Cascade Dynamics Modeling with Attention-based Recurrent Neural Network

Abstract

It is abstract

1 Introduction

The emergence of Social Media platform has revolutionized dissemination of information via its great ease in inforamtion delivery, accessing and filtering. In Social Media, pieces of information, posted by users spontaneously, propagate along social relationships between users, explict or implict, forming cascade dynamics. Modeling cascade dynamics is the fundamental to understand information propagation and launch series of social applictions, i.e., viral marketing, popularity prediction and rumor detection.

The key to cascade dynamics modeling is to find a well-defined function in hypothesis space based on observed cascades. Existed methods for this problem fall into three main paradigms: pairwise, nodewise and eventwise modeling. The majority works in cascade dynamics modeling focus on pairwise modeling, defining the propagation probability of information between all pairs of users [Saito *et al.*, 2008; Goyal *et al.*, 2010; Gomez-Rodriguez *et al.*, 2013]. However, pairwise models suffer severe overfitting and overrepresentation problems especially in sparse social data, proved in [Wang *et al.*, 2015]. Nodewise modeling learn latent user-specific characteristics instead of pairwise manners, effectively combating overfitting and overrepresentation problems. Bourigault et.al. [Bourigault *et al.*, 2016] learn user-specific latent space in Independent Cascade (IC) model. Wang et.al [Wang *et al.*, 2015] capture users' influence and susceptibility in latent space and define propagation probability according to users latent characteristics. Kurashima et.al. [Kurashima *et al.*, 2014] embeded users into low-dimensional visualization space in Continous Time Independent Cascade (CTIC) model. But nodewise models require strong prior knowledge on generation processes of cascades in order to better fit the observations. Recently, eventwise models received great success in modeling sequence data.

Eventwise methods aims to learn history embedding in order to model the generation of next event, e.g., cascade. Manavoglu [Manavoglu *et al.*, 2003] propose users behavior generation method based on maxent and Markov mixture model. Recently, the efficient way of eventwise modeling can be achieved by Recurrent Neural Network (RNN) [Bengio *et al.*, 2003; Goldberg and Levy, 2014; Mikolov *et al.*, 2010; Sundermeyer *et al.*, 2012]. Du et.al. [Du *et al.*, 2016] proposed a Recurrent Marked Temporal Point Process (RMTPP) for event streams. The outputs of hidden layer in Recurrent Neural Network (RNN) represents the embedding of the event histories, then parameterizing the random process. The benefits of eventwise modeling are two folds: 1) avoiding strong prior knowledge on models and networks with respect to different observed cascades; 2) enlarging the functional space when searching the optimal cascade dynamics models, which may have great probability to better model cascade dynamics.

Despite of advantages in eventwise modeling, the traditional sequence models may meet "crossing dependency" problem in cascade dynamics modeling. The crossing dependencies problem is mainly caused by tree structure of propagation. Fig. 1 shows two typical crossing dependency cases in practical. For modeling dependence between 1st and 3rd event, we must use redundant information passing from 2nd event, called "redundant modeling". If we abandon useful information inherited by 3rd event when modeling the 4th event, the generation of 5th event would lose useful information from 3rd event, called "cut-off modeling". Crossing dependency problems limit the efficiency of sequence modeling.

In this paper, we propose a **C**ascade d**Y**namics modeling with **A**ttentio**N**-based RNN, named (CYAN-RNN). We construct a pooling layer above the output of hidden layer in RNN, aggregating event embedding in history. The weights in pooling layer pointing to each historical event embedding refers to connections between current event and history. We choose attention mechanism [Bahdanau *et al.*, 2014] to realize the pooling layer, automatically learned the connection weights. The benifits of our proposed model are three-fold: 1) We propose a eventwise method, using sequence modeling, for cascade dynamics modeling; 2) We point out crossing dependency problem in traditional sequence modeling when model cascade dynamics. Thus, we proposed CYAN-RNN to solve the problem; 3) We conduct experiments on synthetic and real-world datasets to show that our model consistently outperform than previous modeling methods in cascade dynamics modeling.

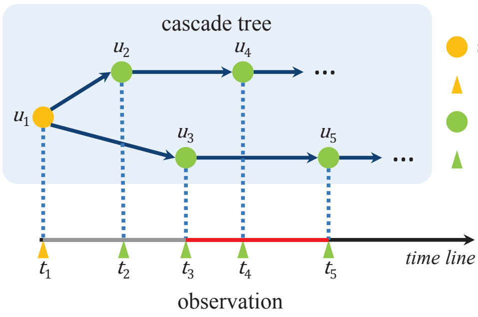


Figure 1: Tree structure of propagation and crossing dependency problems in sequence modeling. For modeling dependence between 1st and 3rd event, we must use redundant information passing from 2nd event, called "redundant modeling". If we abandon useful information inherited by 3rd event when modeling the 4th event, the generation of 5th event would lose useful information from 3rd event, called "cut-off modeling".

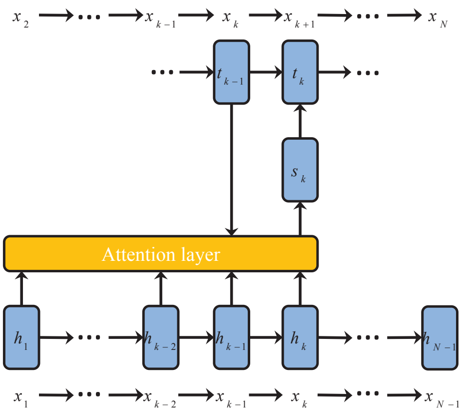
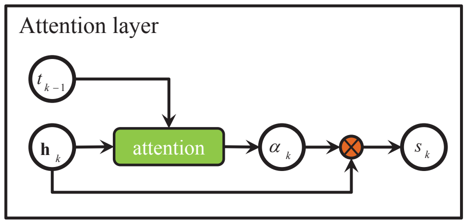
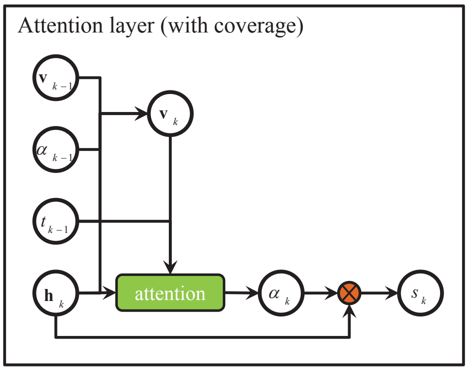


Figure 2:



(a)



(b)

Figure 3:

2 Model

References

[Bahdanau *et al.*, 2014] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

[Bengio *et al.*, 2003] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. A neural probabilistic language model, 2003.

[Bourigault *et al.*, 2016] Simon Bourigault, Sylvain Lamprier, and Patrick Gallinari. Representation learning for information diffusion through social networks: an embedded cascade model. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 573–582. ACM, 2016.

[Du *et al.*, 2016] Nan Du, Hanjun Dai, Rakshit Trivedi, Utkarsh Upadhyay, Manuel Gomez-Rodriguez, and Le Song. Recurrent marked temporal point processes: Embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 1555–1564, New York, NY, USA, 2016.

[Goldberg and Levy, 2014] Yoav Goldberg and Omer Levy. word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method. *arXiv preprint arXiv:1402.3722*, 2014.

[Gomez-Rodriguez *et al.*, 2013] Manuel Gomez-Rodriguez, Jure Leskovec, and Bernhard Schölkopf. Modeling information propagation with survival theory. In *Proceedings of the 30th International Conference on Machine Learning*, pages 666–674, 2013.

[Goyal *et al.*, 2010] Amit Goyal, Francesco Bonchi, and Laks VS Lakshmanan. Learning influence probabilities in social networks. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 241–250. ACM, 2010.

[Kurashima *et al.*, 2014] Takeshi Kurashima, Tomoharu Iwata, Noriko Takaya, and Hiroshi Sawada. Probabilistic latent network visualization: Inferring and embedding diffusion networks. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1236–1245, New York, NY, USA, 2014. ACM.

[Manavoglu *et al.*, 2003] Eren Manavoglu, Dmitry Pavlov, and C. Lee Giles. Probabilistic user behavior models. In *IEEE International Conference on Data Mining*, pages 203–210, 2003.

[Mikolov *et al.*, 2010] Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernockỳ, and Sanjeev Khudanpur. Recurrent neural network based language model. In *INTERSPEECH*, volume 2, page 3, 2010.

[Saito *et al.*, 2008] Kazumi Saito, Ryohei Nakano, and Masahiro Kimura. Prediction of information diffusion probabilities for independent cascade model. In *Knowledge-Based Intelligent Information & Engineering Systems*, pages 67–75, 2008.

[Sundermeyer *et al.*, 2012] Martin Sundermeyer, Ralf Schlüter, and Hermann Ney. Lstm neural networks for language modeling. In *INTERSPEECH*, pages 194–197, 2012.

[Wang *et al.*, 2015] Yongqing Wang, Huawei Shen, Shenghua Liu, and Xueqi Cheng. Learning user-specific latent influence and susceptibility from information cascades, 2015.