Learning the Joint Representation of Heterogeneous Temporal Events for Clinical Endpoint Prediction

Ming Zhang
School of Electronics Engineering and Computer Science
mzhang@net.pku.edu.cn
http://net.pku.edu.cn/dlib/mzhang

北京大学

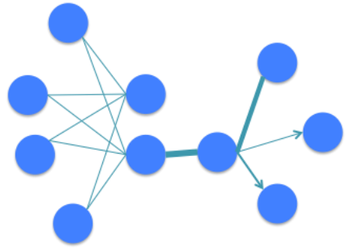
http://net.pku.edu.cn/dlib/

http://csrankings.org/

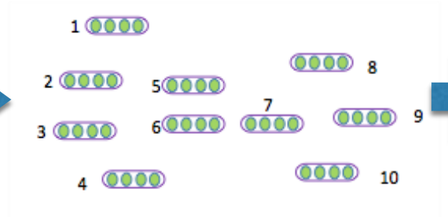
Ongoing Projects

- Co-PI, Multi-source and Heterogeneous Big Data Fusion Method Research Based on Whole Process Smart Health Management Decision, NSFC Key project, 2017-2020.
- PI, Machine Learning Models Based on Knowledge Graph and Deep Neural Network, Beijing Science and Technology Commission, 2018-2019.
- PI, A Knowledge Graphs Assisted General Framework to Construct Automatic Human-Computer Dialogue Systems for Vertical Domains, NSFC, 2018-2021
- PI, Research on user retention in massive open online courses, NSFC, 2015-2018

Learning Representations of Large-Scale Networks



Network



Node representations

- Node Classification
- Node Clustering
- Link Prediction
- Recommendation
- …

- E.g., Facebook social network -> user representations (features)-> friend recommendation

High-dimensional Data

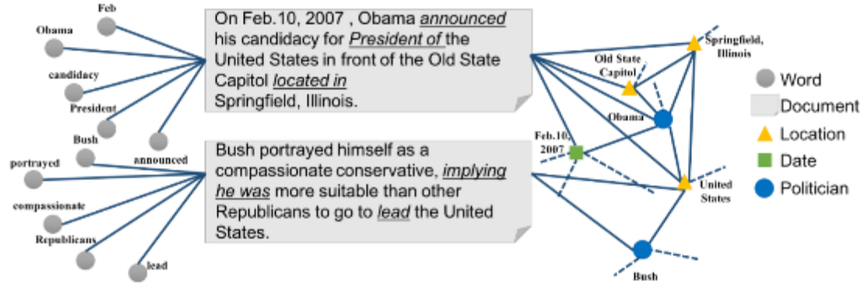
Networks

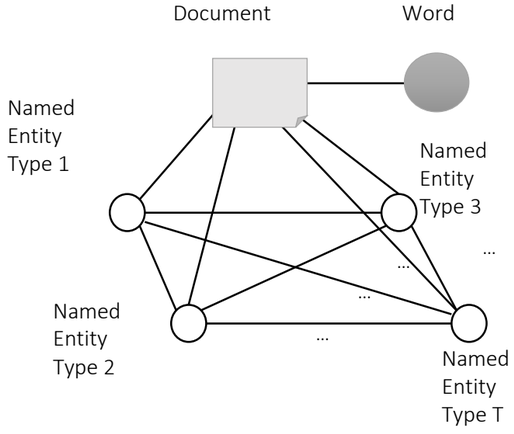
2D/3D Layout

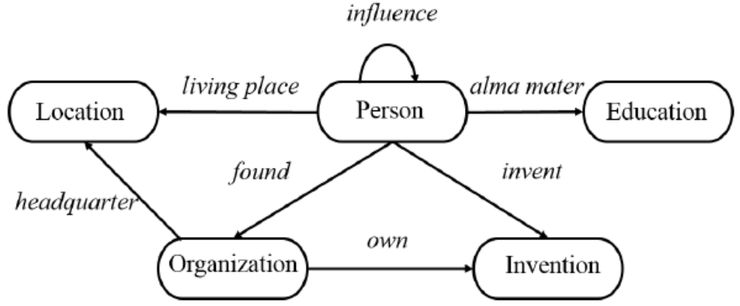
Network Embedding

- Jian Tang, Meng Qu, Mingzhe Wang, **Ming Zhang**, Jun Yan, Qiaozhu Mei. LINE: Large-scale Information Network Embedding. WWW 2015, 1067-1077. **citations 747**
- Jian Tang， Zhaoshi Meng， XuanLong Nguyen， Qiaozhu Mei， **Ming Zhang**， Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis， The 31st International Conference on Machine Learning (**ICML2014**)， **Best paper**， 2014.6.21-2014.6.26， **citations 123**
- Jian Tang， Jingzhou Liu， **Ming Zhang**， Qiaozhu Mei， Visualizing Large-scale and High-dimensional Data， **WWW 2016**. 04.11-2016.04.15， **Best paper runner-up， citations 92**
- Jian Tang, Ming Zhang, Qiaozhu Mei: **One theme in all views: modeling consensus topics in multiple contexts.** KDD 2013: 5-13， **citations 39**
- Meng Qu, Jian Tang, Jingbo Shang, Xiang Ren, Ming Zhang, Jiawei Han: **An Attention-based Collaboration Framework for Multi-View Network Representation Learning.** CIKM 2017: 1767-1776， **citations 8**

World Knowledge Representation： Heterogeneous Information Network (HIN)







Incorporating World Knowledge to Heterogeneous Information Networks

[1] Chenguang Wang, Yizhou Sun, Yanglei Song, Jiawei Han, Yangqiu Song, Lidan Wang, and **Ming Zhang**: RelSim: RelSim: Relation Similarity Search in Schema-Rich Heterogeneous Information Networks. Proc. 2016 SIAM Int. Conf. on Data Mining **(SDM'16)**.**citations 11**

[2] Chenguang Wang, Yangqiu Song, Haoran Li, **Ming Zhang**, and Jiawei Han: Text Classification with Heterogeneous Information Network Kernels. Proc. 2016 AAAI Conf. on Artificial Intelligence **(AAAI'16)**.**他34**

[3] Chenguang Wang, Yangqiu Song, Haoran Li, **Ming Zhang**, and Jiawei Han: KnowSim: A Document Similarity Measure on Structured Heterogeneous Information Networks. Proc. of 2014 IEEE Int. Conf. on Data Mining **(ICDM'15)**.**citations 32**

[4] Chenguang Wang, Yangqiu Song, Ahmed El-Kishky, Dan Roth, **Ming Zhang**, and Jiawei Han: Incorporating World Knowledge to Document Clustering via Heterogeneous Information Networks. Proc. 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining **(KDD'15)**.**citations 30**

[5] Chenguang Wang, Yangqiu Song, Dan Roth, Chi Wang, Jiawei Han, Heng Ji, and **Ming Zhang**: Constrained Information-Theoretic Tripartite Graph Clustering to Identify Semantically Similar Relations. Proc. 2015 Int. Joint Conf. on Artificial Intelligence **(IJCAI'15)**.**citations 13**

[6] Yangqiu Song, Chenguang Wang, Ming Zhang and Hailong Sun: Spectral Label Refinement for Noisy and Missing Text Labels. Proc. 2015 AAAI Conf. on Artificial Intelligence **(AAAI'15)**.**citations 6**

[7] Chenguang Wang, Yangqiu Song, Dan Roth, Ming Zhang, Jiawei Han: **World Knowledge as Indirect Supervision for Document Clustering.** TKDD 11(2): 13:1-13:36 (2016).**citations 4**

[9] He Jiang, Yangqiu Song, Chenguang Wang, Ming Zhang, Yizhou Sun: **Semi-supervised Learning over Heterogeneous Information Networks by Ensemble of Meta-graph Guided Random Walks.** IJCAI 2017: 1944-1950.**citations 6**

[8] Chenguang Wang, Yangqiu Song, Haoran Li, Yizhou Sun, Ming Zhang, Jiawei Han: **Distant Meta-Path Similarities for Text-Based Heterogeneous Information Networks.** CIKM2017: 1629-1638.**citations 6**

Learning the Joint Representation of Heterogeneous Temporal Events for EHR

- LuchenLiu, Jianhao Shen, Ming Zhang, Zichang Wang, Jian Tang: Learningthe Joint Representation of Heterogeneous Temporal Events for Clinical EndpointPrediction. AAAI 2018
- LuchenLiu, Haoran Li, Jianhao Shen, Ming Zhang, Zichang Wang, Jian Tang: Modeling Temporal Events with Heterogeneous Attributes by Extracting Latent Groups for Clinical Outcome Prediction. AAAI 2019 in submission
- LuchenLiu, Jianhao Shen, Ming Zhang, Zichang Wang, Jian Tang: Learning Hierarchical Representations of Heterogeneous Event Sequences. JAMIA in submission

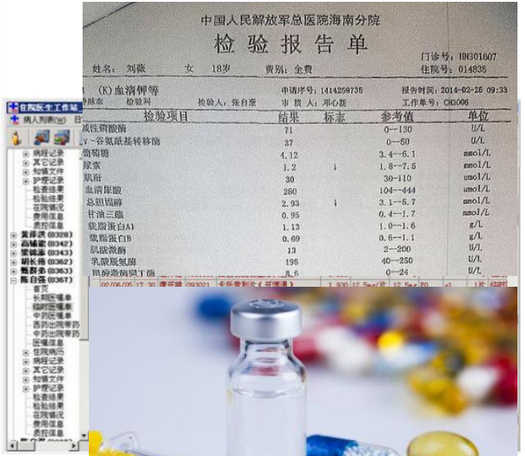
Outline

- Background——endpoint prediction in EHR
- Heterogeneous temporal events
- Model
- Experiments

Outline

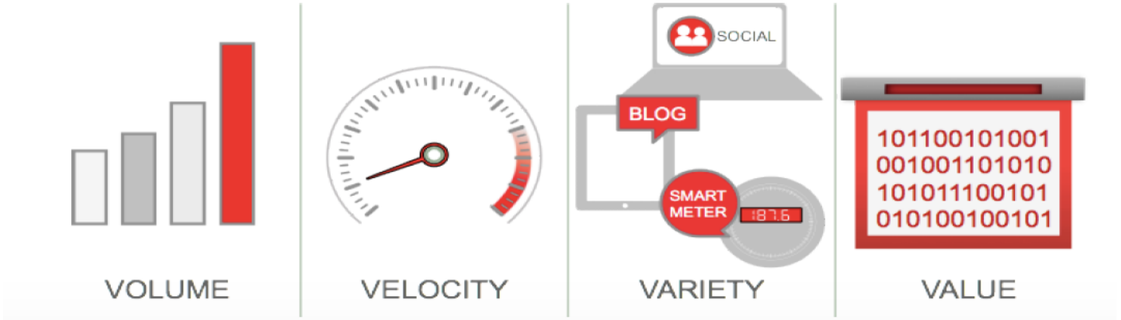
- Background——endpoint prediction in EHR
- Heterogeneous temporal events
- Model
- Experiments

Bacl



HER (**E**lectronic **H**ealth **R**ecord)

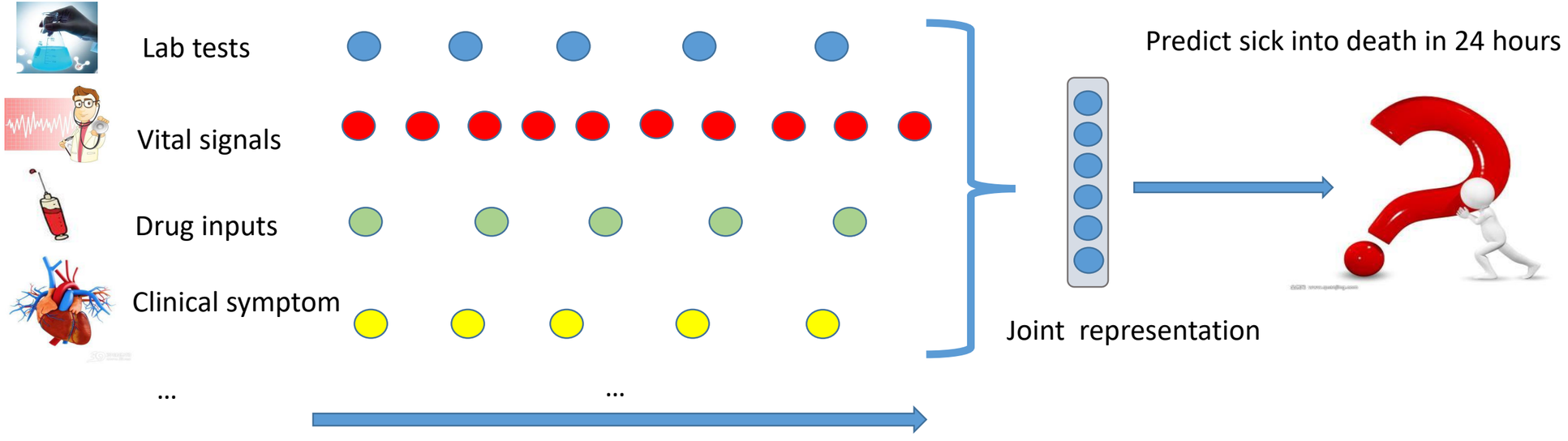
Background



- Challenges in today's healthcare system —— limited analysis ability for Big Data
  - Lab test results ≠ disease mechanism
  - Statistical significance ≠ clinical significance
    - A certain medicine can reduce the risk of heart attack by 34%
    - Real clinical risk are reduced by 1.4%（N Engl J med. 1987 Nov 12; 317(20):1237-45）
- The value of Medical Big Data
  - HER (**E**lectronic **H**ealth **R**ecord) ～ Clinical Endpoint (the target outcomes, e.g. death, symptom…)

Endpoint prediction based on EHR

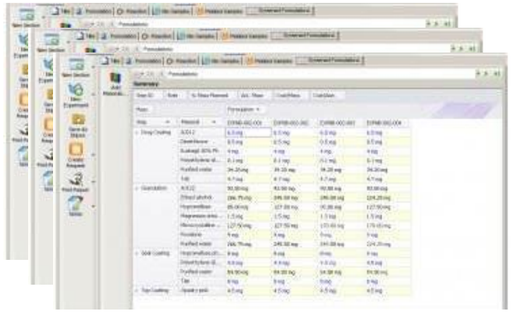
- EHR events embedding
  - Clinical events → patient status representation reflect the disease mechanism
- clinical endpoint prediction
  - Clinical endpoint prediction → personalized diagnosis decision



Outline

- Background——endpoint prediction in EHR
- Heterogeneous temporal events
- Model
- Experiments

Heterogeneous temporal events



Thousands of event types

event

Lab tests

Vital signals

Drug inputs

Clinical symptom

...

...

Drug input

Solution type

dosage

rate

Lab test

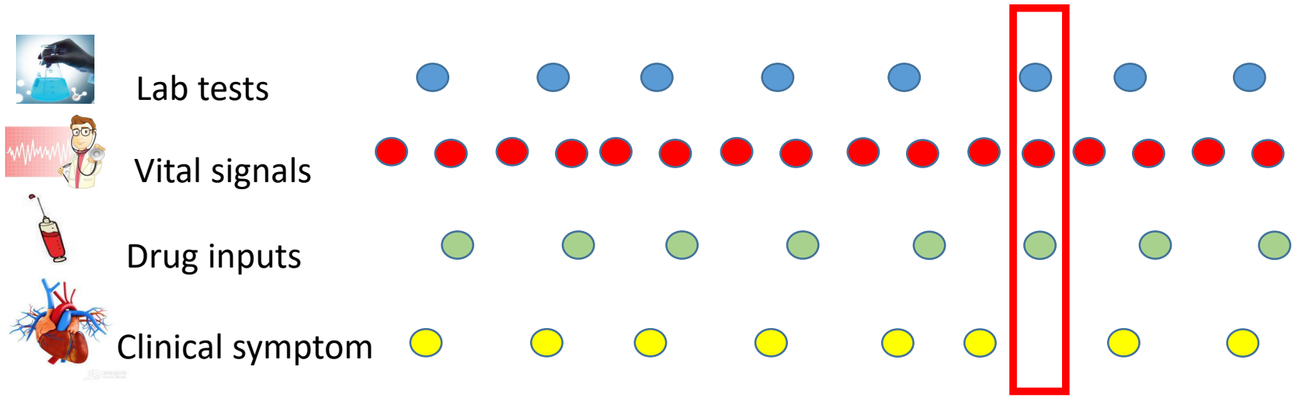
Abnormal or normal

Result index

The frontier of heterogeneous temporal events
----multiple type with multi-scale rate and multiple dependency

Multi-scale record rate

Thousands of
event type

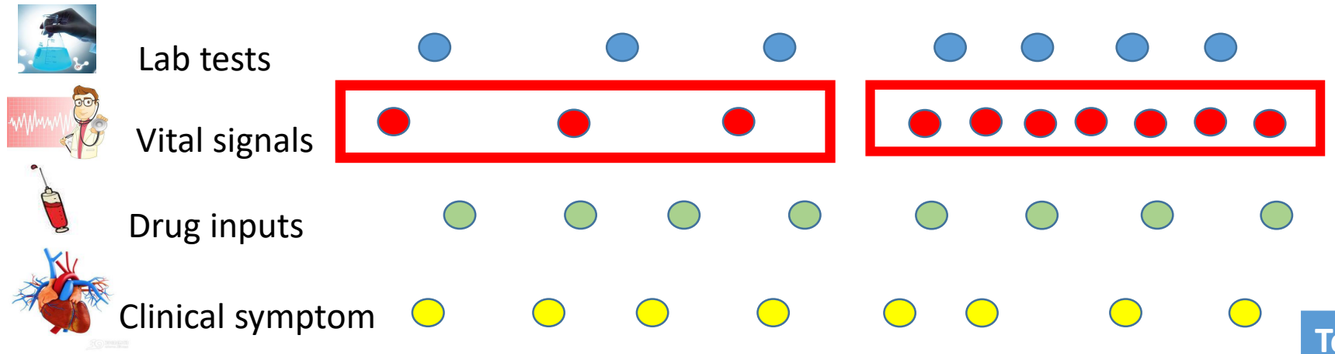


Temporal Dependency of events
Event Correlation at a certain time
Frequency variation
Periodic Pattern of multiple events

The frontier of heterogeneous temporal events
----multiple type with multi-scale rate and multiple dependency

Multi-scale record rate

Thousands of
event type

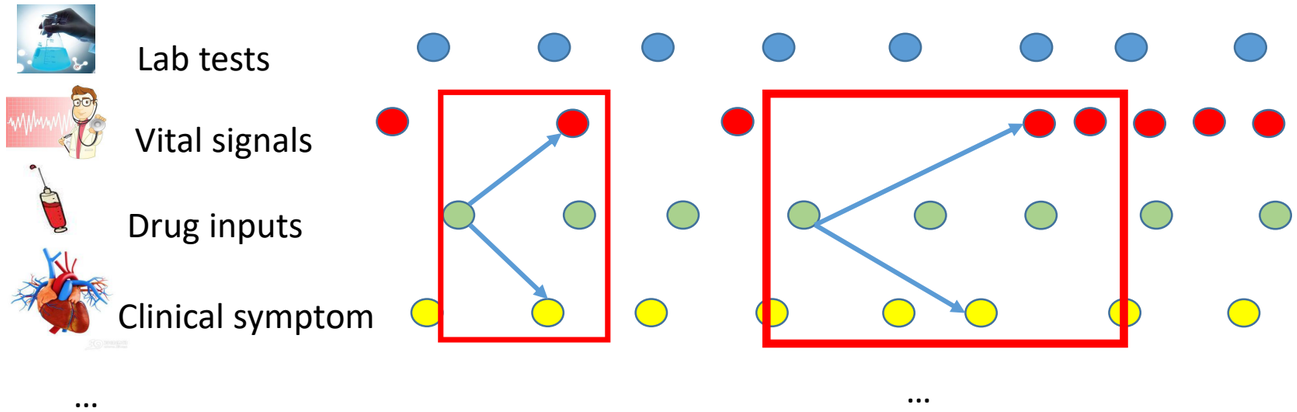


Temporal Dependency of events
Event Correlation at a certain time
Frequency variation
Periodic Pattern of multiple events

The frontier of heterogeneous temporal events ----multiple type with multi-scale rate and multiple dependency

Multi-scale record rate

Thousands of event type



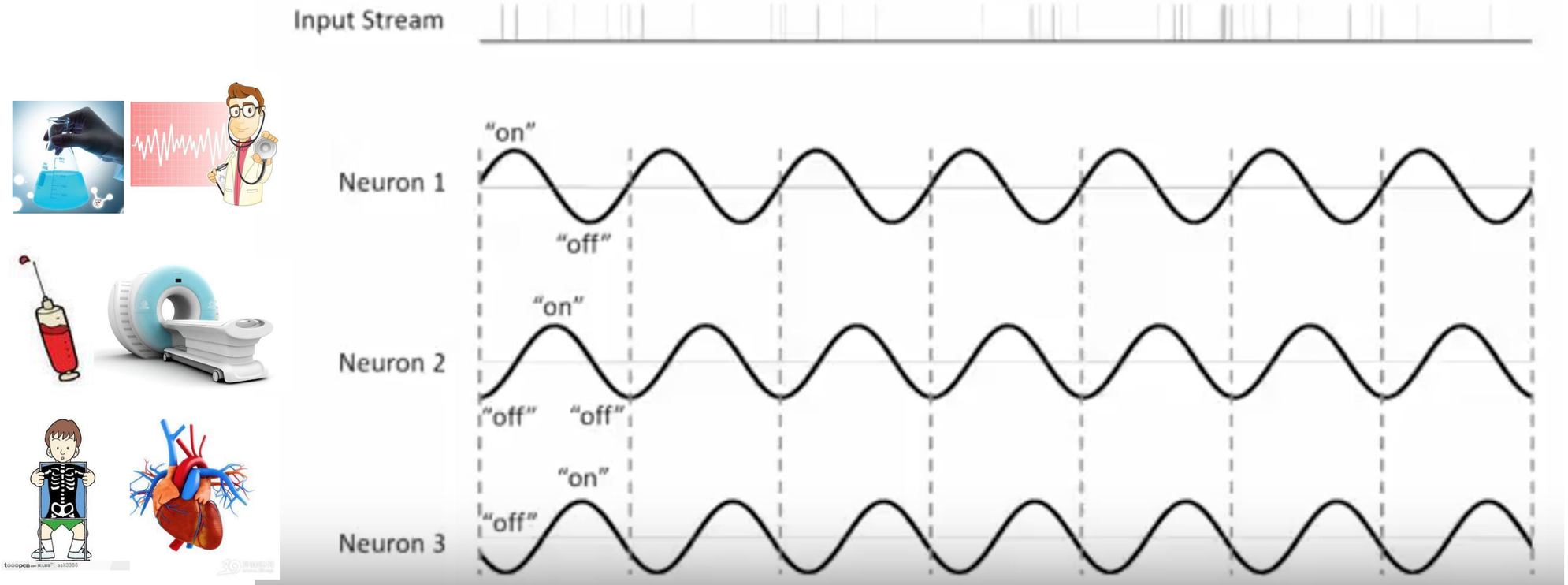
Temporal Dependency of events
Event Correlation at a certain time
Frequency variation
Periodic Pattern of multiple events

Outline

- Background——endpoint prediction in EHR
- Heterogeneous temporal events
- Model —— Heterogeneous Event LSTM
- Experiments

Idea: work in cooperation

Asynchronously tracing important information of related events



Event gate

Asynchronously tracing important information of related events

- Event gate to decide whether or not to update hidden state

$$\tilde{c}_l = \boldsymbol{f}_l \circ \boldsymbol{c}_{l-1} + \boldsymbol{i}_l \circ \tanh\left(W_{cx}\boldsymbol{x}_l + W_{ch}\boldsymbol{h}_{l-1} + \boldsymbol{b}_c\right)$$
$$\boldsymbol{c}_l = \boldsymbol{j}_l \circ \tilde{c}_l + (1 - \boldsymbol{j}_l) \circ \boldsymbol{c}_{l-1}$$

- Event gate is controlled by the event type and time

$$\boldsymbol{j}_{s,t} = \boldsymbol{e}_s \circ \boldsymbol{k}_t$$

$$\boldsymbol{e}_s = \sigma(W_{em}\tanh\left(W_{ms}\boldsymbol{s} + \boldsymbol{b}_m\right) + \boldsymbol{b}_e)$$

- Each neuron of the C vector refers to the status of a set of related **events** at a certain record **rate**

Heterogeneous Event LSTM(HE-LSTM)

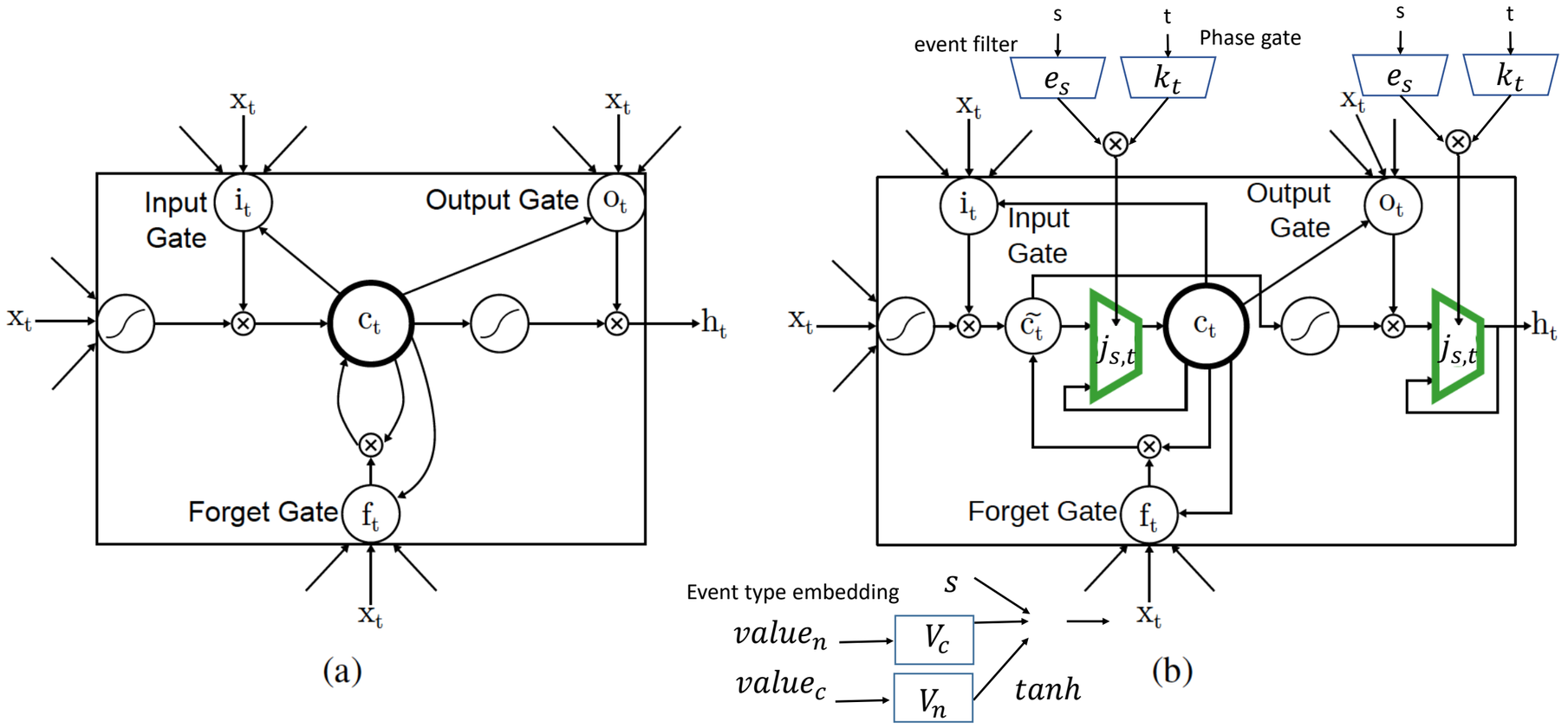
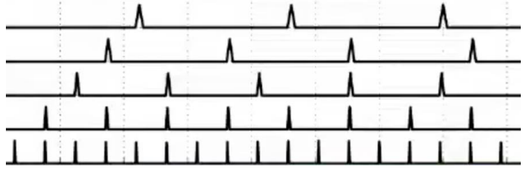
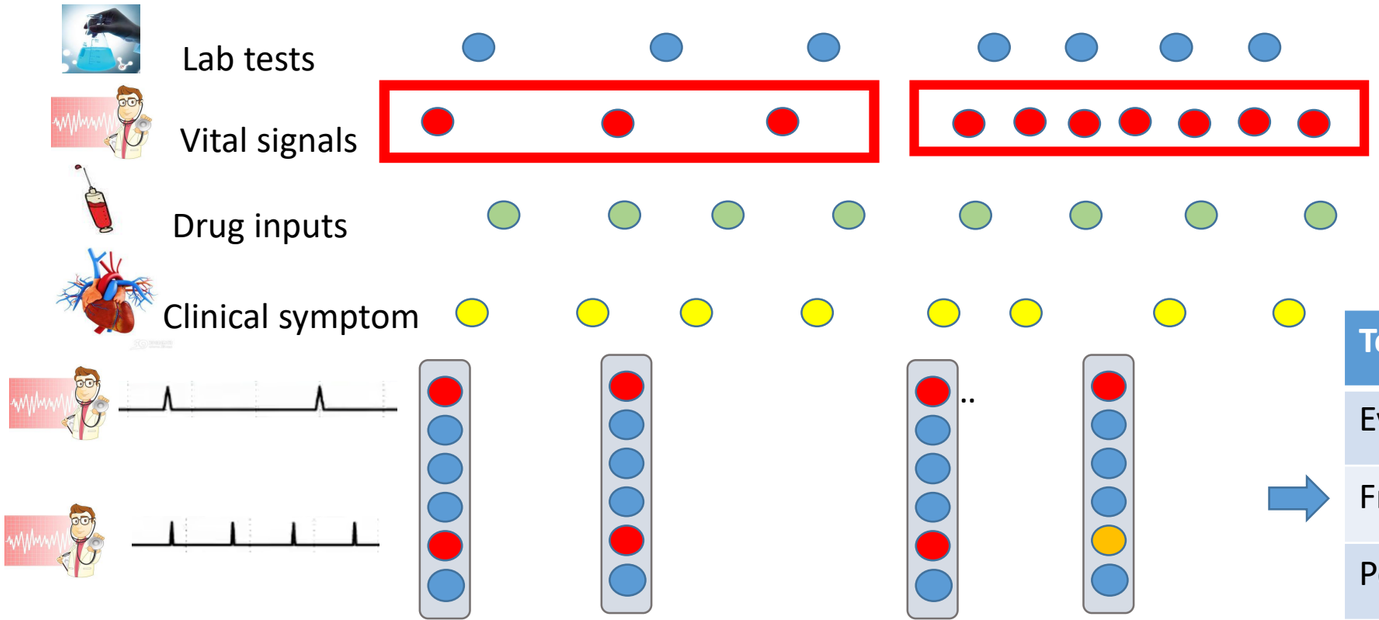


Illustration of the multiple dependency of related events

Multi-scale record rate



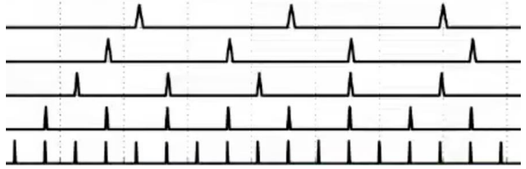
Thousands of event type



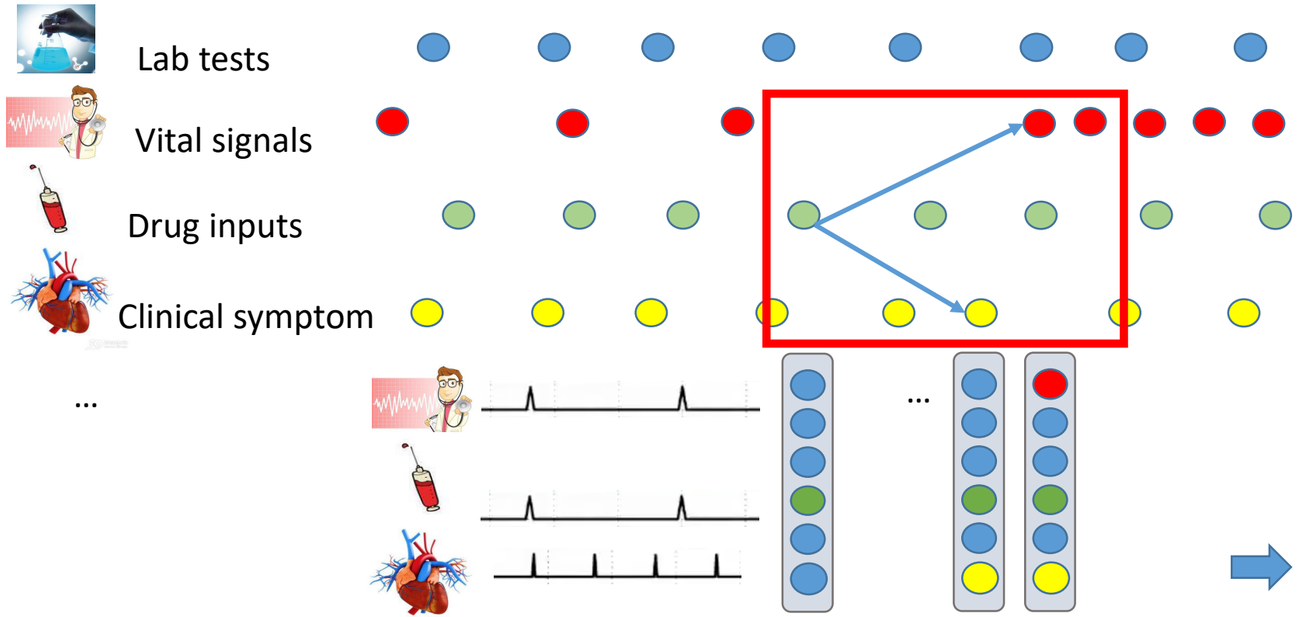
Temporal Dependency of events
Event Correlation at a certain time
Frequency variation
Periodic Pattern of multiple events

Illustration of the multiple dependency of related events

Multi-scale record rate



Thousands of
event type



Temporal Dependency of events
Event Correlation at a certain time
Frequency variation
Periodic Pattern of multiple events

Outline

- Background——endpoint prediction in EHR
- Heterogeneous temporal events
- Model
- Experiments

Performance on two prediction tasks

Methods	death		lab test	
	AUC	**AP**	**AUC**	**AP**
Independent LSTM	0.8771 ± 0.0005	0.5573 ± 0.0006	0.7196 ± 0.0006	0.2969 ± 0.0008
Independent LSTM(shared weight)	0.8064 ± 0.0005	0.5301 ± 0.0006	0.5308 ± 0.0005	0.1098 ± 0.0005
Phased LSTM	0.8474 ± 0.0005	0.4900 ± 0.0075	0.7722 ± 0.0007	0.3575 ± 0.0026
Clock-work RNN	0.8400 ± 0.0001	0.7181 ± 0.0003	0.6516 ± 0.0002	0.2208 ± 0.0003
RETAIN	0.8967 ± 0.0011	0.5808 ± 0.0114	0.7325 ± 0.0022	0.3096 ± 0.0052
LSTM + event embedding & attr encoding	0.9466 ± 0.0002	0.7445 ± 0.0007	0.7231 ± 0.0028	0.3021 ± 0.0014
HE-LSTM	$\mathbf{0.9516 \pm 0.0003}$	$\mathbf{0.7687 \pm 0.0011}$	$\mathbf{0.7987 \pm 0.0008}$	$\mathbf{0.3914 \pm 0.0013}$

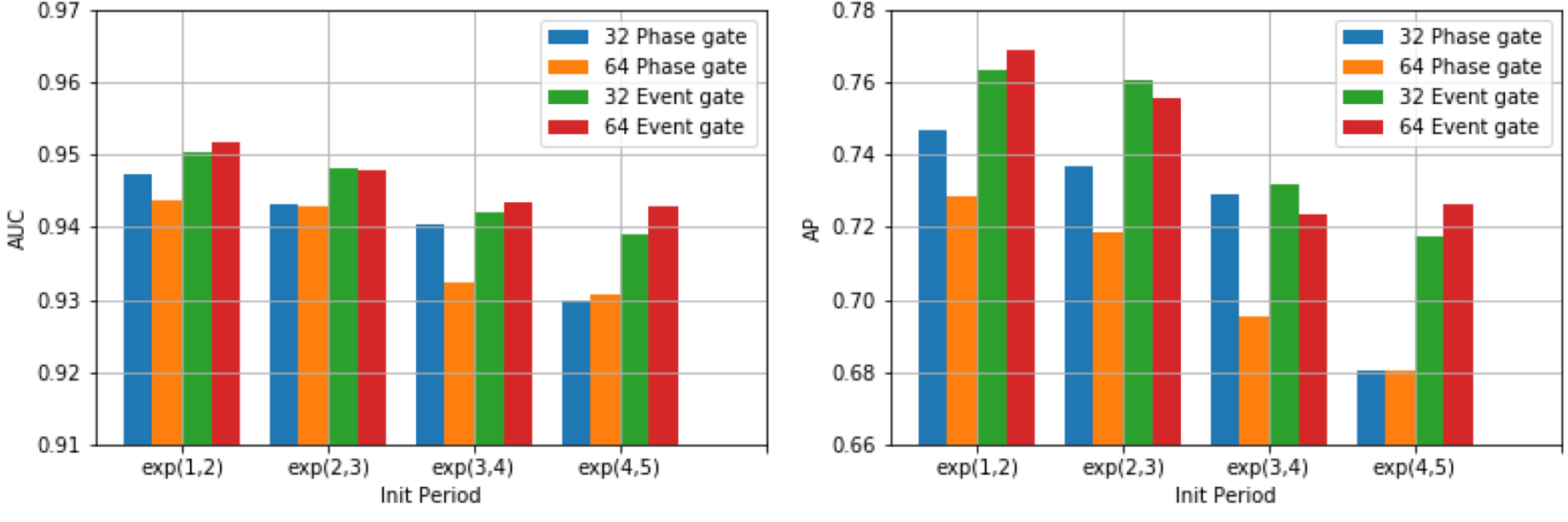
Performance with different settings of event gates

- The phase gate helps to achieve a fast convergence

Methods		Phase gate	Event filter	Event gate
death prediction	AUC(1st epoch)	0.9301	0.9105	**0.9370**
	AUC	0.9471	**0.9518**	0.9516
	AP(1st epoch)	0.6856	0.6048	**0.7094**
	AP	0.7467	0.7679	**0.7687**
	Entropy(1st epoch)	0.1561	0.1835	**0.1479**
	Entropy	0.1369	0.1301	**0.1297**
abnormal lab test prediction	AUC(1st epoch)	0.7050	0.6747	**0.7275**
	AUC	0.7945	0.7559	**0.7987**
	AP(1st epoch)	0.2752	0.2403	**0.2965**
	AP	0.3875	0.3410	**0.3914**
	Entropy(1st epoch)	0.3373	0.3448	**0.3298**
	Entropy	0.3019	0.3178	**0.3003**

Different initial periods

- Robust to different initial period by adding event filter



Conclusion

- The clinical endpoint prediction task based on EHR data
- The representation learning problem of heterogeneous temporal events consists of asynchronous clinical records from multiple sources.
- We propose a novel model called HE-LSTM
  - Modeling the multi-scale sampling rates of different kinds of events and their temporal dependency.

THANK YOU