Bioinformatics
doi.10.1093/bioinformatics/xxxxxx
Advance Access Publication Date: Day Month Year
Manuscript Category

Doc-Start

A Probabilistic Approach for Collective Similarity-based Drug-Drug Interaction Prediction

Dhanya Sridhar [1,*], Shobeir Fakhraei [1,2] and Lise Getoor[1]

[1]Computer Science Department, University of California Santa Cruz, Santa Cruz, CA, 95050, United States
[2]Computer Science Department, University of Maryland, College Park, MD, 20740, United States

*dsridhar@ucsc.edu

Associate Editor: Jonathan Wren

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: As concurrent use of multiple medications becomes ubiquitous among patients, it is crucial to characterize both adverse and synergistic interactions between drugs. Statistical methods for prediction of putative drug-drug interactions can guide *in-vitro* testing and cut down significant cost and effort. With the abundance of experimental data characterizing drugs and their associated targets, such methods must effectively fuse multiple sources of information and perform inference over the network of drugs.
Results: We propose a probabilistic approach for jointly inferring unknown drug-drug interactions from a network of multiple drug-based similarities and known interactions. We use the highly scalable and easily extensible probabilistic programming framework *Probabilistic Soft Logic*. We compare against two methods including a state-of-the-art drug-drug interaction prediction system across three experiments and show best performing improvements of more than 50% in AUPR over both baselines. We find five novel interactions validated by external sources among the top-ranked predictions of our model.
Availability: Final versions of all datasets and implementations will be made publicly available.
Contact: dsridhar@ucsc.edu

1 Introduction

Increasingly, patients use multiple pharamceutical drugs simultaneously to treat their illnesses. Interactions between drugs can result in reduced efficacy of one or more drugs, and in some cases, even deletrious side-effects. The risk of adverse effects is higher in demographics like the elderly that commonly take multiple medications at once. On the other hand, certain drugs interact to produce synergistic effects that are more effective in combatting diseases like cancer (Nahta *et al.*, 2004; Chou, 2010). Crowther *et al.* (1997) characterizes a drug-drug interaction (DDI) as a drug effect that is greater or less than expected in the presence of another drug. While it remains crucial to verify potential DDIs *in vitro*, it is prohibitively expensive to exhaustively test all possible interactions. Therefore, computational modeling and predictive methods provide a viable way to identify the most salient potential interactions for downstream experimental validation (Zhang *et al.*, 2009).

Interactions between drugs are classified as *pharmacokinetic* and *pharmacodynamic*. Computational and mathematical modeling methods rely on current understanding of the mechanisms underlying each of these types of interactions and are specific to each interaction type. A pharmacokinetic interaction with a drug affects the process by which the other drug is absorbed, distributed, metabolized or excreted in the body (Crowther *et al.*, 1997). On the other hand, drugs acting on the same receptor, site of action, or physiological system constitutes a pharmacodynamic interaction. While pharmacokinetic interactions are usually associated with an adverse or exaggerated response, pharmacodynamic interactions are implicated in both synergistic and detrimental effects. Many pharmacokinetic interactions are facilitated by the enzyme family *Cytochrome P450* (CYP) and extensive but incomplete knowledge of its mechanisms have been used for computational modeling of pharmacokinetic interactions (Ekins and Wrighton, 2001). Similarly, prior work has applied mathematical modeling of known drug response mechanisms to simulate and predict pharmacodynamic interactions (Jonker *et al.*, 2005; Jin *et al.*, 2011).

In contrast to computational modeling, statistical and predictive methods leverage data and evidence from related experiments as domain knowledge and biological priors. Recent advancements in high-throughput experimentation have generated a wealth of biological characterizations of drug compounds and their target genes (Fakhraei *et al.*, 2015). A key challenge for statistical models of drug-drug interactions, or the closely

related problem of drug-target interactions, lies in fusing or combining information from multiple data sources. Much related work has developed ways of computing similarity scores between drugs or pairs of drugs to be used as features for machine learning classifiers (Cheng and Zhao, 2014; Gottlieb *et al.*, 2012; Atias and Sharan, 2011; Vilar *et al.*, 2014, 2013). Sophisticated algorithms such as restricted Boltzmann machines and matrix factorization are especially effective in combining two types of similarities by learning latent representations of the entities (Wang and Zeng, 2013; Cao *et al.*, 2015; Gönen, 2012), however, they do not inherently support multiple similarities in the same model. Statistical methods for DDI prediction are more generalizable as they do not rely on extensive expert knowledge of each mode of interaction. Although Park *et al.* (2015); Huang *et al.* (2013) apply their predictive methods only to pharmacodynamic interactions, statisicals models can be easily extended to both types of interactions. To the best of our knowledge, Gottlieb *et al.* (2012) present state-of-the-art results for drug-drug interaction prediction of both pharmacokinetic and pharmacodynamic interactions with their INDI system for combining similarity measures to use features for a local logistic regression classifier.

However, these similarity-based methods neglect the structural information encoded in the biological network of drugs and their interactions. Two general types of approaches have been studied for adding network information to similarity-based features: methods that compute additional network-based features and methods that perform inference directly over the structure of the network. We refer to these as network-similarity methods and network-based inference methods, respectively. Both kinds of approaches begin by formulating a graph of drugs and their interactions. Network-similarity methods proceed by computing *relational* features, based on the local neighborhoods of drugs such as neighborhood overlap and other well-studied network attributes (Cheng and Zhao, 2014; Cao *et al.*, 2015; Huang *et al.*, 2013). The relational features supplement the similarity information given as input to a classifier. In contrast, network-based inference methods reason over the graph structure when predicting interactions.

Multiple network-based inference approaches have been introduced for the closely related problem of drug-target interaction prediction. Bleakley and Yamanishi (2009) formulate the problem of inferring missing links in a bipartite graph of drugs and targets, and introduce a model that uses local bipartite structure for prediction. Cheng *et al.* (2012); Mei *et al.* (2013) similarly leverage local bipartite topology for inference and Park *et al.* (2015) introduce a random walk approach for reasoning over the network of drugs and targets. However, local network-based features cannot enforce global constraints based on the full graph of entities. Given local relational features, current network-based inference methods follow traditional machine learning algorithms in assuming the instances to be independent and identically distributed. In recent work, Fakhraei *et al.* (2014, 2013) improve upon existing bipartite drug-target interaction prediction approaches using the probabilistic programming framework *Probabilistic Soft Logic* (PSL) to jointly classify all interactions, fusing similarity relations and global network information.

In this work, we formalize the problem of network-based drug-drug interaction prediction using multiple similarity relations. We *collectively* predict drug-drug interactions, considering statistical dependencies between predictions along with knowledge of observed interactions using Probabilistic Soft Logic. We apply our collective approach to predict DDIs on three kinds of interactions: (1) CYP-related interactions (2) non-CYP related interactions (3) general interactions documented by Drugbank (Wishart *et al.*, 2006). For all settings, we evaluate our collective approach against two non-collective methods including state-of-the-art INDI system of Gottlieb *et al.* (2012) and a non-collective PSL model. Our model achieves statistically significant improvement up to 5% in area under the ROC (AUC) results from Gottlieb *et al.* (2012). We further assess area under the precision-recall curve (AUPR) for all methods and show that our collective DDI prediction approach significantly outperforms the state-of-the-art baseline method by up to 50%. Finally, we present important novel DDIs predicted by our approach that are validated in literature.

2 Materials

We use two datasets for our experimental evaluation. The first dataset, released by Gottlieb *et al.* (2012), includes pairwise interactions between 807 drugs, with the drug IDs anonymized . We constructed the second dataset by extracting interactions from Drugbank for the 315 drugs used by Fakhraei *et al.* (2014); Perlman *et al.* (2011), where Drugbank IDs are provided for additional validation. The following section described interaction types and similarities used in these datasets.

2.1 Drug Interaction Data

For the *first* dataset, Gottlieb *et al.* (2012) download 10,702 interactions from DrugBank and 70,099 interactions listed as moderate or high from Drugs.com website(Wishart *et al.*, 2006). The dataset contains two types of interactions: (1) CYP-related interactions (CRDs), where both drugs are metabolized by the same cytochrome P450 (CYP) enzyme (2) non-CYP-related interactions, where no CYP is shared between the drugs (NCRDs). After filtering and processing, the final dataset includes 10,106 CRD and 45,737 NCRD DDIs (Gottlieb *et al.*, 2012) across 807 drugs.

For the *second* dataset, we download interactions from DrugBank version 4.3 for the 315 drugs used by Perlman *et al.* (2011); Fakhraei *et al.* (2014). We cross referenced Drugbank IDs released for the 315 drugs to extract the listed drug interactions, resulting in 4293 known interactions.

2.2 Drug Similarity Data

Both datasets contain seven drug-drug similarities. Four of these similarity measures are drug-based: Chemical-based, Ligand-based, Side-effect-based, Annotation-based. Three similarities are between drug targets and computed by aggregating over known targets for the drugs: Sequence-based, PPI network-based, and Gene Ontology-based. In the first dataset, Gottlieb *et al.* (2012) average maximal similarities between the associated targets for drugs that have more than one target. In the second dataset, we average over all possible pairwise similarities between target genes for drugs that have multiple targets.

Following section provides a brief description of the methods in Gottlieb *et al.* (2012); Perlman *et al.* (2011) for similarity extraction:

Chemical-based is the Jaccard similarity, or closely related Dice similarity, of molecular fingerprints from pairs of drugs. Molecular fingerprints are retrieved from cheminformatics toolkits such as chemical development kit (CDK) (Steinbeck *et al.*, 2006) or RDKit using canonical SMILES[1]. Fingerprinting methods represent molecules as bit strings for fast similarity computation and are grouped into hashing-based and structural methods. Hashed fingerprints such as the Daylight method rely on hash functions to represent linear substructures of molecules as bit strings. Structural fingerprints such as MACCS, Atom-Pair, Morgan and Feature-Based Morgan methods use features of molecular substructures to compute bit strings. The Jaccard and Dice similarity scores between two sets X and Y are defined as

$$\text{Jaccard}(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}, \text{Dice}(X,Y) = \frac{2|X \cap Y|}{|X| + |Y|}$$

[1] Simplified Molecular Input Line Entry Specification

We obtain all fingerprints described above from RDKit and additionally, the hashed fingerprint from CDK computed with default values as used by Gottlieb *et al.* (2012). In our experiments, we use the hashed fingerprint from CDK after comparing performance of all fingerprinting methods on development data.

Ligand-based is the Jaccard similarity between the corresponding sets of protein-receptor families for each drug pair. The protein-receptor is obtained from the similarity ensemble approach (SEA) search tool (Keiser *et al.*, 2009) Drugs' canonical SMILES compared with a collection of ligands[2].

Side-effect-based is the Jaccard similarity score between common side-effects for each pair of drugs.

Annotation-based is the Resnik semantic similarity (Resnik *et al.*, 1999) of Drugs' ATC codes mapped to the World Health Organization ATC classification system (Skrbo *et al.*, 2003).

Sequence-based is the Smith-Waterman sequence alignment score between the corresponding drug targets (proteins). They are normalized via dividing the pairwise score by the geometric mean of the alignment scores of each sequence against itself, suggested in Bleakley and Yamanishi (2009).

Protein-protein interaction network-based is the distance between pairs of corresponding drug targets using their corresponding proteins in the human protein-protein interactions network via an all-pairs shortest path algorithm.

Gene Ontology-based is the Resnik semantic similarity (Resnik *et al.*, 1999) between Gene Ontology annotations of drugs' corresponding targets.

For more detailed descriptions of these similarities, refer to Perlman *et al.* (2011); Gottlieb *et al.* (2012).

3 Methods

3.1 The Problem of Drug-Drug Interaction Prediction

We consider the problem of inferring new edges in a partially observed graph of interactions between drug vertices by leveraging multiple known similarity relations between vertices. We are given a set of drugs $D = \{D_1 \dots D_n\}$. We observe a set of interaction edges between the drugs denoted by $n \times n$ interaction matrix I where $I_{ij} = 1$ indicates an interaction between d_i and d_j and is 0 indicates an unobserved or missing edge. Additionally we are given a set of $n \times n$ drug-drug similarity relations encoded by tables $\{M_1 \dots M_k\}$ where $M_{l_{ij}} \in [0, 1]$ and indicates similarity between d_i and d_j according to biological similarity l.

We define a drug network as a multigraph $G = (V, \mathbf{E})$ where $V = D$ is the vertex set of drugs and $\mathbf{E} = \{M_1 \dots M_k\} \cup I$ is the collection of multiple edge types given by the similarity relations and the interaction matrix I. The drug-drug interaction prediction problem is to use all the information encoded in G to predict the unobserved interaction edges between drug vertices in G.

3.2 Collective Probabilistic Reasoning for Network-based Inference of Interactions

Given all the information G, we want to infer interaction values for missing edges $U = \{(d_i, d_j) | I_{ij} = 0\}$. Many techniques have been studied for inference of missing links but here we focus on the intersection of two well known approaches: network-based methods and collective probabilistic methods. Generally, network-based inference techniques make use of the structure of G by considering the local neighborhoods for each d_i and d_j in edges we want to infer. For example, network-based methods might include set similarity of the neighbors of d_i and d_j along with the local edge similarites encoded in G. Collective probabilistic methods learn joint distributions $P(U, G)$ to infer the most probable joint assignment to all edges in U thereby leveraging statistical dependencies between prediction targets as well as the observations in G. Network-based collective methods combine the two techniques by parametrizing $P(U, G)$ according to structural features of G. Collective prediction methods have been shown to work well in the closely related setting of drug-target interaction prediction (Fakhraei *et al.*, 2014). Below we describe hinge-loss Markov random fields and probabilistic soft logic, a framework for performing network based collective inference, and describe our model for drug-drug interaction prediction.

3.2.1 Hinge-loss Markov Random Fields and Probabilistic Soft Logic

Our model for collective drug-drug interaction prediction uses a special class of Markov random field (MRF). In this section, we review the foundations of our model and describe its use in interaction prediction between drugs.

Probabilistic Soft Logic (PSL) uses declarative, first order logic-like syntax to template for a special class of Markov random field (MRF) model known as hinge-loss MRFs (HL-MRFs). HL-MRFs admit efficient, scalable and exact maximum a posteriori (MAP) inference (Bach *et al.*, 2015). These models are defined over continuous random variables, which provide a natural interpretation for real-valued similarities. MAP inference in HL-MRFs is a convex optimization problem over these variables. Formally, a hinge-loss MRF defines a joint probability density function of the form

$$P(\mathbf{Y}|\mathbf{X}) = \frac{1}{\mathcal{Z}} \exp\Big(-\sum_{r=1}^{M} \lambda_r \phi_r(\mathbf{Y}, \mathbf{X})\Big), \quad (1)$$

where the entries of target variables $\mathbf{Y}$ and observed variables $\mathbf{X}$ are in $[0, 1]$, λ is a vector of weight parameters, $\mathcal{Z}$ is a normalization constant, and

$$\phi_r(\mathbf{Y}, \mathbf{X}) = (\max\{l_r(\mathbf{Y}, \mathbf{X}), 0\})^{\rho_r} \quad (2)$$

is a *hinge-loss potential* specified by a linear function l_r and optional exponent $\rho_r \in \{1, 2\}$. Given a collection of logical implications based on domain knowledge described in PSL and a set of observations from data, the rules are instantiated, or grounded out, with known entities in the dataset. Each instantiation of the rules maps to a hinge-loss potential function as in Equation 2, and the potential functions define an HL-MRF model.

To illustrate modeling in PSL, we consider a prototypical similarity based rule that encourages transitive closure for link prediction between entities a, b, c:

$$\text{Similar}(a, b) \wedge \text{Link}(b, c) \rightarrow \text{Link}(a, c)$$

where *link* represents the continuous target variable for a link prediction task and *similar* is a continuous observed variable. The convex relaxation of this logical implication for continuous truth values is

$$\max(\text{similar}(a, b) + \text{link}(b, c) - \text{link}(a, c) - 1, 0)$$

and can be understood as its *distance to satisfaction*.

MAP inference minimizes the weighted, convex distances to satisfaction to find an consistent joint assignment for all the target variables.

[2] A substance that binds with a biomolecule to serve a biological purpose.

Higher rule weights induce higher penalties for violating the rule increasing its relative importance to other rules. Weights are learned from data through maximum likelihood estimation using training data and the structured perceptron algorithm. Exact MAP inference is performed on the learned model to find the most likely assignments for variables using the consensus based ADMM algorithm. PSL supports latent variable modeling with additional EM-based learning algorithms. For a full description of PSL, see Bach *et al.* (2015). Thus, PSL rules encode the domain knowledge that leads to a consistent assignment to all target variables. HL-MRFs have achieved state-of-the-art performance in many domains including the collective drug-target interaction prediction task (Fakhraei *et al.*, 2014). The open source PSL software can be downloaded from the website (`http://psl.umiacs.umd.edu/`).

3.2.2 Collective Drug-drug Interaction Prediction PSL Model

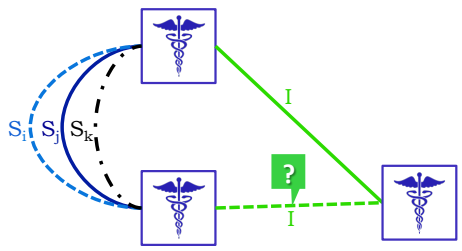

Fig. 1: Triad-based drug-drug interaction prediction rules.

The rules of a PSL model capture beliefs or knowledge about the problem domain. For the drug-drug interaction domain encoded by drug network G, we assert that a drug is likely to be involved in an interaction if it is similar to another drug that is a known interactor. To model the notion of similarity, we are interested in fusing multiple sources of drug similarity. We make this concrete in the full set of rules for drug-drug interaction prediction shown in figure 2.

$$w_1 : \text{Sim}_{Chemical}(D_1, D_2) \wedge \text{Interacts}(D_2, D_3) \rightarrow \text{Interacts}(D_1, D_3)$$
$$w_2 : \text{Sim}_{Ligand}(D_1, D_2) \wedge \text{Interacts}(D_2, D_3) \rightarrow \text{Interacts}(D_1, D_3)$$
$$\ldots$$
$$w_7 : \text{Sim}_{GO}(D_1, D_2) \wedge \text{Interacts}(D_2, D_3) \rightarrow \text{Interacts}(D_1, D_3)$$

Fig. 2: PSL model for collective drug-drug interaction prediction.

where we have one rule for each drug similarity described in section 2, resulting in seven rules. We represent the prediction target with the $\text{Interacts}(D_1, D_3)$ predicate. Given a set of drugs d_1, d_2, and d_3 with known interaction between d_2 and d_3, the rule results in groundings of the form:

$$w_1 : \text{Sim}_{Chemical}(d_1, d_2) \wedge \text{Interacts}(d_2, d_3) \rightarrow \text{Interacts}(d_1, d_3)$$
$$w_2 : \text{Sim}_{Chemical}(d_2, d_3) \wedge \text{Interacts}(d_3, d_1) \rightarrow \text{Interacts}(d_2, d_1)$$

Fig. 3: Small subset of ground PSL rules.

We exclude multiple symmetric groundings for ease of exposition. The ground rules illustrate the propagation of similarity information between target variables as the prediction of $\text{Interacts}(d_1, d_3)$ informs the assignment of $\text{Interact}(d_2, d_1)$. Following Fakhraei *et al.* (2014), we refer to these as 'triad rules' as they encourage triangle completion, or triadic closure. Figure 1 shows a schematic overview of the triad rules. The predicted interaction edge provides evidence for other inferences, resulting in a flow of information throughout the network. This form of collective prediction leverages the full structure of the drug network graph G while combining multiple sources of similarity information. To fully evaluate the impact of joint prediction, we describe below two baseline methods that work non-collectively and assume independence between predicted interactions.

3.3 Comparison Methods

We compare against two non-collective methods including the state-of-the-art INDI framework for inferring interactions between drugs (Gottlieb *et al.*, 2012). We describe each of these below.

3.3.1 State-of-the-art INDI method

Gottlieb *et al.* (2012) introduce the INDI framework for novel drug-drug interaction prediction. They introduce a method for computing similarity scores between target interaction edges to known interaction edges based on the given drug-drug similarities. For each target drug-pair, each pairwise combination of similarities is considered for computing the similarity score to the most similar known drug interaction. The procedure effectively performs nearest neighbor search using different similarity distance measures. Each score is then used as a feature to train a logistic regression classifier. We refer to Gottlieb *et al.* (2012) for full details.

3.3.2 Non-collective PSL Model

To quantify the effect of collective prediction, we also use a non-collective PSL model that considers the dependencies between the target interactions and observed interactions only, as in the INDI method. Formally, we modify the triad rules above as follows: where we introduce the Interacts_{Obs} predicate to limit the triadic closure of predicted interactions to known interactions only.

$$w_1 : \text{Sim}_{Chemical}(D_1, D_2) \wedge \text{Interacts}_{Obs}(D_2, D_3) \rightarrow \text{Interacts}(D_1, D_3)$$
$$\ldots$$
$$w_7 : \text{Sim}_{GO}(D_1, D_2) \wedge \text{Interacts}_{Obs}(D_2, D_3) \rightarrow \text{Interacts}(D_1, D_3)$$

Fig. 4: Non-collective PSL model for drug-drug interaction prediction.

3.4 Experimental Protocol

In order to validate our collective drug-drug interaction prediction method and compare against state-of-the-art methods, we perform experiments on the two drug interaction networks described in section 3. For each dataset, we perform ten-fold cross-validation across all pairs of interactions. We use eight folds as interaction evidence or observations, one fold as training labels to learn weights for the rules, and the final fold as a held-out test set. All similarities between drugs are used as evidence, or features, for the models. The similarity distributions are highly left-skewed, which is problematic for the soft truth interpretation used by PSL, as values below 0.5 do not highly affect the inference. We transform all similaritity values between drugs by taking the cube-root to normalize the distributions and allow for proper interpretation by PSL.

We compute area under the precision-recall curve (AUPR) for the positive class, area under the ROC (AUC) and F1 score on the test set. Link prediction tasks usually suffer from class imbalance as true positive links are sparse compared to true negatives. Related work on general link prediction and DDI prediction report AUC because it is more robust to the skewness than metrics such as accuracy. However, AUC is still sensitive to the high number of negative examples. For practical downstream biological validation of predicted DDIs, it is more important to have a reliable ranking of candidate positive interactions. The precision-recall curve better captures the effectiveness of models at discriminating true

positive examples. F1 score is another measure of classification accuracy and can interpreted as a weighted average of precision and recall. Since PSL outputs real-valued truth scores and logistic regression produces class probabilities, we threshold the values to {0,1} to compute the F1 score. We perform grid search over a range of threshold values between [0,1] to obtain best-performing thresholds.

We implement the INDI feature computation method in `Matlab` by extending a related implementation of the computation for the drug-target interaction prediction setting (Fakhraei *et al.*, 2014; Perlman *et al.*, 2011). We use the logistic regression classifier provided in the `glmfit` package with default settings. For our models, we use the open-source PSL framework. We run 700 iterations of the structured voted-perceptron weight learning algorithm in PSL and use default settings for the ADMM inference algorithm. We will make all code and datasets publicly available.

3.4.1 Blocking Methods for PSL

In a drug network with n drugs and n^2 interactions where PSL considers dependencies between pairs of interactions, the computational complexity reaches $O(n^4)$, which quickly becomes expensive for large networks. To make the approach scalable, we employ a common techniques to block unimportant links from being grounded out by the model. In the PSL triad rule setting, for each similarity i, we limit the possible $\text{Similar}_i(D_1, D_2)$ edges that are considered for each drug D_1. By blocking on the similarity links, we restrict the grounding of all possible triads to only the ones that are most likely.

To block similarities in the grounded out PSL models, for each drug, we perform nearest neighbor search to pick the top 15 most similar other drugs as evidence for $\text{Sim}_i(D_1, D_2)$. In the first drug dataset, for the CRD interaction experiments, we use a more restrictive blocking method to induce more sparsity since CRD interactions are rarer. When searching for the 15 nearest neighbors for each drug, we restrict ourselves to those drugs that have appeared in at least one observed interaction in the full network. In this sparser setting, some drugs may not appear in any $\text{Sim}_i(D_1, D_2)$ groundings. For those drugs, we additionally retrieve five most similar other drugs using standard nearest neighbor search and include the pairs as evidence for $\text{Sim}_i(D_1, D_2)$. Fakhraei *et al.* (2014) provide more comprehensive analysis on techniques for blocking.

4 Results

4.1 Comparison to State-of-the-art Baselines

We compare our proposed collective PSL approach for DDI prediction to two baselines including the state-of-the-art INDI system with 10-fold cross-validation experiments. We apply the three methods to each fold and report average and standard deviations of our chosen metrics for each model. We refer to the INDI system as INDI, the non-collective PSL baseline as NC-PSL, and collective PSL model as PSL. Tables 1-3 present average and standard deviation for area under the precision-recall curve (AUPR) and area under the ROC (AUC) from cross-validation experiments on the three interaction types from two datasets: (1) CYP-related interactions (CRD) from Drugs.com and Drugbank (Gottlieb *et al.*, 2012) (2) Non-CYP-related interactions (NCRD) from Drugs.com and Drugbank (Gottlieb *et al.*, 2012) (3) General interactions from DrugBank. Bolded results highlight statistically significant improvement over both baselines with $\alpha = 0.05$. Figures 5, 6 and 7 show precision-recall curves of all methods plotted for interaction type settings (1), (2) and (3) respectively. Additionally, to assess the benefit of fusing multiple similarities, we compare against our collective PSL model implemented with single similarities. Table 4 shows AUPR for the collective PSL model for single similarities across interaction type settings (1), (2) and (3).

Table 1. Average AUPR, AUC and F1 scores (with best threshold t indicated), and standard deviation for 10 fold CV comparing all DDI prediction models for CRD interactions from dataset 1.

Method	**AUPR-Pos**	**AUROC**	**F1**
INDI	0.15 ± 0.007	0.92 ± 0.003	0.24 ± 0.005 (t = 0.1)
NC-PSL	0.15 ± 0.01	0.91 ± 0.004	0.23 ± 0.01 (t = 0.8)
PSL	**0.34 ± 0.02**	**0.96 ± 0.003**	**0.4 ± 0.02** (t **= 0.3**)

Table 2. Average AUPR, AUC and F1 scores (with best threshold t indicated), and standard deviation for 10 fold CV comparing all DDI prediction models for NCRD interactions from dataset 1.

Method	**AUPR-Pos**	**AUROC**	**F1**
INDI	0.64 ± 0.01	0.95 ± 0.003	0.63 ± 0.01 (t = 0.35)
NC-PSL	0.70 ± 0.006	0.96 ± 0.001	0.62 ± 0.01 (t = 0.9)
PSL	**0.78 ± 0.006**	**0.97 ± 0.001**	**0.70 ± 0.01** (t **= 0.3**)

Table 3. Average AUPR, AUC and F1 scores (with best threshold t indicated), and standard deviation for 10 fold CV comparing all DDI prediction models for general interactions from dataset 2.

Method	**AUPR-Pos**	**AUROC**	**F1**
INDI	0.47 ± 0.04	0.91 ± 0.01	0.51 ± 0.03 (t = 0.2)
NC-PSL	0.56 ± 0.04	0.95 ± 0.006	0.6 ± 0.03 (t = 0.5)
PSL	**0.69 ± 0.02**	**0.96 ± 0.006**	**0.67 ± 0.02** (t **= 0.4**)

Table 4. Average AUPR and standard deviation for 10 fold CV for single similarity collective DDI prediction models across all interaction types

Similarity	**CRD**	**NCRD**	**General**
ATC	0.18 ± 0.01	0.73 ± 0.01	0.68 ± 0.02
Chemical	0.32 ± 0.02	0.58 ± 0.01	0.46 ± 0.04
Distance	0.31 ± 0.03	0.63 ± 0.004	0.35 ± 0.04
Gene Ontology	0.33 ± 0.02	0.63 ± 0.004	0.39 ± 0.04
Ligand	0.18 ± 0.01	0.67 ± 0.01	0.37 ± 0.03
Sequence	0.29 ± 0.02	0.63 ± 0.004	0.37 ± 0.04
Side Effect	0.30 ± 0.01	0.56 ± 0.01	0.51 ± 0.03

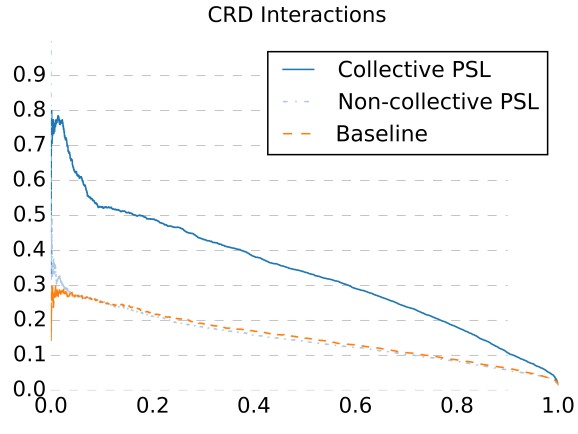

Fig. 5: Precision-recall curves comparing all DDI prediction models on CRD Interactions dataset.

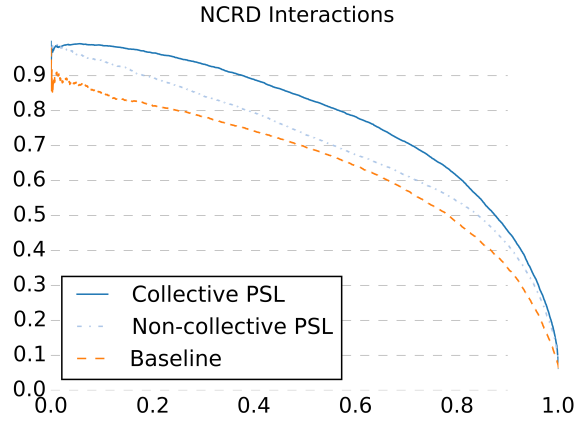

Fig. 6: Precision-recall curves comparing all DDI models on NCRD Interactions dataset.

Table 5. Top ranked PSL model predictions for interactions unknown in DrugBank

Rank	Drug Bank IDs	Drug Bank IDs
1	**DB00870; DB01418**	**Suprofen and Acenocoumarol**
2	DB01067; DB00839	Glipizide and Tolazamide
3	**DB01297; DB00806**	**Practolol and Pentoxifylline**
4	DB00870; DB00806	Suprofen and Pentoxifylline
5	DB00272; DB01232	Betazole and Saquinavir
6	**DB00870; DB01032**	**Suprofen and Probenecid**
7	DB00939; DB01418	Meclofenamic acid and Acenocoumarol
8	**DB00414; DB01032**	**Acetohexamide and Probenecid**
9	DB01297; DB01392	Practolol and Yohimbine
10	**DB01097; DB01262**	**Leflunomide and Decitabine**

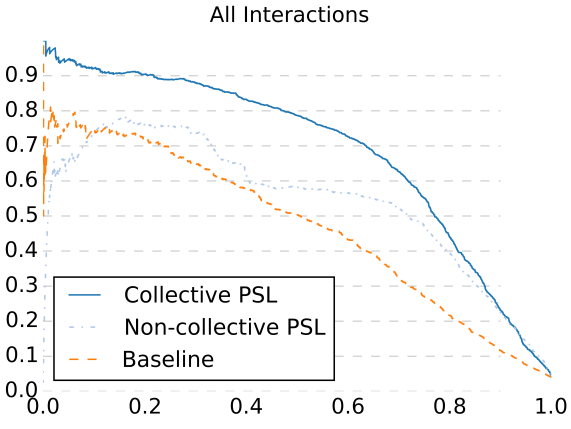

Fig. 7: Precision-recall curves comparing all DDI models on general interactions dataset.

Gottlieb *et al.* (2012) report AUC results consistent with our evaluation of the INDI system as given in tables 1 and 2. Our collective PSL model statistically significantly outperforms both baselines in AUC, AUPR and F1-score for all three interaction type prediction experiments. For AUPR, in the best case CRD interactions setting, our collective model improves up to 50% in AUPR over the state-of-the-art INDI system and non-collective PSL model, from 0.15 to 0.34. For the NCRD and general interactions, the collective PSL approach sees gains of up to 20% in AUPR over both baselines. The collective model improves up to 0.05 in AUC over the state-of-the-art INDI method, with AUC as high as 0.97 for the NCRD interaction setting, signficantly improving over the 0.95 achieved by the INDI system. For F1-score, our collective model improves close to 50% over the INDI and non-collective baselines for the CRD setting and up to 30% for NCRD and general interaction settings. Interestingly, the non-collective PSL method performs at least as well as the INDI system in setting (1) and for settings (2) and (3), significantly outperforms the INDI system in AUPR and AUC. This improvement by the non-collective PSL model demonstrates the method's effectiveness in combining multiple similarities as well as or better than the state-of-the-art similarity combination technique used by INDI. The gains achieved by the fully collective PSL model highlights the benefits of joint inference over the full drug-drug interaction network. Additionally, for all interaction settings, the multiple similarity collective approach significantly improves in AUPR over all individual similarity collective models. This result supports the findings of Fakhraei *et al.* (2014); Gottlieb *et al.* (2012) that multiple similarities benefit performance of both drug-target and drug-drug interaction prediction tasks.

4.2 Validation of Unseen Interaction Predictions

In order for statistical methods to be useful for domain experts, predictive models should produce highly probable novel interactions for subsequent in vitro testing. Thus, following Gottlieb *et al.* (2012); Fakhraei *et al.* (2014); Bleakley and Yamanishi (2009), we compare top-ranked, unseen DDI predictions produced by our collective PSL model with evidence from medical and biological data sources. These predictions are novel with respect to Drugbank interactions used as training data and validate the ability of our collective approach to produce salient interaction predictions given observations.

For this experiment, we use predicted drug-drug interactions from non-anonymized dataset 2 to cross-reference in literature. From our 10-fold cross validation experiments, we output the predictions and filter out those that are not present in Drugbank as verified interactions. We rank these new predictions and consider the top 10 interactions as shown in table 5. Bolded rows indicate drug pairs that are verified by literature or another database as interactors, or have substantive supporting evidence for potential interaction. We use the Interactions Checker tool provided by drugs.com (`http://drugs.com`) for validation, as these interactions were not used to train any of our models. Additionally, Drugbank provides the *BioInteractor* tool that uses drug-target, -enzyme and -transporter associations to predict highly probable interactions that are not included in the main database. Our collective PSL approach highly ranks five interactions that are substantiated by Interactions Checker or BioInteractor. Some interactions involve the following four drugs that are no longer FDA approved or used outside of the United States: Suprofen, Acenocoumarol (used worldwide but not in U.S.), Practolol and Acetohexamide. Because these drugs are presently less well-studied and documented, they arise naturally as test cases for our validation study.

The top predicted interaction is between Suprofen, a non-steroidal anti-inflammatory drug, and Acenocoumarol, an anticoagulant. BioInteractor characterizes the effect of Suprofen on Acenocoumarol as a CYP mediated pharmacokinetic interaction. The sixth most highly ranked prediction involving Suprofen and Probenecid, a uricosuric agent used to treat gout, is also classified by BioInteractor as a CYP related interaction. Acetohexamide is in the sulfonylurea class of compounds used to treat type-II diabetes and is predicted by our model to interact with Probenecid, which is highly protein-bound. Interactions Checker characterizes this particular interaction as enhancing the hypoglycemic effects of sulfonyureas when taken together. The risk of using Probenecid together with Acetohexamide is high with the elderly, who are commonly treated simultaneously for gout and diabetes.

The collective PSL model also ranks Decitabine, used to treat Leukemia, and Leflunomide, used for rheumatoid arthritis treatment, as interactors. Interactions Checker indicates a major interaction between Leflunomide and Decitabine in conjunction since both are immunosuppressants and can have additive effects to increase risk of serious infection. Ranked third, Pentoxifyline, a vasodilator and anti-inflammatory used to improve blood circulation, is predicted to interact with Practolol, a beta-blocked formerly used to treat cardiac arrhythmias. Although Drugs.com does not list this particular interaction, the Interactions Checker lists moderate interaction between Propranolol, beta-blocker now used in place of Practolol, and Pentoxifyline. The prediction of an effect on Pentoxifyline by a drug chemically similar to Propranolol also demonstrates the effectiveness of the PSL triad rules. The propagation of likely interaction across drugs that are similar is also evident in the second ranked prediction of interaction between Glipizide and Tolazamide. Both are sulonyureas like Acetohexamide and are used to treat type-II diabetes. Though the drugs deliver similar responses, currently there is no strong evidence of their interaction.

We compare the predictions in Table 5 to the top ten novel interactions predicted by the INDI system. In contrast, only three out of the ten predictions made by INDI can be verified by BioInteractor or Interactions Checker: (1) Ciprofloxacin and Lomefloxacin (rank 2) (2) Methotrexate and Lomefloxacin (rank 4) (3) Mifepristone and Lomefloxacin (rank 6). There are no overlaps with the predictions ranked highly by PSL. Lomefloxacin and Ciprofloxacin both fight infection, Methotrexate treats cancers and Mifepristone ends pregancy. Interestingly, both the collective PSL model and the INDI system predict interactions involving major cancer drugs, Decitabine and Methotrexate, respectively.

5 Discussion

In this work, we formulate the problem of collective drug-drug interaction prediction. We introduce a joint probabilistic approach using the PSL framework to fuse multiple sources of similarity information together with domain-knowledge of the network structure for this domain. The originality of this work lies in proposing and experimentally validating a highly scalable, collective probabilistic approach for DDI prediction that is easily extensible with different sources of information and similarity measures. We evaluate our approach on two datasets containing three types of interactions, including one extracted for this work with known Drugbank IDs for additional validations. We perform ten-fold cross-validation on all settings and see that our collective PSL model significantly outperforms two other similarity-based methods, including the state-of-the-art INDI system, on two important metrics for link prediction, AUPR and AUC. Our best performing PSL model improves more than 50% upon AUPR of both baselines and achieves a best AUC of 0.97. Moreover, the non-collective similarity-based method implemented in PSL also significantly outperforms INDI in two settings and performs comparably to INDI in other settings. This result also highlights the effectiveness of PSL as an extensible framework for similarity-based reasoning that enjoys the benefits of collective inference shown by the first result. Furthermore, the top then predictions of our best performing collective PSL methods contain five interactions that are unseen in Drugbank but substantiated by Drugs.com and the BioInteractor tool on Drugbank. This result signifies the usefulness of our collective approach for producing high-quality predictions that can be verified experimentally downstream. Another benefit of our collective PSL method is scalability and speed. The focus of the INDI method is combining similarities by computing interaction edge-based similarity score using a nearest-neighbor search approach. This feature computation is a computationally expensive procedure, requiring $O(n^4)$ passes over the drug entities. For a dataset containing 807 drugs, this computation takes approximately 12 hours on average per fold on a single 32GB machine with 4 cores. The comparable non-collective PSL model introduced in this work takes approximately 1 hour for a round of weight learning and inference per fold on the same machine. The collective PSL model completes computation for a fold in approximately 7 hours. The PSL framework admits highly efficient, polynomial-time inference and here, we further reduce computational complexity by blocking unnecessary groundings of the model. Scalability is crucial for link prediction tasks in increasingly massive biological networks, as new drugs are frequently introduced.

The task of DDI prediction is closely related to problems of predicting drug side effects, drug adverse reactions, and synergistic drug pairs. In fact, predicting synergistic drug interactions is just a specific subtask of the DDI prediction problem. Our collective approach for similarity-based reasoning in networks can be applied and generalized to all these related settings.

6 Acknowledgements

We would like to thank Evan Paull, Vladislav Uzunangelov and Joshua Stuart for their valuable feedback and insightful discussions. *Funding:* This work was supported by the National Science Foundation [IIS1218488].
Conflicts of Interest: None declared.

References

Atias, N. and Sharan, R. (2011). An algorithmic framework for predicting side effects of drugs. *J. Comput. Biol.*, **18**, 207–218.

Bach, S. H., Broecheler, M., Huang, B., and Getoor, L. (2015). Hinge-loss markov random fields and probabilistic soft logic. *arXiv preprint arXiv:1505.04406*.

Bleakley, K. and Yamanishi, Y. (2009). Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*, **25**, 2397–2403.

Cao, D.-S., Xiao, N., Li, Y.-J., Zeng, W.-B., Liang, Y.-Z., Lu, A.-P., Xu, Q.-S., and Chen, A. (2015). Integrating multiple evidence sources to predict adverse drug reactions based on a systems pharmacology model. *CPT: Pharmacometrics Syst. Pharmacol.*, **4**, 498–506.

Cheng, F. and Zhao, Z. (2014). Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *Journal of the American Medical Informatics Association*, **21**, e278–e286.

Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., Zhou, W., Huang, J., and Tang, Y. (2012). Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol*, **8**, e1002503.

Chou, T.-C. (2010). Drug combination studies and their synergy quantification using the chou-talalay method. *Cancer research*, **70**, 440–446.

Crowther, N. R., Holbrook, A. M., Kenwright, R., and Kenwright, M. (1997). Drug interactions among commonly used medications. chart simplifies data from critical literature review. *Canadian Family Physician*, **43**, 1972.

Ekins, S. and Wrighton, S. A. (2001). Application of in silico approaches to predicting drug–drug interactions. *J. Pharmacol. Toxicol. Methods*, **45**, 65–69.

Fakhraei, S., Raschid, L., and Getoor, L. (2013). Drug-target interaction prediction for drug repurposing with probabilistic similarity logic. In *ACM SIGKDD 12th International Workshop on Data Mining in Bioinformatics (BIOKDD)*.

Fakhraei, S., Huang, B., Raschid, L., and Getoor, L. (2014). Network-based drug-target interaction prediction with probabilistic soft logic. *Comput. Biol. Bioinf.*, **11**, 775–787.

Fakhraei, S., Onukwugha, E., and Getoor, L. (2015). Data analytics for pharmaceutical discoveries. In *Healthcare Data Analytics*, pages 599–623. CRC Press.

Gönen, M. (2012). Predicting drug–target interactions from chemical and genomic kernels using bayesian matrix factorization. *Bioinformatics*, **28**, 2304–2310.

Gottlieb, A., Stein, G. Y., Oron, Y., Ruppin, E., and Sharan, R. (2012). Indi: a computational framework for inferring drug interactions and their associated recommendations. *Mol. Syst. Biol.*, **8**, 592.

Huang, J., Niu, C., Green, C. D., Yang, L., Mei, H., and Han, J. (2013). Systematic prediction of pharmacodynamic drug-drug interactions through protein-protein-interaction network. *PLoS Comput Biol*, **9**, e1002998.

Jin, G., Zhao, H., Zhou, X., and Wong, S. T. (2011). An enhanced petri-net model to predict synergistic effects of pairwise drug combinations from gene microarray data. *Bioinformatics*, **27**, i310–i316.

Jonker, D. M., Visser, S. A., van der Graaf, P. H., Voskuyl, R. A., and Danhof, M. (2005). Towards a mechanism-based analysis of pharmacodynamic drug–drug interactions in vivo. *Pharmacol. Ther.*, **106**, 1–18.

Keiser, M. J., Setola, V., Irwin, J. J., Laggner, C., Abbas, A. I., Hufeisen, S. J., Jensen, N. H., Kuijer, M. B., Matos, R. C., Tran, T. B., *et al.* (2009). Predicting new molecular targets for known drugs. *Nature*, **462**, 175–181.

Mei, J.-P., Kwoh, C.-K., Yang, P., Li, X.-L., and Zheng, J. (2013). Drug–target interaction prediction by learning from local information and neighbors. *Bioinformatics*, **29**, 238–245.

Nahta, R., Hung, M.-C., and Esteva, F. J. (2004). The her-2-targeting antibodies trastuzumab and pertuzumab synergistically inhibit the survival of breast cancer cells. *Cancer research*, **64**, 2343–2346.

Park, K., Kim, D., Ha, S., and Lee, D. (2015). Predicting pharmacodynamic drug-drug interactions through signaling propagation interference on protein-protein interaction networks. *PloS one*, **10**, e0140816.

Perlman, L., Gottlieb, A., Atias, N., Ruppin, E., and Sharan, R. (2011). Combining drug and gene similarity measures for drug-target elucidation. *J. Comput. Biol.*, **18**, 133–145.

Resnik, P. *et al.* (1999). Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *J. Artif. Intell. Res.(JAIR)*, **11**, 95–130.

Skrbo, A., Begović, B., and Skrbo, S. (2003). [classification of drugs using the atc system (anatomic, therapeutic, chemical classification) and the latest changes]. *Med. Arh.*, **58**, 138–141.

Steinbeck, C., Hoppe, C., Kuhn, S., Floris, M., Guha, R., and Willighagen, E. L. (2006). Recent developments of the chemistry development kit (cdk)-an open-source java library for chemo-and bioinformatics. *Curr. Pharm. Des.*, **12**, 2111–2120.

Vilar, S., Uriarte, E., Santana, L., Tatonetti, N. P., and Friedman, C. (2013). Detection of drug-drug interactions by modeling interaction profile fingerprints. *PloS one*, **8**(3), e58321.

Vilar, S., Uriarte, E., Santana, L., Lorberbaum, T., Hripcsak, G., Friedman, C., and Tatonetti, N. P. (2014). Similarity-based modeling in large-scale prediction of drug-drug interactions. *Nature protocols*, **9**(9), 2147–2163.

Wang, Y. and Zeng, J. (2013). Predicting drug-target interactions using restricted boltzmann machines. *Bioinformatics*, **29**(13), i126–i134.

Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., and Woolsey, J. (2006). Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.*, **34**(suppl 1), D668–D672.

Zhang, L., Zhang, Y. D., Zhao, P., and Huang, S.-M. (2009). Predicting drug–drug interactions: an fda perspective. *The AAPS journal*, **11**(2), 300–306.