**用户文档**

您好，十分感谢您抽出时间完成我们的用户实验！

我们建议您留出约**1个小时**的空余时间，从阅读本用户文档开始进行实验。若您中途有事，建议您在**做完每一个大部分之后再暂停答题**，有时间后再继续开始答题，注意：中途请**不要关闭网页**，否则信息将不做保存，谢谢！

接下来我将从3个方面向您介绍，以便您能更好地理解与进行实验。

**一、实验组成**

数据集：

我们使用了3个数据集，分别是：

|  |  |  |
| --- | --- | --- |
| Southern Women | DBLP | Co-Author Network |
| 记录了18名女性在9个月的14次非正式社会活动中的社会互动 | 记录了70位作者在30篇期刊论文中的合作关系 | 记录了学术出版物中114个作者之间的合作关系 |

实验共分为7个部分：

第一部分为背景信息，目的是了解您的年龄、性别、专业以及对可视化熟悉程度。

第二部分为“实验1”，共6个问题，前3个问题分别对应1个数据集，后3个问题是请您对实验1的难度、自己的完成速度、自己的完成正确率进行评价。实验1的前3个问题探讨凹度，我们将在后续为您详细讲解。

第三部分为“实验2”，共6个问题，前3个问题分别对应1个数据集，后3个问题是请您对实验2的难度、自己的完成速度、自己的完成正确率进行评价。实验2的前3个问题探讨平面度，我们将在后续为您详细讲解。

第四部分为“实验3”，共6个问题，前3个问题分别对应1个数据集，后3个问题是请您对实验3的难度、自己的完成速度、自己的完成正确率进行评价。实验3的前3个问题探讨覆盖率，我们将在后续为您详细讲解。

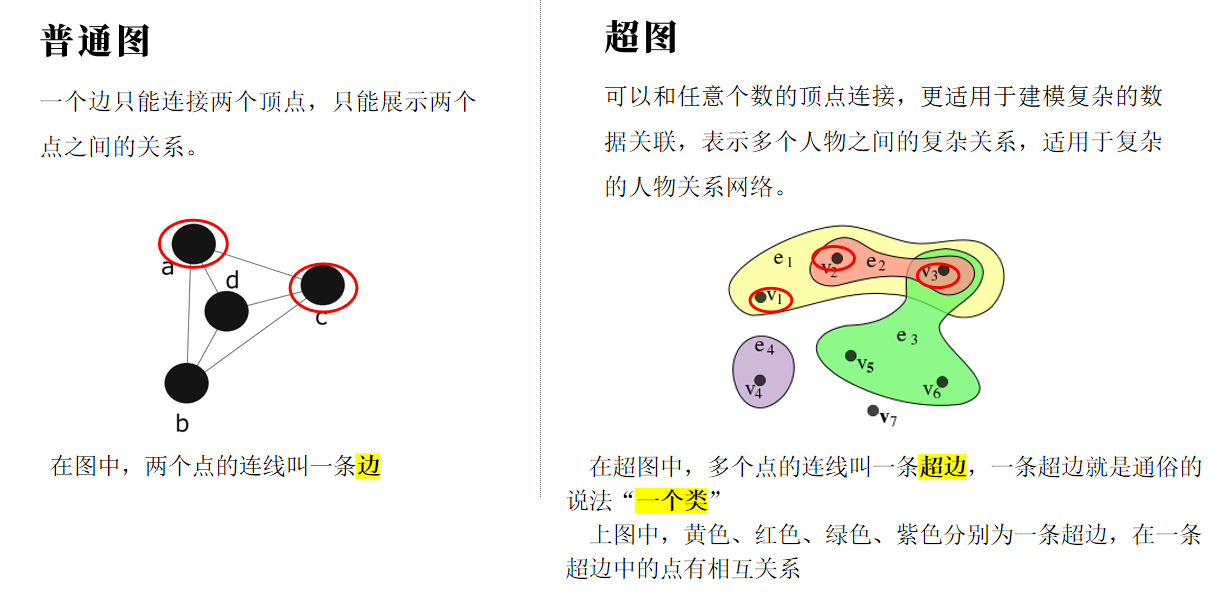
第五部分为“实验4”，共6个问题，前3个问题中每2个对应1个数据集，后3个问题是请您对实验4的难度、自己的完成速度、自己的完成正确率进行评价。实验4的前3个问题探讨节点间连线距离的均匀程度，我们将在后续为您详细讲解。

第六部分为“实验5”，共39个问题。前12个问题“实验5-1-1”到“实验5-1-12”对应于第一个数据集，中间12个问题“实验5-2-1”到“实验5-2-12”对应于第二个数据集，后12个问题“实验5-3-1”到“实验5-3-12”对应于第三个数据集，最后3个问题是请您对实验5的难度、自己的完成速度、自己的完成正确率进行评价。实验5的前36个问题旨在探究用户对数据类别的认知能力，我们将在后续为您详细讲解。

第七部分为问卷调查，共9个问题。问题1到问题3各对应1个数据集，希望您按照喜爱程度进行排名，问题4到问题8是向用户询问各评估因素的重要性，问题9希望用户给出在实验过程中，所关注的超图中的其他因素。

**二、超图的介绍**

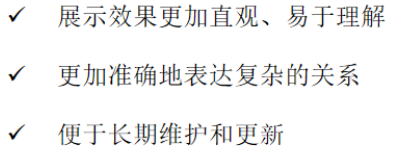
* **超图是什么？**



需要注意的是：判断一个点是否属于这一条超边，需要看这个点是否完全在这条超边内。如下图是截取的部分超图，请问红圈中的PA和SA在一条超边中吗？答案是否定的，因为如果SA也在这条棕色的超边中，那么SA所在的这一块会向左下方凸起，把SA完全包容在里面，这是超图绘制算法的规则。



* **超图的好处有哪些？**



* **举个使用超图的例子吧！**

**用超图表示作品中的人物与章节的关系**

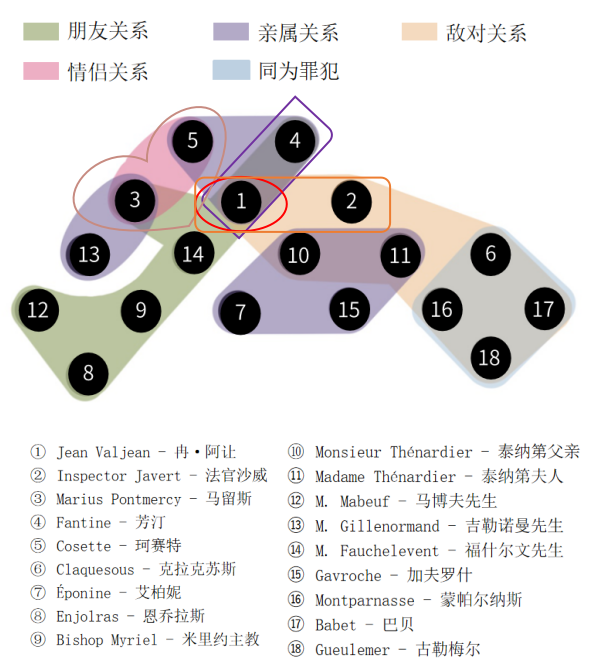
该方法可以揭示《悲惨世界》中一些有趣的信息和模式，例如：

【顶点1】冉·阿让(Jean Valjean)是超图中很重要的顶点，他与其他顶点有最多的连接，他也是唯一一个在所有场景中都出现过的顶点。

【顶点1和2】法官沙威(Javert) 是冉·阿让最强大的敌人，他们之间有一个肉色超边，表示他们之间的长期对抗。

【顶点1和4】芳汀(Fantine) 是冉·阿让最重要的朋友之一，她是唯一一个与冉·阿让有两个不同类型的超边(友谊和亲属)的顶点。

【顶点3和5】马留斯(Marius)和珂赛特(Cosette)是故事中的一对恋人，他们之间有一个粉色的超边，表示他们之间的爱情。



**三、实验所关注的因素的介绍**

1. **凹度——实验1**

**凹度**：指每条超边的非凸的平均程度（“凹”的反义词是“凸”）。凹度越小，即超边非凸的程度越低，超图就越美观。下图图1中，右图中超边非凸的程度低，凹度比左图小，更加美观。

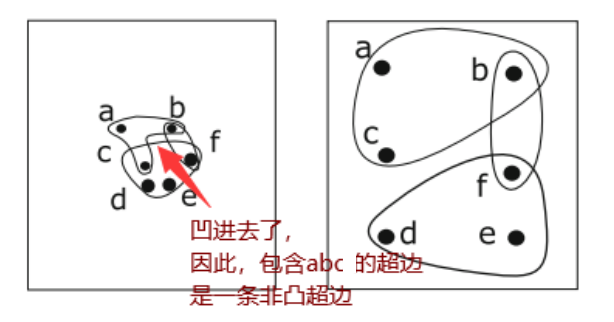


图1

需要注意的是：用户在回答问题时，关注的是每条超边的凹度，而非整体形状的凹度。如下图图2中，整体形状近似为圆、凹度很低，但其中的每条边凹度都较大。

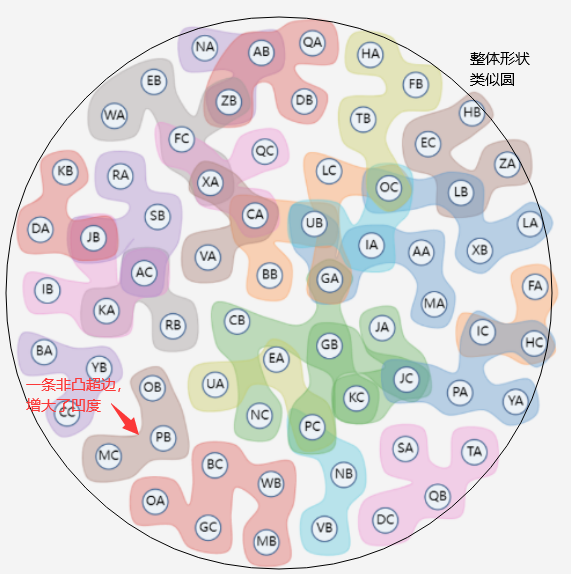


图2

提示：在实验过程中，您可能会遇到凹度相近让您难以排序的情况，请无需非常精确地排序，因为我们看重的往往是排序后首尾两端的结果。

1. **平面度——实验2**

**平面度**：指超边相交的数量。平面度越小，就越有利于避免杂乱，从而避免了超图所代表的关系的模糊性。

需要注意的是：如下图图3，红色框里，绿色段超边覆盖了粉红色超边，这不算相交（因为这是数据集定义的关系，不是通过布局算法可以改变的）；黑色框里，多条超边汇聚于点“NORA”，这不算相交；紫色框里，绿色段和棕色段超边有重叠部分，且不属于上面2种情况，这**算相交**。

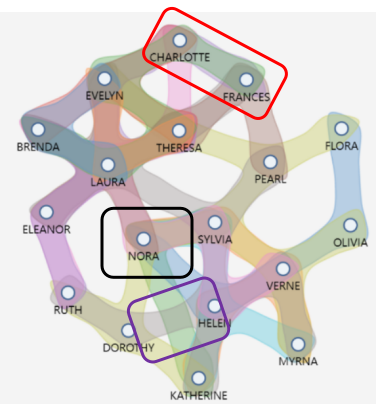


图3

1. **覆盖率——实验3**

**覆盖率**：指超图对画布的利用率。

需要注意的是：超图中间的空白部分也算作超图对画布的利用，如下图图4中红圈内的DB和LC之间有缝隙，但也算作利用了画布（我们将其视作是由布局算法的排斥力导致的）。

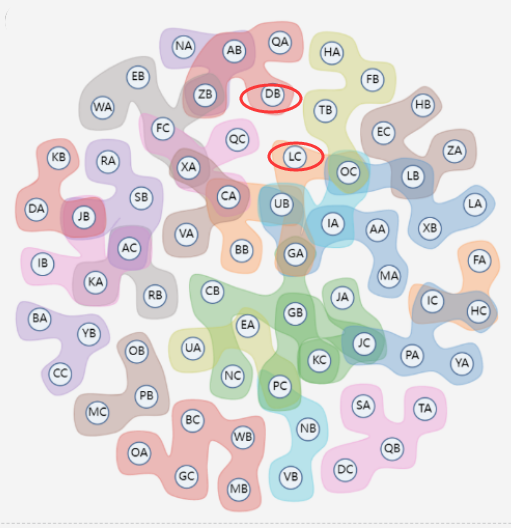


图4

1. **节点间连线距离——实验4**

**节点间连线距离**：其均匀程度主要看各节点间的边长是否均匀。

需要注意的是：请注意我们探究的是连线距离，而不是节点距离。如图5，节点距离较为均匀，因为每个节点间距离较为一致，如点MC到周围的PA和HC距离相近；但是节点间连线距离均匀程度却很低，因为红框中连接着MC和OB两点的棕色边极长（注意PA、SA、TA都不属于棕色超边）。

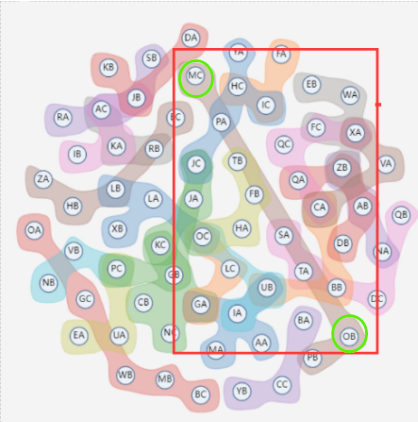


图5

1. **对数据类别的认知能力——实验5**

我们会根据列出具体的问题，请用户判断两个点是否在同一个类中。

需要注意的是，由于第二个和第三个数据集点数较多，因此我们将探究的两个点“LB”和“XB”用青蓝色方框标出，如下图图6所示。

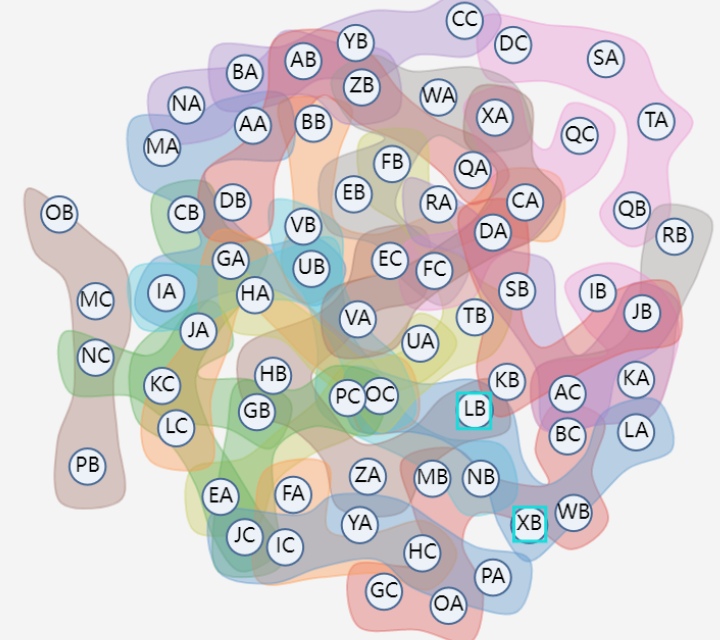


图6