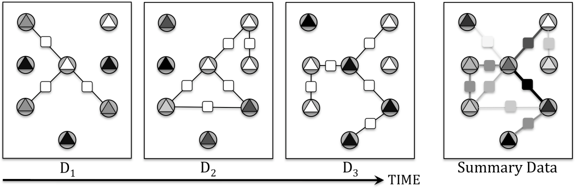
Temporal Relational Classifiers: Evolving Nodes, Edges, and Attributes

Ryan A. Rossi

Department of Computer Science
Purdue University



PURDUE UNIVERSITY

Motivation

- Majority of research focuses on modeling static snapshots...
- Intuition is the significance of an observation decays over time (many examples).
- *Autocorrelation* $\rightarrow$ *Temporal Autocorrelation*
- Emphasize more influential attributes, relationships, and nodes

Approach

Temporal Relational Classification Framework

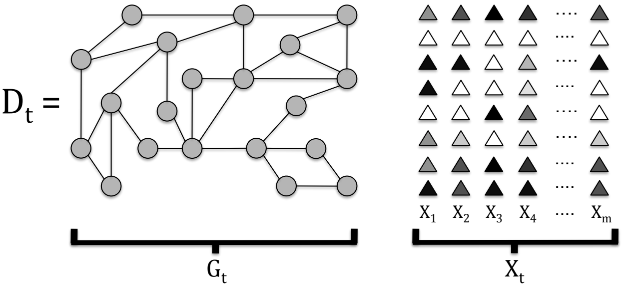
A novel framework for modeling the complete-set of temporal dynamics that can serve as a basis for including temporal dynamics into various relational learning tasks or other statistical relational models.

Validation

To validate our framework we demonstrate its significance in relational classification where the performance of arbitrary classifiers are drastically improved.

Approach

- $G = (E, V)$ and $G_t = \{G_1, G_2, ..., G_t\}$
- $X_m^g = \{X_1, X_2, ..., X_m\}$ and $X_i = X_{i_1}, X_{i_2}, ..., X_{i_t}$
- $D = \langle G, X \rangle$ and $D_t = \langle G_t, \ \{X_{1t}, X_{2t}, ..., X_{mt}\}\rangle$



Graph Summarization

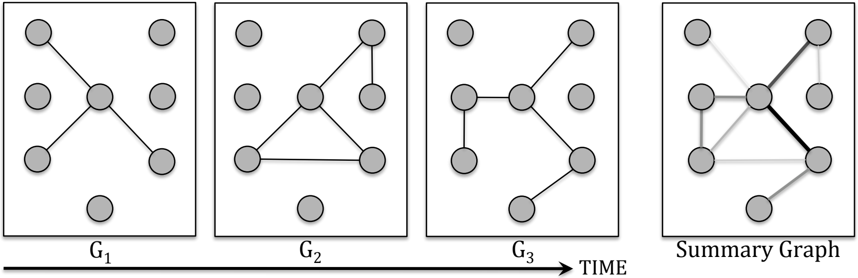


Figure: Graph Summarization: Temporally evolving links.

More recent observations are given a higher weight or probability while observations further in the past are assigned substantially lower weights.

Attribute Summarization

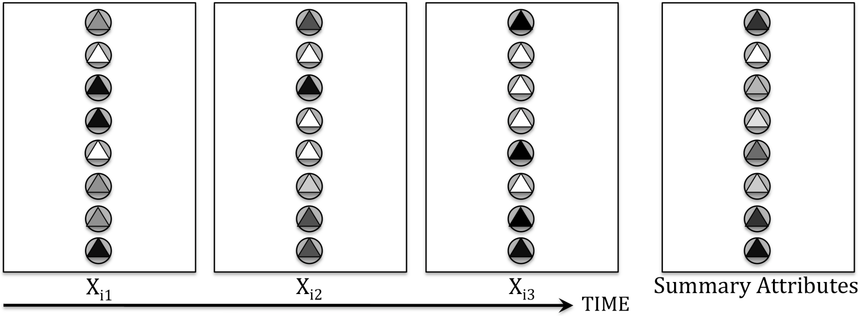


Figure: Attribute Summarization: Temporally evolving attributes.

Data Summarization

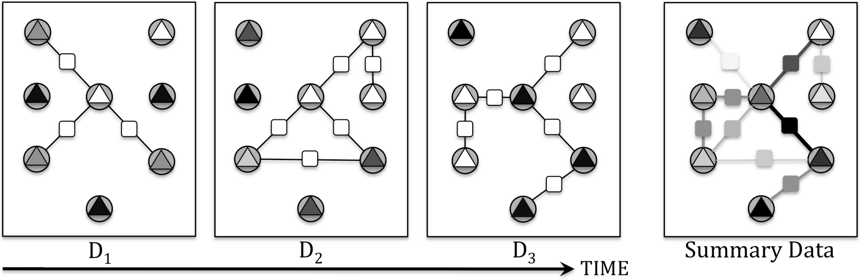


Figure: Combined Data Summarization: Temporally evolving links and attributes.

$$\mathcal{D}^{\mathbf{S}} = \langle G^S, \{X_1^S, X_2^S, ..., X_m^S\} \rangle$$

$$\mathcal{W} = \langle \mathcal{W}^{\mathbf{E}}, \mathcal{W}^{\mathbf{X}} \rangle$$

PURDUE UNIVERSITY

Kernel Smoothing

$$\mathcal{W}_{\mathbf{t}}^{\mathbf{S}} = \sum_{i=1}^{t} K(\mathcal{D}_i; t, \theta, \lambda)$$

Arbitrary Kernel Function

- **Exponential Kernel**
- Inverse Linear Kernel
- Linear Kernel
- ...

Temporal Relational Classification

Any relational model can be easily extended.

Relational Bayes Classifier

$$P(C_t^i | \mathbf{X}, R) \propto \prod_{X_m \in \mathbf{X}^{\mathbf{G(i)}}} \prod_{x \in X_m^i} w_i^{t(X_m)} \cdot P(X_m^i = x | C) \times$$

$$\prod_{j \in R} \prod_{X_k \in \mathbf{X}^{\mathbf{G(j)}}} w_{ij}^{t(E)} \times \prod_{x \in X_k^j} w_j^{t(X_k)} \cdot P(X_k^j = x | C) \times P(C)$$

Relational Probability Trees

Dynamically proportionalizes the relational data creating binary splits.
Modified the aggregate functions to allow for weights.

Evaluation

- TVRC $\rightarrow$ Fully-joint Temporal Classifier
- TVRC Baseline $\rightarrow$ Temporal Edge Strengths
- Window Model $\rightarrow$ Uses the immediate past information for prediction ignoring all the other historical data
- Snapshot Model $\rightarrow$ Uses a graph $G_{\leq t}$ which includes all objects and links up to and including t

Datasets

PyComm	CORA
Developers: 185 Emails: 13181 Bugs Msgs: 69435 Teams: 18	Papers:16,153 References: 29,603 Authors: 21,976
Time Window: Feb2007-May2008	Time Window: 1981-1998

Table: Datasets used for Empirical Evaluation.

- Time Window depends on the domain (Any temporal granularity)

Dataset	Prediction task
PyComm	*Effectiveness* (is developer productive or not)
Cora	Paper *topic* (is ML paper or not)

Table: Classification Tasks

k-fold Cross-validation

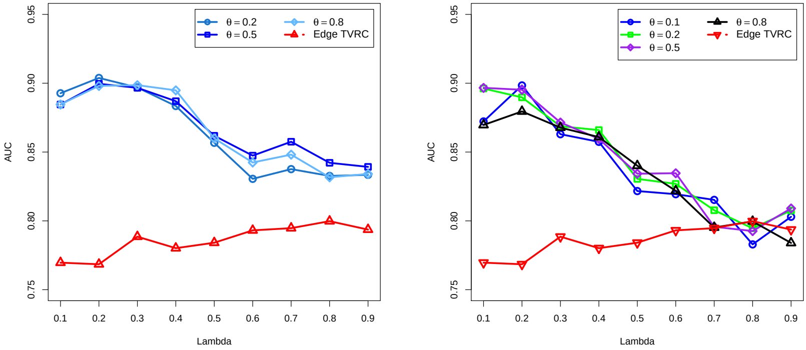
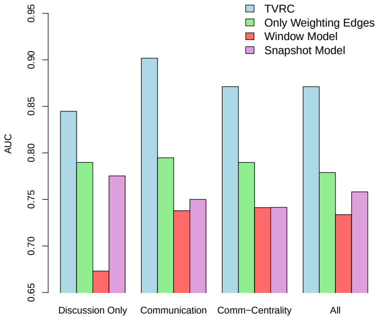


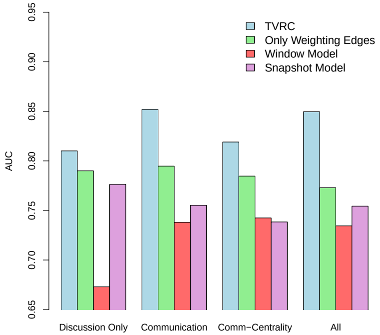
Figure: Selection of $\theta = 0.1$ and $\lambda = 0.2$ using cross-validation.

PURDUE UNIVERSITY

Average AUC Values



(a) CV $\theta = 0.1$ and $\lambda = 0.2$

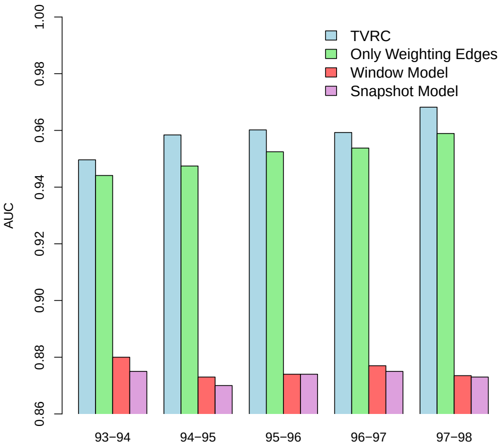


(b) $\theta = 0.7$ and $\lambda = 0.7$

Figure: Average AUC values of the models

PURDUE UNIVERSITY

Cora Evaluation



Validated Framework on Relational Classification!

Defined a unifying framework for including the complete-set of temporal dynamics into statistical relational models.

Validation

- The results indicate the necessity and signifiance of modeling the temporal dynamics.
- Our TVRC framework can be useful in other areas of relational learning
  - Diffusion
  - Graph Generation
  - Sampling
  - Extrapolations
  - Community/Group Discovery
  - Temporal Centrality, Patterns, and other Measures

Other Working Directions

Incorporating Temporal Latent Textual Information

- Discovery of Evolutionary Patterns from Latent Textual Information
  - Labeling Links with Latent Information (Latent Link Semantics)
  - Applied not only for natural languages, but could be used to incorporate other types of latent information such as the hidden nature of protein iteractions using the protein sequences to extract the latent structure

Temporal Cross-Validation

Traditional Cross-Validation procedure no longer makes sense applied to temporal relational data

Other Working Directions

Distributed Teams

- Revisiting: Predicting Individual Effectiveness in Distributed Teams
  - Incorporating Latent Textual Information
  - Defined various types of attributes such as Team Membership, Topic, Centrality, Performance, Communication
  - Evaluate: Intrinsic $\Leftrightarrow$ Relational $\Leftrightarrow$ Temporal
  - Single Models
  - Ensemble Models
  - Randomization: Evaluate Individual/Group Feature Significance

Questions?