Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 120 (2017) 104–111

Procedia Computer Science

www.elsevier.com/locate/procedia

9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 24-25 August 2017, Budapest, Hungary

Synthesis of new materials by using fuzzy and big data concepts

Babanli M. B.[a]*

[a]*Azerbaijan State Oil and Industry University, Baku, Azerbaijan Republic*

Abstract

Classical approaches to material synthesis rely on experimental results obtained by academicians and practitioners during a long period. However, a huge amount of data containing complex relationships mandates the use of computer-guided methods of material synthesis. Big data driven knowledge based models and other computational schemes allows to alternate hard experimental works. Such approaches provide ability of systematic and computationally effective analysis in order to predict composition, structure and related properties of new materials.

In this paper we apply the fuzzy set theory to knowledge mining from big data on material characteristics. We propose fuzzy clustering-generated If-Then rules as a basis for computer synthesis of new materials. Validity of the proposed approach is verified on an example of prediction properties of Ti-Ni alloy.

Peer-review under responsibility of the scientific committee of the 9th International Conference on Theory and application of Soft Computing, Computing with Words and Perception.

Keywords: Material engineering; big data; fuzzy number; if-then rules; alloy.

1. Introduction

Development new materials is one of important tasks of theoretical and practical interest. Traditionally, this task is implemented mainly on the basis of intensive (and sometimes 'ad-hoc') experiments which are time and recourse consuming, or even not practically implementable. Nowadays, it is well-understood that more systematic and effective approaches are needed which are based on computer-guided synthesis of materials. Such approaches rely on data-driven mathematical models and knowledge base obtained from big data previously collected during

* Corresponding author.
E-mail address: mustafababanli@yahoo.com

1877-0509
Peer-review under responsibility of the scientific committee of the 9th International Conference on Theory and application of Soft Computing, Computing with Words and Perception.
10.1016/j.procs.2017.11.216

intensive experiments. In view of this, Nicola Nosengo studied the use of computers and databases for developing new materials in his paper published in Nature (Nosengo, 2016). Nosengo overviews the present state-of-the art of this field. He notes that scientists deal with using databases on existing materials and their basic properties. Such databases would help to infer information about new materials with predefined properties. For example, a team at Haverford College proposed an approach to develop algorithms for computerized predictions of reaction outcomes to obtain new materials.

In this stream, a very important paradigm to be used is machine learning. Paul Raccuglia and colleagues (Racuglia et al. (2106)) considered machine-learning algorithms to discover new hybrid materials. One the insight is that the field of compound materials is a large undiscovered field with lack of systematic knowledge. As the authors note, machine-learning algorithms are good alternative to trial-and-error techniques, it allows o handle more systematically big experimental data to uncover knowledge on important relationship between material compositions and related properties. The proposed study involves the use of cheminformatics techniques for property descriptions. The obtained results show that the proposed approach outperforms classical human strategies and allows inferring new hypotheses in the considered phenomena.

In another paper published in Nature they propose an algorithm to computational analysis of chemical reactions to uncover synthesizable materials (Collins et al. (2017)). The algorithm allows to analyze a large diversity of combinations of molecules and atoms.

One important issue related to big data based computerized material synthesis is that experimental data are not perfect. They include measurement errors, partially reliable information, imprecise evaluations, incomplete information on experiment conditions, qualitative information etc. Imperfect information and complexity of relation between composition and properties creates problems for classical computation based methods and mandates the use of soft computing approaches for material synthesis. A systematic review of the use of soft computing approaches for material engineering is conducted by Odejobi and Umoru (2009). The main paradigms are fuzzy logic (Zadeh (1975), Aliev (2001)) and artificial neural networks (Pedrycz and Peters J.F. (1998)). It is shown that the use of fuzzy logic for properties modeling and prediction helps to express development process better and interpret results easier (Babanli and Huseynov (2016), Chen (1997), Cheng et al. (2008)). Sometimes, it allows to obtain more information than artificial neural networks do. In turn, artificial neural networks have significantly better learning abilities which are important when dealing with huge variants of complex material compositions (Col et al. (2007)).

The use of hybrid soft computing approaches (Hancheng et al. (2002), Chen et al. (2003)) due to combination of advantages of different paradigms may potentially provide better results for material engineering. The strategies applied are ANN-fuzzy models (57% works on material engineering, fuzzy-genetic and neuro-genetic (18%) and neuro-fuzzy-genetic (7%) approaches. Odejobi and Umoru (2009) introduce systematic basics for formalizing methodology of SC application to ME.

In this paper we propose fuzzy If-Then rules based model to predict properties of new materials. The model is constructed on the basis of fuzzy clustering of big data on dependence between material composition and related properties. The motivation to use fuzzy model is inspired by necessity to construct an intuitively well-interpretable development strategy from imperfect and complex data.

The paper is organized as follows. In Section 2 we provide the necessary concepts used in the paper including a fuzzy number, fuzzy clustering, fuzzy If-Then rules, big data etc. Section 3 is devoted to a problem of computer synthesis of new materials on the basis of fuzzy If-Then rules. In Section 4 a solution approach for the problem stated in Section 3. In Section 5 we consider application of the proposed approach to synthesis of Ti-Ni alloy with predefined properties. Section 6 concludes.

2. Preliminaries

Definition 1. Type–1 Continuous fuzzy number (Aliev, 2001). A fuzzy number is a fuzzy set A on $\mathcal{R}$ which possesses the following properties: a) A is a normal fuzzy set; b) A is a convex fuzzy set; c) α-cut of A , A^{α} is a closed interval for every $\alpha \in (0,1]$; d) the support of A, $\text{supp}(A)$ is bounded.

Definition 2. Big data. Big data provide ability to infer laws, knowledge from large sets of data with low information specificity to perform prediction of outputs.