上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Are We Still Friends: Kernel Multivariate Survival Analysis

Shiyu Liang, Ruotian Luo, Ge Chen, Songjun Ma, Weijie Wu, Li Song, Xiaohua Tian, Xinbing Wang

Presenter: Shiyu Liang

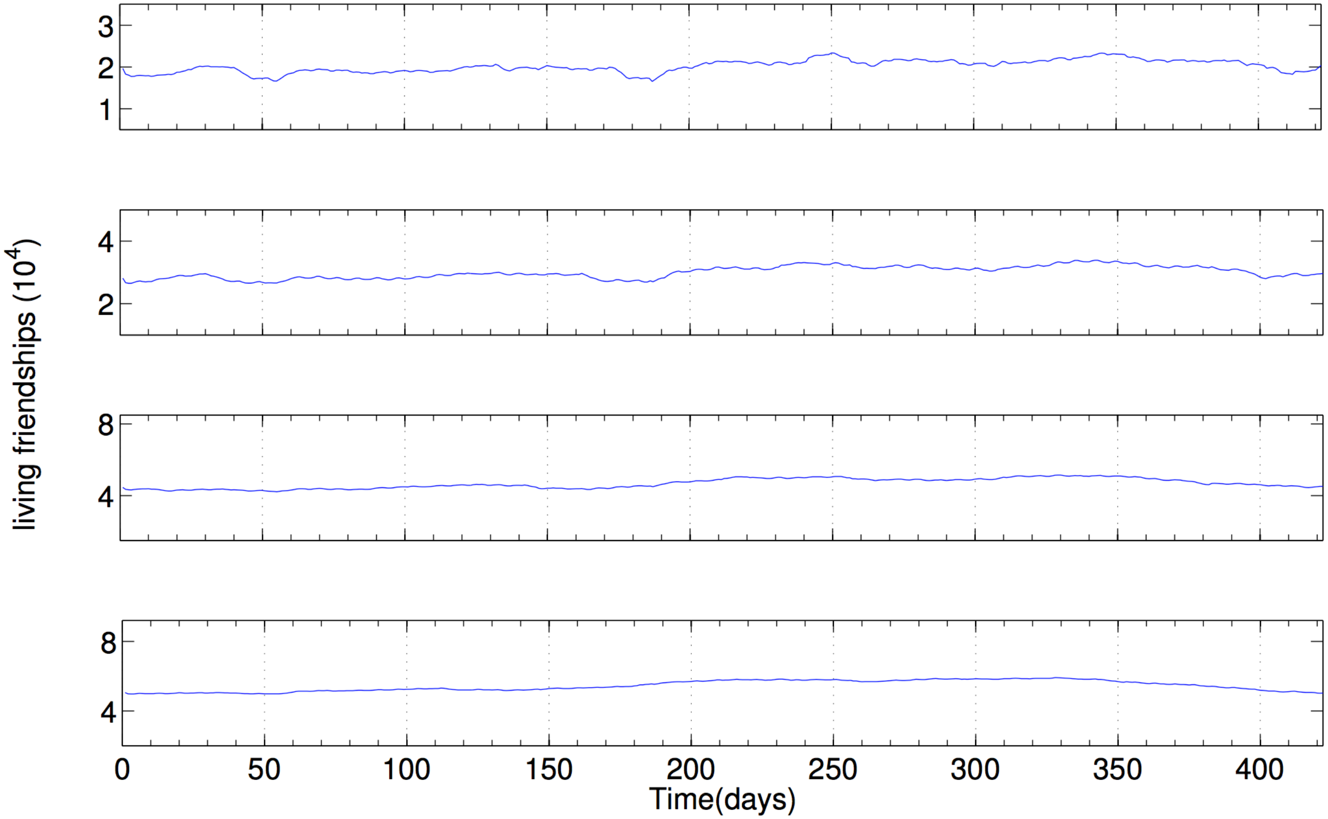
Outline

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

- Introduction
- Data Preprocessing
- System Model
- Experimental Evaluation
- Conclusion

Background

- Friendships Evolution
  - Various factors are conducive to *forming* a friendship
  - Some friendships *deteriorate* and *vanish*
  - Some extinct ones even *rehabilitate* after a long interval
  - *The evolution makes Online Social Network in equilibrium*



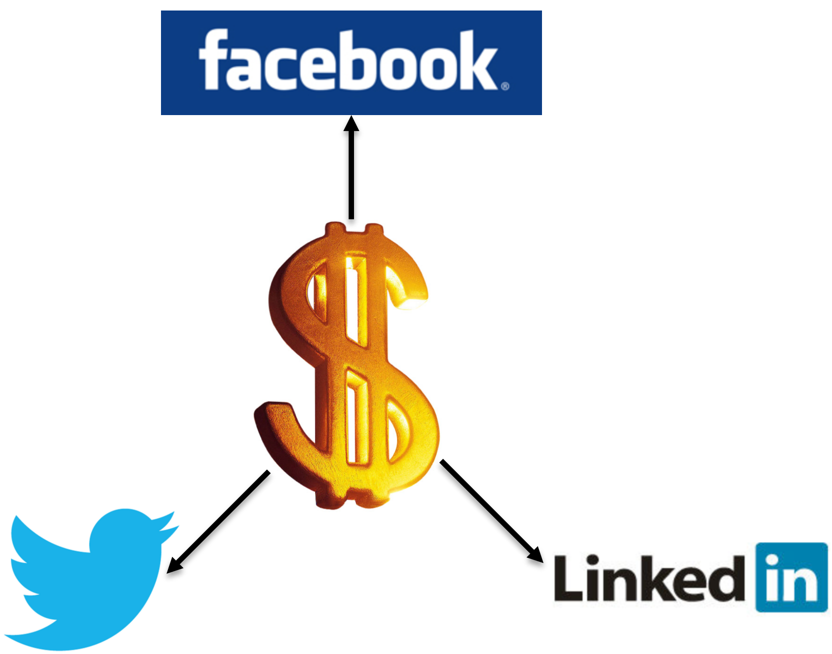
Main Problem

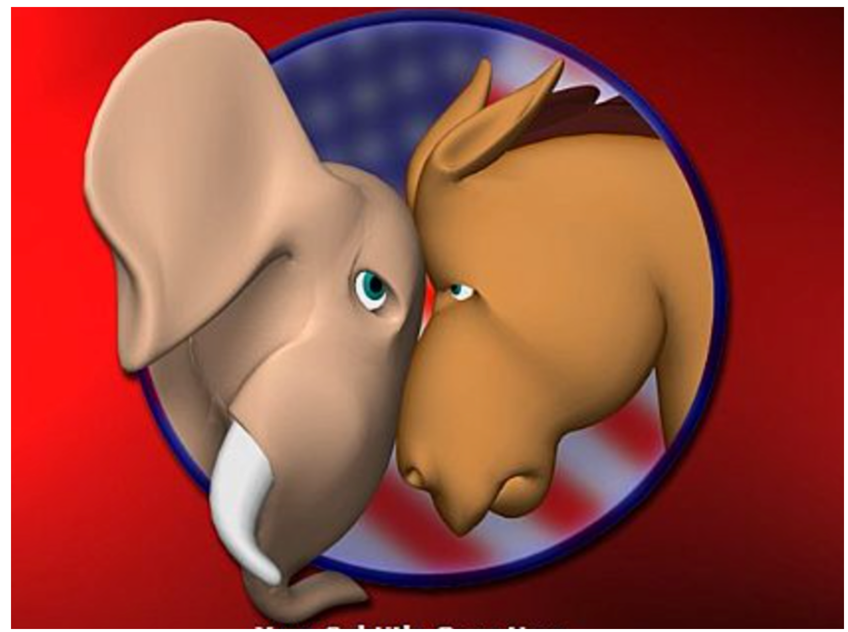
To be or NOT to be, that is the question

Motivation

- Online Service Providers
  - Avoid losing subscribers and make profits

- Political Propaganda
  - Enhance online influence to obtain more votes





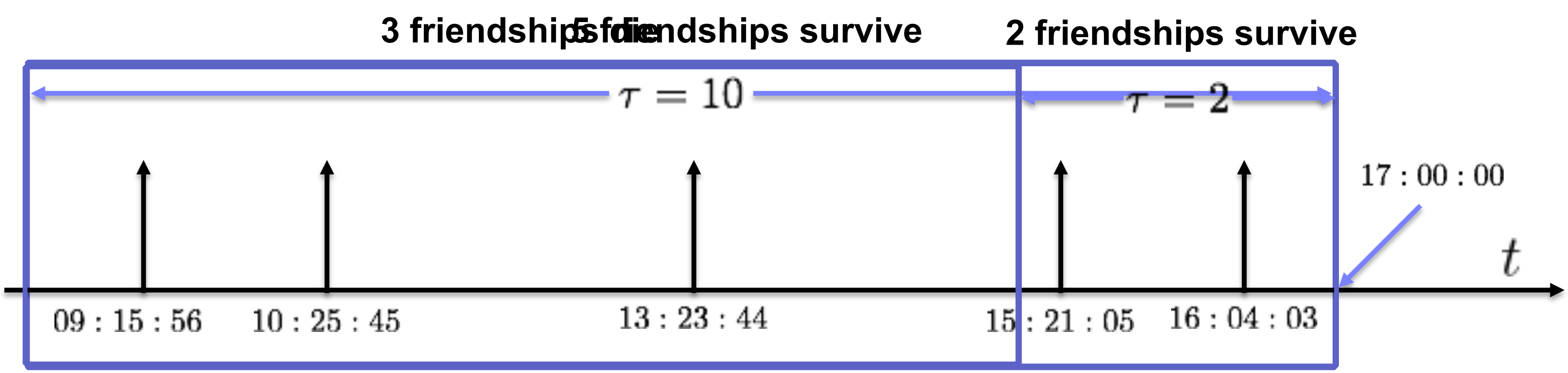
Difficulties

Sporadically Discrete Data

- Users exchange messages *randomly and discretely.*
- *Moving Average* should be adopted.
- *Choosing* an appropriate smoothing length is essential.

Specific Example

- An arrow indicates that two friends are exchanging online messages.
- We compare smoothing length $\tau = 2$ hours and $\tau = 10$ hours.



Difficulties

- Choosing Model
  - Using some *static factors* to predict the friendship state.
  - *Static factors* could be gender, mutual friends, mutual preference, etc.
  - The *multivariate survival model* is chosen in this work.

- **High Complexity Learning Methods**
  - Some learning algorithms have the computational complexity of $O(n^3)$.
  - Our data set contains 40,601 online users with 273,053 online friendships.
  - We need **6.24 years** just train our model !
  - Methods of lowing down the computational complexity are pivotal.

Overview of Results

Data Preprocessing

- Hypothesis testing for finding ***Appropriate Smoothing Length***

Survival Model

- Incorporating social explanatory variables in model
- Applying ***Kernel Learning***
- Imposing ***Sparsity*** and lowing down the training complexity

Model Evaluation

Outline

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

- Introduction
- **Data Preprocessing**
- System Model
- Experimental Evaluation
- Conclusion

Moving Average Process

- We will reject the null hypothesis (*order is m*) with probability of type-1 error equal to if θ $\left|\boldsymbol{\sigma}^T \boldsymbol{\Psi}^{-1}; \boldsymbol{\sigma}\right| \geq l^2$ l is chosen so that

$$\frac{1}{(2\pi)^{m/2}} \int_0^l exp(-\frac{r^2}{2}) S(r) dr \geq 1 - \theta$$

where $S(r)dr$ is the spherically symmetric volume in a m-dimensional Euclidian space.

- After constructing a series of hypothesis testing, we could determine a region where all the values are adequate candidates for the order.

Outline

- Introduction
- Data Preprocessing
- **System Model**
  - Multivariate Survival model
  - Incorporating Kernel Learning
  - Imposing Sparcity
- Experimental Evaluation
- Conclusion

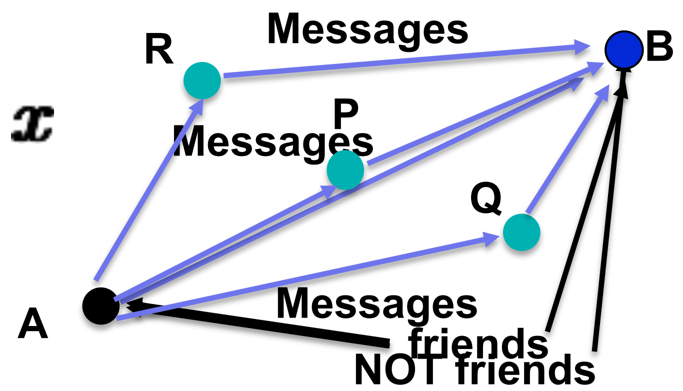
Multivariate Survival Model

- Why is this model appropriate?
  - Commonly used for precisely *predicting* longevity of patients.
  - Considering *confluent* effects of *various* medicine on patients.

- Premise
  - Merely considering a *homogeneous* population of online friendships, each having a ‘surviving time’.
  - *Homogeneous*: the surviving time of each friendship *relies essentially on* proposed social explanatory variables, while other factors will be *neglected*.

Incorporating Social Features

SHANGHAI JIAO TONG UNIVERSITY



- Social explanatory variable vector x
  - Strong indirect relationship
  - Mutual acquaintances
  - Gender
  - All above should be *normalized* before training

- Data likelihood Function

$$P(D) = \prod_{i \in U} f(t_i) \prod_{i \in C} F(t_i)$$

Data likelihood

Surviving time for i-th friendship

Find optimal parameters to maximize it !

Index set of uncensored data

Index set of censored data

Incorporating Kernel Learning

- Mapping function $\phi(\boldsymbol{x})$
  - Transform the input space X into the feature space F $\phi(\boldsymbol{x}) : X \to F$.
  - The inner product of two mapping function is the kernel function.

$$k(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{x} \cdot \boldsymbol{x}' + 1$$

  - Incorporating kernel function in survival model.

- Kernel Learning

$$L(\boldsymbol{\alpha}, b) = \sum_{i=1}^{l} \left[t_i exp \left\{ \sum_{j=1}^{l} \alpha_j k(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right\} \right] - \sum_{i \in U} \left(\sum_{j=1}^{l} \alpha_j k(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right)$$

Find optimal $(\boldsymbol{\alpha}, b)$ to minimize it !

Imposing Sparcity

$$L(\boldsymbol{\alpha}, b) = \sum_{i=1}^{l} \left[t_i exp \left\{ \sum_{j=1}^{l} \alpha_j k(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right\} \right] - \sum_{i \in U} \left(\sum_{j=1}^{l} \alpha_j k(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right)$$

computation complexity → $O(l^3) \approx 6.24\ years$

memory requirement → $O(l^2) \approx 145.8\ GB$

$[K]_{270,000 \times 270,000}$ ⟶ $[\hat{K}]_{4 \times 4}$

$k(\boldsymbol{x}, \boldsymbol{x}')$ ⟶ $\hat{k}(\boldsymbol{x}, \boldsymbol{x}')$

$(\boldsymbol{\alpha}, b)$ ⟶ $(\boldsymbol{\beta}, b)$

$$L(\boldsymbol{\beta}, b) = \sum_{i=1}^{l} \left[t_i exp \left\{ \sum_{j=1}^{M} \beta_j \hat{k}(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right\} \right] - \sum_{i \in U} \left(\sum_{j=1}^{M} \beta_j \hat{k}(\boldsymbol{x_j}, \boldsymbol{x_i}) + b \right)$$

Find optimal $(\boldsymbol{\beta}, b)$ to minimize it !

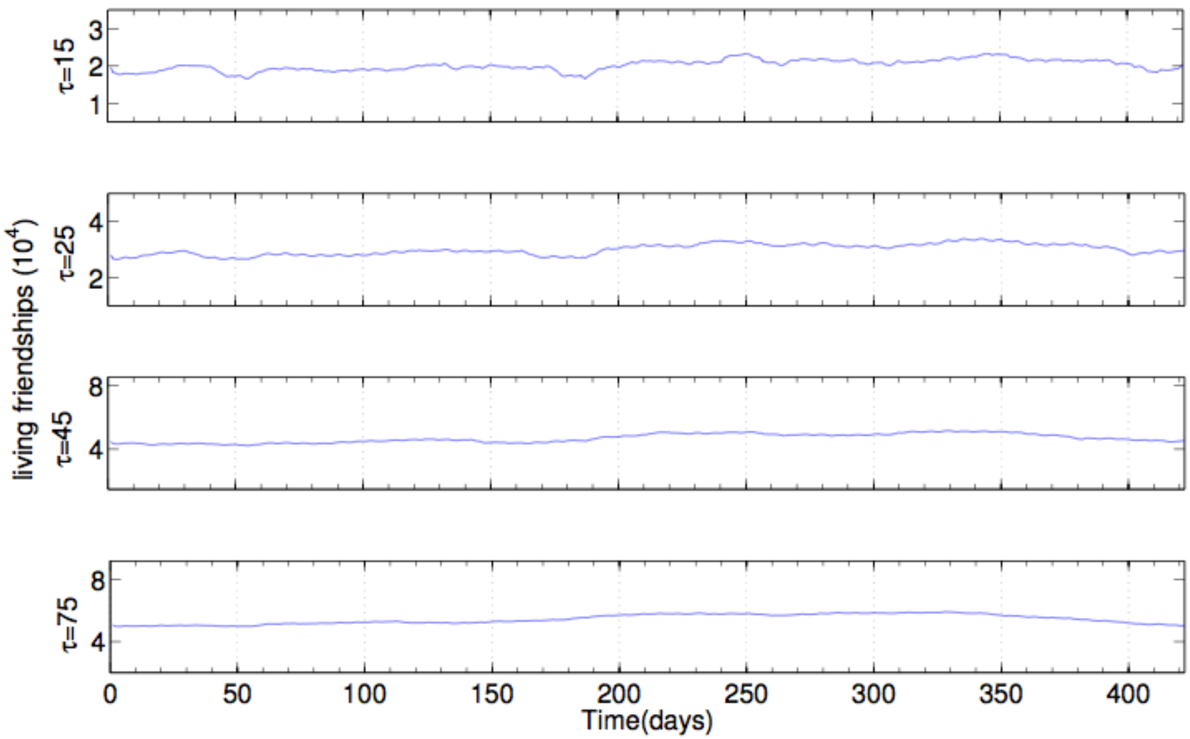
Outline

- Introduction
- Data Preprocessing
- System Model
- **Experimental Evaluation**
- Conclusion

Experimental Evaluation

SHANGHAI JIAO TONG UNIVERSITY

- Online Social Network Equilibrium



Existing different levels of stability

Good for evaluating our model!

Experimental Evaluation

- Criterion

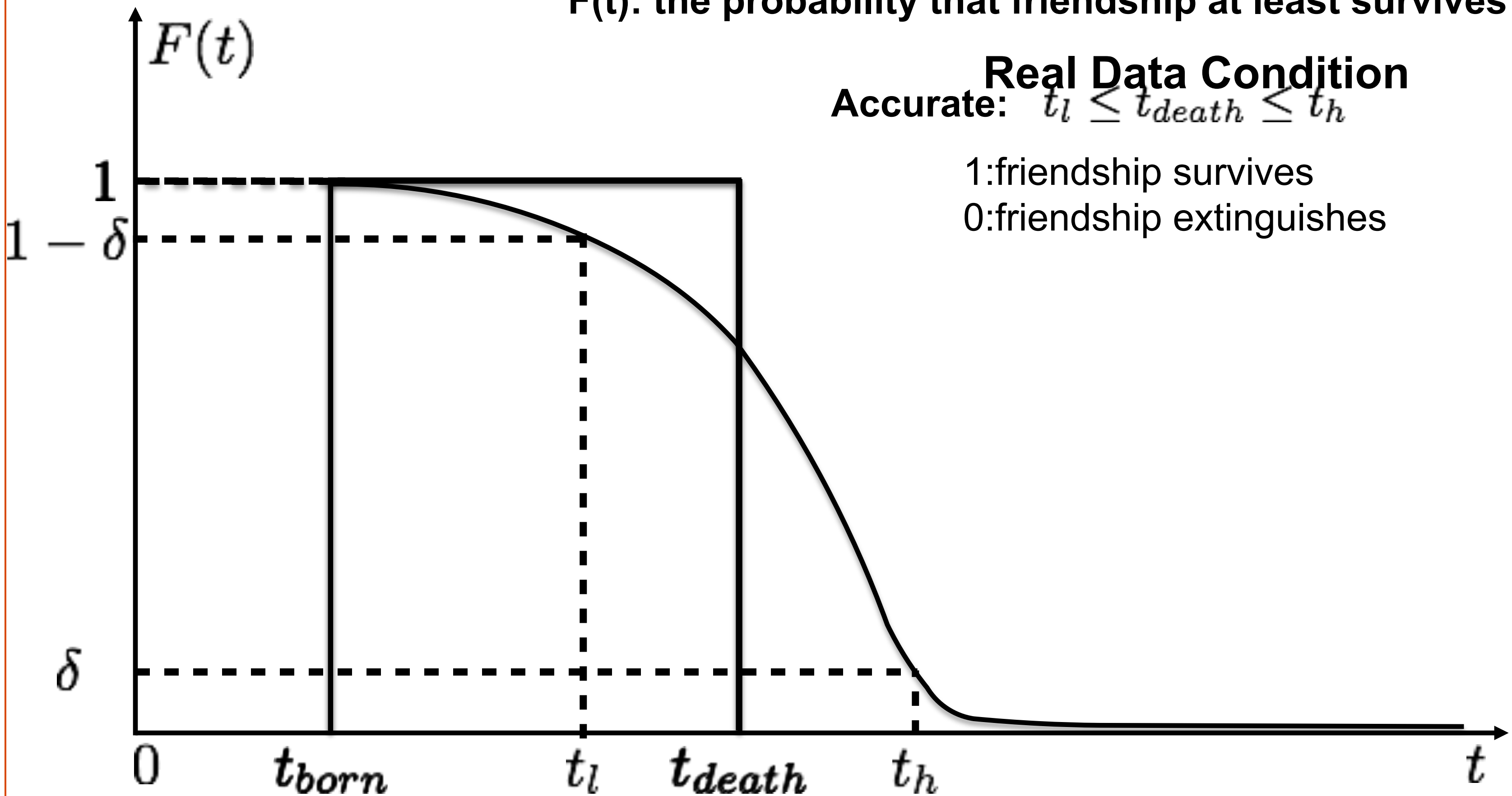
Survival Model

F(t): the probability that friendship at least survives till t

Real Data Condition

Accurate: $t_l \leq t_{death} \leq t_h$

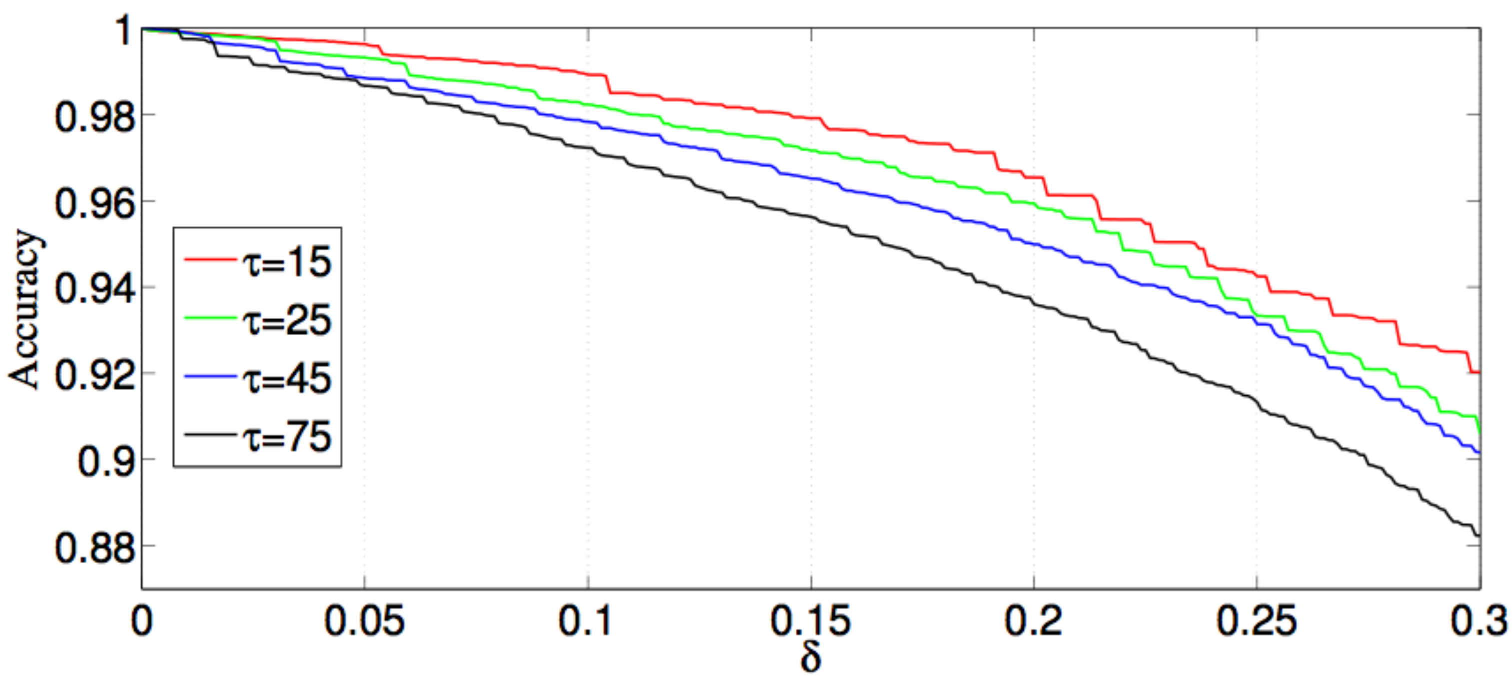
1:friendship survives
0:friendship extinguishes



Experimental Evaluation

Model Accuracy

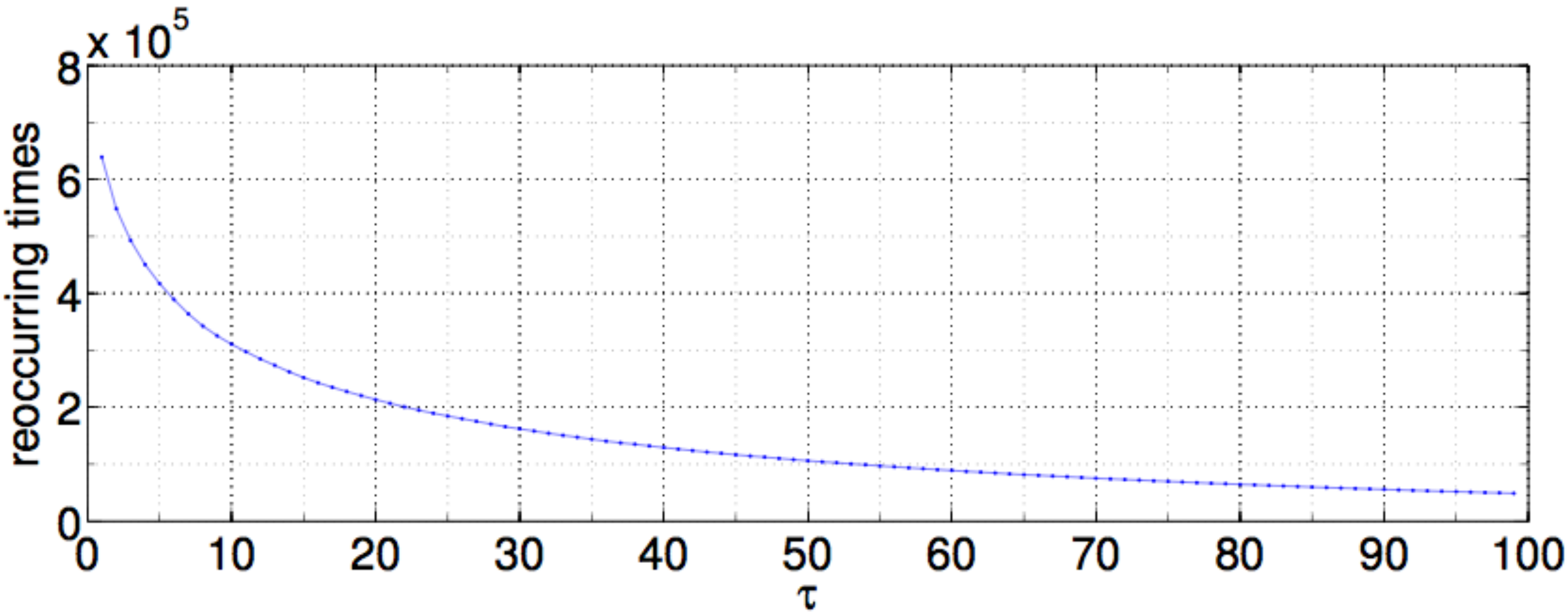
- Accuracy stays high while δ changes severely.
- Model performs better when τ has smaller value.



Experimental Evaluation

❑ Friendship Reoccurrence

- ➢ Smaller τ leads to more reoccurrence.
- ➢ Avoiding huge amount of reoccurrence is necessary.



Outline

Conclusion

- Moving Average Process
  - Analyzed for choosing appropriate smoothing length.
  - Handling trade-off between *higher accuracy* and *less reoccurrence*.
- Multivariate Survival Model
  - Predicting surviving time with *high accuracy* by incorporating kernel learning methods.
  - *Little cost* in training model by imposing sparsity.

Thank you!