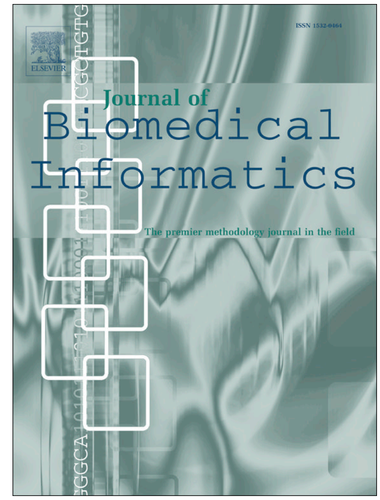

Modeling Asynchronous Event Sequences with RNNs

Stephen Wu, Sijia Liu, Sunghwan Sohn, Sungrim Moon, Chung-il Wi, Young Juhn, Hongfang Liu

PII: S1532-0464(18)30099-6
DOI: https://doi.org/10.1016/j.jbi.2018.05.016
Reference: YJBIN 2986

To appear in: *Journal of Biomedical Informatics*

Received Date: 26 June 2017
Revised Date: 10 May 2018
Accepted Date: 26 May 2018

Please cite this article as: Wu, S., Liu, S., Sohn, S., Moon, S., Wi, C-i., Juhn, Y., Liu, H., Modeling Asynchronous Event Sequences with RNNs, *Journal of Biomedical Informatics* (2018), doi: https://doi.org/10.1016/j.jbi.2018.05.016

Modeling Asynchronous Event Sequences with RNNs

Stephen Wu, Sijia Liu, Sunghwan Sohn, Sungrim Moon, Chung-il Wi, Young Juhn, Hongfang Liu

Abstract

Sequences of events have often been modeled with computational techniques, but typical preprocessing steps and problem settings do not explicitly address the ramifications of timestamped events. Clinical data, such as is found in electronic health records (EHRs), typically comes with timestamp information. In this work, we define event sequences and their properties: *synchronicity*, *evenness*, and *co-cardinality*; we then show how asynchronous, uneven, and multi-cardinal problem settings can support explicit accountings of *relative time*. Our evaluation uses the temporally sensitive clinical use case of pediatric asthma, which is a chronic disease with symptoms (and lack thereof) evolving over time. We show several approaches to explicitly incorporate relative time into a recurrent neural network (RNN) model that improve the overall classification of patients into those with *no asthma*, those with persistent *asthma*, those in long-term *remission*, and those who have experienced *relapse*. We also compare and contrast these results with those in an inpatient intensive care setting.

1. Introduction

Sequences of events have often been modeled with techniques such as Hidden Markov Models (HMMs), Conditional Random Fields (CRFs), Recurrent Neural Networks (RNNs), and their derivatives. Here, we zoom in on the typical temporal assumptions being made in these types of models, and provide evaluations on a use case of medical information (asthma diagnoses) that could intuitively benefit from more explicit accountings of event timing.

We will define event sequences and differentiate between *synchronous* v. *asynchronus*, *co-cardinal* v. *multi-cardinal*, and *even* v. *uneven* pairs of such sequences (see §3.1). In particular, where most methods transform event sequences to be synchronous, co-cardinal, and even, we will consider the ramifications of using event sequences that are either synchronous, co-cardinal, but *un*even; or synchronous, *multi*-cardinal, and *un*even.

To evaluate the representational power and computational tractability

Version: 1.20
Date: April 6, 2018
Contact: wu.stephen.t@gmail.com