**论文题目：Big Graph Search: Challenges and Techniques**

作者：Shuai Ma, Jia Li, Chunming Hu, Xuelian Lin, and Jinpeng Huai

Frontiers of Computer Science, 10(3): 387-398 (2016)

DOI: https://doi.org/10.1007/s11704-015-4515-1

图的表达能力非常强，近年来受到工业界和学术界的共同关注。大图搜索指的是一种概念上非常广泛的图的查询，有着广泛的应用。

我们首先给出一个简单的形式化定义。图搜索：给定一个模式图（pattern graph）和一个数据图（data graph）：（1）判断是否“匹配”；或者（2）从中找出所有跟“匹配”的子图。注1：这里图是由顶点集和边集组合而成，而顶点和边上通常会有标签标注相关信息。注2：图搜索的定义包含了两类查询，第一类查询是布尔查询，即需要回答“是”或者“否”的查询；第二类查询返回结果时需要利用第一类查询，两者之间有着紧密的关系。此外，模式图通常比较小，仅仅包含几个或者几十个顶点；而数据图通常较大，甚至包含以“亿”为数量级的顶点和边。此外，不同的“匹配”语义表示不同类型的图搜索,包括:最短路径/距离，子图同构，图同态及其扩展，图模拟及其扩展，图关键字搜索和紧邻查询等。

大图搜索的挑战主要来自于三个方面，也即FAE法则。在大量、动态和不确定图数据中，

（1）友好性(Friendliness)：如何提供友好的图搜索界面？如何以友好的方式提供“图搜索”的查询界面？如，关键字的搜索模式非常友好，直接让用户输入模式图非常不友好，如何提供方便的方式隐式表达模式图。

（2）准确性(Accuracy)：如何搜索“信息”更准？如，加强用户意图理解（融合用户的行为特征），融合知识图谱，基于知识/用户意图的查询转换等。

（3）高效性(Efficiency)：如何搜索“信息”更快，如查询近似技术和数据近似技术等。

大图搜索的关键技术包括，查询近似、数据近似、分布式处理、增量计算等。

作者简介：

马帅，北京航空航天大学计算机学院教授，主要研究方向为数据库理论与系统、大数据等。



李佳，北京航空航天大学计算机学院博士，主要研究方向为图匹配和社交推荐等。



胡春明，北京航空航天大学软件学院教授，主要研究方向为计算机软件与理论、分布式系统等



林学练，计算机学院副研究员，主要研究方向为时序数据挖掘、分布式计算等，



怀进鹏，中国科学院院士，北京航空航天大学计算机学院教授。主要研究方向为网络化软件技术和系统研究工作等