Time-Evolving Relational Classification and Ensemble Methods

Ryan Rossi and Jennifer Neville

Purdue University,
West Lafayette, IN 47906, USA
{rrossi,neville}@purdue.edu

Abstract. Relational networks often evolve over time by the addition, deletion, and changing of links, nodes, and attributes. However, accurately incorporating the full range of temporal dependencies into relational learning algorithms remains a challenge. We propose a novel framework for discovering *temporal-relational representations* for classification. The framework considers transformations over *all* the evolving relational components (attributes, edges, and nodes) in order to accurately incorporate temporal dependencies into relational models. Additionally, we propose *temporal ensemble methods* and demonstrate their effectiveness against traditional and relational ensembles on two real-world datasets. In all cases, the proposed temporal-relational models outperform competing models that ignore temporal information.

1 Introduction

Temporal-relational information is present in many domains such as the Internet, citation and collaboration networks, communication and email networks, social networks, biological networks, among many others. These domains all have attributes, links, and/or nodes changing over time which are important to model. We conjecture that discovering an accurate *temporal-relational representation* will disambiguate the true nature and strength of links, attributes, and nodes. However, the majority of research in relational learning has focused on modeling static snapshots [2, 6] and has largely ignored the utility of learning and incorporating temporal dynamics into relational representations.

Temporal relational data has three main components (attributes, nodes, links) that vary in time. First, the attribute values (on nodes or links) may change over time (e.g., research area of an author). Next, links might be created and deleted throughout time (e.g., host connections are opened and closed). Finally, nodes might appear and disappear over time (e.g., through activity in an online social network).

Within the context of evolving relational data, there are two types of prediction tasks. In a *temporal* prediction task, the attribute to predict is changing over time (e.g., student GPA), whereas in a *static* prediction task, the predictive attribute is constant (e.g., paper topic). For these prediction tasks, the space of temporal-relational representations is defined by the set of relational

P.-N. Tan et al. (Eds.): PAKDD 2012, Part I, LNAI 7301, pp. 1–13, 2012.

elements that change over time (attributes, links, and nodes). To incorporate temporal information in a representation that is appropriate for relational models, we consider two transformations based on *temporal weighting* and *temporal granularity*. Temporal weighting aims to represent the temporal influence of the links, attributes and nodes by decaying the weights of each with respect to time, whereas the choice of temporal granularity restricts attention to links, attributes, and nodes within a particular window of time. The optimal temporal-relational representation and the corresponding temporal classifier depends on the particular temporal dynamics of the links, attributes, and nodes present in the data, as well as the network domain (e.g., social vs. biological networks).

In this work, we address the problem of selecting the most optimal temporal-relational representation to increase the accuracy of predictive models. We consider the full space of *temporal-relational representations* and propose **(1)** a temporal-relational classification framework, and **(2)** a set of temporal ensemble methods, to leverage time-varying links, attributes, and nodes in relational networks. We illustrate the different types of models on a variety of classification tasks and evaluate each under various conditions. The results demonstrate the flexibility and effectiveness of the temporal-relational framework for classification in time-evolving relational domains. Furthermore, the framework provides a foundation for automatically searching over temporal-relational representations to increase the accuracy of predictive models.

2 Related Work

Recent work has started to model network dynamics in order to better predict link and structure formation over time [7, 10], but this work focuses on unattributed graphs. Previous work in relational learning on attributed graphs either uses static network snapshots or significantly limits the amount of temporal information incorporated into the models. Sharan et al. [18] assumes a strict representation that only uses kernel estimation for link weights, while GA-TVRC [9] uses a genetic algorithm to learn the link weights. SRPTs [11] incorporate temporal and spatial information in the relational attributes. However, the above approaches focus only on one specific temporal pattern and do not consider different temporal granularities. In contrast, we explore a larger space of temporal-relational representations in a flexible framework that can capture temporal dependencies over *links*, *attributes*, and *nodes*.

To the best of our knowledge, we are the first to propose and investigate *temporal-relational ensemble methods* for time-varying relational classification. However, there has been recent work on relational ensemble methods [8, 14, 15] and non-relational ensemble methods for evolving streams [1]. Preisach et al. [14] use voting and stacking methods to combine relational data with multiple relations. In contrast, Eldardiry and Neville [8] incorporates prediction averaging in the collective inference process to reduce both learning and inference variance.

3 Temporal-Relational Classification Framework

Below we outline a temporal-relational classification framework for prediction tasks in dynamic relational networks. Relational data is represented as an attributed graph $D = (G, \mathbf{X})$ where the graph $G = (V, E)$ represents a set of N nodes, such that $v_i \in V$ corresponds to node i and each edge $e_{ij} \in E$ corresponds to a link (e.g., email) between nodes i and j. The attribute set:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X^V} = [X^1, X^2, ..., X^{m_v}], \\ \mathbf{X^E} = [X^{m_v+1}, X^{m_v+2}, ..., X^{m_v+m_e}] \end{pmatrix}$$

contains m_v observed attributes on the nodes ($\mathbf{X^V}$) and m_e observed attributes on the edges ($\mathbf{X^E}$). Dynamic relational data evolves over time by the addition, deletion, and changing of nodes, edges, and attributes. Let $D_t = (G_t, \mathbf{X}_t)$ refer to the dataset at time t, where $G_t = (V, E_t)$ and $\mathbf{X}_t = (\mathbf{X}_t^V, \mathbf{X}_t^E)$. In our classification framework, we consider relational data observed over a range of timesteps $t = \{1, ..., T\}$ (e.g., citations over a period of years, emails over a period of days). Given this time-varying relational data, the task is to learn a model to predict either a static attribute Y or a dynamic attribute at a particular timestep Y_t, while exploiting both the relational and temporal dependencies in the data.

We define our temporal-relational classification framework with respect to a set of possible transformations of links, attributes, or nodes (as a function of time). The temporal weighting (e.g., exponential decay of past information) and temporal granularity (e.g., window of timesteps) of the links, attributes and nodes form the basis for any arbitrary transformation with respect to the temporal information (See Table 1). The discovered temporal-relational representation can be applied for mining temporal patterns, classification, and as a means for constructing temporal-ensembles. An overview of the temporal-relational representation discovery is provided below:

Table 1. Temporal-Relational Representation

1. For each Relational Component
   – Links, Attributes, or Nodes
2. Select the Temporal Granularity
   ⋆ Timestep t_i
   ⋆ Window $\{\mathbf{t_j}, \mathbf{t_{j+1}}, ..., \mathbf{t_i}\}$
   ⋆ Union $T = \{t_0, ..., t_n\}$
3. Select the Temporal Influence
   ⋆ Weighted
   ⋆ Uniform
   Repeat steps 1-3 for each component.
4. Select the Relational Classifier
   ⋆ Relational Bayes Classifier (RBC)
   ⋆ Relational Probability Trees (RPT)

	Uniform			**Weighting**		
	Timestep	**Window**	**Union**	**Timestep**	**Window**	**Union**
Edges						
Attributes						
Nodes						

Table 1 provides an intuitive view of the possible temporal-relational representations. For instance, the TVRC model is a special case of the proposed framework where the links, attributes, and nodes are unioned and the links are weighted. Below we provide more detail on steps 2-4.

3.1 Temporal Granularity

Traditionally, relational classifiers have attempted to use all the data available in a network [18]. However, since the relevance of data may change over time (e.g., links become stale), learning the *appropriate* temporal granularity (i.e., range of timesteps) can improve classification accuracy. We briefly define three general classes for varying the temporal granularity of the links, attributes, and nodes.

1. **Timestep.** The timestep models only use a single timestep t_i for learning.
2. **Window.** The window models use a sliding window of (multiple) timesteps $\{t_j, t_{j+1}, ..., t_i\}$ for learning. When the size of window is varied, the space of possible models in this category is by far the largest.
3. **Union.** The union model uses all previous temporal information for learning at time t_i, i.e., $T = \{0, ..., t_i\}$.

The timestep and union models are separated into distinct classes for clarity in evaluation and for understandability in pattern mining.

3.2 Temporal Influence: Links, Attributes, Nodes

We model the influence of relational components over time using temporal weighting. Specifically, when considering a temporal dataset $D_t = (G_t, \mathbf{X}_t)$, we will construct a weighted network $G_t = (V, E_t, W_t^E)$ and $\mathbf{X}_t = (\mathbf{X}_t^V, \mathbf{X}_t^E, W_t^X)$. Here W_t refers to a function that assigns weights on the edges and attributes that are used in the classifiers below.

Initially, we define $W_t^E(i,j) = 1$ if $e_{ij} \in E_t$ and 0 otherwise. Similarly, we define $W_t^X(x_i^m) = 1$ if $X_i^m = x_i^m \in \mathbf{X_t^m}$ and 0 otherwise. Then we consider two different approaches to revise these initial weights:

1. **Weighting.** These temporal weights can be viewed as probabilities that a relational component is still active at the current time step t, given that it was observed at time $(t-k)$. We investigated three temporal weighting functions:
   - *Exponential Kernel.* The exponential kernel weights the recent past highly and decays the weight rapidly as time passes [3]. The kernel function K_E for temporal data is defined as: $K_E(D_i; t, \theta) = (1-\theta)^{t-i}\theta W_i$
   - *Linear Kernel.* The linear kernel decays more grdually and retains the historical information longer: $K_L(D_i; t, \theta) = \theta W_i(\frac{t_* - t_i + 1}{t_* - t_o + 1})$
   - *Inverse Linear Kernel.* This kernel lies between the exponential and linear kernels when moderating historical information: $K_{IL}(D_i; t, \theta) = \theta W_i(\frac{1}{t_i - t_o + 1})$

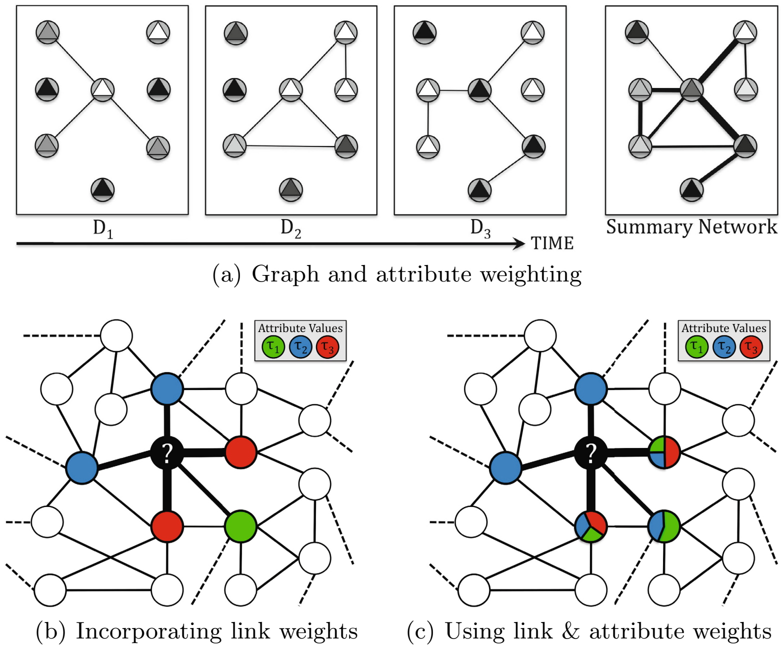

Fig. 1. **(a)** Temporally weighting the attributes and links. **(b)** The feature calculation that includes only the temporal link weights. **(c)** The feature calculation that incorporates *both* the temporal attribute weights and the temporal link weights.

2. **Uniform.** These weights ignore the temporal influence of a relational component, and weight them uniformly over time, i.e., $W_t^E(i,j) = 1$ if $e_{ij} \in E_{t'} : t' \in T$ and 0 otherwise. A relational component can be assigned uniform weights within the selected temporal granularity or over the entire time window (e.g., traditional classifiers assign uniform weights, but they don't select the appropriate temporal granularity).

We note that different weighting functions can be chosen for different relational components (edges, attributes, nodes) with varying temporal granularities. For instance, the temporal influence of the links might be predicted using the exponential kernel while the attributes are uniformly weighted but have a different temporal granularity than the links.

3.3 Temporal-Relational Classifiers

Once the temporal granularity and temporal weighting are selected for each relational component, then a temporal-relational classifier can learned. In this work, we use modified versions of the RBC [13] and RPT [12] to model the transformed temporal-relational representation. However, we note that any relational model that can be modified to incorporate node, link, and attribute weights is suitable for this phase. We extended RBCs and RPTs since they are interpretable, diverse, simple, and efficient. We use k-fold x-validation to learn the "best" model. Both classifiers are extended for learning and prediction over time.

Weighted Relational Bayes Classifier. RBCs extend naive Bayes classifiers [5] to relational settings by treating heterogeneous relational subgraphs as a homogeneous set of attribute multisets. The weighted RBC uses standard maximum likelihood learning. More specifically, the sufficient statistics for each conditional probability distribution are computed as weighted sums of counts based on the link and attribute weights. More formally, for a class label C, attributes $\mathbf{X}$, and related items R, the RBC calculates the probability of C for an item i of type $G(i)$ as follows:

$$P(C^i|\mathbf{X}, R) \propto \prod_{X_m \in \mathbf{X}^{\mathbf{G(i)}}} P(X_m^i|C) \prod_{j \in R} \prod_{X_k \in \mathbf{X}^{\mathbf{G(j)}}} P(X_k^j|C)P(C)$$

Weighted Relational Probability Trees. RPTs extend standard probability estimation trees to a relational setting. We use the standard learning algorithm [12] except that the aggregate functions are computed after the appropriate links and attributes weights are included for the selected temporal granularity (shown in Figure 1). For prediction, if the model is applied to predict attribute Y_t at time t, we first calculate the weighted data D_t . Then the learned model from time $(t-1)$ is applied to D_t. The weighted classifier is appropriately augmented to incorporate the weights from D_t.

4 Temporal Ensemble Methods

Ensemble methods have traditionally been used to improve predictions by considering a weighted vote from a set of classifiers [4]. We propose temporal ensemble methods that exploit the *temporal dimension of relational data* to construct more accurate predictors. This is in contrast to traditional ensembles that do not explicitly use the temporal information. The *temporal-relational classification framework* and in particular the temporal-relational representations of the time-varying links, nodes, and attributes form the basis of the temporal ensembles (i.e., as a wrapper over the framework). The proposed temporal ensemble techniques are drawn from one of the five methodologies described below.

1. **Transforming the Temporal Nodes and Links:** The first method learns an ensemble of classifiers, where each of the classifiers are learned from, and then applied to, link and node sets that are sampled from each discrete timestep according to some probability. This sampling strategy is performed after selecting a temporal weighting and temporal granularity, and transforming the data to the appropriate temporal-relational representation. We note that the sampling probabilities for each timestep can be modified to bias the sampling toward the present or the past.

2. **Sampling or Transforming the Temporal Feature Space:** The second method transforms the temporal feature space by localizing randomization (for attributes at each timestep), weighting, or by varying the temporal granularity of the features, and then learning an ensemble of classifiers with different feature sets. Additionally, we might use only one temporal weighting function but learn models with different decay parameters or resample from the temporal features.
3. **Adding Noise or Randomness:** The third method is based on adding noise along the temporal dimension of the data, to increase generalization and performance. Specifically, we randomly permute the nodes feature values across the timesteps (i.e., a nodes recent behavior is observed in the past and vice versa) or links between nodes are permuted across time, and then learn an ensemble of models from several versions of the data.
4. **Transforming the Time-Varying Class Labels:** The fourth method introduces variance in the data by randomly permuting the previously learned labels at t-1 (or more distant) with the true labels at t, again learning an ensemble of models from several versions of the data.
5. **Multiple Classification Algorithms and Weightings:** The fifth method constructs and ensemble by randomly selecting from a set of classification algorithms (i.e., RPT, RBC, wvRN, RDN), while using the same temporal-relational representation, or by varying the representation with respect to the temporal weighting or granularity. Notably, an ensemble that uses both RPT and RBC models significantly increases accuracy, most likely due to the diversity of these temporal classifiers (i.e., correctly predicting different instances). Additionally, the temporal-classifiers might be assigned weights based on assessment of accuracy from cross-validation (or a Bayesian model selection approach).

5 Methodology

For evaluating the framework, we use both static (Y is constant over time) and temporal prediction tasks (Y_t changes over time).

5.1 Datasets

PyComm Developer Communication Network. We analyze email and bug communication networks extracted from the python-dev mailing list archive (www.python.org) for the period 01/01/07−09/30/08. The network consists of 13181 email messages, among 1914 users. Bug reports were also extracted and used to construct a *bug* discussion network consisting of 69435 bug comments among 5108 users. The size of the timesteps are three months. We also extracted text from emails and bug messages and use it to dynamically model topics between individuals and teams. Additionally, we discover temporal centrality attributes (i.e., clustering coefficient, betweenness). The prediction task is whether a developer is effective (i.e., if a user closed a bug in that timestep).

Cora Citation Network. The Cora dataset contains authorship and citation information about CS research papers extracted automatically from the web. The prediction tasks are to predict one of seven machine learning papers and to predict AI papers given the topic of its references. In addition, these techniques are evaluated using the most prevalent topics its authors are working on through collaborations with other authors.

5.2 Temporal Models

The space of temporal-relational models are evaluated using a representative sample of classifiers with varying temporal weightings and granularities. For every timestep t, we learn a model on D_t (i.e., some set of timesteps) and apply the model to D_{t+1}. The utility of the temporal-relational classifiers and representation are measured using the area under the ROC curve (AUC). Below, we briefly describe a few classes of models that were evaluated.

- **TENC:** The TENC models predict the temporal influence of both the links and attributes [16].
- **TVRC:** This model weights only the links using all previous timesteps.
- **Union Model:** The union model uses all links and nodes up to and including t for learning.
- **Window Model:** The window model uses the data D_{t-1} for prediction on D_t (unless otherwise specified).

We also compare simpler models such as the RPT (relational information only) and the DT (non-relational) that ignore any temporal information. Additionally, we explore many other models, including the class of window models, various weighting functions (besides exponential kernel), and built models that vary the set of windows in TENC and TVRC.

6 Empirical Results

In this section, we demonstrate the effectiveness of the temporal-relational framework and temporal ensemble methods on two real-world datasets. The main findings are summarized below:

⋆ Temporal-relational models significantly outperform relational and non-relational models.

⋆ The classes of temporal-relational models each have advantages and disadvantages in terms of accuracy, efficiency, and interpretability. Models based strictly on temporal granularity are more interpretable but less accurate than models that *learn* the temporal influence. The more complex models that combine both are generally more accurate, but less efficient.

⋆ *Temporal ensemble methods* significantly outperform non-relational and relational ensembles. In addition, the temporal ensembles are an efficient and accurate alternative to searching over the space of temporal models.

6.1 Single Models[1]

We evaluate the temporal-relational framework using single-models and show that in all cases the performance of classification improves when the temporal dynamics are appropriately modeled.

Temporal, Relational, and Non-Relational Information. The utility of the temporal (TVRC), relational (RPT), and non-relational information (decision tree; DT) is assessed using the most primitive models. Figure 2 compares TVRC with the RPT and DT models that use more features but ignore the temporal dynamics of the data. We find the TVRC to be the simplest temporal-relational classifier that still outperforms the others. Interestingly, the discovered topic features are the only additional features that improve performance of the DT model. This is significant as these attributes are discovered by dynamically modeling the topics, but are included in the DT model as simple non-relational features (i.e., no temporal weighting or granularity).

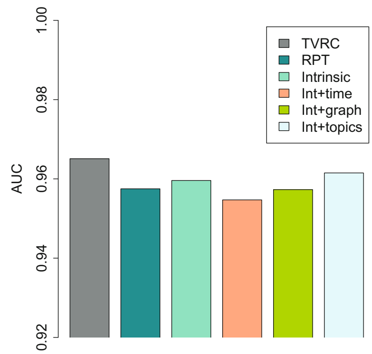

Fig. 2. Comparing a primitive *temporal* model (TVRC) to competing relational (RPT), and non-relational (DT) models

Exploring Temporal-Relational Models. We focus on exploring a representative set of temporal-relational models from the proposed framework. To more appropriately evaluate the models, we remove highly correlated attributes (i.e., that are not necessarily temporal patterns, or motifs), such as "assignedto" in the PyComm prediction task. In Figure 3, we find that TENC outperforms the other models over all timesteps. This class of models are significantly more complex than TVRC since the temporal influence of both links and attributes are learned.

We then explored learning the appropriate temporal granularity. Figure 3 shows the results from two models in the TVRC class where we tease apart the superiority of TENC (i.e., weighting or granularity). However, both TVRC models outperform one another on different timesteps, indicating the necessity for a more precise temporal-representation that optimizes the temporal granularity by selecting the appropriate decay parameters for links and attributes

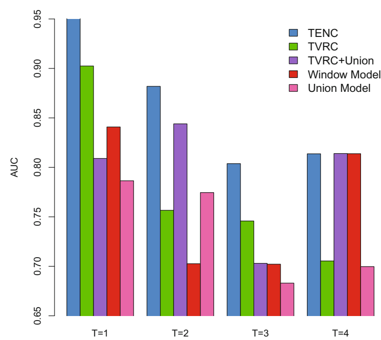

Fig. 3. Exploring the space of temporal relational models. Significantly different temporal-relational representations from the proposed framework are evaluated.

[1] For brevity, some plots and comparisons were omitted [17].

(i.e., TENC). Similar results were found using CORA and other base classifiers such as RBC. Models based strictly on varying the *temporal granularity* were also explored. More details can be found in [17].

6.2 Temporal-Ensemble Models

Instead of directly learning the optimal temporal-relational representation to increase the accuracy of classification, we use *temporal ensembles* by varying the relational representation with respect to the temporal information. These ensemble models reduce error due to variance and allow us to assess which features are most relevant to the domain with respect to the relational or temporal information.

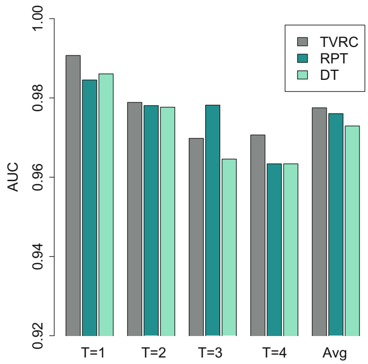

Fig. 4. Comparing temporal, relational, and traditional ensembles

Temporal, Relational, and Traditional Ensembles. We first resampled the instances (nodes, links, features) repeatedly and then learn TVRC, RPT, and DT models. Across almost all the timesteps, we find the temporal-ensemble that uses various temporal-relational representations outperforms the relational-ensemble and the traditional ensemble (see Figure 4). The temporal-ensemble outperforms the others even when the minimum amount of temporal information is used (e.g., time-varying links). More sophisticated temporal-ensembles can be constructed to further increase accuracy. We have investigated ensembles that use significantly different temporal-relational representations (i.e., from a wide range of model classes) and ensembles that use various temporal weighting parameters. In all cases, these ensembles are more robust and increase the accuracy over more traditional ensemble techniques (and single classifiers). Further, the average improvement of the temporal-ensembles is significant at $p < 0.05$ with a 16% reduction in error, justifying the proposed temporal ensemble methodologies.

In the next experiment, we construct ensembles using the feature classes. We use the primitive models (with the transformed feature space) in order to investigate (more accurately) the most significant feature class (communication, team, centrality, topics) and also to identify the minimum amount of temporal information required to outperform relational ensembles.

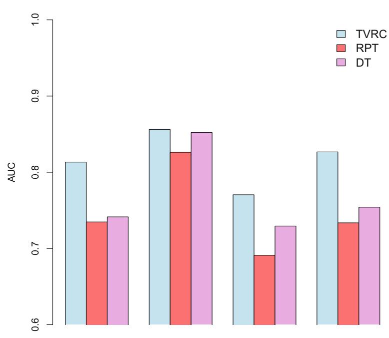

Fig. 5. Comparing attribute classes w.r.t. temporal, relational, and traditional ensembles

In Figure 5, we find several striking temporal patterns. First, the team features are localized in time and are not changing frequently. For instance, it is unlikely that a developer changes their assigned teams and

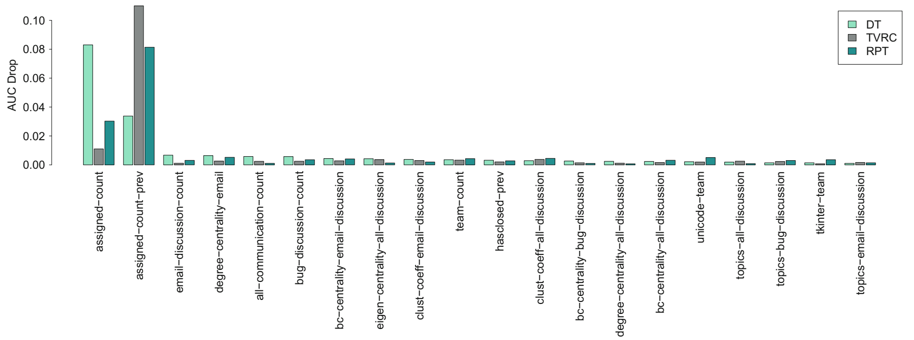

Fig. 6. Randomization. The significant attributes used in the *temporal ensemble* are compared to the relational and traditional ensembles. The change in AUC is measured.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
dev	logged	gt	code	test
wrote	patch	file	object	lib
guido	issue	lt	class	view
import	bugs	line	case	svn
code	bug	os	method	trunk
pep	problem	import	type	rev
mail	fix	print	list	modules
release	fixed	call	set	build
tests	days	read	objects	amp
work	created	socket	change	error
people	time	path	imple	usr
make	docu	data	functions	include
pm	module	error	argument	home
ve	docs	open	dict	file
support	added	windows	add	run
module	check	problem	def	main
things	doc	traceback	methods	local
good	doesnt	mailto	exception	src
van	report	recent	ms	directory

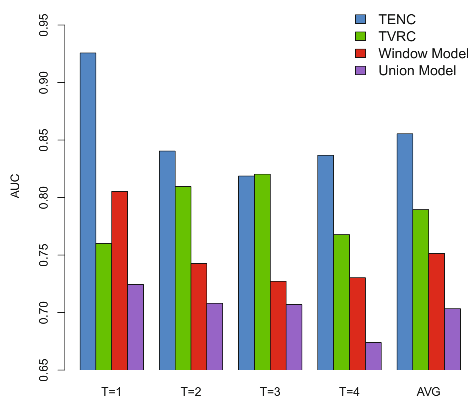

Fig. 7. Evaluation of temporal-relational classifiers using only the latent topics of the communications to predict effectiveness. LDA is used to automatically discover the latent topics as well as annotating the communication links and individuals with their appropriate topic in the temporal networks.

therefore modeling the temporal dynamics only increases accuracy by a relatively small percent. However, the *temporal-ensemble* is still more accurate than traditional ensemble methods that ignore temporal patterns. This indicates the robustness of the temporal-relational representations. More importantly, the other classes of attributes are evolving considerably and this fact is captured by the significant improvement of the temporal ensemble models. Similar performance is also obtained by varying the temporal granularity (see previous examples).

Randomization. We use randomization to identify the significant attributes in the *temporal-ensemble models*. Randomization provides a means to rank and eliminate redundant attributes (i.e., two attributes may share the same significant temporal pattern). We randomize each attribute in each timestep and measure the change in AUC. The results are shown in Figure 6.

We find that the basic traditional ensemble relies on "assignedto" (in the current time step) while the temporal ensemble (and even less for the relational ensemble) relies on the previous "assignedto" attributes. This indicates that relational information in the past is more useful than intrinsic information in the present—which points to an interesting hypothesis that a colleagues behavior (and interactions) precedes their own behavior. Organizations might use this to predict future behavior with less information and proactively respond more quickly. Additionally, the topic attributes are shown to be the most useful for the temporal ensembles (Fig. 7), indicating the utility of using topics to understand the context and strength of relationships.

7 Conclusion

We proposed and validated a framework for temporal-relational classifiers, ensembles, and more generally, representations for temporal-relational data. We evaluated an illustrative set of temporal-relational models from the proposed framework. Empirical results show that the models significantly outperform competing classification models that use either no temporal information or a very limited amount. The proposed temporal ensemble methods (i.e., temporally sampling, randomizing, and transforming features) were shown to significantly outperform traditional and relational ensembles. Furthermore, the temporal-ensemble methods were shown to increase the accuracy over traditional models while providing an efficient alternative to exploring the space of temporal-models. The results demonstrated the effectiveness, scalability, and flexibility of the temporal-relational representations for classification and ensembles in time-evolving domains. In future work, we will theoretically analyze the framework and the proposed ensemble methods.

Acknowledgments. This research is supported by NSF under contract number SES-0823313 and by the Department of Defense through a National Defense Science & Engineering Graduate Fellowship (NDSEG). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements either expressed or implied, of the NSF or the U.S. Government.

References

1. Bifet, A., Holmes, G., Pfahringer, B., Kirkby, R., Gavaldà, R.: New ensemble methods for evolving data streams. In: SIGKDD, pp. 139–148 (2009)
2. Chakrabarti, S., Dom, B., Indyk, P.: Enhanced hypertext categorization using hyperlinks. In: SIGMOD, pp. 307–318 (1998)
3. Cortes, C., Pregibon, D., Volinsky, C.: Communities of Interest. In: Hoffmann, F., Adams, N., Fisher, D., Guimarães, G., Hand, D.J. (eds.) IDA 2001. LNCS, vol. 2189, pp. 105–114. Springer, Heidelberg (2001)

4. Dietterich, T.G.: Ensemble Methods in Machine Learning. In: Kittler, J., Roli, F. (eds.) MCS 2000. LNCS, vol. 1857, pp. 1–15. Springer, Heidelberg (2000)
5. Domingos, P., Pazzani, M.: On the optimality of the simple bayesian classifier under zero-one loss. Machine Learning 29, 103–130 (1997)
6. Domingos, P., Richardson, M.: Mining the network value of customers. In: SIGKDD, pp. 57–66 (2001)
7. Dunlavy, D., Kolda, T., Acar, E.: Temporal link prediction using matrix and tensor factorizations. TKDD 5(2), 10 (2011)
8. Eldardiry, H., Neville, J.: Across-model collective ensemble classification. AAAI (2011)
9. Güneş, İ., Çataltepe, Z., Öğüdücü, Ş.G.: GA-TVRC: A Novel Relational Time Varying Classifier to Extract Temporal Information Using Genetic Algorithms. In: Perner, P. (ed.) MLDM 2011. LNCS, vol. 6871, pp. 568–583. Springer, Heidelberg (2011)
10. Lahiri, M., Berger-Wolf, T.: Structure prediction in temporal networks using frequent subgraphs. In: CIDM, pp. 35–42 (2007)
11. McGovern, A., Collier, N., Matthew Gagne, I., Brown, D., Rodger, A.: Spatiotemporal Relational Probability Trees: An Introduction. In: ICDM, pp. 935–940 (2008)
12. Neville, J., Jensen, D., Friedland, L., Hay, M.: Learning relational probability trees. In: SIGKDD, pp. 625–630 (2003)
13. Neville, J., Jensen, D., Gallagher, B.: Simple estimators for relational Bayesian classifers. In: ICML, pp. 609–612 (2003)
14. Preisach, C., Schmidt-Thieme, L.: Relational ensemble classification. In: ICDM, pp. 499–509. IEEE (2006)
15. Preisach, C., Schmidt-Thieme, L.: Ensembles of relational classifiers. KIS 14(3), 249–272 (2008)
16. Rossi, R., Neville, J.: Modeling the evolution of discussion topics and communication to improve relational classification. In: SOMA-KDD, pp. 89–97 (2010)
17. Rossi, R.A., Neville, J.: Representations and ensemble methods for dynamic relational classification. CoRR abs/1111.5312 (2011)
18. Sharan, U., Neville, J.: Temporal-relational classifiers for prediction in evolving domains. In: ICML (2008)