1．为什么要图分割？

随着图规模和复杂度越来越大，图计算变得越来越困难。为了解决这个问题，我们将图分割在多个分区上，使得它们上的计算可以大部分在本地并且并行执行。

2. 图分割的目标

1) 边的切断越少越好，因为边的两端顶点需要通信，通信代价随断边数量增大而增加。

2）各个分区之间的顶点数量大约一致。

满足以上两点的图分割，我们就称之为平衡图分割。

3. 目前平衡图分割采用的方式

因为平衡图分割是一个NP问题，采用的都是启发式算法[1]。

根据图加载的方式可以将图分割分为：离线分割、在线分割和动态分割[2]。其中离线分割需要加载整个图之后才能进行图分割；在线分割可以将图的结点以流的形式连续加载并分配其到所在分区，因此图的大小和分区的数量上是可以扩展的；动态分割不仅可以将结点动态加载并分配其所在分区，还可以随着图的变化在任何给定时间通过重新分割来改变分区的数量和平衡限制，同时具有较小的迁移成本。

在线分割和动态分割，相对于离线分割的难度在于：

1. 顶点以任意顺序流入，分区时并不能知道图的全局信息
2. 对每个顶点要实时处理，时延小

最先进的在线分割算法是由Stanton和Kliot[3]提出，他们提出了一种自然而简单的只依赖于图结构的启发式算法，相对于基于散列的分割方法和METIS，分割效果有很大提升。 另外，Charalampos E. Tsourakakis等人[4]提出了一个新颖的one-pass流图分割算法，该算法统一了两个看似正交的启发式算法：将新到达的顶点放置在具有最大数量邻居结点的分区中或者具有最小数量的非邻居结点的分区中。相对于METIS，平衡分割时间更短，效果更好。

更一般的，对于顶点和边不断地添加和删除的动态图来说，平衡图分割是更加困难的。Luis M. Vaquero等人[5]提出了一种高度可扩展的自适应分区策略，该分区策略是基于迭代的顶点迁移，仅依赖于本地信息。在3个真实的场景下，相对于基于散列的分割方法，图计算的执行时间减少50%以上。

Ioanna Filippidou和Yannis Kotidis[2]也提出了一种基于精简生成树结构的图分割算法，它不仅能对任意的演变图进行图分割，还允许不同的应用按需来调整分区，如最小化边切割和跨分区平衡加载等。相对于现有的图分割算法，它具有较低的时间和空间复杂度。

[1]A. Konstantin and H. Racke. Balanced graph partitioning. In *SPAA ’04*, pages 120–124, 2004.

[2]Ioanna Filippidou and Yannis Kotidis. Online and On-demand Partitioning of Streaming Graphs. In 2015 IEEE International Conference on Big Data (Big Data), 2015.

[3] Isabelle Stanton and Gabriel Kliot. Streaming Graph Partitioning for Large Distributed Graphs. In *KDD '12, pages 1222-1230,* 2012*.*

[4]Charalampos E. Tsourakakis, Christos Gkantsidis, Bozidar Radunovic and Milan Vojnovic. FENNEL Streaming Graph Partitioning for Massive Scale Graphs. In *WSDM '14*, pages 333-342. 2014.

[5] Luis M. Vaquero, Felix Cuadrado, Dionysios Logothetis and Claudio Martella. Adaptive Partitioning of Large-Scale Dynamic Graphs. In *SOCC '13*, article No. 35, 2013.