Neurocomputing 132 (2014) 85–96

Contents lists available at ScienceDirect

Neurocomputing

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

Binary social impact theory based optimization and its applications in pattern recognition

Martin Macaš [a,*], Amol P. Bhondekar [b], Ritesh Kumar [b], Rishemjit Kaur [b], Jakub Kuzilek [a], Václav Gerla [a], Lenka Lhotská [a], Pawan Kapur [b]

[a] Czech Technical University, Prague, Czech Republic
[b] Central Scientific Instruments Organisation, Chandigarh, India

ARTICLE INFO

Article history:
Received 5 October 2012
Received in revised form
20 February 2013
Accepted 8 March 2013
Available online 12 November 2013

Keywords:
Optimization
Feature selection
Swarm
Social impact

ABSTRACT

The human opinion formation can be understood as a social approach to optimization. In the real world, the opinions on different issues encode a "candidate solution", which is evaluated by a complex and unknown fitness function. The computer models of such processes can be easily modified by introducing a fitness value, which leads to novel family of optimization techniques. This paper demonstrates how the novel algorithms can be derived from opinion formation models and empirically demonstrates their usability in the area of binary optimization. Particularly, it introduces a general SITO algorithmic framework and describes four algorithms based on this general framework. Recent applications of these algorithms to pattern recognition in electronic nose, electronic tongue, new born EEG and ICU patient mortality prediction are discussed. Finally, an open source SITO library for MATLAB and JAVA is introduced.

1. Introduction

This paper connects computational psychology with mathematical optimization and proposes an optimization framework inspired by models of opinion formation (OF). The OF models are the sociopsychological tools that try to describe how individuals change their opinions under the influence of others. Various models have been proposed in the past. For instance, Axelrod's Culture Model [1] postulates that individuals who are similar to each other are likely to interact and become even more similar. Another example is the research on social influence that studies attitudes, social cognition and persuasion. For example, Latané's Dynamic Theory of Social Impact describes how the people are influenced by a group. A description of its computer simulation can be found in [2]. Also, a whole family of opinion formation models has been introduced by Galam [3].

Despite the main purpose of the OF models, there are some studies that noticed the analogy between the OF and optimization and tried to apply the models in the optimization domain. Kennedy and Eberhart showed that Axelrod's simple social principle of "imitation of your betters" is able to find its way through a complex search space and thus optimize an arbitrary fitness function [4]. The Particle Swarm Optimization (PSO) algorithm is inspired by the social behavior of animals [5] and it can also be understood as an imitation of human decision making processes. Recently, Shi [6] used a model of brainstorming scheme and created a new optimizer. Macaš et al. modified the Nowak's, Szamrej's and Latané's simulation [2] by introducing fitness function into its mechanism and found the ability of such modification to solve complex optimization problems [7]. The resulting method has been called the Social Impact Theory based Optimizer (SITO) and here it is referred to as an Original Distance based Social Impact Theory based Optimization (oSITO) algorithm, as it is one instance of a wide family of OF inspired optimizers. Apart from the general algorithmic framework of SITO, its variants and their performance on benchmark experiments have also been discussed here.

The main goal of this paper is to show the new direction in the research of the optimization metaheuristics – research on the OF based methods. To support this effort, we also summarize four novel and successful practical applications of the SITO methods. Most of these applications are from the area of feature subset selection, although in the first two applications, the underlying approaches are also adopted for optimization of feature extraction or classification model selection.

The structure of this paper is as follows. Section 2 describes two opinion formation models namely, Nowak–Szamrej–Latané (NSL) and Galam. Further, in Section 3, these models have been modified by the addition of the fitness function in order to be used in optimization domain. Also, the optimizer variants based on NSL

* Corresponding author. Tel.: +420 737853299.
E-mail address: mmacas@seznam.cz (M. Macaš).

0925-2312/$ - see front matter
http://dx.doi.org/10.1016/j.neucom.2013.03.063

models have been discussed. In Section 4, results obtained by the SITO variants and other well established optimizers on benchmark problems have been compared. Some most promising applications in the area of pattern recognition are summarized in Section 5. Further, an open source library of SITO, developed in MATLAB and JAVA, has been described in Section 6 followed by conclusions in Section 7.

2. Opinion formation models

The central idea of the underlying optimizers is the analogy between general optimization processes and processes of decision making in human societies. One of the possible outputs of any decision making process is an opinion and the process is referred to as OF. There are many OF models in computational psychology. In this section, the social impact based models, here called Nowak–Szamrej–Latané models (NSL), are described that are the central paradigm for the SITOs. We also discuss the Galam model [3] – a different OF model that even though not based on the social impact offers an alternative topology used in one of our novel algorithms.

2.1. Nowak–Szamrej–Latané models

A very important class of OF models has been developed by Andrzej Nowak, Jacek Szamrej and Bibb Latané. These NSL models are based on the notion of social impact. Social impact is any of the great variety of changes in physiological states and subjective feelings, motives, emotions, cognitions, beliefs, values and behavior that occur in an individual, human or animal, as a result of the real, implied, or imagined presence or actions of other individuals [8]. Latané's Dynamic Social Impact Theory [9] tries to describe and predict the diffusion of beliefs through social systems. It views society as a self-organizing complex system composed of interacting individuals each obeying simple principles of social impact. It states that the likelihood that a person will respond to social influence will increase with three factors – strength, immediacy and number.

Strength is a property of influencing individuals which says how important the influencing individual is to an influenced individual.
Immediacy represents the spatial closeness of the influencing individuals from the influenced individual.
Number describes how many individuals are influencing the current one.

This structure was first simulated by Nowak et al. [2]. The simulations represented each individual having four parameters: the individual's attitude, two indicators of *strength* (persuasiveness and supportiveness) and its location in the social structure. The organization of the simulation is as follows.

Each individual i holds a binary opinion $s_i \in \{0,1\}$. Each individual has assigned two strength factors: persuasiveness P_i, the ability to persuade people with opposing beliefs to change their minds, and supportiveness S_i, the ability to provide social support for people with similar beliefs. It was assumed in Nowak et al. [2] that $P_i \in \langle 0;100\rangle$ and $S_i \in \langle 0;100\rangle$ are reassigned randomly after each attitude change.

The concept of *immediacy* is established by organizing the individuals into a spatial grid of cells. The immediacy of two individuals i and j is represented as inverted Euclidean physical distance $1/\delta_{ij}$ between the corresponding cells. Moreover, each individual i has assigned its social neighborhood $\mathcal{N}_i$ – a set of individuals that can influence the individual i.

At each iteration, each individual combines partial social influences into one total social impact. Originally, each individual was able to influence each other. Thus, the social neighborhood was the whole population. At each iteration, the neighbors of i are divided into two disjoint subsets, persuaders $\mathcal{P}_i$ with opinion opposite to s_i and supporters $\mathcal{S}_i$ with the same value of the opinion. The total social impact I_i that ith individual experiences from its social neighborhood $\mathcal{N}_i$ is [10]

$$I_i = I_{Pi} - I_{Si}. \tag{1}$$

The first term of the right hand side of Eq. (1) represents the pressure in favor of the opinion change. It is the total persuasive impact I_{Pi} on a single individual i whose sources are individuals with different opinions:

$$I_{Pi} = g_{Ip}\left[\sum_{j \in \mathcal{P}_i} \frac{g_P(P_j)}{g_\delta(\delta_{ij})}\right], \tag{2}$$

where g_{Ip} denotes the form of the persuasive impact function, g_P is a persuasiveness scaling function and g_δ is some increasing function of the distance.

The second term of the right hand side of Eq. (1) represents the pressure to keep the current opinion. It is the total supportive impact on a single individual i whose sources are individuals sharing the opinion with individual i

$$I_{Si} = g_{Is}\left[\sum_{j \in \mathcal{S}_i} \frac{g_S(S_j)}{g_\delta(\delta_{ij})}\right], \tag{3}$$

where g_{Is} denotes the form of the supportive impact function and g_S is the supportiveness scaling function. The particular forms of the functions $g_P, g_S, g_{Ip}, g_{Is}, g_\delta$ differ from one NSL model to another. If the persuasive impact I_{Pi} is greater than the supportive impact I_{Si}, the total social impact I_i is positive, the influence of the individuals that hold different opinions is greater, and the individual i changes its opinion:

$$s_i(t+1) = \begin{cases} 1 - s_i(t) & \text{if } I_{Pi} > I_{Si} \ (i.e.\ I_i > 0), \\ s_i(t) & \text{otherwise}. \end{cases} \tag{4}$$

Additionally, an influence that can model some individual preference toward one of the opinions (i.e. an influence of some external source) can be introduced in the model. Although there are many types of NSL models [10], we present particular examples of the general model that do not assume any additional influence. These examples are as follows:

Example 1: The first simulations have been performed in [2]. The model has 2D rectangular grid, $g_\delta(\delta_{ij}) = \delta_{ij}^2$ for $i \neq j$, uniformly distributed persuasiveness and supportiveness, persuasiveness and supportiveness scaling functions $g_P(z) = z$, $g_S(z) = z$ and impact functions $g_{Ip}(z) = z/\sqrt{|\mathcal{P}|}$ and $g_{Is}(z) = z/\sqrt{|\mathcal{S}|}$ were used. The impacts described in Eqs. (2) and (3) then lead to

$$I_{Pi} = \frac{1}{\sqrt{|\mathcal{P}_i|}} \sum_{j \in \mathcal{P}_i} \frac{P_j}{\delta_{ij}^2} \tag{5}$$

and

$$I_{Si} = \frac{1}{\sqrt{|\mathcal{S}_i|}} \sum_{j \in \mathcal{S}_i} \frac{S_j}{\delta_{ij}^2} + \beta^2 S_i, \tag{6}$$

where the distance δ_{ii} would cause a singularity of the model, which is solved by redefining $\delta_{ii} \neq 0$ and establishing parameter $\beta = 1/\delta_{ii}$. This parameter is often called self-supportiveness. Moreover, we must define $I_{Si} = 0$ if $\mathcal{S}_i = \varnothing$ and $I_{Pi} = 0$ if $\mathcal{P}_i = \varnothing$.

Nowak et al. [2] observed that "several macrolevel phenomena emerged from the simple operation of the