

Drug repositioning through pharmacological spaces integration based on networks projections

M. Re¹, M. Mesiti², G. Valentini¹ ✉

¹Department of Computer Science, University of Milano, Italy

²Department of Informatics and Communication, University of Milano, Italy

Motivations

Drug development is a costly and failure-prone process and, in recent years, pharmaceutical industry has experienced a difficult period whereby productivity has not kept pace with increases in research and development costs. As a consequence, quite recently research efforts focused on a novel paradigm for drug development, named drug repurposing, to discover novel pharmacological applications of existing drugs. Computational approaches for drug repositioning focused mainly on small-scale applications, such as the analysis of specific classes of drugs or drugs for specific diseases. Large-scale applications, involving a relatively large number of drugs and diseases, count only a few examples. Despite the availability of many drug repositioning methods, they all suffer from a serious limitation: the inference task is performed in an inhomogeneous similarity space induced by the relationships existing between drugs and a second type of entity (e.g. disease, target, or ligand set), thus making difficult the integration

of multiple sources of biomolecular and chemical data into a homogeneous pharmacological space.

Methods

To overcome this limitation we propose a general framework based on bipartite networks projections for the construction of homogeneous pharmacological spaces. The nature of these network structured projected spaces allows the application of prediction algorithms to homogeneous pharmacological spaces and improves the integration of different chemical, biomolecular and clinical sources of information. At the core of the proposed approach there is the notion of homogeneous pharmacological similarity space defined as a collection of similarities between drugs induced by common relationships between drugs and a second type of suitable entities (i.e. drug-protein target). The reconciliation between these heterogeneous similarity spaces is performed by means of a network projection operation enabling the reduction of a network composed by two types of nodes (i.e. drugs and

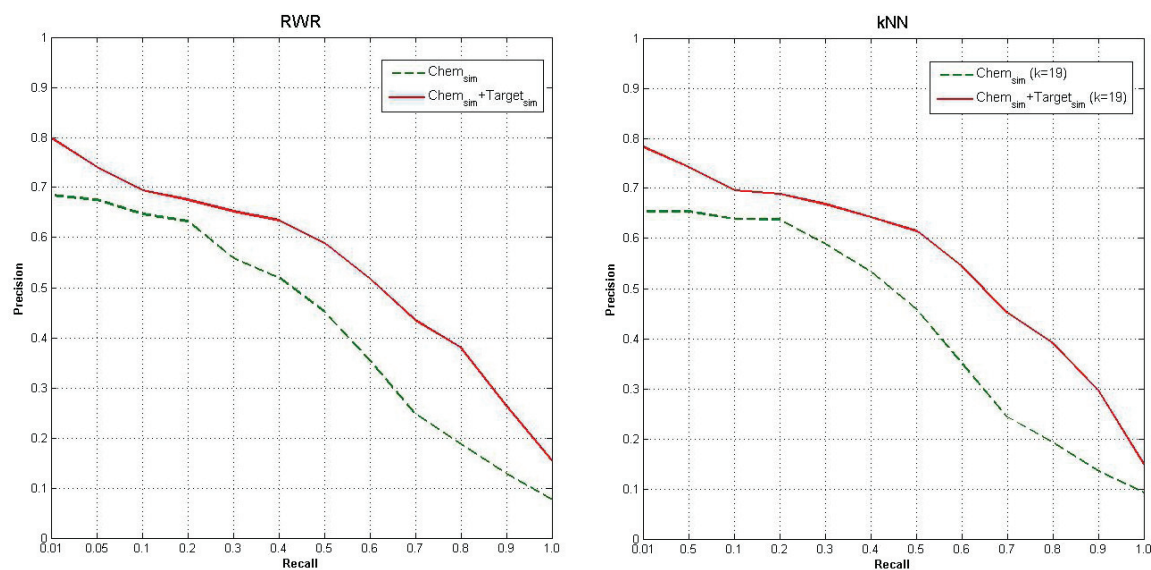


Fig.1 : Precisions at fixed recall levels, with the chemical similarity and chemical/target similarity pharmacological networks. Random Walks with Restart (RWR), our proposed method using the kNN score (kNN). Results are averaged across all the considered therapeutic DrugBank classes.

drug targets) to a network composed only by drugs. A key feature of the proposed framework is its ability to integrate networks of different sizes, enabling the combination of both high and low coverage networks and resulting into a progressively enriched pharmacological similarity network. We also propose a novel and very fast kernelized semi-supervised network based method for ranking drugs according to their likelihood to belong to a given therapeutic category.

Results

We evaluated the proposed approach by integrating two pharmacological similarity spaces accounting, respectively, for chemical similarity and drug-targets interaction similarity, in order to rank about 1300 U.S. Food and Drug Administration (FDA) approved drugs according to DrugBank 3.0 therapeutic categories. The experimental setup is based on a canoni-

cal 5-fold cross validation scheme repeated 10 times. The analysis of the precision at fixed recall levels (see Figure), shows that the integration of pharmacological spaces constructed through the proposed network projections significantly enhances the results obtained with different network-based ranking methods. Moreover our proposed kernelized semi-supervised method for ranking drugs according to a given therapeutic category is at least comparable in terms of AUC and precision at fixed recall, and orders of magnitude faster than state-of-the-art ranking methods. Despite a thorough analysis of the results relative to each therapeutic category is out of the scope of this preliminary investigation, the analysis of the top ranked false positives predicted in three drug categories shows that our proposed approach can be successfully applied to discover potential drug candidates for novel therapeutic indications.