Drug target identification by maximizing information flow

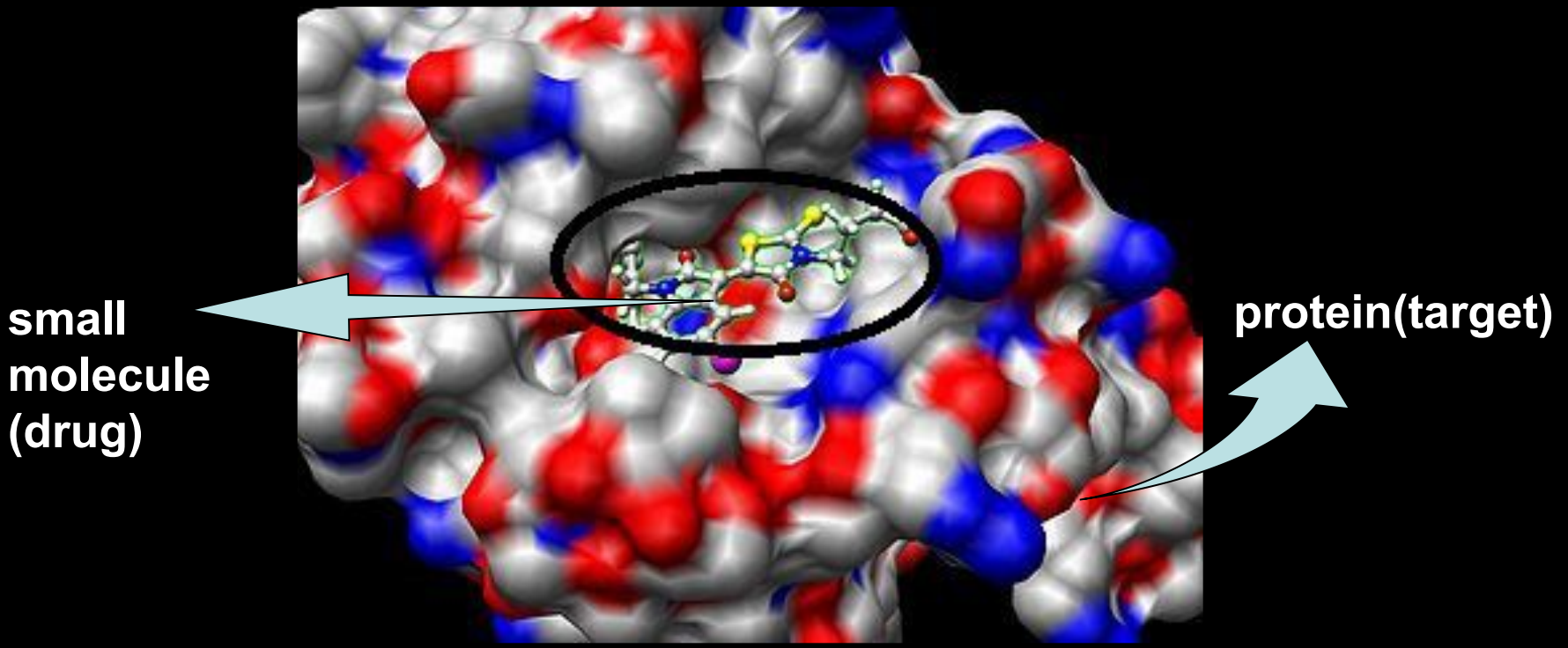
Bingjie (Helen) Zhou

under direction of
Dr. Shao Li
Bioinformatics Division and Center for Synthetic and Systems Biology
Tsinghua university

Research Science Initiative-Tsinghua
July 25,2015

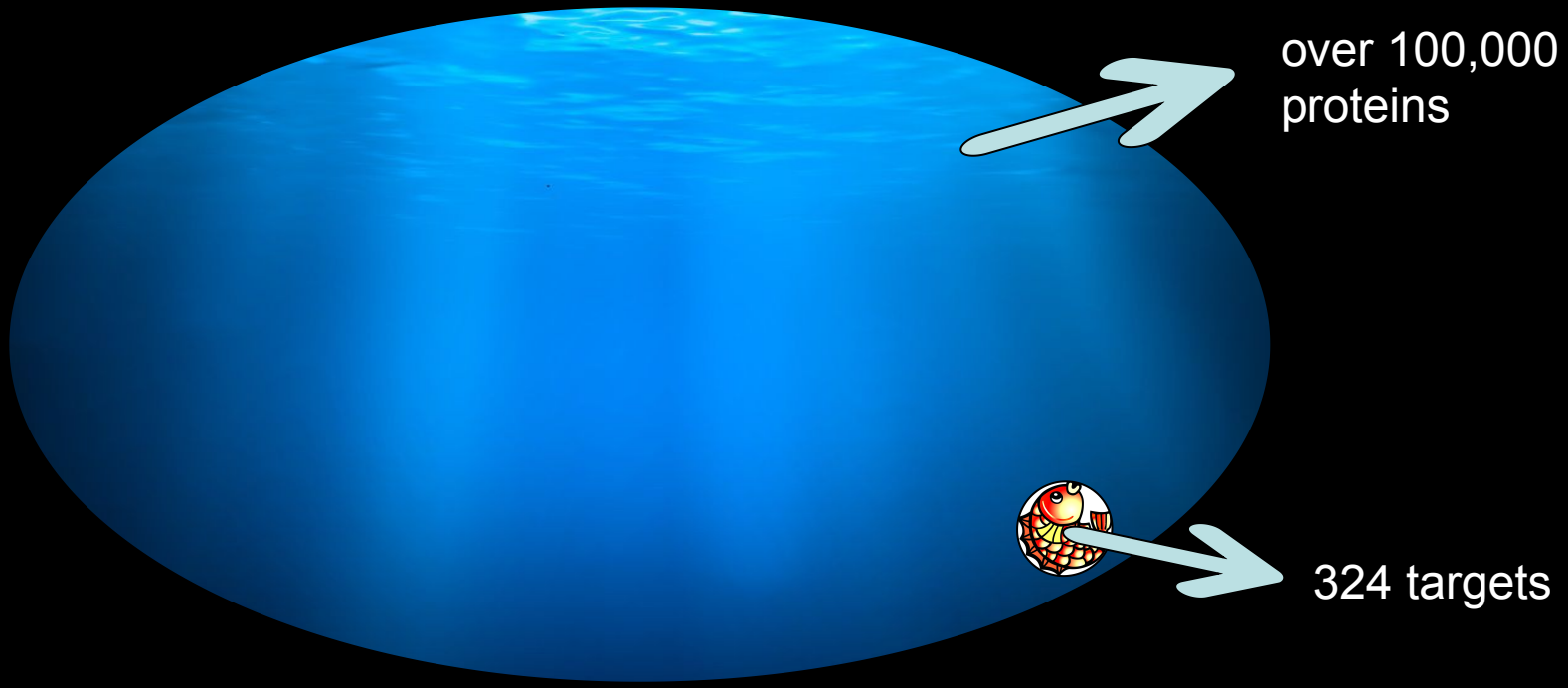
Drug target

Native **protein** in the body whose activity is modified by a drug resulting in a specific effect



Current situation

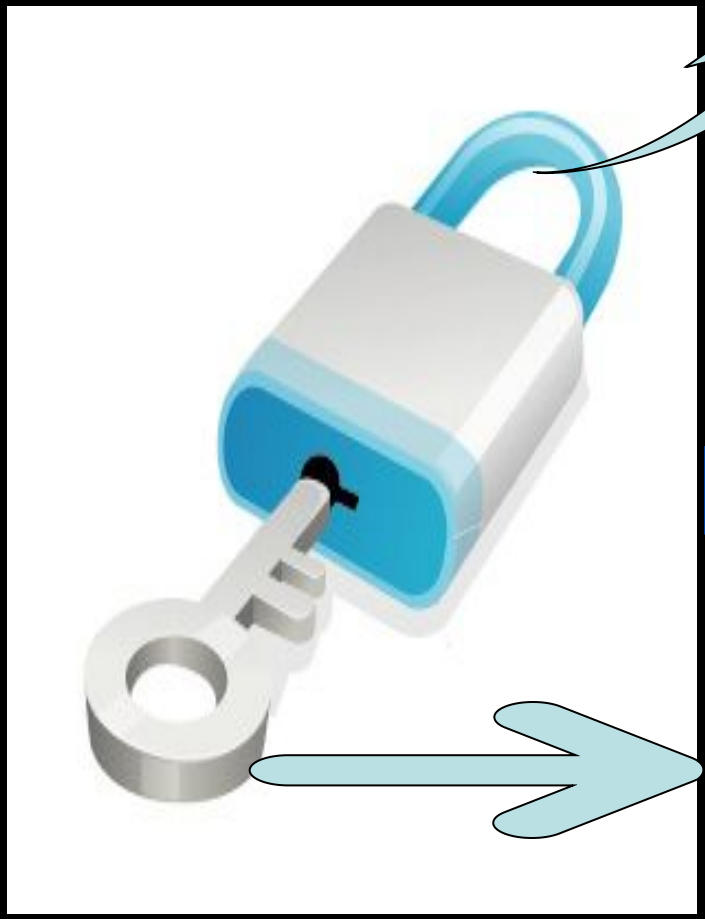
Up to now, only 324 targets have been known for clinical drugs.

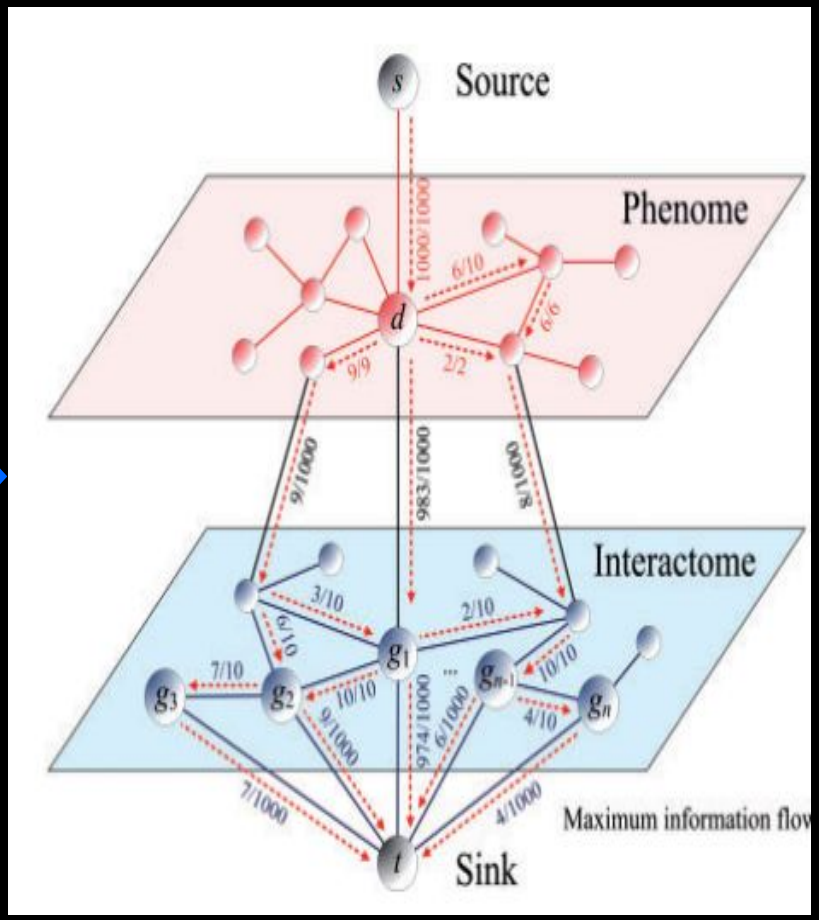


Application

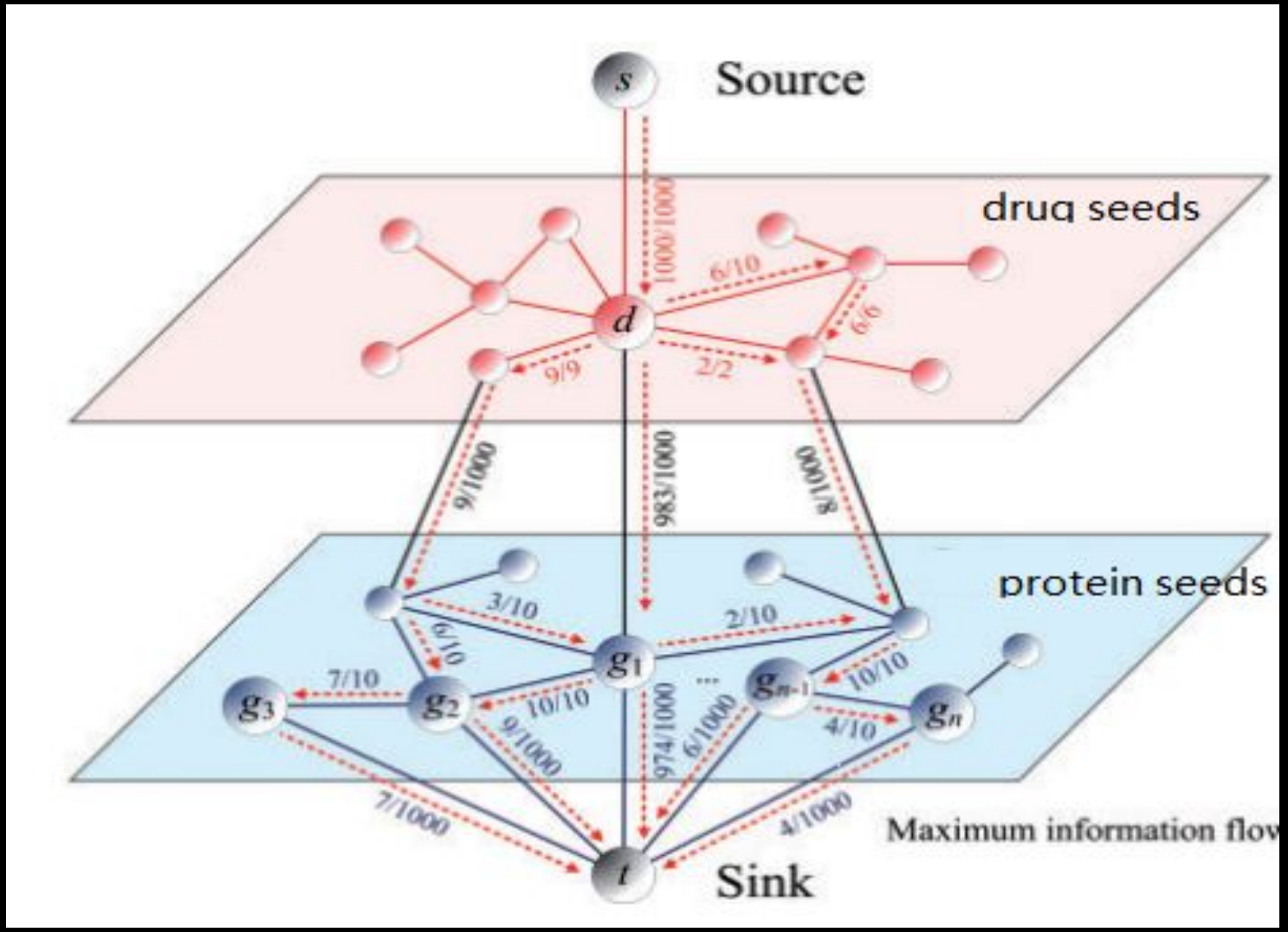
- Identify new drug targets efficiently
- New applications of existing drugs/create new drugs
- Understand side effects and toxicity
- Reduction of time and experiments costs in drug development

Network





Yong Chen et al.(2011)
Uncover disease genes by maximizing information flow in the phenome–interactome network ISMB,2011,i167–i176



Yong Chen et al.(2011)Uncover disease genes by maximizing information flow in the phenome–interactome network ISMB,2011,i167–i176

Collect Data

- Drug-drug similarity
- Protein-protein interaction
- Drug-protein interaction

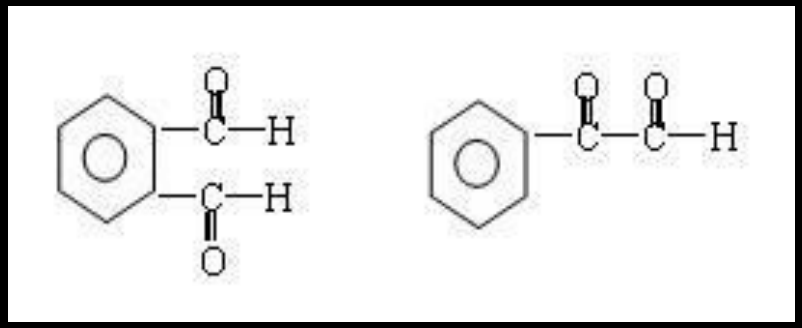
Data resources

Type of interaction	Drug-drug similarity	Drug-target interaction	Protein-protein interaction
Data source	www.drugbank.ca	www.string-db.org	www.drugbank.ca
Website version	Version 4.3	Version 10	Version 4.3

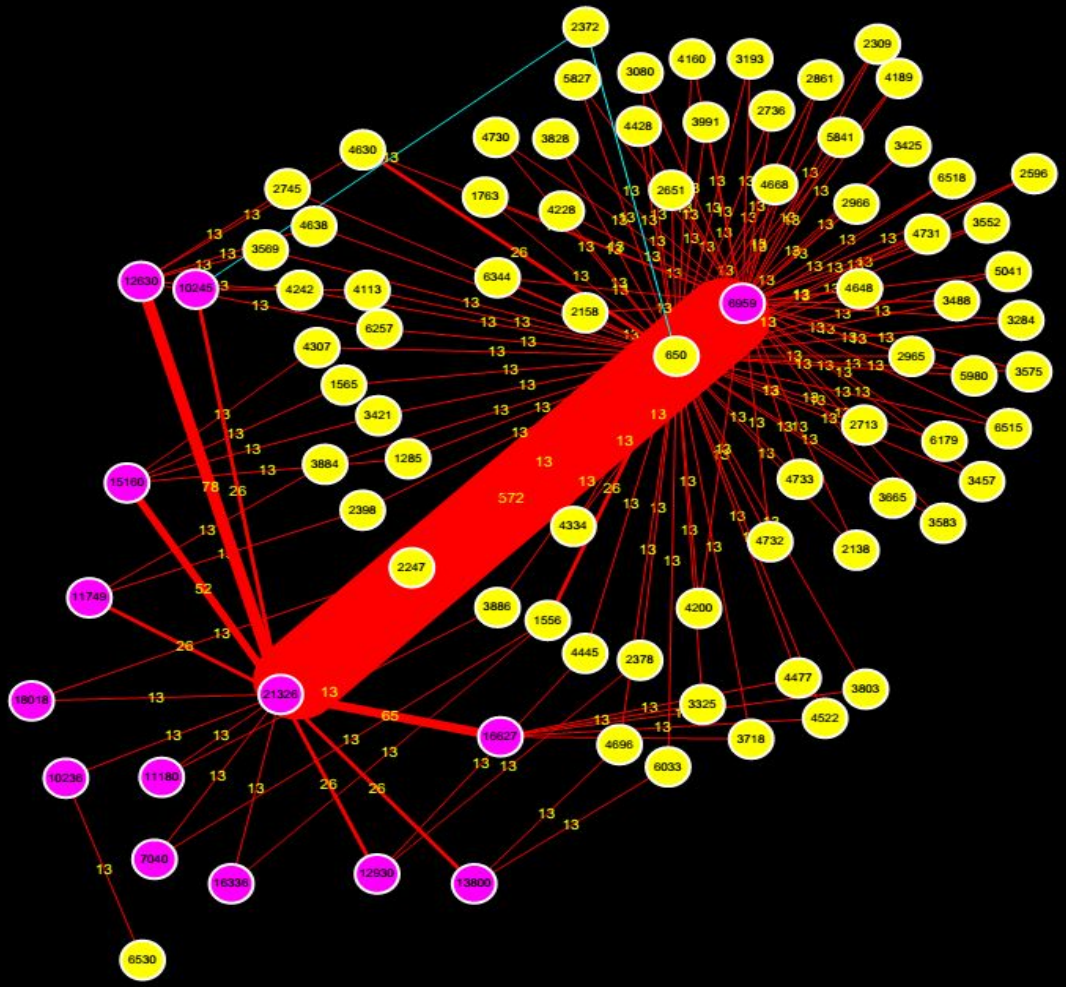
Drug-drug similarity

Assumption: drugs with more similar chemical structure have stronger associations

```
DB00150 N[C@@H](Cc1=CNC2=CC=CC=C12)C(O)=O
DB00151 N[C@@H](CS)C(O)=O
DB00152 CC1=C(CCO)SC=[N+]1CC1=CN=C(C)N=C1N
DB00153 CC(C)[C@@H](C)\C=C\[C@@H](C)[C@@]1([H])CC[C@@]2([H])\C(CCC[C@]12C)=C\C=C1\C
DB00154 CCCCC\C=C/C\C=C/C\C=C/CCCCCCC(O)=O
DB00155 N[C@@H](CCCNC(N)=O)C(O)=O
DB00156 C[C@@H](O)[C@H](N)C(O)=O
DB00158 NC1=NC(=O)C2=NC(CNC3=CC=C(C=C3)C(=O)N[C@@H](CCC(O)=O)C(O)=O)=CN=C2N1
DB00159 CC\C=C/C\C=C/C\C=C/C\C=C/C\C=C/CCCC(O)=O
DB00160 C[C@H](N)C(O)=O
DB00161 CC(C)[C@H](N)C(O)=O
DB00162 C\C(=C/CO)\C=C\C=C(/C)\C=C\C1=C(C)CCCC1(C)C
DB00163 CC(C)CCC[C@@H](C)CCC[C@@H](C)CCC[C@]1(C)CCC2=C(O1)C(C)=C(C)C(O)=C2C
DB00165 CC1=C(O)C(CO)=C(CO)C=N1
DB00166 OC(=O)CCCCC1CCSS1
```

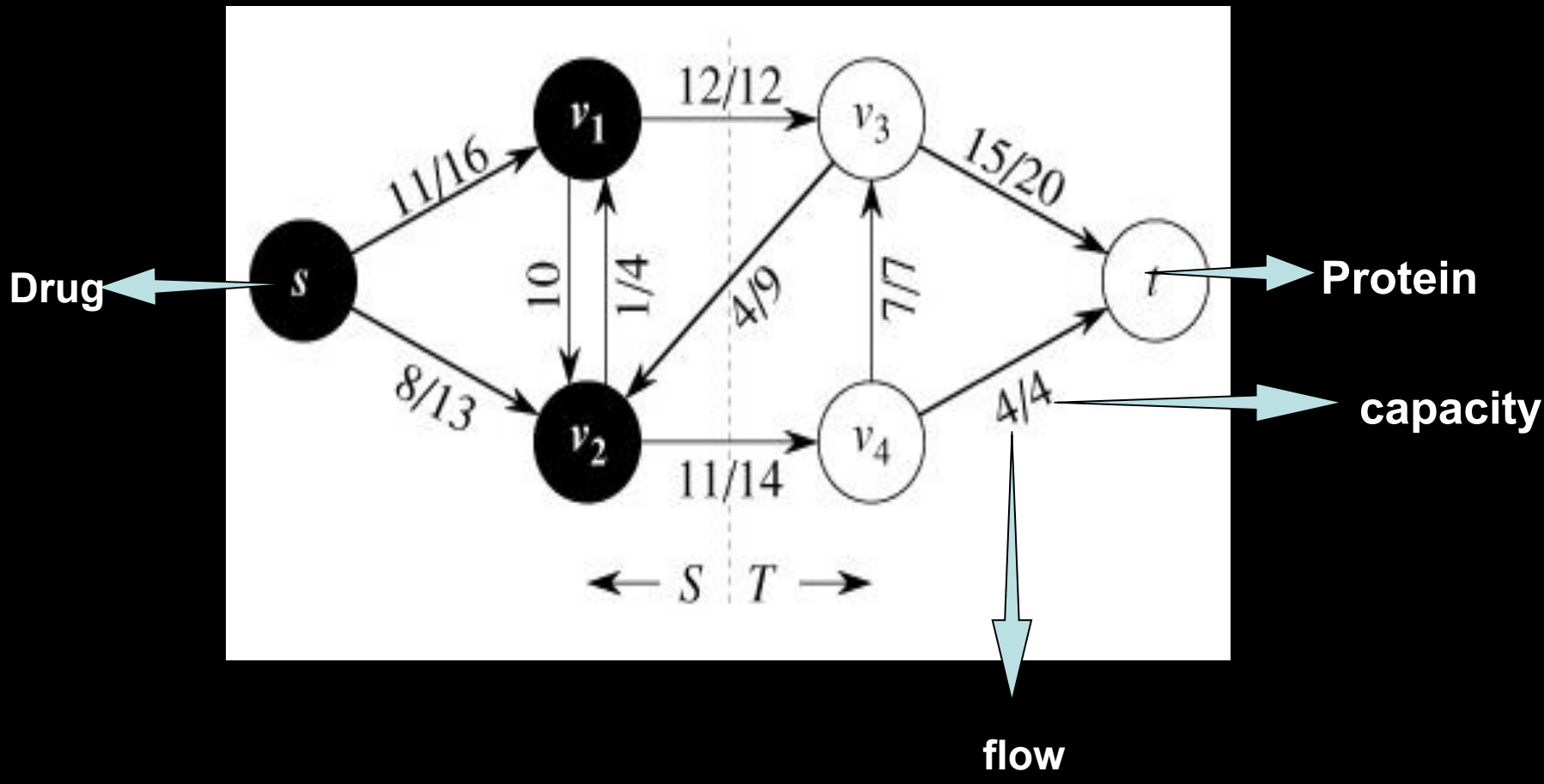


- 6810 drugs
- 14651 proteins
- 23177836 pairs of drug-drug similarities
- 650466 pairs of protein-protein interaction
- 5305 pairs of known drug-target links

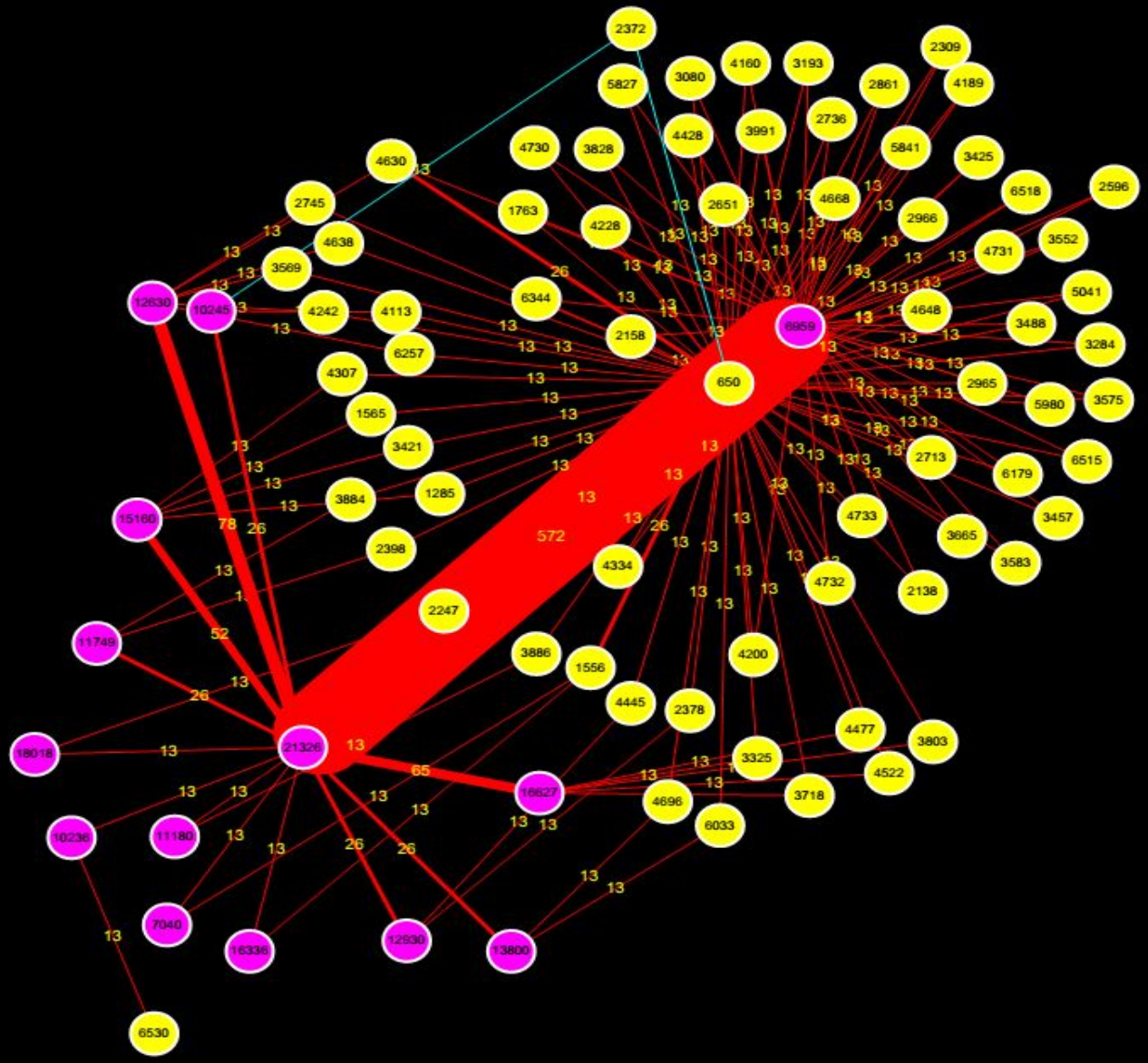


MAXIF

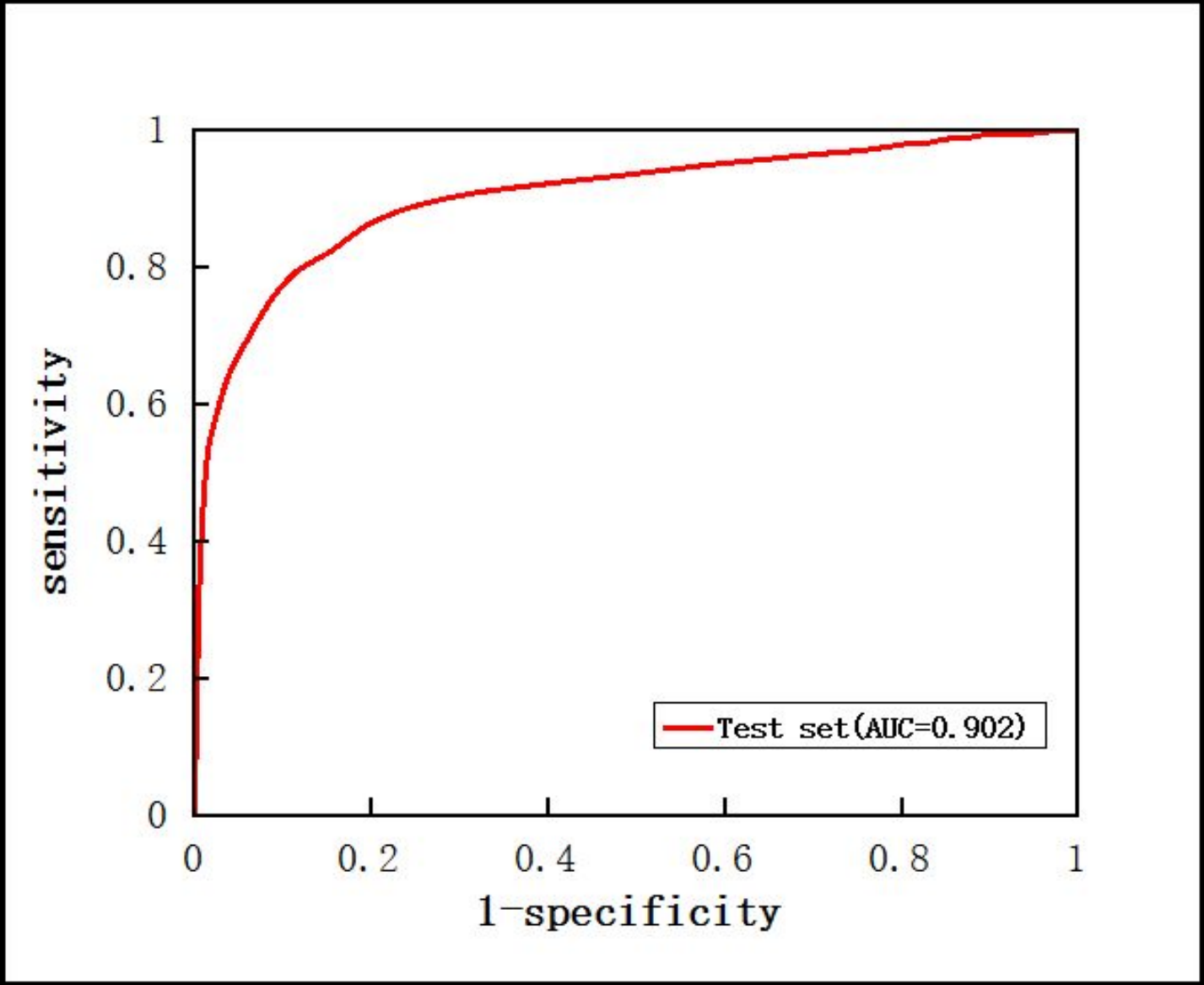
Maximizing information flow



Leave-one-out cross validation



Leave-one-out validation result



Conclusion

MAXIF has high level of sensitivity and specificity.Thus, the method could be applied to

- Predict drug targets efficiently
- Develope new applications of drugs
- Understand side effects
- Reduce time and cost of experiments

Discussion

- Other functional interaction between drugs
- Different levels of association between proteins
- Try different coefficients for different types of interaction

Acknowledgement

Thanks to

- Ms.Ana and Mr.Chris
- My mentor Dr.Shao Li
- my comentor Mr.Zu and labmates
- RSI
- CEE
- Tsinghua university

Thank you for your attention