectives is an effective approach to make the collectives to be more accurate. The second criterion, named *independence*, means that an individual solution must be provided independently of others in the collective [5,21,27,37,41,42]. This criterion has been proven helpful in avoiding the phenomenon of the so-called correlated errors among individual solutions [14,25,28]. By satisfying *decentralization*, we have in mind that collective members are able to specialize and draw on local knowledge [42]. This criterion will ensure the collective members act freely and independently of one another. The last criterion, *aggregation*, is often considered as suitable methods for combining individual solutions to form a collective one. In [24], the authors have shown that the composite face that is combined using a computer program, is more attractive than almost all of the individual faces.

Given that a set of autonomous members (e.g. humans or agent systems) are asked for being given their solutions to a problem in the real world. These members may have different backgrounds, knowledge bases; therefore, their solutions can be different from each other. We assume that these solutions have the unified representation (we used multi-dimensional vectors for such representation). Notice that, in this work, we only concentrate on the kind of problems in which their proper solutions exist independently of individual solutions such as the problem of predicting the outcomes of future events or forecasting the weather state of a specific region. In fact, with this kind of problems, the proper solutions are not known when the members are asked for being given their solutions. Therefore, there appears that the individual solutions reflect the proper solution to some degree because of incompleteness and uncertainty. Taking into account these reasons, we named it as *objective case* to differentiate it from *subjective case* in which the proper solutions are dependent on individual solutions. From a collective of individual solutions, we need to obtain a common one (called *consensus*) that can be considered as the representative of the collective as a whole. In [31], the author has worked out many consensus-based algorithms for determining such a representative. The general procedure of consensus determination is described in Fig. 1.

In the previous work [31,32] we have formally proved some theorems related to the relationship between the consensus and the individual solutions in a collective. That is the consensus is better than the worst one in the collective. If all solutions to the same degree reflect the proper solution, then the consensus is the best one. In [35] the simulation results have shown that the cardinality of a collective is positively associated with the quality of its consensus. With the impact of diversity, most publications have mainly focused on the diversity in the composition of collective members [1,4,7,19,38]. According to these works, a collective composing of more diverse members can outperform that of less diverse members. Besides, the findings in [37,42] have shown that the diversity of individual solutions also play an effective role in leading to a better collective performance. It is because this kind of diversity is helpful in avoiding the phenomenon of correlated errors among individual solutions [14,25,28]. In this paper, we will investigate the impact of this kind of diversity on collective performance by taking into account collective cardinality. The functions defined in [31] will be used for measuring the diversity level of a collective. These functions are not only based on the distances from the consensus to individual solutions but also based on the distances between solutions in a collective. Moreover, the collective performance is measured by taking into account:

- the distance from the consensus to the proper solution (*Diff*)
- the ratio of the collective error and individual errors (*QIC*)

To this end, first, we will simulate collectives with the same cardinality but with different levels of diversity to determine the impact of diversity on collective performance. Then, we focus on how expanding the cardinality of a collective by taking into consideration the diversity. We will prove some theorems related to the positive impact of the diversity on the collective performance. With some restriction, the hypothesis *"the more diverse the collective, the higher the collective performance"* will be formally proved.