Applied Energy 206 (2017) 1022–1039

Contents lists available at ScienceDirect

Applied Energy

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

# Energy prediction using spatiotemporal pattern networks

Zhanhong Jiang[a], Chao Liu[a], Adedotun Akintayo[a], Gregor P. Henze[b,c], Soumik Sarkar[a,*]

[a] Department of Mechanical Engineering, 2025 Black Engineering, Iowa State University, Ames, IA 50011, USA
[b] Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA
[c] National Renewable Energy Laboratory, Golden, CO 80401, USA

## HIGHLIGHTS

- Novel data-driven spatiotemporal pattern network (STPN) to predict energy production/consumption.
- xD-Markov machine learnt to capture causal dependencies between dynamic sub-systems.
- Validated by wind turbine power prediction and non-intrusive load monitoring (NILM).
- STPN captures salient spatiotemporal features and achieves high-accuracy prediction.
- STPN and STPN plus convex programming outperform state-of-the-art techniques in NILM.

## ARTICLE INFO

*Keywords:*
Wind power
Symbolic dynamical filtering
Spatiotemporal pattern network
Probabilistic finite state automata
NILM

## ABSTRACT

This paper presents a novel data-driven technique based on the spatiotemporal pattern network (STPN) for energy/power prediction for complex dynamical systems. Built on symbolic dynamical filtering, the STPN framework is used to capture not only the individual system characteristics but also the pair-wise causal dependencies among different sub-systems. To quantify causal dependencies, a mutual information based metric is presented and an energy prediction approach is subsequently proposed based on the STPN framework. To validate the proposed scheme, two case studies are presented, one involving wind turbine power prediction (supply side energy) using the Western Wind Integration data set generated by the National Renewable Energy Laboratory (NREL) for identifying spatiotemporal characteristics, and the other, residential electric energy disaggregation (demand side energy) using the Building America 2010 data set from NREL for exploring temporal features. In the energy disaggregation context, convex programming techniques beyond the STPN framework are developed and applied to achieve improved disaggregation performance.

## 1. Introduction

Energy prediction problems are essential for operating, monitoring, and optimizing (in terms of efficiency and cost) diverse energy systems, from the supply side (e.g., wind energy, solar energy, power systems, storage) to the demand side (e.g., load monitoring, usage of electric vehicles, building energy management). Numerous studies are being carried out in terms of predicting energy generation/consumption using time-series data [1–6]. For instance, Kalman filtering, wavelet packet transforms, and least squares support vector machines are used to predict wind power performance [4,5], while an analog ensemble method is applied to forecast solar power [3]. Liu et al. [2] predicts remaining state of charge of electric vehicle batteries based on predictive control theory. Hybrid genetic algorithms and Monte Carlo simulation approaches are applied to predict energy generation and consumption in net-zero energy buildings [6]. For modern energy systems, a large number of subsystems is usually involved, for example, hundreds of wind turbines are closely collocated in a wind farm where the wind resource is similar and the conditions of each are analogous in terms of the power transmission to the power system. As a result, there is a relationship among the wind turbine outputs, and the characteristics of their spatial interactions can be potentially applied for prediction [7] and design optimization. The prediction approaches discussed above can be viewed as methods of exploring temporal relationships. Spatial and temporal relationship widely exists in energy systems [8–11], yet spatiotemporal features are less commonly leveraged for energy prediction problems. The exploration of such spatiotemporal features has been shown to be efficient in wind speed forecasting problems [12,10,13].

To facilitate energy prediction for systems with both spatial and

* Corresponding author.
*E-mail address:* soumiks@iastate.edu (S. Sarkar).

http://dx.doi.org/10.1016/j.apenergy.2017.08.225
Received 18 January 2017; Received in revised form 31 July 2017; Accepted 27 August 2017

temporal characteristics, probabilistic graphical models (PGM) may be employed as the spatiotemporal interactions are naturally suited for graph representation and can be evaluated by the associated probabilities. PGM encompasses a variety of models described by conditional dependence structures, so-called graphs, including Bayesian networks and undirected/directed Markov networks, and can be used to deal with dynamical systems and relational data [14]. Bayesian networks are a type of PGM that capture causal relationships using directed edges [14], where the overall joint probability distribution of the network nodes (variables) is computed as a product of the conditional distributions (factors) defined by the nodes in the network. However, prediction problems are not straightforward for Bayesian networks, as they only encode node-based conditional probabilities, and the approximation of the joint distribution using node-based structures is often intractable [15]. This is because a certain directed acyclic graphical structure may not allow for easy and exact computation of certain probabilities related to inference questions.

Markov models, as a class of statistical models, have been widely applied to different domains, e.g., natural language processing and speech recognition [16]. These models are shown to be efficient in identifying the probabilistic dependencies among random variables in both a directed and undirected manner. Hidden Markov Models (HMMs) have been particularly successful for learning temporal dynamics of an underlying process [17]. Several modifications for HMMs have been proposed, such as integrated HMM (IHMM) [18] which integrates several parameters into three hyper-parameters to model countably infinite hidden state sequences. Integrated hierarchical HMM (IHHMM) [19] extends HMMs to an infinite number of hierarchical levels, and [20] applied a forward-backward algorithm to reduce model complexity through the order of operations. However, Markov Models with hidden states usually rely on iterative learning algorithms that may be computationally expensive. To alleviate such issues, symbolic dynamic filtering (SDF) was proposed [21,22] based on the concepts of symbolic dynamics and probabilistic finite state automata (PFSA). Several improvements related to coarse graining of continuous variables [23], state splitting and merging techniques for PFSA [24], efficient inference algorithms [25], and hierarchical model learning [26] have been proposed over the last decade within the SDF framework. SDF has been shown to be extremely efficient for anomaly detection and fault diagnostics of various complex systems, such as gas turbine engines [27], shipboard auxiliary systems [28], nuclear power plants [29], coal gasification systems [30] and bridge monitoring processes [31].

For the purpose of addressing prediction problems in energy systems, this work presents a new data-driven framework (namely spatiotemporal pattern networks, or STPN) to leverage the spatiotemporal interactions of energy systems for prediction. Built on SDF, a STPN aims to capture the spatiotemporal characteristics of complex energy systems, and implements prediction at both spatial and temporal resolutions. For validation purpose, the proposed approach is evaluated on two representative case studies. The first is taken from the energy supply side, wind power prediction in a large-scale wind farm. The second case study is from the energy demand side, energy disaggregation (also as non-intrusive load monitoring (NILM), a well-established problem that involves disaggregating the total electrical energy consumption of a building into its constituent load components without the necessity for extensive metering installations on individual household or appliances [32–34]).

The main reason for choosing both an energy production system and the non-intrusive load monitoring problem on the demand side, is to demonstrate that our proposed method is extremely effective on both sides of the energy meter. Note that as penetration of renewable energy systems increases, prediction accuracy becomes ever more important. This is because without accurate prediction of renewable energy production, it is difficult to control the power distribution, pricing and scheduling of other energy sources. We need insight into the electric load breakdown as well, in order to perform effective demand response and load shaping for peak power reduction. Furthermore, if inexpensive energy disaggregation approaches are widely deployed, we will obtain actionable spatiotemporal information on the types of load components that could respond to local overproduction of renewable energy such as wind power.

**Contributions**: We propose a novel probabilistic graphical modeling framework that can capture causal dependencies (in the sense of Granger causality) among different sub-systems in a large distributed system. The main contribution is that we demonstrate that the proposed data-driven modeling scheme can efficiently learn spatiotemporal characteristics of a distributed energy system in a scalable and computationally efficient manner. The modeling scheme can enable high-accuracy prediction of energy production (for a distributed generation system such as wind farm) and energy consumption (for a complex combination of electrical energy end uses in a building). For wind turbine power prediction, the spatiotemporal characteristics between different wind turbines are identified, while for home energy disaggregation the complex coupled temporal features are revealed. A STPN-based convex programming method is presented in this work in order to improve energy prediction and disaggregation performance. We also compare the performance of our proposed algorithm with other competitive and state-of-the-art data-driven modeling techniques, which clearly demonstrates the significant improvement in accuracy. While energy prediction is critical, the data-driven modeling strategy also opens up many other applications such as performance monitoring, fault diagnostics, control, and optimization in many large energy systems that are difficult to model using traditional physics-based principles.

The remaining sections are outlined as follows. In Section 2, the necessary background of SDF is presented as well as the concepts of a *D*-Markov machine. The prediction approach based on STPN is given in Section 3, and two typical case studies, i.e., supply side (wind turbines) and demand side (NILM), for validating the proposed framework are presented in Sections 4 and 5, respectively. In Section 6, conclusive remarks and future research directions beyond the existing results are offered.

## 2. Symbolic dynamical filtering and *D*-Markov machines

This section gives an essential background on symbolic dynamical filtering necessary to characterize the proposed prediction method. We refer interested readers to [23] for more details. SDF is built upon the relevant concepts of discrete dynamical systems in which discretization and symbolization are critical steps to convert observed continuous data into discrete symbol sequences. Therefore, dynamical systems can be studied in deterministic or probabilistic settings in terms of symbolic space by using language-theoretic approaches, e.g., shift-maps and sliding block codes. The simplest approaches for partitioning are the uniform partitioning and maximum entropy, where these two methods were mainly applied to simple dynamical systems with data of less variance. The state-of-the-art partitioning or discretization approaches include symbolic false nearest neighbor partitioning (SFNNP) [35], wavelet transform [23], and Hilbert-transform-based analytic signal space partitioning (ASSP) [36]. Recently, a supervised partitioning scheme, i.e., maximally bijective discretization (MBD) [23] has been proposed for modeling and analyzing complex dynamical systems. Unlike the other methods, MBD is able to maximally preserve the input-output relationship originating from the continuous domain after discretization in dynamical systems.

After discretization of time-series data in the continuous domain, symbolization is conducted subsequently to establish the *D*-Markov machines. For SDF, a critical assumption is that we can approximate any symbol sequence generated by time series data as a Markov chain of order *D* (which is a positive integer). Therefore, such a Markov chain is called *D*-Markov machine, which is used to establish the model for