Renewable Energy 62 (2014) 484–489

Contents lists available at ScienceDirect

Renewable Energy

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

Semi-automatic construction of a domain ontology for wind energy using Wikipedia articles

Dilek Küçük [a,*], Yusuf Arslan [b]

[a] Power Electronics Department, TÜBİTAK Energy Institute, 06531 Ankara, Turkey
[b] Department of Computer Engineering, Middle East Technical University, 06531 Ankara, Turkey

ARTICLE INFO

Article history:
Received 13 January 2013
Accepted 11 August 2013
Available online 4 September 2013

Keywords:
Wind energy
Wind power
Domain ontology
Text mining
Web mining

ABSTRACT

Domain ontologies are important information sources for knowledge-based systems. Yet, building domain ontologies from scratch is known to be a very labor-intensive process. In this study, we present our semi-automatic approach to building an ontology for the domain of wind energy which is an important type of renewable energy with a growing share in electricity generation all over the world. Related Wikipedia articles are first processed in an automated manner to determine the basic concepts of the domain together with their properties and next the concepts, properties, and relationships are organized to arrive at the ultimate ontology. We also provide pointers to other engineering ontologies which could be utilized together with the proposed wind energy ontology in addition to its prospective application areas. The current study is significant as, to the best of our knowledge, it proposes the first considerably wide-coverage ontology for the wind energy domain and the ontology is built through a semi-automatic process which makes use of the related Web resources, thereby reducing the overall cost of the ontology building process.

1. Introduction

Renewable energy resources (such as wind, solar, and biomass, among others) have emerged as an important engineering topic mainly due to the profound advantages of the corresponding plants over conventional plants (such as fossil-fueled and nuclear). Considering wind energy, the primary benefits of the wind energy plants are that (i) energy production does not lead to environmental pollution and (ii) as wind is abundant in many countries, the financial costs of these plants are comparably lower than that of the conventional plants [1]. Compared to the other renewable resources, wind is considerably variable and hence it is hard to determine optimal places for wind power plant installations and to forecast daily energy productions of the plants precisely. The latter point (the uncertainty of the production) results in particular problems during the integration of these plants to the electrical grid as the electricity transmission system operators should be informed of prospective daily productions in advance to make short-term planning and to ensure that the supply matches the demand [1]. Other directions of research on the topic include predictions for the economic viability of wind power plant projects and wind turbine design [2]. As wind energy keeps increasing its share in the electricity generation (and annually growing at a rate of 30% [3]), research projects targeting at these problems have started to increase (such as [4]).

In this paper, we target at wind power applications from a knowledge-based perspective and propose a wide-coverage domain ontology for wind energy domain.[1] A domain ontology is defined as a reusable vocabulary of concepts, relationships, and activities in the domain along with the governing theories and principles [5]. Various domain ontologies have been reported in the literature [6–10] and there is ongoing research on the topic along with research on the Semantic Web. Although domain ontologies serve various purposes, building them from scratch is too labor-intensive and time-consuming. Hence, we employ a semi-automatic approach to build our wind energy ontology which comprises two main phases: in the first phase, Wikipedia [11] articles on the topic are processed to learn the concepts from highly frequent words and phrases (n-grams) and in the second phase, these concepts are organized with their properties and

* Corresponding author. Tel.: +90 312 2101310; fax: +90 312 2101315.
E-mail addresses: dilek.kucuk@tubitak.gov.tr, dilekkucuk@gmail.com (D. Küçük), yusuf.arslan@ceng.metu.edu.tr (Y. Arslan).

[1] In this study, we utilize the terms *wind power* and *wind energy* interchangeably to denote the electrical power generated from wind through the wind turbines installed on the wind power plants (as they are commonly used interchangeably and they denote wind generated power), although the terms *power* and *energy* are quite distinct terms in physics.

0960-1481/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.renene.2013.08.002

relationships to arrive at the ultimate ontology. Yet, the final form of the ontology is open to extensions as well, that is, other concepts and properties covering more details about the domain can be added to the ontology. It can also be integrated with other related engineering ontologies for better coverage of the domain.

Ontology-based studies on wind power domain are rather limited. We come across only three studies on the topic, the first of which [12] proposes an ontology on wind power plant information and the resulting ontology comprises some generic concepts which are not specific to wind power domain and moreover it basically aims to cover concepts related to the management of wind power plant information. The second study [13] describes an ontology for wind turbines' condition monitoring and is more domain-specific compared to the ontology proposed in Ref. [12]. Yet this study [13] also suffers from low coverage, i.e., it only considers the semantics of wind turbines instead of the whole wind power domain. The last related study [14] utilizes an ontology of facts to be utilized by a multi-agent system which aims to control a photovoltaic (PV)-based microgrid. This ontology covers specific facts or concepts related to agent operations which are far from representing the domain of renewable energy or wind energy. Hence, to the best of our knowledge, in the current study, we propose the first large-scale ontology for the wind power domain. Yet, as will be emphasized in Section 3, this ontology can well be integrated with related engineering ontologies and aligned with the aforementioned ontologies including [12,13].

The rest of the paper is organized as follows: In Section 2, the ontology building process for wind energy domain is presented in details. Section 3 is devoted to prospective integration opportunities with related engineering ontologies and plausible application areas of the proposed ontology. Finally, Section 4 concludes the paper and provides pointers to future work.

2. Building an ontology for wind energy

Our semi-automatic ontology construction process for wind energy domain comprises two main phases:

- Learning the concepts and properties of these concepts from related Wikipedia articles by extracting the frequent phrases as unigrams, bigrams, and trigrams from the article texts. This phase executes in fully automated mode.
- Building the ontology by organizing these components and properties utilizing other written information sources like related international standards and textbooks.

We provide the execution flow of the overall approach with more detailed steps in Fig. 1. In the figure, the resources (input and output) are shown as rectangular shapes, while the functional steps

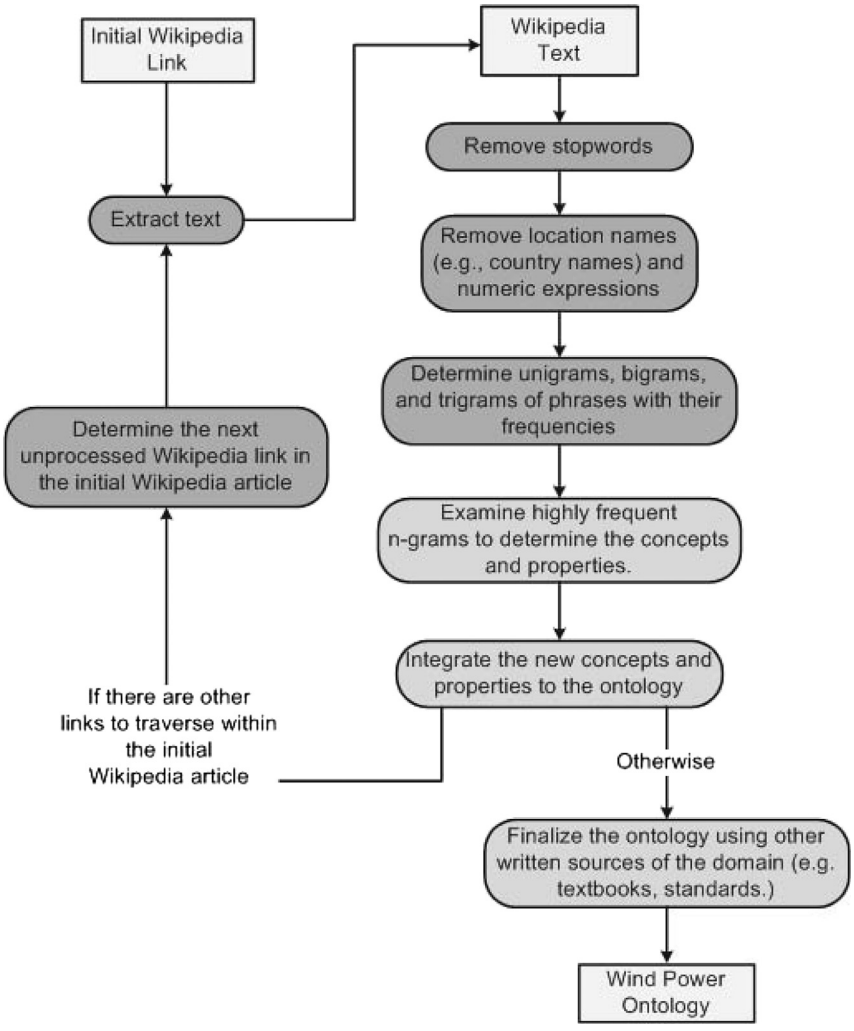

Fig. 1. Execution flow of the employed ontology building approach.