**2012美赛A题：一棵树的叶子**

（数学中国翻译）

)  
“一棵树的叶子有多重？”怎么能估计树的叶子（或者树的任何其它部分）的实际重量？怎样对叶子进行分类？建立一个数学模型来对叶子进行描述和分类。模型要考虑和回答下面的问题：   
! ^3 ^. ^$ E. Y2 l5 m. y9 i • 为什么叶子具有各种形状？  
! G. D, [1 r1 y, K • 叶子之间是要将相互重叠的部分最小化，以便可以最大限度的接触到阳光吗？树叶的分布以及树干和枝杈的体积影响叶子的形状吗？  
5 M\* |\* u/ y+ B+ c) q$ c+ H • 就轮廓来讲，叶形（一般特征）是和树的轮廓以及分枝结构有关吗？  
) l& l- h: d9 P% C) J2 R6 g • 你将如何估计一棵树的叶子质量？叶子的质量和树的尺寸特征（包括和外形轮廓有关的高度、质量、体积）有联系吗？2 S! V- g9 Y1 d( m) q, Z' W  
  
+ U3 z4 X0 o! \_+ O0 a4 V+ ?除了你的一页摘要以外，给科学杂志的编辑写一封信，阐述你的主要发现。

. k& w# R5 n- J0 I

PROBLEM A: The Leaves of a Tree

"How much do the leaves on a tree weigh?" How might one estimate the actual weight of the leaves (or for that matter any other parts of the tree)? How might one classify leaves? Build a mathematical model to describe and classify leaves. Consider and answer the following:

• Why do leaves have the various shapes that they have?

• Do the shapes “minimize” overlapping individual shadows that are cast, so as to maximize exposure? Does the distribution of leaves within the “volume” of the tree and its branches effect the shape?

• Speaking of profiles, is leaf shape (general characteristics) related to tree profile/branching structure?

• How would you estimate the leaf mass of a tree? Is there a correlation between the leaf mass and the size characteristics of the tree (height, mass, volume defined by the profile)?

In addition to your one page summary sheet prepare a one page letter to an editor of a scientific journal outlining your key findings.

**2012美赛B题：沿着“大长河”露营**

9 V% V2 J4 C: x: v3 D k$ E

【数学中国翻译】

, R) P8 H/ E4 R- ]  
  
; e7 Z2 P1 }3 ^) ] t9 Z 游客在“大长河”(225英里)可以享受到秀丽的风光和令人兴奋的白色湍流。这条河对于背包客来说是进不去的， 因此畅游这条长河的唯一办法就是在这条河上露营上几天。这次旅行从开始的下水点到最终结束点，共225英里，且是顺流而下的。乘客可以选择平均4英里/小时的以浆作为动力的橡胶筏或者平均8英里/小时的机动帆船旅行。整个旅行从开始到结束会经历6至18个夜晚。负责管理这条河的政府机构希望到这里的每一次旅行都能够享受到野外经历， 以最少的接触到在河上其它的船只。目前， 每年在六个月期间(一年的其余部分的天气对于河流旅行来说太冷)，共有X次旅行，有Y处露营地，露营地均匀的分布整个河道。由于漂流的受欢迎程度的上升，公园管理者已经被要求允许更多的旅行次数。所以他们想确定怎样可能安排一个最优的混合的旅行方案，不同的时间(单位为夜)和推动方式(马达或浆)，最大限度的利用露营地。换句话说，在长河的漂流季，将会有多少更多的乘船旅行可以加进来?河流的管理者现在雇佣你，为他们提出最佳排程方式和河流承载能力的建议，记住两个露营者不能在同一时间内占据同一个露营地。除了你的一页摘要，准备一页备忘录，对河流的管理者描述你的主要发现。  
6 k+ T' T4 \_/ C& D4 B9

**破案模型**

您的组织，ICM正在调查一个作案阴谋。调查者非常有信心，因为他们知道阴谋集团的几名成员，但他们希望在进行逮捕之前能找出其他成员和领导人。主谋者和所有可能涉嫌同谋的人都以复杂的关系为同一家公司在一个大办公室工作。 这家公司一直快速增长，并在开发和销售适用于银行和信用卡公司的计算机软件方面打出了自己的名气。 ICM最近从一个82个工人的小集体那儿得知了一个消息，他们认为这个消息能将帮助他们在公司里找到目前身份尚不明确的同谋者和未知的领导人的最有可能的人选。由于信息流通涉及到所有的在该公司工作的工人，所以很可能在这次信息流通中有一些（或许很多）已经确定的传播者实际并不涉及阴谋。事实上，他们确定他们知道一些并不参与阴谋的人。  
  
建模工作的目标是确定在这个复杂的办公室里谁是最有可能的同谋。  
一个优先级列表是最理想的，因为ICM可以根据这个来调查，\*\*，和/或询问最有可能的候选人。  
一个划分非同谋者与同谋者的分割线也将是有益的，因为可以对每个组里的人进行清楚的分类。  
如果能提名阴谋的领导人，那对于检察官办公室也是非常有帮助的。  
  
  
在把当前情况下的数据给你的犯罪建模团队之前，你的上司给你以下情形（称为调查EZ），那是她几年前在另一座城市工作时的案例。她对她在简单案件的工作非常自豪，她说，这是一个非常小的，简单的例子，但它可以帮助你了解自己的任务。  
  
她的数据如下：  
她认为是同谋的十人分别为Anne#, Bob, Carol, Dave\*, Ellen, Fred, George\*, Harry, Inez, and Jaye#.（\*表示之前已知的同谋，＃表示事先已知的非同谋者）  
她对她的案件的28个消息记录按照她的分析 依据主题进行了编号。  
Anne to Bob：你今天为什么迟到了？ （1）  
Bob to Carol：这该死的Anne总是看着我。我并没有迟到。 （1）  
Carol to Dave： Anne 和 Bob又再为Bob的迟到吵架了。 （1）  
Dave to Ellen：我今天早上要见你。你什么时候能来？把预算文件顺便带过来。 （2）  
Dave to Fred：我今天随时随地都可以去见你。让我知道什么时候比较好。我需要带预算文件吗？ （2）  
Dave to George：我待会见你---有很多需要谈的。我希望其他人都准备好。获得这项权利？ 很重要。 （3）  
Harry to George：你似乎很紧张。怎么回事？不用担心，我们的预算会好的。 （2）（4）  
Inez to George：我今天真的很累。你呢，还好吗？ （5）  
Jaye to Inez：也不怎么样今天（？）。今天一起去吃午饭怎么样？ （5）  
Inez to Jaye：幸好一切都很平静。我已经精疲力竭，不能做午饭了今天。抱歉！ （5）  
George to Dave：现在来见我！ （3）  
Jaye to Anne：你去吃午饭吗今天？ （5）  
Dave to George：我没法去，现在正要去见Fred。 （3）  
George to Dave：见完他后到我这来。（3）  
Anne to Carol：谁来监督一下Bob？他整天游手好闲的。（1）  
Carol to Anne：别管他。他和George and Dave合作得很好。（1）  
George to Dave：这个很重要。该死的Fred。Ellen怎么样了？ （3）  
Ellen to George：你和Dave谈过了吗？（3）  
George to Ellen：还没。你呢？（3）  
Bob to Anne：我没有迟到。而且你知道我午饭时间都在工作呢。（1）  
Bob to Dave：告诉他们我没有迟到。你了解我的。（1）  
Ellen to Carol：联系Anne安排下个星期的预算会议日程，还有，帮我让George冷静点。（2）  
Harry to Dave：你有没有注意到George今天看上去又很紧张/有压力？（4）  
Dave to George：该死的Harry觉得你很紧张。别让他担心，免得他四处打探。（4）  
George to Harry：我只是工作得太晚，家里又有点问题。不用担心，我很好。（4）  
Ellen to Harry：我忘了今天的会议了，怎么办？Fred会在那的，而且他比我更了解预算。（2）  
Harry to Fred：我觉得明年的预算会让一些人很有压力的。或许你今天该花点时间让大家安心。（2）（4）  
Fred to Harry：我觉得我们的预算很正常，我没觉得会有人感到有压力。（2）  
通信记录结束。  
  
你的上司指出，她只分配和编号了5个不同的消息主题：  
1）Bob的迟到，  
2）预算，  
3）重要的未知的问题，可能是阴谋，  
4）乔治的压力，  
5）午餐和其他社会问题。  
正如看到的消息编码那样，一些消息根据内容有两个主题。  
你的上司按照通信联系和消息类型构造的通信网络分析案件。下图是一个消息网络模型，网络图上注明了消息类型的代码。  
您的上司说，除了已知的同谋George and Dave之外，根据她的分析 Ellen and Carol也被认为是同谋。而且不久后，Bob招认出他确实参与其中，从而希望得到减刑。而对Carol的控告后来被放弃了。  
你的上司至今仍然相当肯定Inez也参与了，但却从未对她立案。  
你的上司建议您的团队，确定有罪的当事人，使像Inez的人不漏网，像Carol的人不被诬陷，从而增加ICM的信用，使像Bob的人不再有获得减刑的机会。  
  
现在的案件：  
你的上司已经把目前的情况下构造成网络状的数据库，它具有和上面相同的结构，只是范围较大。调查者有一些线索表明，一个阴谋正在挪用公司的资金和使用网上诈骗盗窃在该公司做业务的顾客的信用卡资金。  
她给你看的简单案件的小例子，只有10个人（节点），27条边（消息），5个主题，1个可疑/阴谋主题，2个确定的罪犯，2个已知的清白者。而到目前为止，这个新的案件却已经有83个节点，400条边（有些不止涉及1个主题），超过21000个单词的消息记录，15个主题（其中3个已被视为是可疑的），7个已知的罪犯，和8个已知的清白者。这些数据在所附的电子表格文件：names.xls，Topics.xls，Messages.xls中给出。  
names.xls包含办公室的关键节点对应的员工的名字。  
topics.xls包含15个主题的代号及简短说明。  
由于安全和隐私问题，你的团队不会有所有的直接消息记录。   
messages.xls提供传输消息的节点对，和该消息的主题（可能不止一个主题，最多3个主题）。  
为了使信息的沟通更加直观可视，图2提供了员工和消息链接的网络模型。  
在这种情况下，不再像图1那样显示消息的主题。而是在文件Messages.xls里给出主题的数目，并在Topics.xls中给以描述。  
  
要求：  
要求1：到目前为止，已知Jean, Alex, Elsie, Paul, Ulf, Yao, and Harvey是罪犯，Darlene, Tran, Jia, Ellin, Gard, Chris, Paige, and Este不是罪犯。可以的消息主题是7,11和13。关于主题更多的信息在Topics.xls里。  
建立模型和算法，把83个节点按照他是阴谋者的可能性大小排序，并解释你的模型和指标。Jerome, Delores, and Gretchen是该公司的高级经理。如果他们三个人中任何一个涉及阴谋  
这将是非常有益的。  
要求2：优先列表将有神秘变化，如果有新的信息告知我们说主题1也与阴谋有关，而且克里斯是一个阴谋？（即多了两个线索）  
要求3：一个强大的与这个消息流通网络类似的获取和理解文本信息的技术被称为语义网络分析（semantic network analysis）;作为人工智能和计算语言学的方法，它提供了一个结构，并可进行有关知识或语言的推理过程。另一个有关自然语言处理的计算语言学是文本分析text analysis。  
针对我们的破案的情况，解释： 如果你能获得原始消息，那么对信息流量的上下文和内容进行语义和文字分析 对于 帮助你们的团队 开发出更好的模型和办公室人员的分类有多大的帮助和加强作用？  
你有没有使用这些基于文件Topics.xls中的主题描述的功能来提高您的模型？  
要求4：你的完整报告将最终提交给检察官办公室，所以一定要详细、明确地说明您的假设和方法，但不能超过20页。您可以包括你的程序作为单独的文件中的附件使你的论文不超过页面限制，但包括这些程序不是必须的。你的上司希望ICM是世界最好的解决白领、高科技的阴谋罪的机构，并希望您的方法有助于解决重要的世界各地的案件，特别是那些消息流量非常大的数据库（可能有数万的信息和数百万的单词）。她特别要求你在论文中讨论 ：更深入的网络，语义，消息的文本分析内容是如何帮助你的模型和建议的。  
作为给她的报告的一部分，请解释你用到的网络模型技术，以及为什么使用和它们可以怎么被用于任何类型的网络数据库从而来确定，优先级排序，和对相似结点分类的技术的网络模型，而不仅仅是犯罪阴谋和消息数据。比如，给你各种图像或化学数据，其中表明了感染概率和已经确定了的一些受感染的结点，你的方法能用来在生物网络中找到感染或患病的细胞吗？