



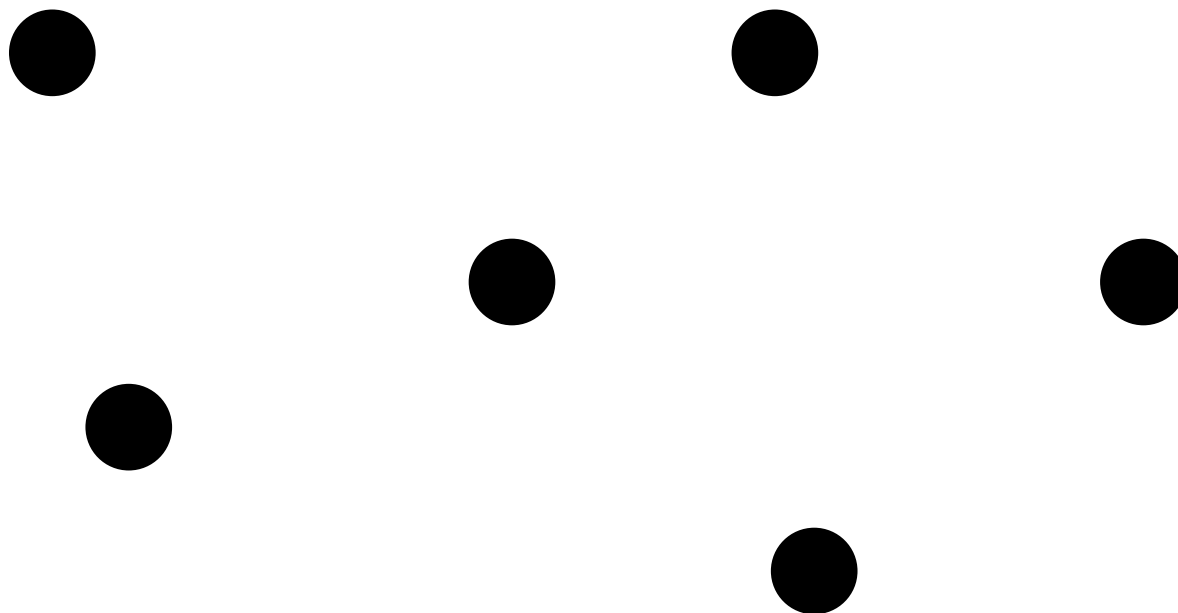
# 一种处理属性图的 深度贝叶斯优化方法

报告人：崔佳旭

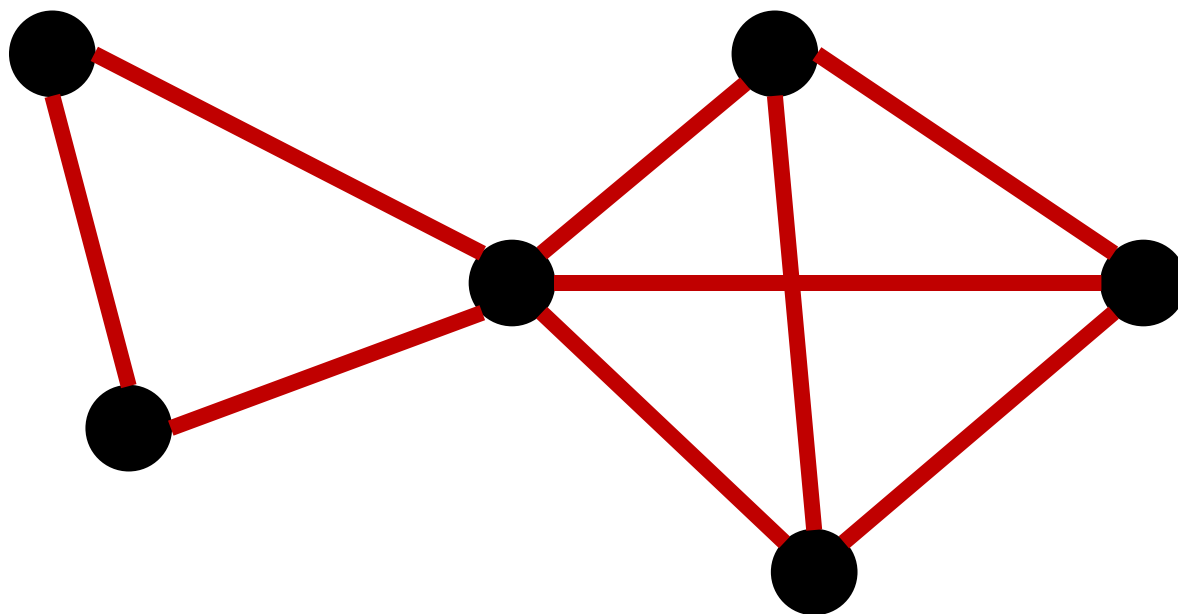
指导教师：杨博 教授

符号计算与知识工程教育部重点实验室  
吉林大学计算机科学与技术学院

# 为什么处理图(Graph)?

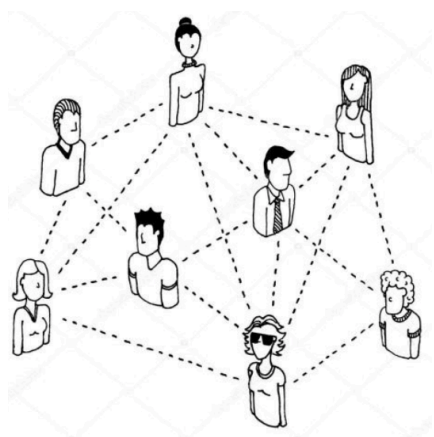


# 为什么处理图(Graph)?

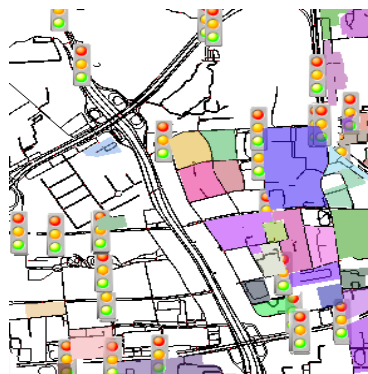


# 为什么处理图？

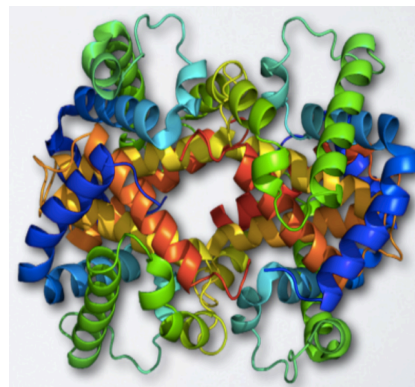
图数据无处不在！



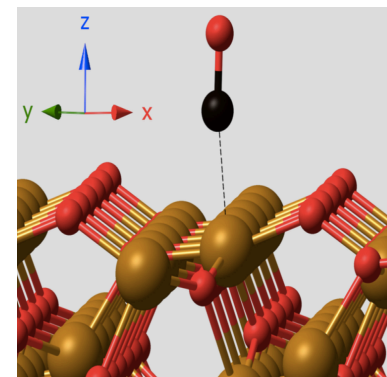
社会网络



交通网络



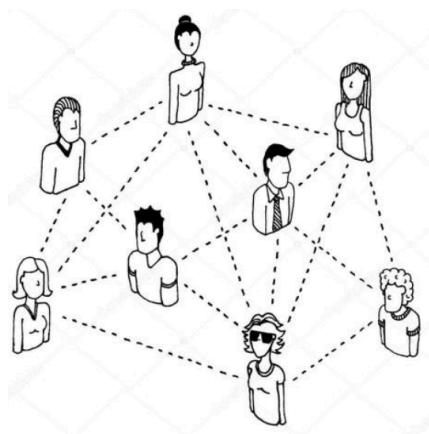
蛋白质结构



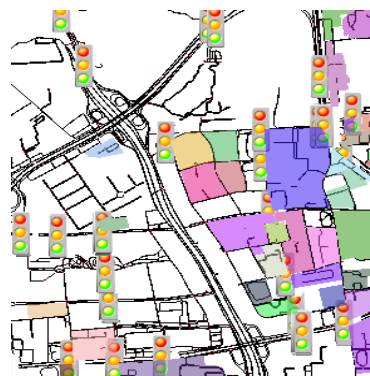
晶体结构

# 为什么是属性图？

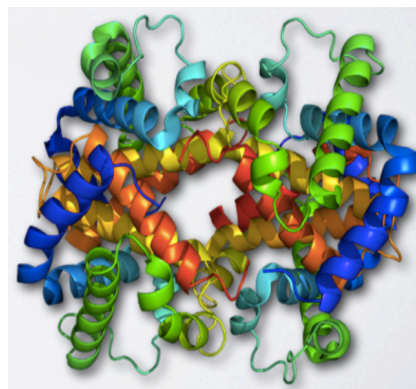
原因1：图数据中通常存在大量可用的上下文信息！



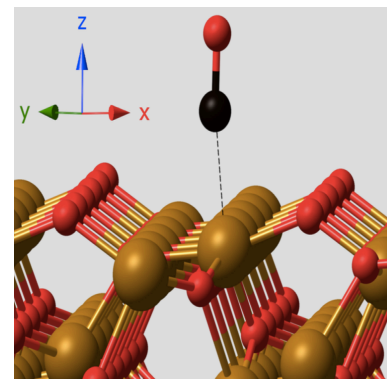
社会网络



交通网络



蛋白质结构



晶体结构

- 用户画像信息
- 连接关系类型
- 建立连接时间
- 用户数量、地点分布
- ...

- 十字路口空间坐标
- 道路长度、车道数、限速
- 路网的面积
- 人口分布
- ...

- 氨基酸类型
- 蛋白质中碳、氢、氧、氮等各元素所含比例
- 重量
- 性质(如：吸收率等)
- ...

- 原子类型
- 原子空间坐标
- 化学键类型
- 晶体质量、性质(如：熔点)
- ...

# 为什么是属性图？

- **原因2：**近期研究表明：

1. 图上的属性与拓扑结构高度相关 (*Zhang et al. 2013*)
2. 对多种网络分析任务具有益处
  - Trust prediction (*Tang et al. 2013*)
  - Network embedding (*Huang et al. 2017*)

# 为什么是属性图？

- **原因2：**近期研究表明：

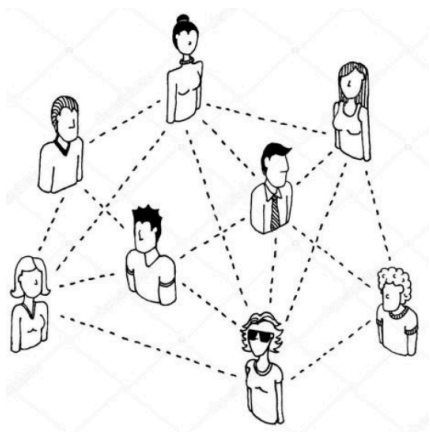
1. 图上的属性与拓扑结构高度相关 (*Zhang et al. 2013*)
2. 对多种网络分析任务具有益处
  - Trust prediction (*Tang et al. 2013*)
  - Network embedding (*Huang et al. 2017*)

因此，我们考虑：

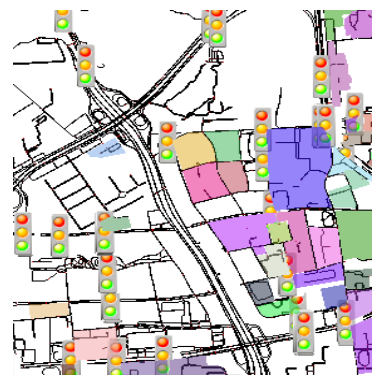
- ① 图上的属性能否帮助促进图结构优化任务？
- ② 如何充分有效地利用这些可用属性？

# 图(网络)结构优化是复杂网络分析中的重要任务

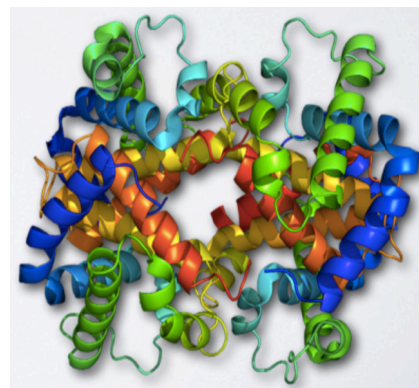
- 什么样的网络结构对应着最优的 **function** ?
- 什么样的网络结构能产生期望的 **dynamics** ?



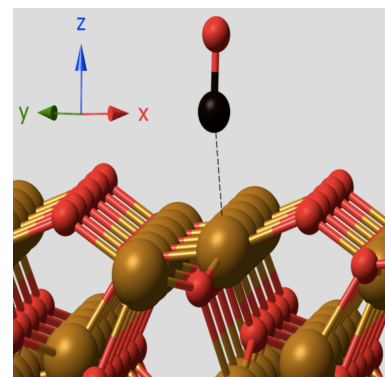
社会网络



交通网络



蛋白质结构



晶体结构

- 最大化信息流
- 发现活跃用户
- 最佳免疫策略

- 改善交通状况

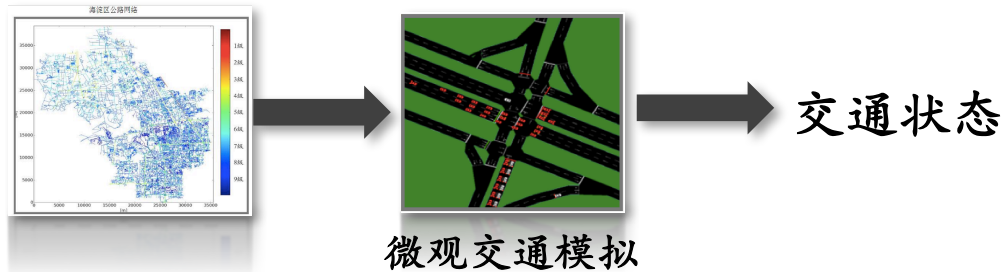
- 提高口感和吸收率

- 寻找高温超导体



# 属性图优化的难点

- ① 图结构优化是NP-hard (Das et al. 2014;Minoux 2015)
- ② 评估复杂图结构好坏的过程通常是黑箱的



- ③ 评估代价高昂
- ④ 搜索空间是非欧的、离散的、巨大(潜在药物分子空间 $10^{23}$ - $10^{60}$  (Polishchuk et al. 2013))
- ⑤ 从属性图中提取合理特征是艰巨的任务

# 已有方法的限制

## 1. 基于进化策略的无模型传统优化技术

- 路网优化(Xiong et al. 1992; Miandoabchi et al. 2010; Farahani et al. 2013)
- 分子发现 (Supady et al. 2015; Rupakheti et al. 2015)
- ...

为保证有效性：需要大量评估个体，以维护种群多样性

# 已有方法的限制

## 2. 基于模型的全局优化技术：贝叶斯优化

- 多针对目标的输入形式是**向量**，如：超参数优化 (Snoek et al. 2012)、机器人控制 (Cully et al. 2015)
- **仅考虑简单结构**，如：考虑输入参数之间关系 (Dalibard et al. 2017; Gardner et al. 2017)
- **仅针对神经网络架构搜索任务**，难以扩展到其他领域 (Kandasamy et al. 2018; Ramachandram et al. 2018; Jin et al. 2018)
- **没有充分利用所有可用属性** (Cui et al. 2018)
- 多用高斯过程做代理函数，模型**推断复杂度较高**(立方阶)

# 已有方法的限制

## 2. 基于模型的全局优化技术：贝叶斯优化

- 多针对目标的输入形式是**向量**，如：超参数优化 (Snoek et al. 2012)、机器人控制 (Cully et al. 2015)
- **仅考虑简单结构**，如：考虑输入参数之间关系 (Dalibard et al. 2017; Gardner et al. 2017)
- **仅针对神经网络架构搜索任务**，难以扩展到其他领域 (Kandasamy et al. 2018; Ramachandram et al. 2018; Jin et al. 2018)
- **没有充分利用所有可用属性** (Cui et al. 2018)
- 多用高斯过程做代理函数，模型**推断复杂度较高**(立方阶)

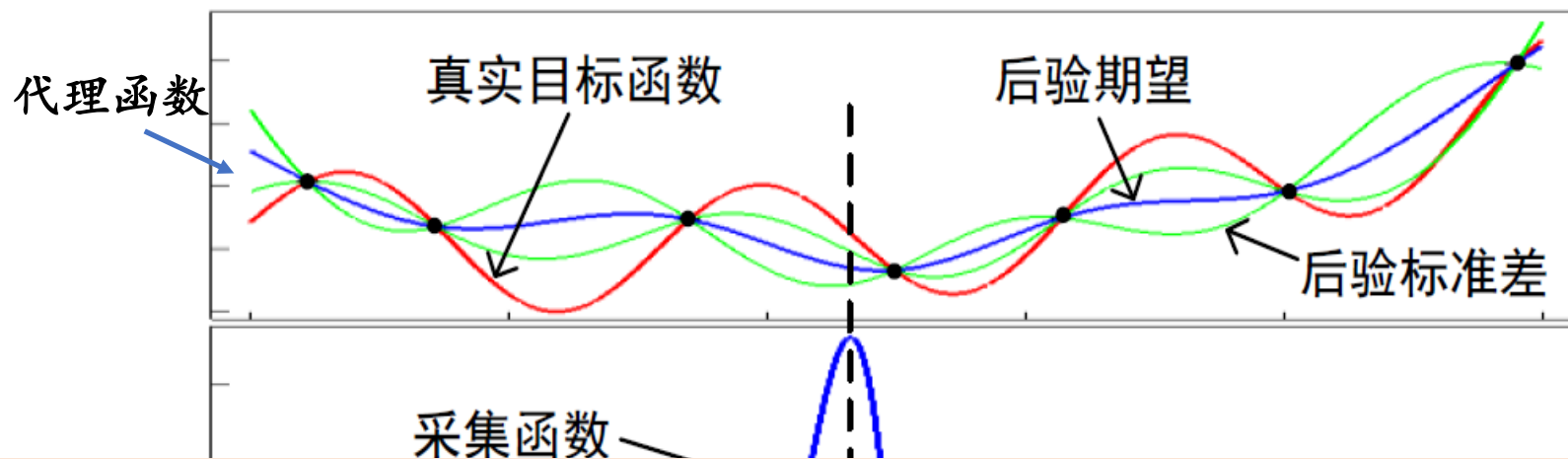
目标：提出一种能够直接处理属性图的、高可扩展的贝叶斯优化方法



# 贝叶斯优化为什么有效？

利用模型集成了先验知识，充分探索和利用了样本间的内在联系

- **代理函数**：替代评估代价大的黑箱函数
- **采集函数**：量化候选样本质量



BO优化过程：利用贝叶斯定理基于观察数据不断更新后验概率模型的过程

(Cui et al. 2018)

模型设计需要具体问题具体分析

## 本文的目标

① 能够直接处理属性图

② 高可扩展性

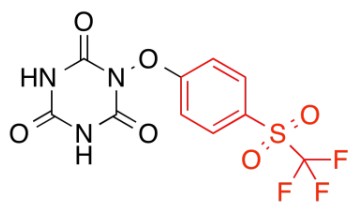
## 解决方案

受图卷积神经网络(Bronstein et al. 2016)启发, 提出一种新奇的深度代理架构

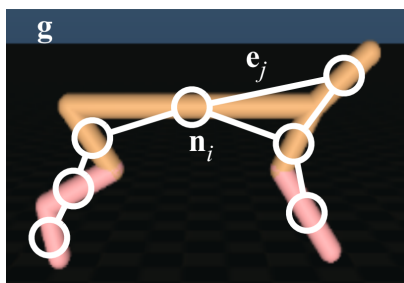
最后一层替换成贝叶斯线性回归

# 为什么选择图卷积神经网络(Graph Convolution Networks)?

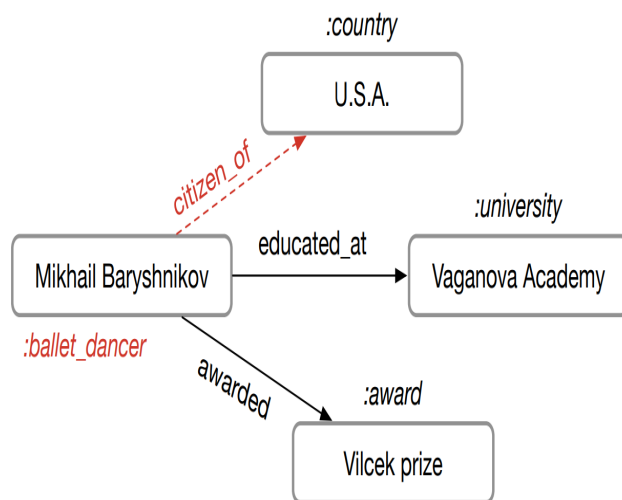
- 网络表示学习的最新技术
- 在许多领域中展现了其强大的应用潜力



化学分子  
(Duvenaud et al. 2015)



复杂动态系统  
(Sanchez-Gonzalez et al. 2018)



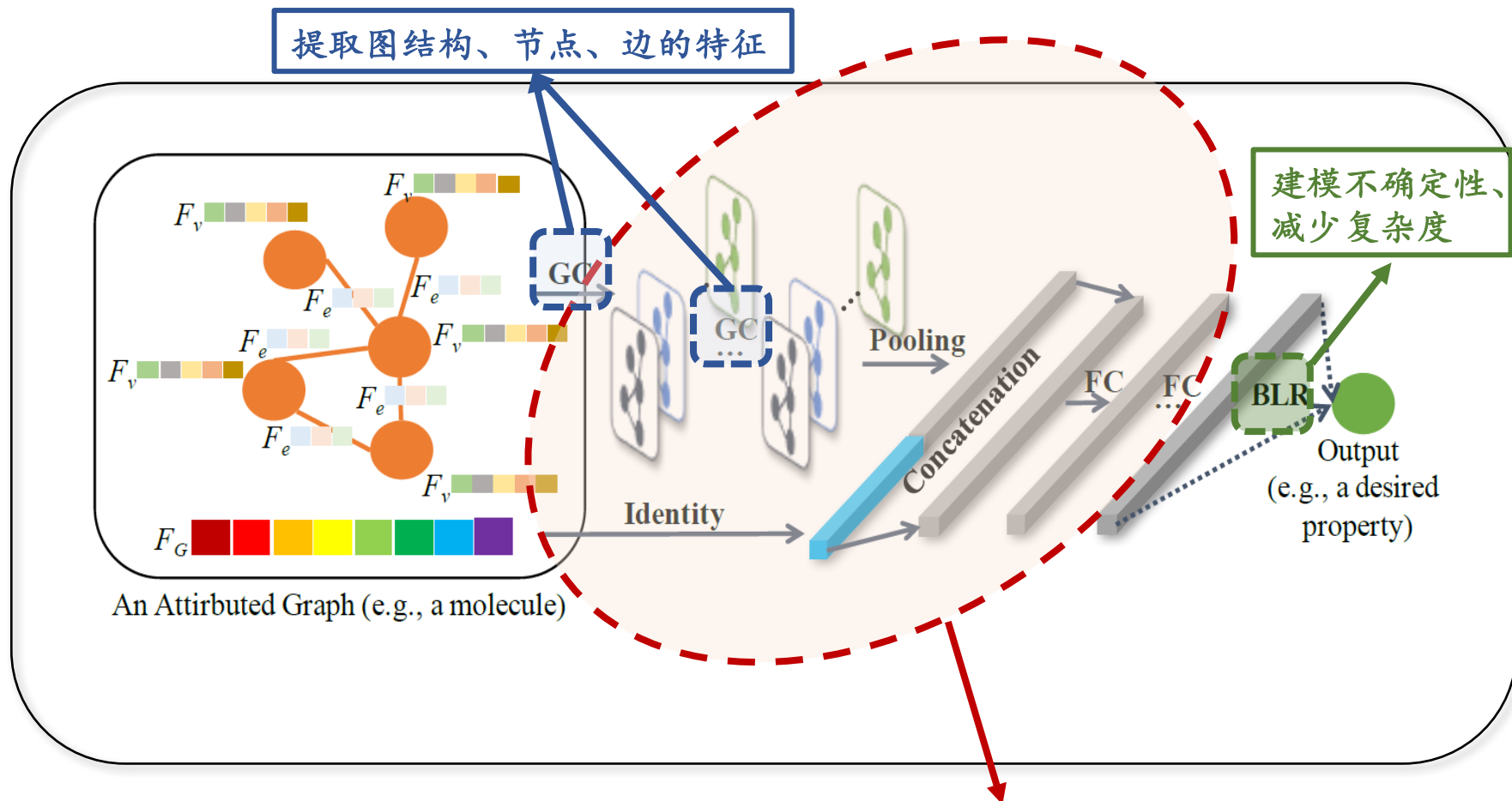
知识图谱  
(Schlichtkrull et al. 2017)



社会网络与推荐系统  
(Ying et al. 2018)

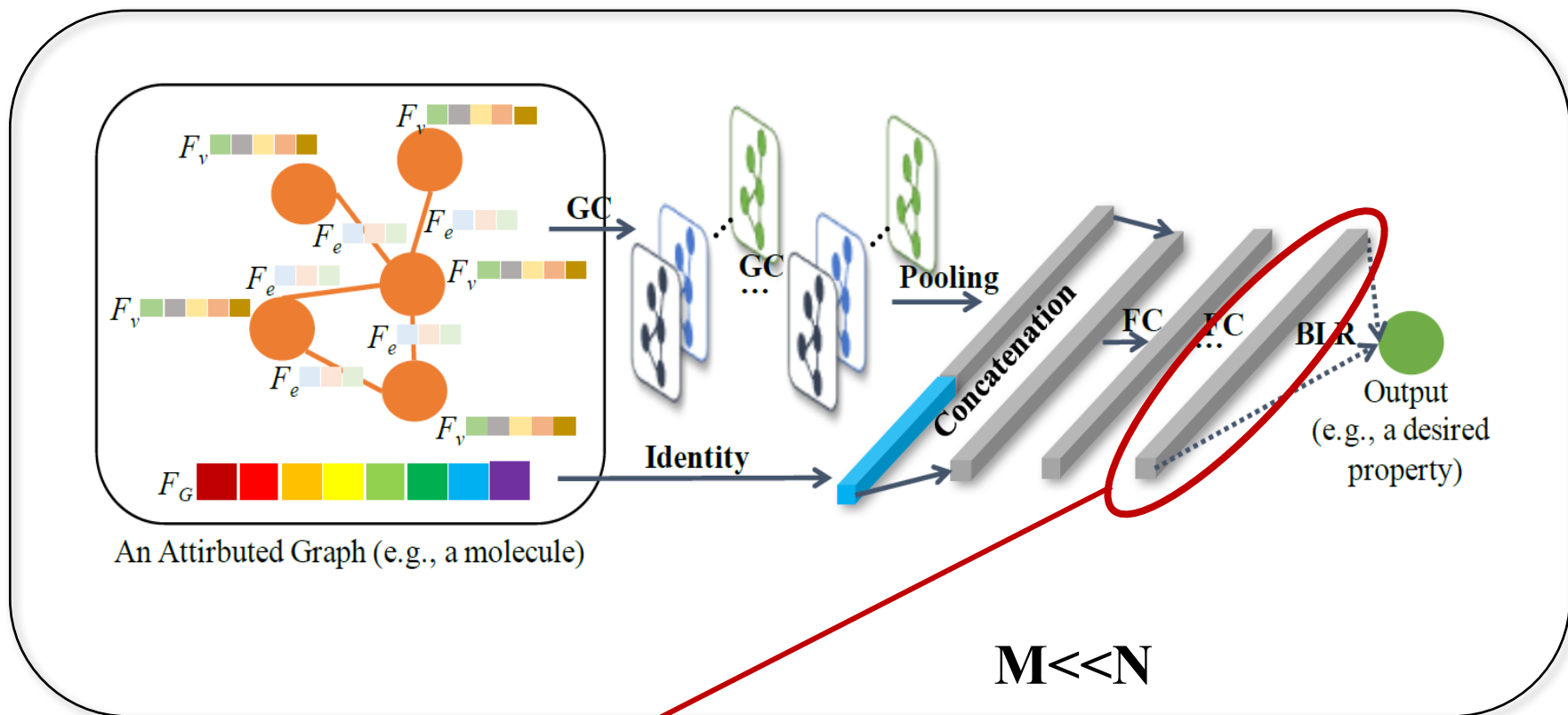


# 提出的深度代理模型架构



可看做为基于神经网络的自**适应基**回归

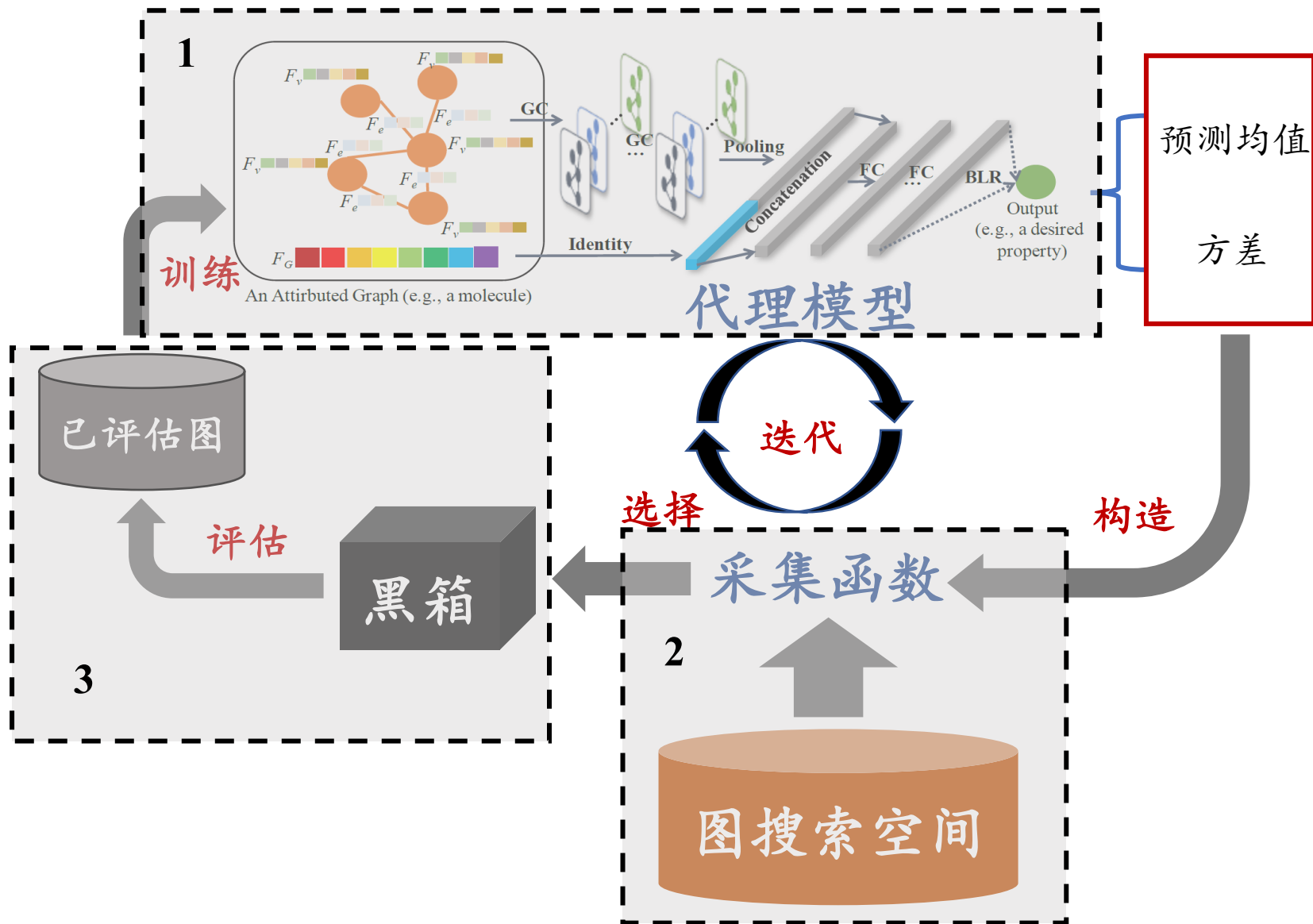
# 提出的深度代理模型架构

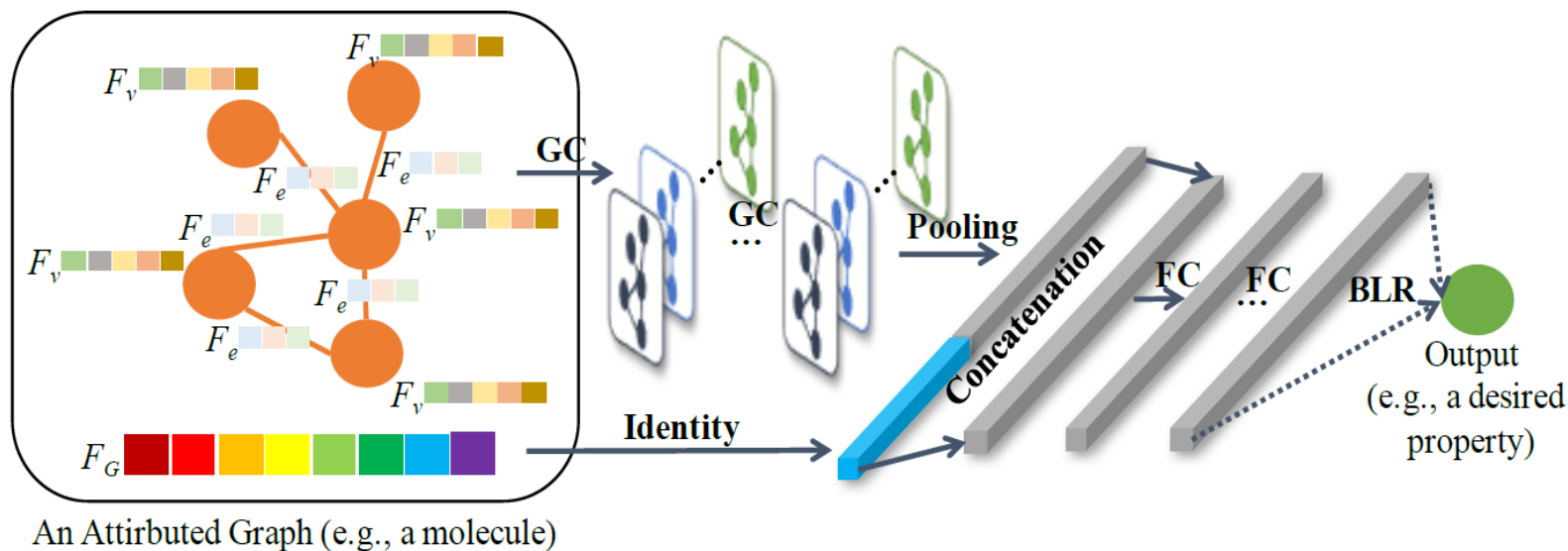


时间复杂度为  $O(M \times M \times N)$ ，而基于高斯过程的复杂度为  $O(N^3)$

样本数

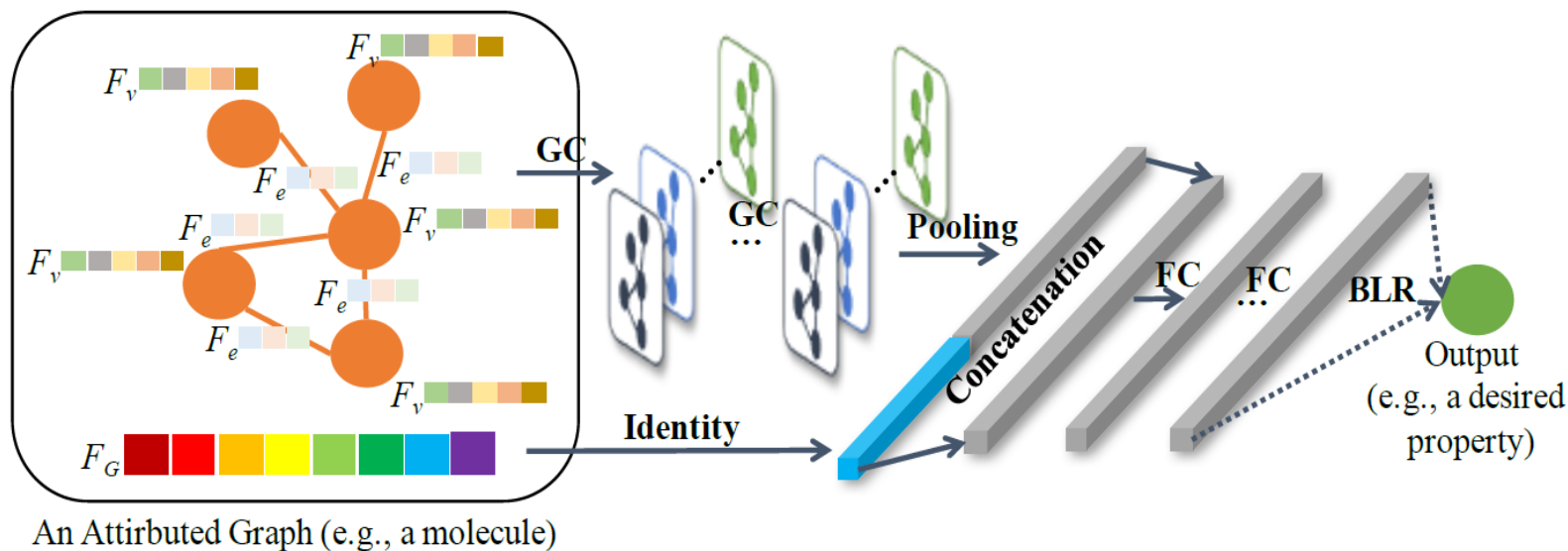
# 深度图贝叶斯优化流程





## 如何确定具体架构？

最直观的想法用待处理的训练数据进行架构搜索



## 如何确定具体架构？

最直观的想法用待处理的训练数据进行架构搜索

少

# 如何确定架构？

- 迁移思想：通过其他领域可用数据进行神经网络架构的优化

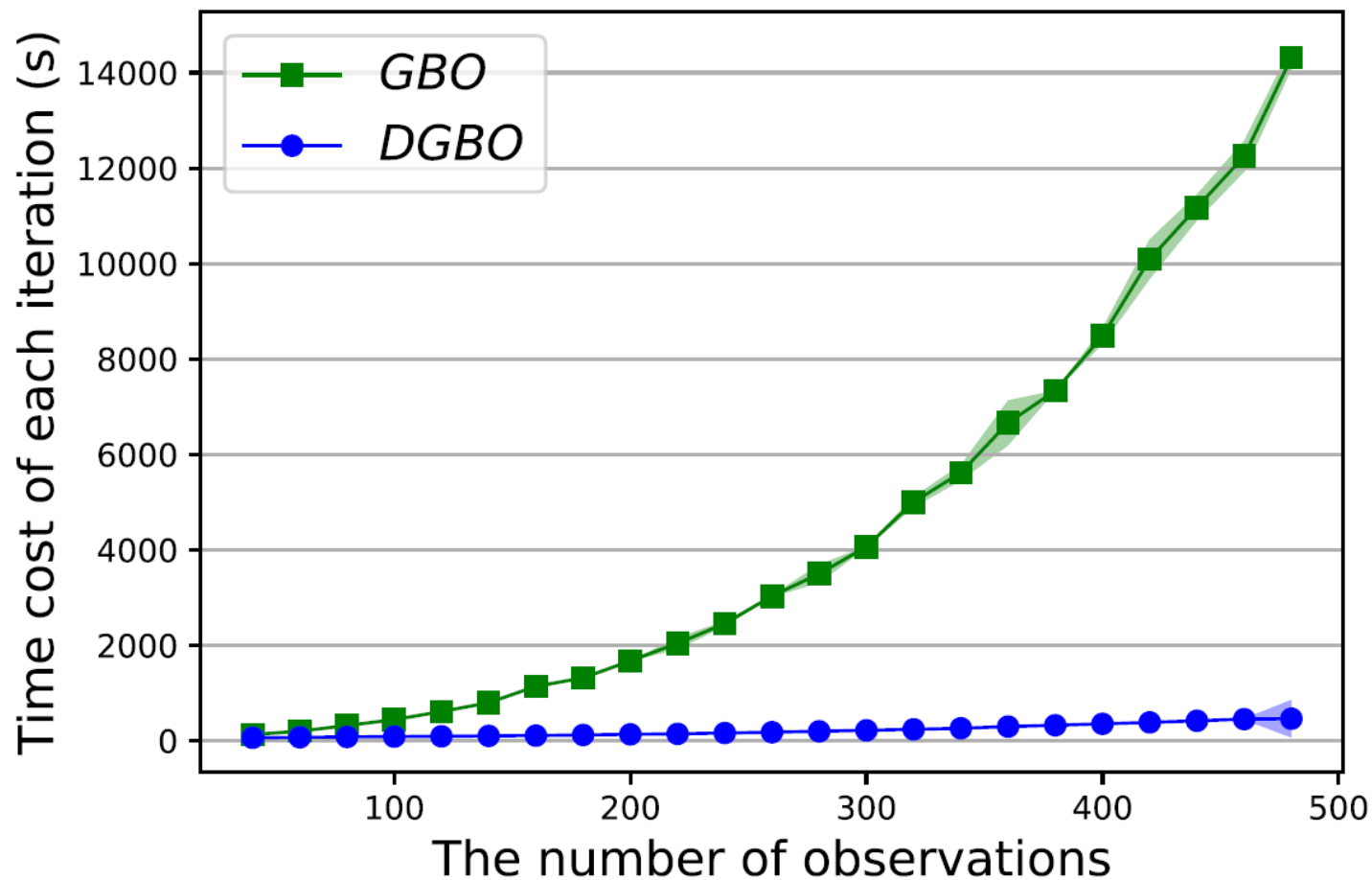
Parameters	Ranges	Optimal
# GC layers	{1, 2, 3, 4, 5}	5
# FC layers	{1, 2, 3, 4, 5}	5
# units of GC	[10, 100]	48
# units of pooling	[10, 100]	50
# units of FC	[10, 100]	45
$\sigma(\cdot)$ of GC	{ <i>ReLU</i> , <i>tanH</i> }	<i>tanH</i>
$\sigma(\cdot)$ of pooling	{ <i>Identity</i> , <i>ReLU</i> , <i>tanH</i> }	<i>Identity</i>
$\sigma(\cdot)$ of FC	{ <i>Identity</i> , <i>ReLU</i> , <i>tanH</i> }	<i>tanH</i>
Learning rate	[1e-4, 1e-1]	1e-4
Dropout	[0, 1]	0.0
Penalty coefficient	[1e-5, 1e-1]	1e-5

Table 1: The optimal surrogate architecture.

# 验证

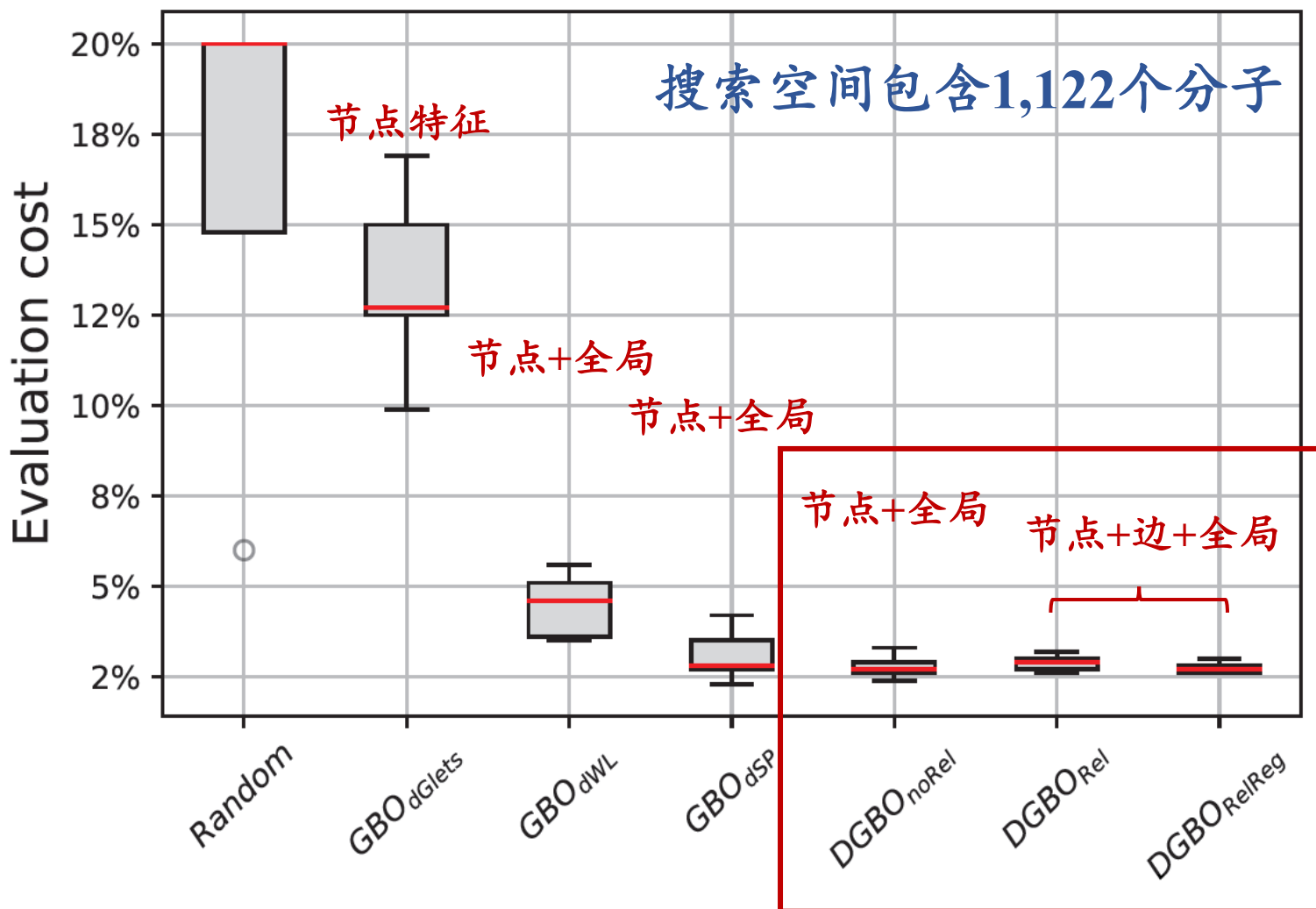
1. 时间复杂度
2. 上下文属性是否益于优化
3. 可扩展性(Scalability)
4. 应用于其他领域(路网优化)

# 验证1：时间复杂度





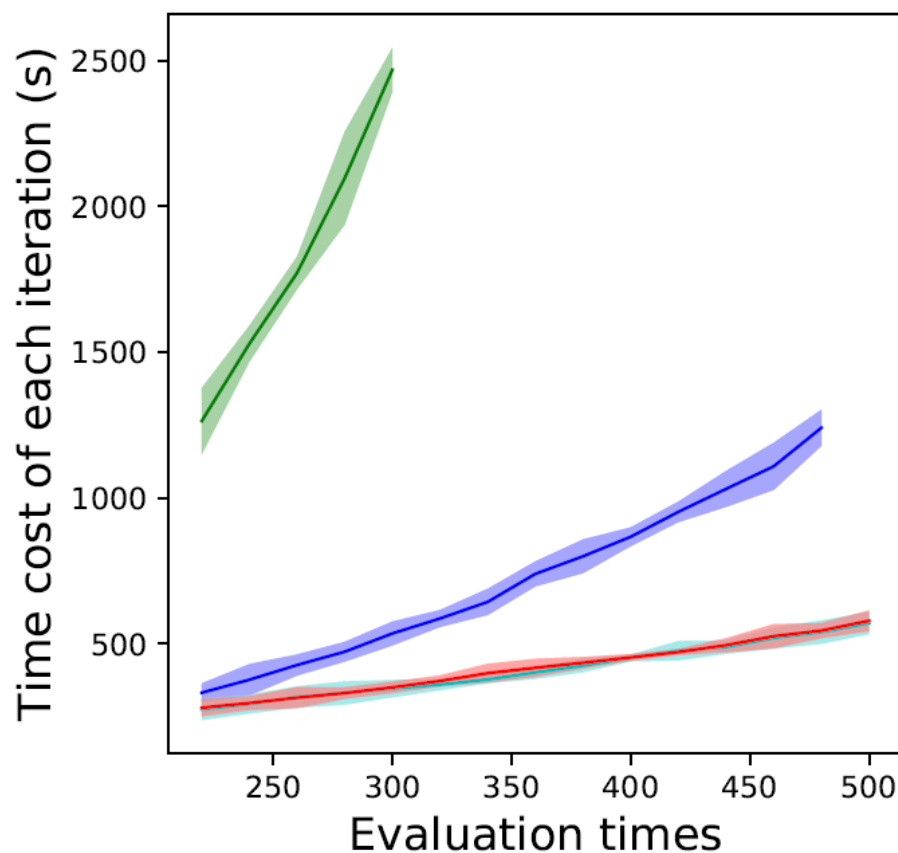
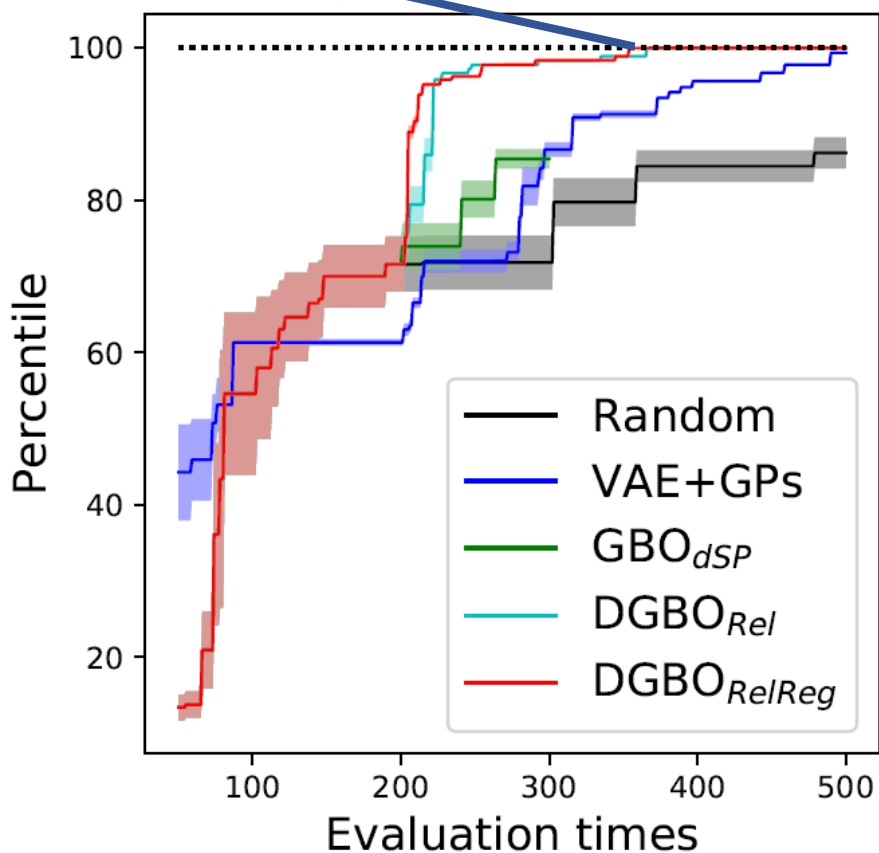
# 验证2：上下文属性益于优化



# 验证3：可扩展性(Scalability)

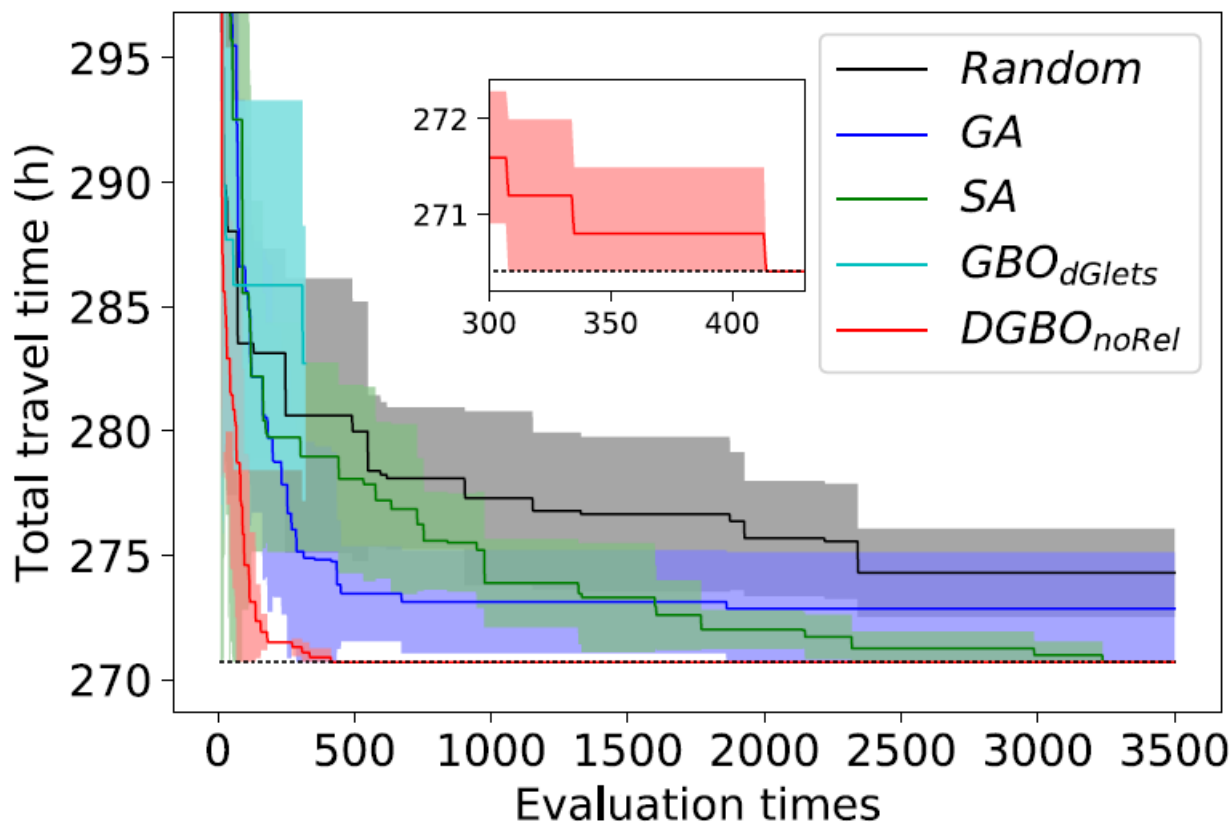
仅1.8%

搜索空间包含20,000个药物分子



# 验证4：应用于其他领域(路网优化)

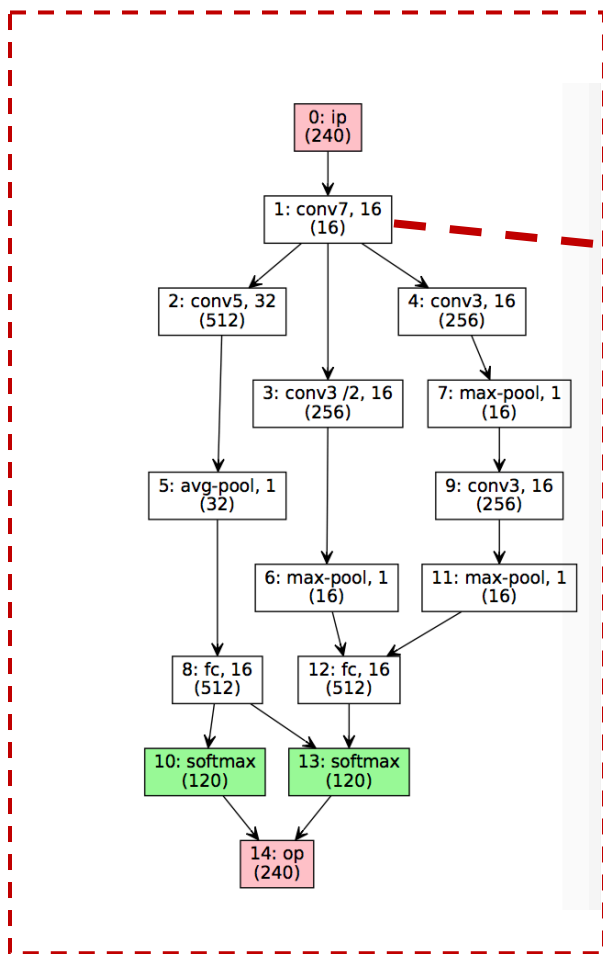
搜索空间共包含32,768个路网



收敛速度快7.7倍

# 实际应用：神经网络架构搜索

我们把神经元网络建模成属性图，节点表示层，边表示数据流



节点属性:

ip,op,conv3, conv5, conv7, res3,res5,res7,max-pool,avg-pool,softmax,卷积通道数,全连接单元数

全局属性:

平均节点的度, 各种类型节点的比例

# 实际应用：神经网络架构搜索

Method	CIFAR10 (error)
Rand	0.1342
EA	0.1411
TreeBO (Jenatton et al. 2017)	0.1533
NASBOT (Kandasamy et al. 2018)	0.1209
<b>Ours</b>	<b>0.1078</b>





# 感谢各位老师和同学！

欢迎大家下载使用杨博老师团队开发的科研助手APP：学术头条



 学术头条  
academic headline

