

文章编号:1007-6735(2011)03-0229-09

关于超网络的一点思考

王众托

(大连理工大学 系统工程研究所, 大连 116085)

摘要: 在探讨超网络的内涵时提到,除了 Nagurney 所着重研究的多层超网络之外,还有一类基于超图(hypergraph)的超网络(hypernetwork)值得关注.超图的特点是图中的边可以连接两个以上的节点,这对研究信息与知识网络的研究与应用会有独特的用途.以知识的超网络表述为例,介绍了基于超图的超网络的简单应用.

关键词: 复杂网络; 超网络; 超图

中图分类号: N 94 **文献标志码:** A

Reflection on supernetwork

WANG Zhong-tuo

(Institute of Systems Engineering, Dalian University of Technology, Dalian 116085, China)

Abstract: In this paper, the concept of hypernetwork based on hypergraph suggested by C. Berge as another kind of supernetwork is introduced. The main feature of hypernetwork is the edge in hypergraph can be connected to more than one nodes. For the investigation of some complex network like telecommunication network, knowledge network, social network, this unique feature may offer more advantages.

Key words: complex work; hypernetwork; hypergraph

现在可以说我们是生活在一个网络世界里,各种设施和事务都是通过网络联系起来的.例如交通运输网络是由道路把站点连接形成的网络.电力网络是由发电厂、变电所和用户通过输电和配电线路连成的网络.计算机网络可以看作是自主工作的计算机通过通信介质如光缆、双绞线、同轴电缆或者无线电波等相互连接形成的网络.神经系统可以看作是大量神经细胞通过神经纤维相互连接形成的网络.上面提到的是有形的网络,还有许多由于社会关

系(业务活动、血缘与人际交往等等)而形成的无形网络.人类社会中的网络的产生和发展推动了社会的进步,例如交通运输系统给人员与货物的流动带来方便,促进了商贸和工农业的发展;信息系统给资金的流转、商务的流通带来很多的方便,提高了生产效率和生活质量.但是另一方面像传染病的流行和网络病毒的传播,电力网的局部故障引起大面积的停电,核辐射的传播给人们的健康带来了威胁,这又是由于网络的存在而带来的危害.因此对自然网络

收稿日期: 2011-06-07

作者简介: 王众托(1928-),男,中国工程院院士.研究方向:系统工程、知识科学与技术、决策分析.

E-mail: wangzt@dlut.edu.cn

和人工网络的认识和干预(构建、运行和防护)就成为科学技术所要解决的问题之一。

典型的网络是由许多“节点”与连接两个节点之间的一些“边”组成的,其中节点用来表示实际系统中不同的个体,连接的边则用来表示两个节点之间具有某种特定的关系。用图形表示就是常见的网络图。在节点之间通过边有网络流(物流、能流、信息流、资金流)在流动。随着网络规模的日益扩大和连接的日益复杂,人们对网络除了从各自专业的角度加以研究外,还关注各类网络在结构和运行上的共性,希望能够用抽象的数学工具来加以描述和分析。从18世纪数学家欧拉对所谓“Konigsberg 七桥问题”的建模和分析,从而开创了数学中图论这一分支以来,在运筹学中的网络系统分析中对最短路径、最大流以及最小费用流等方面,还有对时间流进行分析与设计的网络计划技术等,有了一些解决实际问题的研究。随着交通事业的发展,出现了铁路网、公路网、航空网,信息化的发展产生了各类信息网。网络化的发展,出现了许多复杂的网络,这些网络节点和边的数量众多,结构复杂,例如互联网(Internet)的用户已经是数以亿计,由于缺少统一的管理,时至今日没有任何一个人或组织能够知道互联网上所有的路由器到底是怎么联结在一起的。再就是连接的多样性,有的松散有的紧密,属性也有差异。网络节点和边又都可能具有复杂的动态行为。人们在运用网络满足自己的要求,常常是多目标或多准则的,而网络的存在又提供了更多的选择性,因而使得决策和优化问题更加复杂。生活实践和理论研究都需要对这样的复杂网络(complex network)进行深入的认识与分析。

到目前为止,科学家们还没有给出复杂网络精确严格的定义,但我们已经可以从上面列举的复杂网络的特征来理解它的含义。

一个网络的功能与它的结构密切相关,因此要从研究网络的结构入手。最初对复杂网络的拓扑结构是用具有一定规则的结构表示的,形成所谓规则网络,但其适用范围过窄。20世纪60年代出现的随机图理论,适应了网络的不确定性特点,开创了复杂网络的系统性研究。在这种方法中,两个节点之间是否有边相连不再是确定的事情,而是根据一个概率决定。但大多数实际的复杂网络结构并不是完全规则或完全随机这两个极端,而是要比规则网络和随机网络更为复杂。因此需要探索和发现网络中的一些特殊的性质。当时有两项研究成果使得研究工作

得到了很大的推进,即发现某些网络中的小世界效应(small-world effect)和无标度特性(scale-free property)^[1-2]。科学家在研究节点的度(与该节点连接的边的数目)的分布时,发现大量真实网络的节点度服从幂率分布(连接边多的节点少而连接边少的节点却非常多),而不是像随机网络那样服从钟型分布。我们把节点度服从幂律分布的网络叫做无标度网络(scale-free networks),Barabási 和 Albert 把真实系统通过自组织生成无标度的网络归功于两个主要因素:生长和优先连接,而他们的网络模型(BA网络)正是模拟这两个关键机制设计的^[2-3]。

这些成果使得人们对于实际存在的网络的生成机制、传播机制、茁壮性和脆弱性等有了新的认识。随后有关复杂网络的研究有了长足的进展。

1 超网络

但是在有些情况下,用一般的网络图并不能完全刻画真实世界网络的特征。在研究超大规模的网络系统时,会出现网络中的网络的问题。如果用原来简单或有向图的方法来处理这类问题,就很难理清各个网络之间的关系。因此就出现了超越一般网络的网络系统问题。人们常常把规模巨大、连接复杂、节点具有异质性的网络称为超网络,例如认为互联网是一种超网络,但这只是从常识上的一种称呼,没有明确定义。另有一种对超网络的理解,是认为超网络是网络组成的网络。最早明确提出“超网络”概念的是美国科学家 Nagurney 等^[4-5],她在处理物流网络与信息网络、资金网络相交织的问题时,把高于而又超于现存网络(“above and beyond” existing networks)的网络,称为超网络(supernetwork)。

Nagurney 所研究的超网络具备一种或几种特征:a. 多层特征,例如交通运输网就有物理层、业务层和管理层;信息网络协议也是多层的,层内和层间都有连接。b. 多级特征,例如企业的信息网络有部门、公司、总部等级别,同级和级间都有连接。c. 它的流量可以是多维的,例如铁路、公路、水运和航空都是既有客运又有货运。d. 多属性或多准则的,例如在城市中出行不仅有路径选择,而且有方式(驾车、公交、步行)的选择,运输网络需要同时考虑时间、成本、安全舒适等。e. 存在拥塞性,不仅交通运输网络,而且信息网络也存在拥堵问题。f. 全局优化和个体优化需要协调。

超网络模型可用来描述和表示网络之间的相互

作用和影响. 它可以将许多不同层次、不同标准的决策者之间的关系用关系函数来表示, 在建立关系过程中, 决策者必须付出一定的代价, 如需付出物品、时间、金钱等. 随着关系层的增加, 交易的成本和不确定性就会增加, 因此必须给关系附加描述值. 为了以较小的代价取得最大的利润, 决策者一定会考虑最大关系值, 因此决策者既要找到与网络中其它决策者交易的组合优化, 又要找到关系层的组合优化. 超网络的构架为研究网络之间的相互作用和影响提供了工具. 它可以用一些数学工具对网络上的流量、时间等变量进行定量的分析和计算, 这些数学工具包含: 优化理论、博弈论、变分不等式及可视化工具等. 可以举出下面一些超网络的例子:

a. 供应链与社会网络结合的超网络模型^[5]

这是由社会网络和供应链网络组成的超网络. 这两个网络都由三层决策者构成, 社会网络中的流是各层之间的关系, 供应链网络中的流是产品之间的交易.

b. 电子商务中的供应链超网络模型 这是一个包括 m 个商家和 n 个需求市场的供应链超网络, 其中, 商家与需求市场的交易流程(订货、付款、传递)都在互联网上进行, 脱离了传统的物流配送系统.

c. 回收超网络模型 一个由 $m + n + o$ 个单位组成的回收超网络, 其中包括 m 个可造产品资源地, n 个回收商和 o 个需求生产商, 并任意取第 i 个资源地, 第 j 个回收商和第 k 个需求生产商进行讨论. 有一些载能资源恢复理想中的使用价值会比由原材料生成所需要费用更高, 不值得回收, 因此对它们的处理就是使之销毁并掩埋.

d. 闭环供应链超网络模型 由原材料供应商、生产商、零售商、需求市场和回收商组成的 $I + J + K + M + N$ 个闭环供应链超网络. 其中, 包括 I 个生产商总数, J 个零售商总数, K 个需求市场总数, M 个回收商总数, N 个原材料供应商总数, 并任意取第 i 个生产商, j 个零售商, k 个需求市场, m 个回收商, n 个原材料供应商.

前面的一些超网络是从网络结构来考虑的, 而现实生活中还有另外一些超网络是从联系方式来考虑的. 例如:

a. 社会网络^[6] 在社会网络中, 点表示一个人或一群人, 通常称为行动者, 他们之间如果交往或接触就用一条边将其连结. 这样的连结方式可以是友谊、合作、贸易关系等. 例如贸易交往中, 必须考虑多

于3者之间的关系, 如卖者、买者、经纪人. 其它的例子中, 不仅仅要考虑行动者参加行动的重要性, 还要考虑其它的一些因素, 如参加行动的时间、地点等.

b. 蛋白质网络^[6] 在一个有机体蛋白质水解过程中, 对多蛋白质复合物的系统描述需要蛋白质成员之间有组织的数据. 这个组织最普通的形式是蛋白质之间的相互反应网络和复合物交叉图. 在前一种表示中, 网络的点代表蛋白质而边连着相互反应的蛋白质. 然而, 这个表示并没有考虑到多蛋白质复合物. 在复合物交叉图中, 点表示复合物, 而如果两个复合物之间有一个或多个共同的蛋白质的话, 就用一条链将它们连结. 很明显, 这后一个表示并不能提供关于蛋白质的信息, 因此需要使用超网络来提供蛋白质和多个共同的蛋白质成员之间的信息^[7].

c. 食物网^[6] 在生态学中, 营养关系通常用表示为有向图(子图)的食物网来表示, 其中, 点表示物种而链表示物种之间的营养关系^[8]. 表示食物网的另一种方法是用竞争图 $C(G)$ 来表示, 其中点集与食物网相同, 当且仅当联系的物种在食物网中有共同的猎物时两点之间是连通的^[9]. 在竞争图中, 仅仅知道两个联系的物种之间有共同的猎物, 但并不知道为共同猎物竞争的整个物种群的构成情况. 为了解决这个问题, 需要建立竞争超网络^[10].

d. 计算机网络^[6] 在计算机领域中, 若把多机系统中的处理机看作节点, 处理机的集合作为超网络节点集; 把总线集作为超网络的边集, 于是一个多机系统就抽象为一个超网络.

e. 知识网络^[11] 现在已经存在的互联网和各组织内部的内联网也都可以看作是知识网络, 因为以信息为知识载体, 在网上储存了与流通着大量的知识. 这类知识网络由三层组成: 技术层、知识内容层和人员层(社会网络层).

2 基于超图的超网络

目前还有用超图(Hypergraph)来定义的另一类超网络. 这类超网络的英文名字叫 Hypernetwork, 有别于 Supernetwork. 超图的概念是 Berge^[12] 于1970年提出的, 文献[12-13]第一次系统地建立了无向超图理论. 超图不同于一般图论中的无向或有向图的地方在于: 后者的每一个边只连接两个节点, 而超图中的边可以连接两个以上的节点, 所以称之为超边.

数学上超图的严格定义为:

设 $V = \{v_1, v_2, \dots, v_n\}$ 是一个有限集. 若

(1) $e_i \neq \emptyset (i = 1, 2, \dots, m)$;

(2) $\bigcup_{i=1}^m e_i = V$.

则称二元关系 $H = (E, V)$ 为一个超图. V 的元素 $v_1, v_2, \dots, v_n$ 称为超图的顶点, $E = \{e_1, e_2, \dots, e_m\}$ 是超图的边集合, 集合 $e_i = \{v_{i_1}, v_{i_2}, \dots, v_{i_j}\}$

($i = 1, 2, \dots, m$) 称为超图的边.

超图 H 的对偶 H^* 称为对偶超图. 如果对所有 $j, V_j = \{e_i \mid i \leq m, v_j \in E_i\}, v_j \neq \emptyset, \forall j \cup_j V_j = E$, 那么, 超图 $H^* = (E; V_1, V_2, \dots, V_n)$ 称为 H 的对偶超图. 显然, $(H^*)^* = H$. 例如, 研究图 1 中的 4 个拱门 a_1, a_2, a_3, a_4 . 它们是由总共 8 类构件 $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$ 组成的, 每个拱门所用的构件不同.

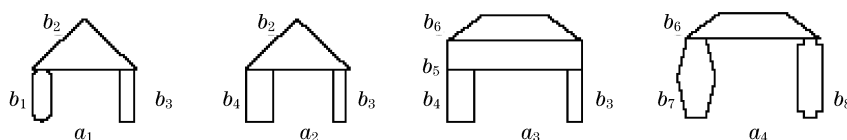


图 1 4 个拱门的图形

Fig.1 Diagram of 4 arches

如果用二分图来表示构建和拱门的关系, 可以画出图 2 来 (见图 2). 这里有两类性质不同的节点, 一类表示拱门, 一类表示构件, 这时节点的同质性失去了, 在处理上就会带来许多不方便.

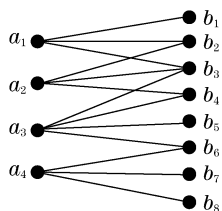


图 2 用二分图来表示 4 个拱门

Fig.2 Representation of 4 arches by Bi-partition graph

如果用超图表示其中的关系, 以各构件 b 为节点, 以 $a_1(b_1, b_2, b_3), a_2(b_2, b_3, b_4), a_3(b_4, b_5, b_6), a_4(b_6, b_7, b_8)$ 为超边, 就得到了如图 3 所示的超图. 这些超边表示出各拱门都是用哪些构件构成的. 假如以各拱门 a 为节点, 以 $b_1(a_1), b_2(a_1, a_2), b_3(a_1, a_2, a_3), b_4(a_2, a_3), b_5(a_3), b_6(a_3, a_4), b_7(a_4), b_8(a_4)$ 为超边, 就得到如图 4 所示的超图. 其中的超边表示各构件都用在哪些拱门上.

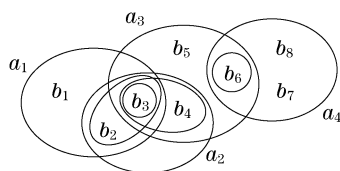


图 3 超图表示的 4 个拱门

Fig.3 Representation of 4 arches by hypergraph

图 3、图 4 两个超图是互为对偶的. 超图在描述多个节点有连接的网络时, 有它的方便之处. 例如在

电话的通话业务中, 每一个用户可以用一个节点来表示. 如果两个人通话, 可以用一般图论中的边来描述. 如果是 3 个人召开电话会议, 用一般图论中的边来描述, 就得像图 5(a) 中那样, 需要 3 个边. 如果用超图来描述, 则如图 5(b) 那样, 使用一个超边就行了.

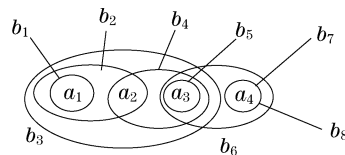


图 4 超图 3 的对偶超图

Fig.4 Duality of Fig.3

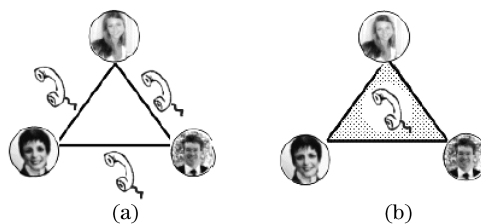


图 5 电话通话的图形表示

Fig.5 Representation of phone communication

从现有研究成果来看, 关于超图理论及其应用的研究是另外一条主线. 这里可以提到的主要有以下几个方面:

关于超图理论本身的研究开始较早, 文献[14-17]分别研究了超图理论的基本概念及基本性质. 文献[18]引入了超通路、超回路、超边分解、分解超图、 k -超树、单向超通路和单向超树等概念, 建立了有向超图理论. 文献[19]研究了超图的 Helly 性质、Ramsey 数、横贯与匹配、分数横贯、着色等. 文献

[20]研究了超图的 t -设计.

基于超图的超网络的应用方面,文献[21]将超图应用在大规模集成电路版图设计中.文献[22]提出一种“基于超图的容错 VLSI 处理阵列”,基于文献[23]的成果,文献[24-25]应用具有最佳连通性超图的一些性质解决了多总线系统的分析和综合问题.文献[26]提出了一种基于优先记发式的超图二划分算法.文献[27]给出了函数依赖集的超图表示,提出了一种基于有向超图的求最小覆盖集的新算法.文献[28]基于超图模型生成了紧的无冗余 XML 文件.文献[29]将超图划分方法应用在 VLSI 领域.文献[30-32]将超图与图象处理相结合,建立了图象邻域超图模型.文献[33]对细胞动态交流系统建立了超图模型.文献[34-35]对数据挖掘中的聚类提出了新的算法且建立了超图模型.文献[36]将超图应用在化学上,利用超图的拓扑指数来识别化学中的化合物分子结构.文献[37]用超网络来辨别 DNA 切块,文献[38]用 DNA 超网络对信息进行存贮和获取.文献[39]用进化超网络对模式进行了分类,建立了进化变阶的超网络模型.超图作为一类特殊的图,不像一般图论那样和易于运用.它的应用也还没有像一般图论那样成熟,但是在运用于复杂的网络系统的描述和分析上,具有很大的潜力.

笔者认为,基于超图的超网络可以认为是广义的超网络中的一个子类,就像以 Nagurney 为代表所研究具有前面所述的 6 种特性的网络,也是超网络的一个子类.因为两者像 Nagurney 所说的,都是高于而又超于现存网络(“above and beyond” existing networks)的网络.

3 基于超图的超网络在知识表述上的应用举例

如何将网络同知识的表示与组织结合起来,成为许多知识科学与知识管理研究者感兴趣的话题之一.经过近年来的探索,知识网络的概念逐步形成.知识网络作为一个概念,早在 1995 年 Beckmann^[40]就提出过,他认为知识网络指的是从事科学知识的生产和传播的机构和活动.超网络和超图在知识的表述和处理(聚类、分类、集成等)中将会有很多的用途.可以用一个简单的例子来加以说明^[11].主题的辨识与提取是知识建模中的重要环节.长期以来认为通过超链可以找到围绕着主题的主题文档类,但这还

不足,应该进一步研究辨识主题或者说生成主题的方法.一类含有同样主题的文档是可以聚类的,但主题是隐含在其中的.有时候一个主题是可以用一个表述能力很强的词语来描述的,也有时候需要一组相关的词语来描述.有的词语有利于检索,但并不一定是好的描述符.这又和具体任务有关,有的是为了寻找一个很狭小的范围内的具体内容,有的是要在一个很宽泛的范围内收集涉及面很广的知识.两者对描述符的要求就不一样.从直觉上来判断,如果一个或几个词语能够明确回答“这个主题是关于什么的”问题,就是好的主题描述符.如果一个或几个词语能够回答“什么是能够得到类似的信息的好的询问词语”,就是好的鉴别符.

为了确定主题的描述符和鉴别符,需要分析词语、文档和主体之间的关系.文献[41]建议使用超图来表示这样的关系.我们可以把一组文档看做一个超图 $H = (T, D)$,其中每一个节点 $t \in T$ 对应于一个词语,每一个超边 $d \in D$ 对应于一个文档.一个超边 d 是一个 T 中元素的多元集,把一个文档抽象地表示为像一个装了词语的口袋.可以把这种超图称为以文档为中心的超图.

对应于每一个以文档为中心的超图 $H = (T, D)$,有一个以词语为中心的超图 $H^* = (D, T)$,它的节点对应于文档而超边对应于词语,超边是文档的多元集.超图 H^* 是超图 H 的对偶超图.图 6 中的超图,表示的是由 3 个文档 A、B、C 和其中包含的词语 1,2,3,4 之间的关系.

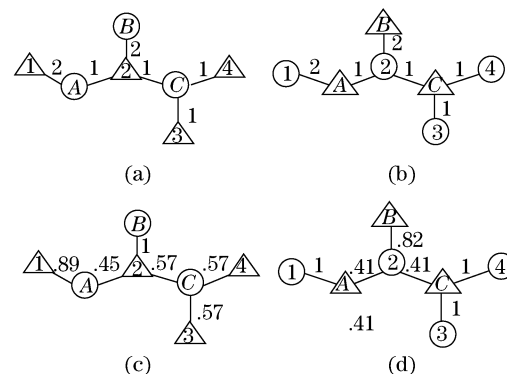


图 6 用超图表示文档(取自于文献[41])

Fig.6 Representation of documents (from [41])

图 6(a)是以文档为中心的超图

$$H = (\{1, 2, 3, 4\}, \{A, B, C\})$$

其中, $A = \{1, 1, 2\}$, $B = \{2, 2\}$, $C = \{2, 3, 4\}$

图 6(b)是它的对偶超图

$$H^* = (\{A, B, C\}, \{1, 2, 3, 4\})$$

其中, $1 = \{A, A\}$, $2 = \{A, B, B, C\}$, $3 = \{C\}$, $4 = \{C\}$.

在图6中, 圆圈表示超边而小三角表示节点. 图6(a)与图6(b)中节点和超边的连接上的数字表示的是该节点在超边中出现的次数. 例如节点1和超边A的连接上标注的2, 表示词语1在文档A中出现2次.

这里这样定义关联矩阵: 对一个由 m 个文档和 n 个词语的以文档为中心的超图来说, 其关联矩阵的 m 行对应于文档, n 列对应于词语. 矩阵的元素

$$H[i, j] = k$$

其中, k 是第 j 列的词语在第 i 行的文档中出现的次数. 我们还可以知道, 对偶超图的关联矩阵是原超图关联矩阵的转置. 可以使用关联矩阵的某些性质来反映词语在文档中的描述能力和鉴别能力. 词语 j 在文档 i 中的描述能力可以用下列函数来衡量

$$\lambda(d_i, t_j) = \frac{H[i, j]}{\sqrt{\sum_{k=0}^{n-1} (H[i, k])^2}}$$

利用这一函数(它的数值在0与1之间)可以构成加权的以文档为中心的超图, 如图6(c)所示, 它可以称为 d -超图. 从图中可以看出, 不同的词语的描述能力是不同的, 例如词语2对文档B的描述能力最强, 词语1对文档A的描述能力比词语2强.

为了度量词语对文档的鉴别能力, 可以使用下列函数

$$\delta(t_i, d_j) = \frac{s(H^T[i, j])}{\sqrt{\sum_{k=0}^{m-1} s(H^T[i, k])^2}}$$

其中的 $s(k)$ 是当 $k > 0$ 时其值为1, 当 $k = 0$ 时其值为0. 这个函数的值也是在0与1之间. 如果第 i 个词语不出现在文档 j 之中, 则函数值为0; 如果第 i 个词语除了出现在文档 j 之外不再出现在其它文档之中, 则该函数值为1, 因而可以说这个词语完全鉴别出该文档. 这个函数与词语在文档中出现的次数无关. 利用这一函数可以构成加权的以词语为中心的超图, 如图6(d)所示, 它可以称为 t -超图. 其中, 词语1完全鉴别出文档A.

对 d -超图与 t -超图来说存在关系

$$\sum_j (\lambda(d_i, t_j))^2 = 1$$

$$\sum_j (\delta(t_i, d_j))^2 = 1$$

上面这种选择最好的描述符合鉴别符的方法仅仅是对文档而言的, 还需要为主题选取好的描述符和鉴别符. 主题总是要涉及多个类似的文档或词语, 因此需要研究文档的相似性和词语的同时出现性.

两个文档 d_i 与 d_j 的相似性可以按照众所熟知的余弦测度来衡量

$$\sigma(d_i, d_j) = \frac{\sum_{k=0}^{n-1} (\lambda(d_i, t_k) \lambda(d_j, t_k))}{\sqrt{\sum_{k=0}^{n-1} (\lambda(d_i, t_k))^2 \sum_{k=0}^{n-1} (\lambda(d_j, t_k))^2}} = \frac{\sum_{k=0}^{n-1} (\lambda(d_i, t_k) \lambda(d_j, t_k))}{\sqrt{\sum_{k=0}^{n-1} (\lambda(d_i, t_k))^2 \sum_{k=0}^{n-1} (\lambda(d_j, t_k))^2}}$$

两个词语 t_i 与 t_j 的同时出现性可用下列测度来衡量

$$\kappa(t_i, t_j) = \frac{\sum_{k=0}^{m-1} (\delta(t_k, d_i) \delta(t_k, d_j))}{\sqrt{\sum_{k=0}^{m-1} (\delta(t_i, d_k))^2 \sum_{k=0}^{m-1} (\delta(t_j, d_k))^2}} = \frac{\sum_{k=0}^{m-1} (\delta(t_i, d_k) \delta(t_j, d_k))}{\sqrt{\sum_{k=0}^{m-1} (\delta(t_i, d_k))^2 \sum_{k=0}^{m-1} (\delta(t_j, d_k))^2}}$$

图7(a)使用 d -超图来阐明文档的相似性. 从图中可以看出, 文档D和E是相似的. 图7(b)使用 t -超图来阐明同时出现性. 可以看出, 词语3和4是同时出现的. 怎样识别一个文档的主题的最好的鉴别符呢? 如果一个词语他所鉴别出来的文档是和某一已知文档相似的, 那么这个词语就是这个文档的好的鉴别符. 如果用公式, 则可用函数

$$\Delta(t_i, d_j) = \sum_{\substack{k=0 \\ k \neq j}}^{m-1} (\delta(t_i, d_k)^2 \sigma(d_k, d_j))$$

表述. 可以把词语 t_i 对文档 d_j 的鉴别能力, 看作是文档 d_j 与由 t_i 鉴别出来的其它文档的相似性的平均值.

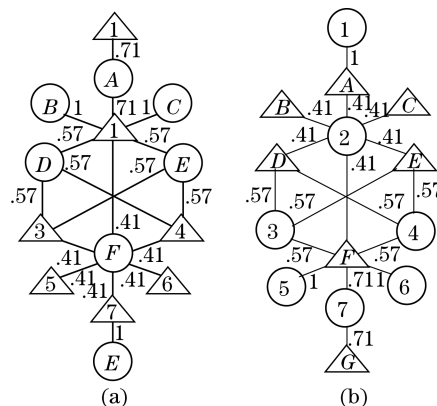


图7 用超图来说明文档的性质图(取自于文献[41])

Fig.7 Explanation of features of documents (from[41])

下面再来看主题的描述符. 前面曾经非正式地说到过, 主题的描述符就是在主题的环境中经常出现的词语. 一个词语在主题中的描述能力可以用前面用到的对文档相似性和文档中词语的描述能力来计算. 这样就能够用函数来度量

$$\Lambda(d_i, t_j) = \begin{cases} 0 & \text{if } \sum_{\substack{k=0 \\ k \neq i}}^{m-1} \sigma(d_i, d_k) = 0 \\ \frac{\sum_{\substack{k=0 \\ k \neq i}}^{m-1} (\sigma(d_i, d_k) \lambda(d_k, t_j)^2)}{\sum_{\substack{k=0 \\ k \neq i}}^{m-1} \sigma(d_i, d_k)} & \text{otherwise} \end{cases}$$

一个词语 t_j 在文档 d_i 的主题中的描述能力, 是词语 t_j 作为与文档 d_i 相似的所有文档的描述符的质量指标. 如果没有其它文档和 d_i 相似, 或者 t_j 没有出现在与 d_i 相似的其它文档之中, 则 t_j 在 d_i 的主题中的描述能力为 0. 在图 7 中, 词语 2, 3, 4 在文档 D 、 E 、 F 的主题中都是好的描述符. 但是 3, 4 是这一主题好的鉴别符, 而 2 却不是. 因为 2 虽然在该主题中, 但却不仅在这一主题中. 在这 3 个文档中, 只有 D 和 E 是聚焦在这一主题上的. 文档 F 包含了这一主题同时出现的词语, 但不是这一主题的独有词语.

以上仅仅是一个简单的例子用来说明这类超网络的应用, 这方面还有很大的发展空间.

4 超网络进一步发展的一点思考

超网络的提出只是近几年的事, 目前它也只是一个概念, 其边界并不十分清晰. 对于一些实际问题, 仅仅在利用变分不等式和超图的基础上建立了一些模型, 提出了一些解法. 对超网络本身还没有众所公认的定义. 对于什么样的网络可以算作超网络也还有一个逐步明确的过程. 但是从上面已经列举的一些事例来看, 由于社会经济和科学技术的网络化的进一步发展, 这类网络问题逐渐呈现, 像对有关复杂网络的研究一样, 人们开始找到了它的一些特征, 以及合适的描述方法.

对照复杂网络的含义, 可以说超网络也是一类特殊的复杂网络. 上世纪末和本世纪初国内外都掀起了对复杂网络研究的高潮. 特别是人们摆脱了还

原论的局限性, 进而开始研究网络各组成部分的相互影响, 这正是超网络的研究内容. 在文献[42]中一批权威作者经过讨论提出了当前网络研究的 10 个主要问题, 其中谈到了有关“网络的网络”问题, 即研究不同性质网络间的相互作用问题. 可见超网络问题将逐步进入网络研究的主流. 特别是在系统工程界近年来开始研究所谓“系统的系统, SoS”, 如果系统可以用网络来描述, 那么“系统的系统”就可以用“网络的网络”也就是超网络来描述. 对超网络的研究, 和对一般的复杂网络的研究一样, 将会有助于理解“复杂系统之所以复杂”这一极其重要的问题, 这也就是说, 超网络问题的研究不仅有其实际意义, 而且也对系统复杂性的探索开辟了一个新的途径. 超网络概念的提出, 使得人们对一些多层、多级、多维网络流、多属性和多准则的网络问题有了新的认识和理解, 但是我们还只是在复杂性的密林中刚刚找到一些小道, 还需要开辟新的道路. 尽管在文献[6]中用到了复杂超网络这一提法, 但依我们的看法, 除了一些极其简单的超网络问题可以凭经验解决外, 大多数超网络都是复杂的超网络, 都可以看作是一类特殊的复杂网络, 这种复杂性主要表现在属性上, 而不是在规模上. 所以不必另加复杂这一定语. 正是由于其属性的特点才使超网络问题需要另辟蹊径专门加以研究.

依笔者的浅见, 对超网络问题的研究首先要考虑的问题, 是探索发现实际生活中的一些网络由于哪些特征才需要, 而且可以把它们看作是超网络问题, 以及怎样建立从概念模型、结构模型乃至数学模型. 现有的超网络分析优化方法还是极其有限的, 需要探寻新的方法. 超网络虽然一时无法得到众所公认的定义, 但是否可以从深入研究其典型特征入手来加以界定. 网络的拓扑结构在网络的稳定性等方面有着非常重要的作用.

在应用方面, 值得考虑的问题是:

a. 复杂网络和超网络概念的提出, 一方面在为客观事物的描述和建模过程中, 填补了离散对象的集总参数模型(利用代数和常微分方程、差分方程)和连续介质对象的分布参数模型(利用数学物理方程)之间的空白(过去像对复杂运输网络这类系统的建模是面临很大困难的). 另一方面, 提供了考虑和分析网络各组成部分之间(特别是层间和级间)以及网络和网络之间关系的方法. 这就为许多复杂系统研究提供了新思路, 随着实际生活中系统的日益复杂和多样化, 将有更多的系统可按照超网络来处理.

b. 在 Nagurney 所说的“高于而又超于现存网络”中所指的现存网络,是节点对应于空间位置、边对应于物理连接(如道路、线缆)的网络,而超乎这样网络的则是一些带有虚拟特点的点、边和流,如知识节点、社会关系、价格流等.因此超网络的概念和处理方法必将在虚实结合的网络研究中显示其独特优势,得到广泛应用,特别是基于超图的超网络.

c. 在解决问题的已有研究基础上,应该在更广泛的运输网络、通信网络、能源网络、金融网络中,考虑更多的风险和其它因素,将超网络模型进一步完善,使之更符合现实.

d. 将超网络进一步应用于物联网系统、知识管理系统^[11]等一些新型系统之中.

参考文献:

- [1] WATTS D J, STROGATES S H. Collective dynamics of “small-world” networks [J]. Nature, 1998, 393: 440 – 442.
- [2] ALBERT R, BARABASI A L. Statistical mechanics of complex networks [J]. Review of Modern Physics, 2002, 74: 47 – 97.
- [3] BARABASI A L. Linked: How Everything Is Connected to Everything Else and What It Means for Science, Business and Everyday Life [M]. Cambridge: Perseus Publishing, Inc, 2002.
- [4] NAGURNEY A, DONG J. Supernetworks: Decision – Making for the Information Age [M]. Cheltenham: Edward Elgar Publishers, 2002.
- [5] NAGURNEY A, WAKOLBINGER T. Supernetworks: An introduction to the concept and its applications with a specific focus on knowledge supernetworks [J]. International Journal of Knowledge Culture and Change Management, 2005, 4: 1 – 16.
- [6] EMESTO ESTRAD A, JUAN A. Rodriguez-velazquez. subgraph centrality and clustering in complex hyper – networks [J]. Physica A, 2006, 364: 581 – 594.
- [7] RAMADAN E, TARAFDAR A, POTHEN A. A hypergraph model for the yeast protein complex network [C] // IPDPS Workshop on High – Performance Computational Biology. Proceedings of the HICOMB 2004. New Mexico: Santa Fe NM, 2004.
- [8] PIMM S L. Food Webs [M]. London: Chapman & Hall, 1982.
- [9] COHEN J E. Interval graphs and food webs, a finding and a problem. Document 17696 – PR [Z]. Santa Monica: RAND CORP, 1968.
- [10] SONNTAG M, TEICHERT H M. Competition hypergraphs [J]. Discrete Appl Math, 2004, 143: 324 – 329.
- [11] 王志平, 王众托. 超网络理论及其应用 [M]. 北京: 科学出版社, 2008.
- [12] BERGE C. Graphs and Hypergraphs [M]. New York: Elsevier, 1973.
- [13] BERGE C. Hypergraphs: The Theory of Finite Sets [M]. Amsterdam: Elsevier Science, 1989.
- [14] 罗静, 毛其智, 党安荣. 基于超图的生态环境修复模型的研究与应用 [J]. 计算机工程与应用, 2007, 43(31): 188 – 191.
- [15] KARONSKI M. A review of random graphs [J]. Journal of Graph Theory, 1982, 6(4): 349 – 389.
- [16] ARIEZRI S, FRACNKDL E. A deletions game on hypergraph [J]. Discrete Applied Mathematics, 1991, 30: 155 – 162.
- [17] ENDRE BOROS (USA). On shift stable hypergraphs [J]. Discrete mathematics, 1991, 87: 81 – 84.
- [18] 黄汝激. 超网络的有向 k 超树分析法 [J]. 电子科学学刊, 1987, 19(3): 244 – 255.
- [19] 贝尔热 C. 超图 – 有限集的组合学 [M]. 卜月华, 张克民译. 南京: 东南大学出版社, 2002.
- [20] TUAN N D. Hypergraphical t – designs [J]. Discrete Mathematics, 2006, 306: 1189 – 1197.
- [21] 黄汝激. 应用超图理论实现有向基本割集矩阵 [J]. 电子科学学刊, 1992, 14(1): 50 – 60.
- [22] HU T C, MOERDER K. 超图中的多端流 [C] // 图论应用新进展, 北京: 中国电子学会, 1989: 235 – 245.
- [23] ROSENBERG A. A hypergraph model for fault-tolerant VLSI processor arrays [J]. IEEE Transaction, Computer, 1985, C – 34(6): 578 – 584.
- [24] 陈廷槐. 超图的连通性和容错多总线系统的设计 [J]. 中国科学 A, 1990(11): 1309 – 1319.
- [25] 高则年. 具有最佳连通性的超图和容错总线系统的设计 [J]. 计算机学报, 1990(11): 878 – 881.
- [26] KAMIDOI Y, WAKABAYASHI S, YOSHIDA N. A fast heuristic algorithm for hypergraph bisection [J]. IEEE ISCAS, 1991(2): 1160 – 1163.
- [27] 郝忠孝. 一种基于超图的最小覆盖集求法 [M]. 计算机研究与发展, 1990(10): 58 – 64.
- [28] WAI YIN MOK, EMBLEY D W. Generating compact redundancy-free XML documents from conceptual-model hypergraphs [J]. IEEE Transactions on Knowledge & Data Engineering, 2006(18): 1082 – 1096.
- [29] KARYPIS G. Multilevel hypergraph partitioning: Applications in VLSI domain [J]. IEEE Transactions on Very Large Scale Integration Systems, 1999 (7): 69 – 70.
- [30] BRETTO A, GILLIBERT L. Hypergraph – based image

- representation[J]. Lecture Notes in Computer Science, 2005, 3435(135): 1 - 11.
- [31] ALAIN B, HOCINE C, DRISS A. Hypergraph Imaging an overview[J]. Pattern Recognition, 2002(35): 651 - 658.
- [32] CHASTEL S, COLANTONI P. Displaying image neighborhood hypergraphs line-graphs[J]. Lecture Notes in Computer Science, 2002(2301): 124 - 135.
- [33] SARKAL S, KUMAR N. Hypergraph models for cellular mobile communications systems[J]. IEEE Transactions on Vehicular Technology, 1998(47): 460 - 471.
- [34] HAN E-H (Sam), GEORGE K. Research Issue on Data Mining and Knowledge Discovery[M]. Navi Mumbai: Bioinfo Publications, 1997.
- [35] OZDAL M, AYKANAT C. Hypergraph models and algorithms for date-pattern-based clustering[J]. Data mining and Knowledge Discovery, 2004(9): 29 - 57.
- [36] ELENA V K, VLADIMIR A S. Application of hypergraph theory in chemistry[J]. Discrete Mathematics, 2001, 235: 365 - 383.
- [37] SEGOVIA-JUAREZ J L, COLOMBANO S, KIRSCHNER D. Identifying DNA splice sites using hypernetworks with artificial molecular evolution[J]. BioSystems, 2007, 87: 117 - 124.
- [38] MAO C, YOKOMORI T. DNA hypernetworks for information storage and retrieval[J]. DNA12, LNCS, 2006, 4287: 298 - 307.
- [39] KIM J K, ZHANG B T. Evolving hypernetworks for pattern classification[C]//IEEE Congress on Evolutionary Computation (CEC), Singapore: 2007: 1856 - 1862.
- [40] BECKMAN M J. Economic models of knowledge networks[C]//BATTEN D, CASTI J. Networks Action. Berlin: Springer-Verlag, 1995: 159 - 174.
- [41] MAGUITMAN A G. Intelligent support for knowledge capture and construction[D]. Indianapolis: Indiana University, USA, 2004.
- [42] BARABASI A L. Virtual round table on ten leading questions for network research[J]. European Physical Journal B, 2004, 38(2): 143 - 145.

作者介绍

王众托 中国工程院院士, 系统工程与管理工程专家. 1951年毕业于清华大学电机系. 曾任大连理工大学管理学院院长、中国系统工程学会副理事长. 现任大连理工大学教授、知识科学与技术研究中心主任. 20世纪50年代至60年代, 从事自动化专业建设、自动控制理论与计算机应用方面的学科引进和研究, 长期从事教材的编写与组织, 做过许多艰苦的开拓工作. 曾研究开发过我国自主研发的石油管卷管机的自动控制系统, 以及我国第一台电渣重熔的交流控制系统. 70年代后期开始从事系统工程专业与学位建设, 是我国系统工程研究生教育与学科创建人之一. 在学科创建初期就开展网络计划技术的新方法(决策关键路线法)与应用、基于虚拟装置的炼油企业生产计划与调度系统分析、元决策理论与应用、智能型交互式集成化决策支持系统等研究, 取得显著经济效益. 曾在维也纳国际应用系统分析研究所(IIASA)任研究员, 主持国际合作项目, 最早开发了用于宏观经济发展研究的智能型交互式集成化决策支持系统, 开我国决策支持系统研究之先河. 曾获国家级奖励两项, 部委级奖励6项. 撰写出版14种教材与专著, 翻译9种经典著作和教材. 在国内外发表140余篇学术论文与科学报告. 现从事所倡导的知识系统工程研究与应用.