

树结构信息网络在经济学中的应用

冯时 (2019012342, fengs19@mails.tsinghua.com)

摘要：树结构是一种非常基本却有着很强实用价值的结构，其连通性以及没有回路的特征保证了在使用树结构信息网络传递信息时，信息不会丢失且很难产生阻塞。在第一部分中，本文将介绍树结构信息网络的基本性质；第二部分中本文将介绍利用树结构设计的决策树模型在经济学中的应用以及树结构信息网络在算法经济学和博弈论中的应用；在第三部分中，本文将简要讲述树结构信息网络在处理复杂因果关系时的不足之处，并给出一个基于超图的改进模型。

树结构信息网络的简单介绍

世间万物存在即有其道理，树也不例外，是树的结构决定了它能屹立百年而不倒。具体来说，每一棵树都有一个唯一的根，养分从树根传输到树的枝干，再从树的枝干传输到树更小的枝干，最后传输到树的每一片树叶。树的结构保证了养分能通过枝干传输到树的每一个角落。同时也保证了养分的传输是单向的，减小了产生阻塞或浪费的风险。

我们利用图论语言来对树结构进行抽象化并严格地阐述树的性质。在图论中，树结构是最基本的结构之一，也是最先被研究的结构之一。1857 年，英国数学家 Arthur Cayley 在研究理论化学时第一次提出了树的概念，其定义如下^①：

树是一种无向图，其任意两个顶点之间存在唯一一条路径

首先，根据树的定义，任意两点之间存在一条唯一路径，因此树是连通的。这也就意味着从树的根出发，信息可以传递到整棵树的每一个角落。第二条性质是树没有回路。反之若有回路，取回路上两点，则这两点之间有两条不同的路径，与树的定义矛盾。这也就意味着，大多数情况下，信息只有一种方式进行传递。正是由于这两条性质，树结构信息网络受到了人们的重视。如图 1、图 2 所示。

^① Diestel and Reinhard, *Graph Theory (3rd ed.)*, Berlin, New York: Springer-Verlag, 2015.

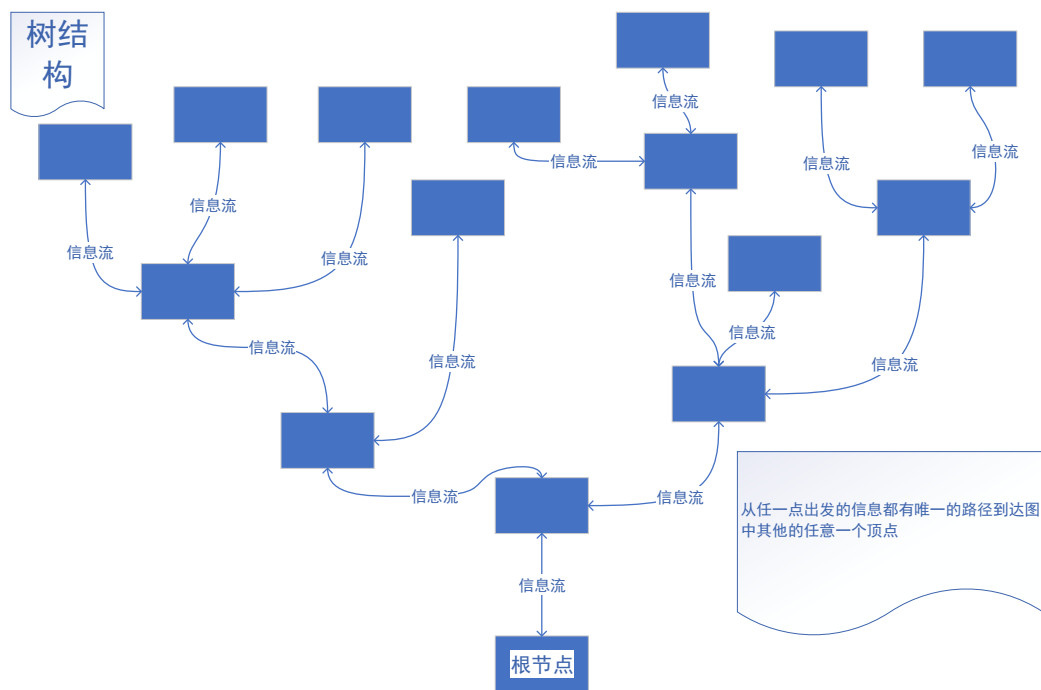


图1 树结构信息网络信息传输示意

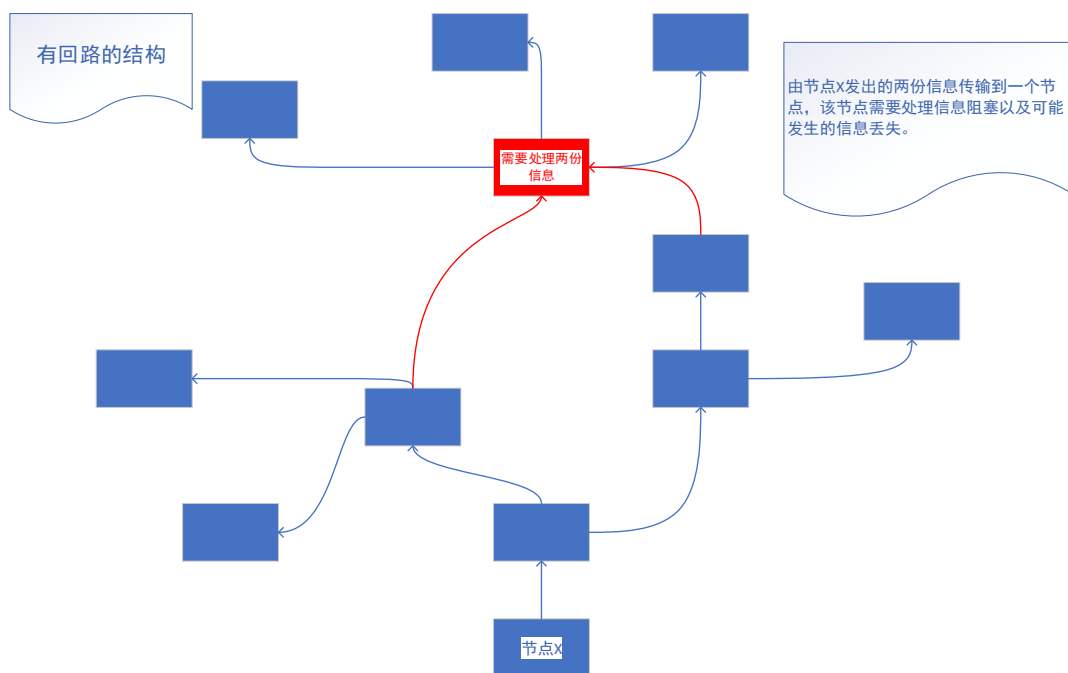


图2 有回路的结构可能产生信息阻塞及丢失

如果把力也看作是一种信息，树结构信息网络则能带来良好的力学性质：不仅可以传递从上到下的力，而且由于传递路径是唯一的，受力分析也不会过于复杂。

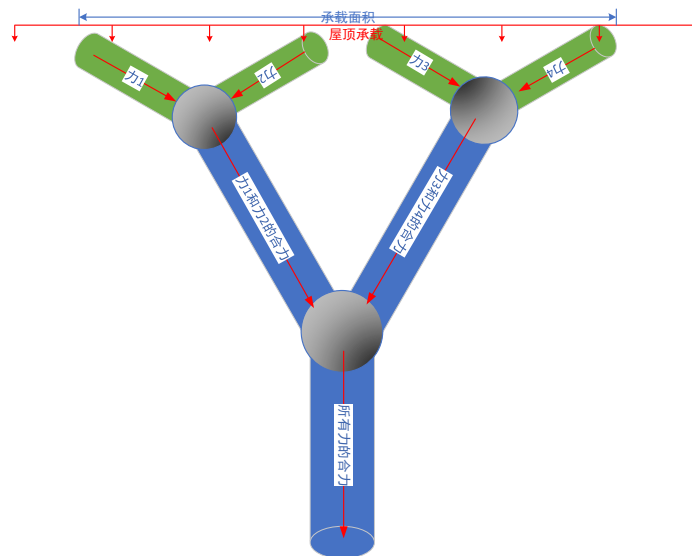


图3 树结构的力学性质（力的传递）

如图3所示，屋顶下方四个受力点所受的力1、力2、力3和力4最终都会被传递到底部的根节点。这样相比传统的柱式结构，承载的面积就明显增大了。这在建筑学中有着极为广泛的应用，比如应用在斯图加特机场 T3 候机楼上的树结构承载柱^②。进一步，如果对树结构进行分层，我们就可以更好地利用树的传递性来表示一些信息，节点与节点之间信息的传递可以上升为层次之间信息的传递，如图4所示。

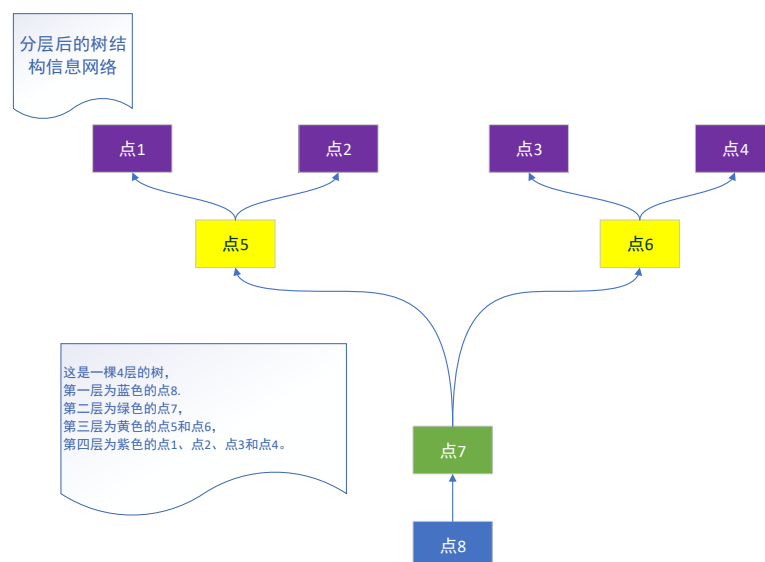


图4 树结构信息网络的分层

^② 参见陈俊，张其林，谢步瀛：《树状柱在大跨度空间结构中的应用》，《钢结构》，2010年第25卷第3期，第1页。

用这样一棵四层的树来表达信息的传输过程有着很好的直观性，同时可以很好的表达出节点之间的内在关联。比如因特网应用程序就可以用一棵四层的树来表示，如图 5 所示^③。

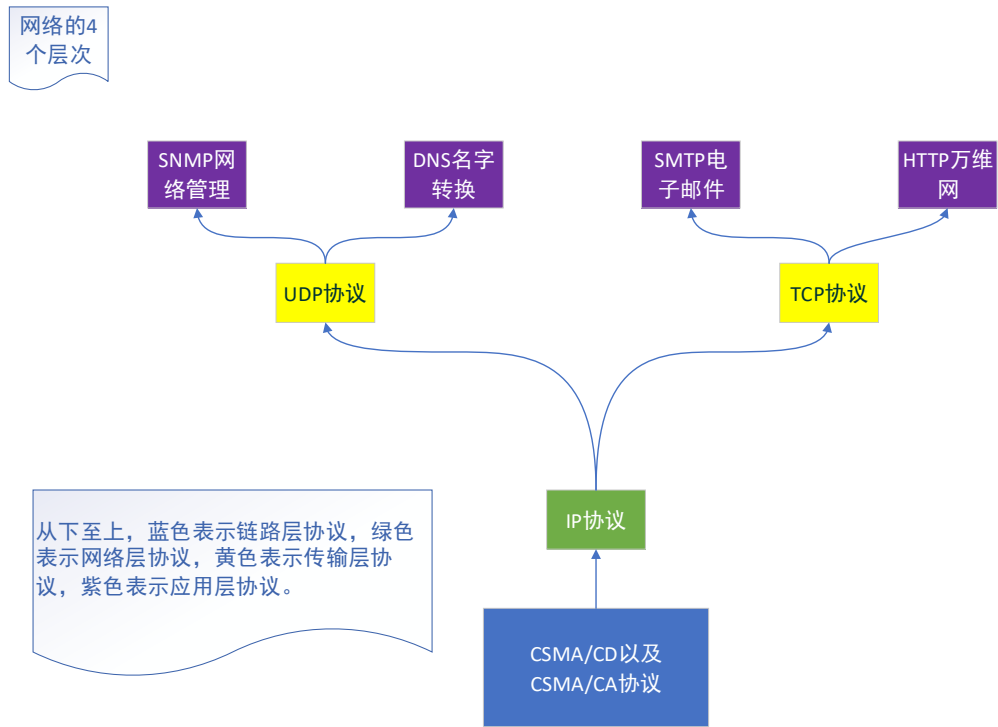


图 5 因特网应用程序的树结构信息网络表示

最后，我们总结一下树结构信息网络的性质。首先从任何一个节点出发，信息可以传输到整个网络的每个角落并且到每个节点的路径都是唯一的，这带来了树结构信息网络信息不易丢失且传输过程较为简洁的优点。另外树结构信息网络可以进行分层，使得信息的传递更为直观。

树结构信息网络在经济学中的应用

利用树结构信息网络可以构造出机器学习中一个很重要的模型——决策树模型。这是一种模拟人的思维过程的模型^④，在经济学中有很多应用，比如在证券市场上的证券选择问题中的应用。文[6]中作者就利用决策树模型给出了在非负约束条件下证券选择问题的一个解

^③ 参见 J. Glenn Brookshear Dennis Brylow, 《计算机科学概论》，刘艺等译，北京：人民邮电出版社，2017 年，第 132-135 页。

^④ 参见周志华：《机器学习》，北京：清华大学出版社，2016 年。

决方案。与此同时，在进行供应链的金融风险评估时，决策树模型也能发挥很大的作用，如图 6 所示^⑤。

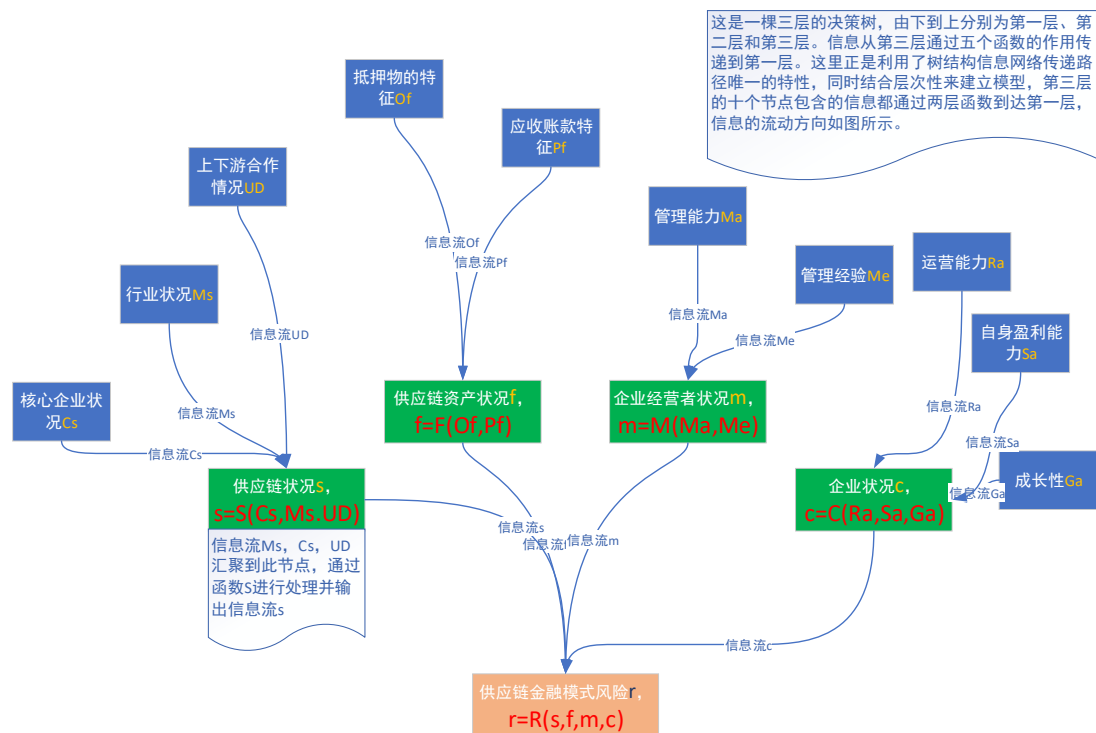


图 6 供应链金融风险评估决策树

其中 $s=S(Cs, Ms, UD)$ ， $f=F(Of, Pf)$ ， $m=M(Ma, Me)$ ， $c=C(Ra, Sa, Ga)$ 以及 $r=R(s, f, m, c)$ 。这是一棵三层的决策树，我们需要将大量的数据代入到决策树中，利用拟合的方法逆向推出函数 S、F、M、C 和 R。这样一来我们只需要拥有第三层的数据就可以通过这几个函数推导出一个理论上的风险值。决策树利用树结构信息网络信息传递路径唯一的特点汇总信息，同时建立了良好的层次性，从而使得问题变得直观且易于处理。决策树模型在经济学中的更多应用可以参见文[8]和文[9]，其中包含了一些有关信用评级的问题。

除了决策树模型，在另一个经济学领域算法经济学中，树结构信息网络则常常用来描述一些现实问题，比如发表在《Science》上的经典论文《The diffusion of microfinance》^⑥就描述了一个关于信息传递效率的问题。文章将印度农村发布通知的过程抽象为在简单图中寻找合适的节点发出信息的操作。从这个节点发出的信息经过传播，最终会得到一个树结构信息

^⑤ 参见王琪：《基于决策树的供应链金融模式信用风险评估》，《新金融》，2010 年第 3 期，第 38-41 页。

^⑥ 参见 Banerjee Abhijit, Chandrasekhar Arun G, Duflo Esther and Jackson Matthew O, "The diffusion of microfinance", *Science*, vol. 6144, no. 341. (July 2013)。

网络，如图 7 所示。

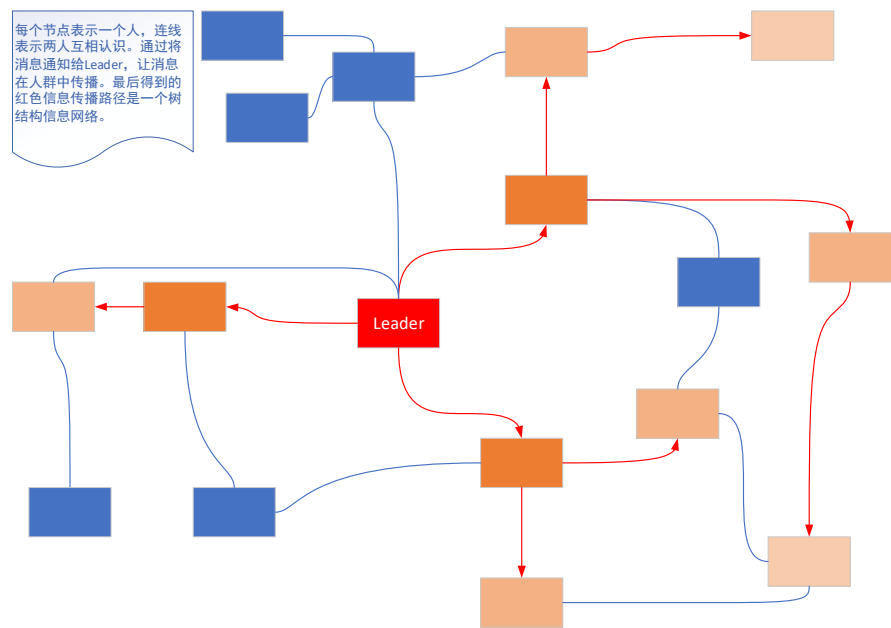


图 7 信息传递模型

如何选择合适的 Leader 使得信息传递范围更广，就是文章最后解决的问题。由此可见树结构信息网络是对现实问题进行建模的，因为在没有必要考虑一个人重复获取通知的情况（一个消息不必通知给一个人两遍所以没有回路），树结构信息传递路径唯一的特征大大简化了对问题的分析与处理（这样就只需要观察得到的树结构信息网络的范围了）。

在博弈论中，也常常利用树结构信息网络建立博弈树模型，利用贝叶斯纳什均衡解决实际问题。树结构的层次性使得我们可以直观地表示一个博弈问题。而又由于树结构信息网络中，信息传递的路径是唯一的，所以我们可以利用回溯法对博弈问题进行逆推分析。无论是决策树，现实中的信息传递问题，还是博弈树，其结构本质上都是树结构信息网络，而且都是树结构没有回路的特点以及良好的层次性，才带来了这些模型的优势。

树结构信息网络的不足以及改进

树结构信息网络实际上也有很多不足，最值得关注的一点就是图中不能有圈，于是解决很多问题的时候就产生了阻碍。比如有关社交媒体的管理问题就是一个很难用树结构信息网络描述的问题，这个问题的因果关系图的结构中出现了圈。具体来讲，政府很难了解机器学习算法的内在机制，导致很难对社交媒体进行合理的管控。这也就意味着供给侧，需求侧和算法制造商甚至可能是一家企业，而企业内部的交易合作就会形成圈，使得有效地监管变得

困难。如图 8 所示。

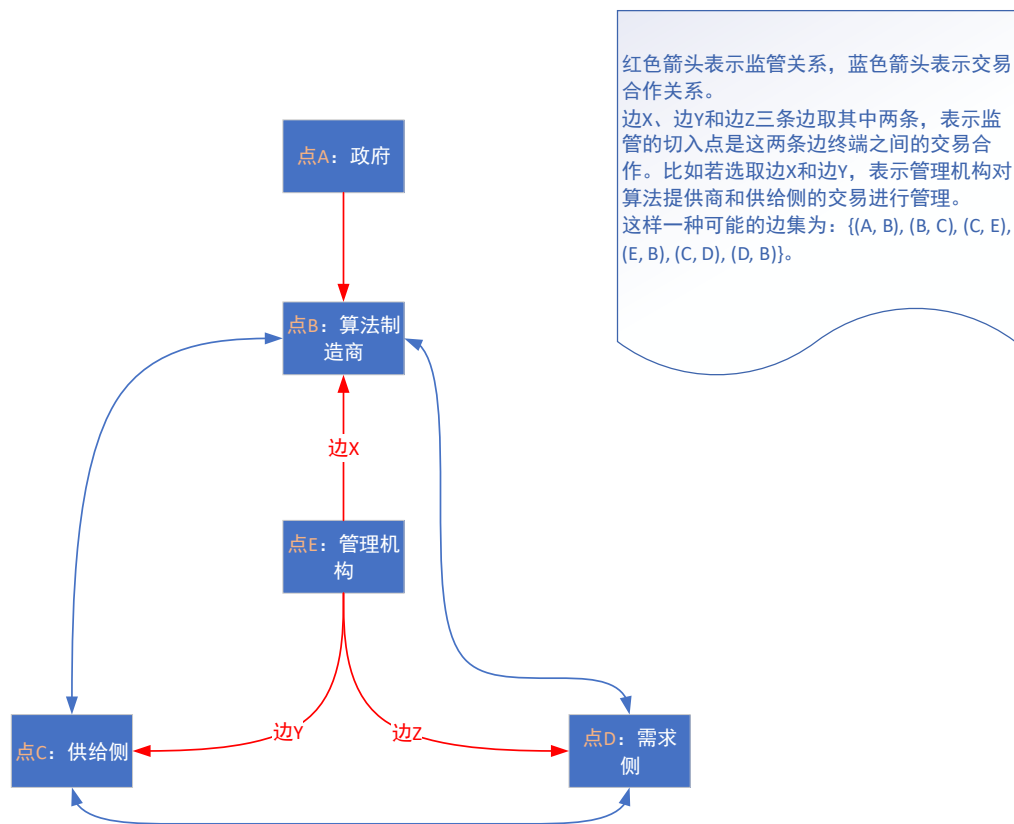


图 8 社交媒体管控模型的简单图表示

如果不能选择合理的切入点处理社交媒体与政府的关系，将会产生不可估量的舆论影响。这个时候树结构信息网络不再适用，但我们仍然希望得到的模型中信息的传递路径是唯一的。我们可以尝试对模型进行改进。具体来说，我们可以将树结构信息网络推广到超图，构造无圈超图信息网络。在普通的树结构信息网络中，我们记连接 X、Y 的边为(X,Y)。而在超图中，我们可以有连接任意数量的点的边，比如(A,B,C,D)。信息仍然只能通过边进行传递，比如我们想把信息从 A 传输到 C，那么我们就需要利用边(A,B,C,D)。为了实现类似树结构信息网络没有回路的性质，我们定义无圈超图中同样不能有回路，比如图 9 中的结构。

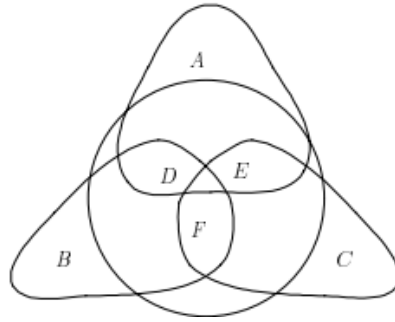


图9 无圈超图^⑦

无圈超图明确的定义可以参见文[12]，有了无圈超图之后，对社交媒体的管控问题可以进行如图10所示的理解。

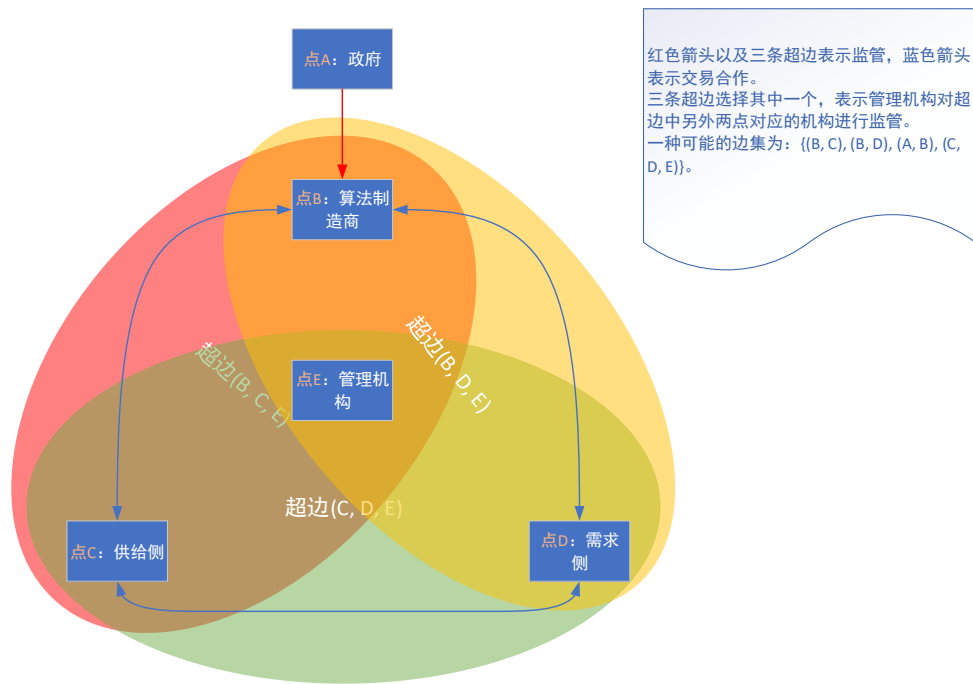


图10 社交媒体管控模型的无圈超图表示

这样一来，就解决了我们在利用机器学习算法训练有回路的网络时可能产生的死循环问题，保证了信息流的传输路径是唯一的。无圈超图也还存在着许多的问题有待解决，比如在顶点数量过大时不易建立无圈超图信息网络模型。但是这不能否定无圈超图在有关信息传输的问题中的价值。总而言之，树结构信息网络的性质是有很高的利用价值的。在建立经济学模型的时候，我们可以根据实际情况，合理利用树结构信息网络解决问题。

^⑦ 图9来自 Catriel Beeri, Ronald Fagin, David Maier, and Mihalis Yannakakis, "On the Desirability of Acyclic Database Schemes", *Journal of the ACM*, vol. 30, no. 3 (July 1983), pp. 479-513.

参考文献

- [1] Diestel and Reinhard, *Graph Theory (3rd ed.)*, Berlin, New York: *Springer-Verlag*, 2015.
- [2] 陈俊, 张其林, 谢步瀛: 《树状柱在大跨度空间结构中的研究与应用》, 《钢结构》, 2010 年第 25 卷第 3 期, 第 1 页。
- [3] J. Glenn Brookshear Dennis Brylow, 《计算机科学概论》, 刘艺等译, 北京: 人民邮电出版社, 2017 年, 第 132-135 页。
- [4] 周志华: 《机器学习》, 北京: 清华大学出版社, 2016 年。
- [5] 李双双, 刘海军, 郭盼盼: 《一种新的证券选择的决策树模型》, 《应用数学进展》, 2018 年第 9 期, 第 1166-1173 页。
- [6] 唐小我, 傅庚, 曹长修: 《非负约束条件下组合证券投资决策方法研究》, 《系统工程》, 1994 年第 6 期, 第 23-29+38 页。
- [7] 王琪: 《基于决策树的供应链金融模式信用风险评估》, 《新金融》, 2010 年第 3 期, 第 38-41 页。
- [8] 石振华: 《基于决策树算法的银行信用评分》, 《赤峰学院学报(自然科学版)》, 2012 年第 1 期, 第 75-76 页。
- [9] 沈术: 《决策树算法在农村信用社农户信用评级中的应用》, 工程硕士学位论文, 湖南大学软件学院, 2013 年。
- [10] Banerjee Abhijit, Chandrasekhar Arun G, Duflo Esther and Jackson Matthew O, “The diffusion of microfinance”, *Science*, vol. 6144, no. 341. (July 2013).
- [11] 汤振宇, 陈曜: 《我国商业银行博弈分析》, 《国际金融研究》, 2001 年第 4 期, 第 23-28 页。
- [12] Catriel Beeri, Ronald Fagin, David Maier, and Mihalis Yannakakis, “On the Desirability of Acyclic Database Schemes”, *Journal of the ACM*, vol. 30, no. 3 (July 1983), pp. 479-513.