

基于博弈论的 P2P 系统分析

张 怡,刘高嵩,李章华,刘轲平

(中南大学 信息科学与工程学院,湖南 长沙 410083)

摘 要: P2P 系统近几年来在网络上有飞速的发展,然而随着用户对网络和 P2P 系统的了解,使得 P2P 系统中大量存在欺骗行为,从而导致效率下降。文中以解决 P2P 系统中的欺骗为目的,通过引入博弈论思想从宏观和微观两方面对 P2P 系统进行分析:宏观上通过建立数学模型分析 P2P 系统效率下降甚至崩溃的原因;微观上,通过合理的论证,提出一种以牙还牙机制对 P2P 系统的欺骗问题进行解决。

关键词: 对等网;博弈论;纳什均衡;以牙还牙

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2007)08-0026-03

Analysis on P2P System Based on Game Theory

ZHANG Yi, LIU Gao-song, LI Zhang-hua, LIU Ke-ping

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: P2P system develops very fast these years, in the same time people find how to cheat in the P2P system, so the efficiency of the system comes down. Uses ideas from game theory to study P2P system from macro-way and micro-way: from macro-way, build a mathematic model to find the cause of low efficiency and even collapse of some P2P system; from micro-way, use "return like for like" to help P2P system to restrict cheatings.

Key words: peer-to-peer network; game theory; Nash equilibrium; return like for like

1 博弈论的引入

为了从理论上对 P2P 系统进行研究,必须对 P2P 系统^[1]设计的一些问题和研究对象进行合理的假设和抽象,首先从简单的两个节点 A, B 发生交互谈起, A 和 B 都有两个策略——合作或欺骗可以选,且各自的选择不受对方的影响,如表 1 所示。

表 1 博弈矩阵

	B	
	合作	欺骗
A	合作	$(u-v, u-v)$
	欺骗	$(u, -v)$

结合实际,该矩阵表示 A, B 都遵守 P2P 规则需要付出的代价位 $v > 0$, 同时享受系统提供的服务可以得到的收益为 u , 所以当 A, B 交互时都采取合作, 则两人的收益都为 $u-v$, 记作表中 $(u-v, u-v)$ 。这里 P2P 系统能够存在的一个默认前提是 $u > v$, 否则这样的系统不能给参与者带来好处, 也就没有办法吸引

人们参与而失去存在的价值。博弈论^[2]中纳什均衡点是指, 处于该点对应的策略组合, 使得任意一方单独改变自己的策略都无法提高自身的收益, 可以看到在上面 A, B 两点的博弈中策略组合(欺骗, 欺骗)是唯一的纳什均衡点^[3], 因为 A, B 任意一人把自己的策略从欺骗变为合作都会使自身收益由 0 变为 $-v < 0$ 。

2 P2P 系统的宏观分析

下面结合实际情况对 P2P 系统中的参与者进行分类。

1) 合作者(n_1): 在每次和其他节点进行信息交流的时候总表现为合作。在现实中服务器、种子以及不懂欺骗技巧的使用者属于这个类型, 记作 n_1 。

2) 精明欺骗者(n_2): 在每次和其他节点进行信息交流的时候总希望通过自身对 P2P 系统具备的相对较高的认知和使用技巧不择手段提高自身收益。

3) 随机者(n_3): 采取合作还是欺骗取决于其他与 P2P 系统无关的因素。在现实生活中网络的突然故障就造成这样的节点。这里为了简化假设他们都以 p 的概率采取合作策略($0 < p < 1$), 而以 $1-p$ 的概率采取欺骗策略。

收稿日期: 2006-11-08

基金项目: 国家自然科学基金资助项目(50405034)

作者简介: 张 怡(1981-), 男, 湖南人, 硕士研究生, 研究方向为对等网络、组播; 刘高嵩, 副教授, 研究方向为软件工程、网络技术。

假设 P2P 系统初始状态为 S_0 , 第 i 轮的状态为 S_i , 合作者的数目为 $S_i(n_1)$, 欺骗者的数目为 $S_i(n_2)$, 随机者的数目为 $S_i(n_3)$ 。从第 i 轮到第 $i+1$ 轮对应增加的用户数目分别为 $S_{i \rightarrow i+1}(\Delta n_1)$, $S_{i \rightarrow i+1}(\Delta n_2)$, $S_{i \rightarrow i+1}(\Delta n_3)$ 。第 i 轮参与交互的节点总数为: $M_i = S_i(n_1) + S_i(n_2) + S_i(n_3)$, 则每个参与交互的节点在第 i 轮遇到采取合作策略的节点的概率为(忽略自身影响):

$$P_c^i = \frac{S_i(n_1) + pS_i(n_3)}{M_i} \quad (1)$$

遇到采取欺骗策略的节点的概率(忽略自身影响):

$$P_d^i = \frac{S_i(n_2) + (1-p)S_i(n_3)}{M_i} \quad (2)$$

这样可以得到从第 i 轮到第 $i+1$ 轮各类参与者的数量有以下的递推关系式:

$$\begin{cases} S_{i+1}(n_1) = S_i(n_1) + S_{i \rightarrow i+1}(\Delta n_1) \\ S_{i+1}(n_2) = S_i(n_2) + S_{i \rightarrow i+1}(\Delta n_2) \\ S_{i+1}(n_3) = S_i(n_3) + S_{i \rightarrow i+1}(\Delta n_3) \end{cases} \quad (3)$$

假设单个 n_1, n_2, n_3 型用户在第 i 轮的收益分别为 $W_i(n_1), W_i(n_2), W_i(n_3)$, 结合收益矩阵则:

$$\begin{cases} W_i(n_1) = (u-v)P_c^i - vP_d^i \\ W_i(n_2) = uP_c^i \\ W_i(n_3) = (u-v)pP_c^i + u(1-p)P_c^i - vpP_d^i \end{cases} \quad (4)$$

因为 P_c^i, P_d^i 和 p 都是大于 0 小于 1 的, 且 $0 < v < u$, 所以很容易得到 $W_i(n_1) < W_i(n_2)$, 将上面的 $W_i(n_3)$ 化简得 $W_i(n_3) = uP_c^i - pv(P_c^i + P_d^i)$, 有 $W_i(n_1) < W_i(n_3) < W_i(n_2)$

从上面的不等式可以看到在每一轮的交互中合作者的收益总数比欺骗者的收益少, 这样假设考虑 T 轮交互有以下结论:

$$\sum_{i=1}^{T} W_i(n_1) < \sum_{i=1}^{T} W_i(n_3) < \sum_{i=1}^{T} W_i(n_2) \quad (5)$$

上面的结论表示一个很不幸的结果, 如果系统中人都是追求自身利益最大化, 他们就会想办法去学习欺骗, 因为欺骗总是比合作收益大, 这样就会导致越来越多的合作者向欺骗者转化。然而这样会不会使他们真正的收益增大呢?

现在分析第 i 轮系统所有节点的总受益:

$$W_i = S_i(n_1)W_i(n_1) + S_i(n_2)W_i(n_2) + S_i(n_3)W_i(n_3) \quad (6)$$

把式(1)、(2)、(4)代入到式(6)并化简得:

$$W_i = (u-v)[S_i(n_1) + pS_i(n_3)] \quad (7)$$

从式(7)可以看到系统在第 i 轮的总收益只取决于系统中采取合作策略的节点数量。而 i 是任意轮也就是说系统中采用合作策略的用户数量越少那么系统的收益越少, 极端情况下只存在 n_2 型用户则此时的系统收益为零。

综合上面的结论可以看到, P2P 系统在不制约欺骗的情况下, 由于追求个人最大收益的驱使, 将导致 n_1 型用户向 n_2 型用户变迁, 最终的结果是导致系统的总收益下降, 系统总收益下降说明交互量在减少, 交互量减少最后导致的结果是系统吸引新的节点加入能力减弱, 甚至出现大量的用户由于受骗而离开系统, 这样最终导致系统坍塌。这正像 Napster^[4], Gnutella^[5] 等一代 P2P 系统走向消亡。

最理想情况下希望设计一个 P2P 系统, 宏观上使得下面不等式成立:

$$\sum_{i=1}^{T} W_i(n_1) > \sum_{i=1}^{T} W_i(n_3) > \sum_{i=1}^{T} W_i(n_2) \quad (8)$$

然而在一个开放、公平而且不具备强有力的惩罚机制的系统中要设计弱者战胜强者的规则是不可能的, 所以无法达到不等式(8), 唯一可以做的是尽量削弱精明欺骗者和合作者之间技术的差异, 使他们的收益非常接近。这里 P2P 系统不具备强有力的惩罚机制是因为任何通过欺骗获利的都可以通过重新加入系统而逃避它应受的惩罚。上面提到的惩罚目前基本是采用差异性服务来实现的, 有兴趣的读者可以阅读相关文献。

3 P2P 系统的微观分析

由于在宏观上很难取得突破, 下面采用微观分析, 研究各类型用户组成的 T 轮两两博弈情况, 收益参照“表 1 博弈矩阵”。

(1) 合作者 vs 精明欺骗者。在合作者和精明欺骗者之间的 T 轮博弈中, 总收益为 $T(u-v)$, 欺骗所得收益为 Tu , 合作所得收益为 $-Tv$ 。精明欺骗者收益 Tu , 合作者收益 $-Tv$ 。

(2) 合作者 vs 随机者。两者之间 T 轮博弈, 总收益为 $2Tp(u-v) + T(1-p)(u-v) = T(1+p)(u-v)$, 欺骗产生收益为 $Tu(1-p)$, 合作产生的收益为 $2Tp(u-v) - T(1-p)v$, 合作者收益为 $Tp(u-v) - Tv(1-p)$, 随机者所得收益为 $Tp(u-v) + Tu(1-p)$ 。

(3) 随机者 vs 精明欺骗者。两者之间 T 轮博弈, 总收益为 $Tp(u-v)$, 欺骗产生的收益为 Tpu , 合作产生的收益为 $-Tpv$, 随机者收益为 $-Tpv$, 精明欺骗者收益为 Tpu 。

(4) 两合作者。两合作者之间的 T 轮博弈, 由于双方一直采用合作, 所以总收益为 $2T(u-v)$, 欺骗产生的收益为 0, 合作产生的收益为 $2T(u-v)$, 两参与者收益都为 $T(u-v)$ 。

(5) 两精明欺骗者。 T 轮都采用欺骗, 所以总收益为 0, 欺骗收益为 0, 合作收益为 0, 两人参与者收益都为 0。

(6) 两随机者。 T 轮都是 p 概率合作, 则有总收益为 $2Tp^2(u-v) + Tp(1-p)(u-v)$, 欺骗产生收益为 $Tp(1-p)u$, 合作产生的收益 $2Tp^2(u-v) - Tp(1-p)v$ 。

从以上(1)~(6)的微观分析可以看出, 不含欺骗时系统效率最高, 也最公平, 具体情况参考(4); 一旦存在欺骗与合作, 采取合作态度坚定的节点收益明显小于采取欺骗态度的, 参考(2), (3); 同时也可以看出总收益只与采取合作策略的次数有关, 极端情况可以看出(5)两个过分精明的人在一起做出不精明的事。

下面讨论一种直接有效、简单易行的方法来制约欺骗, 它的名字叫做以牙还牙, 通俗的理解是人家对我做什么, 我就对人家做什么。为了说明博弈怎么进行, 需要更精确的表述如下, 以牙还牙这个策略在开始一阶段博弈采取合作, 以后则模仿对手在前一阶段的行为。

引入以牙还牙机制的 P2P 系统, 在编写代码时, 每个节点自动写入以牙还牙的策略不需要使用者具备任何欺骗技巧。这样组成系统的是大量的以牙还牙节点, 同时也不可避免有少量系统安排的合作者, 由于网络故障引发的随机者和深入研究该系统的精明欺骗者, 下面再次从微观上对他们的 T 轮博弈行为进行分析。

(7) 以牙还牙 vs 精明欺骗。当考虑一个以牙还牙用户和一个企图欺骗的精明用户进行 T 轮的博弈, 当精明用户一直选欺骗, 则结果只能欺骗一次, 总收益为 $u-v$, 欺骗收益为 u , 合作收益为 0, 以牙还牙用户收益为 $-v$, 精明欺骗者收益为 u 。而精明欺骗者也可以选这样的策略, 前面 $(T-1)$ 次交互选择合作, 最后一次选择欺骗。这样两个人总收益为 $(2T-1)(u-v)$, 欺骗收益为 u , 合作的收益为 $(2T-2)u - (2T-1)v$, 以牙还牙用户收益为 $(T-1)u - Tv$, 精明欺骗者收益 $Tu - (T-1)v$ 。精明欺骗者可选择两个策略比较, 第二个策略比第一个策略的收益要高 $(T-1)(u-v)$, 所以他将采用第二策略。

(8) 以牙还牙 vs 合作者。在以牙还牙用户和一直采取合作策略的用户, 由于双方不存在欺骗所以总收益为 $2T(u-v)$, 欺骗产生的收益为 0, 合作产生的收益为 $2T(u-v)$, 以牙还牙用户收益为 $T(u-v)$, 合作者收益为 $T(u-v)$ 。

(9) 以牙还牙 vs 随机者。在以牙还牙用户和采用 p 概率合作的随机者进行 T 轮博弈, 总收益为 $(2Tp+1)(u-v)$, 以牙还牙者产生的收益为 $Tp(u-v) - v$, 而随机者产生的收益为 $Tp(u-v) + u$ 。

(10) 两以牙还牙者。由于在未被先欺骗的情况下以牙还牙一直采取合作, 所以 T 轮总收益为 $2T(u-v)$, 欺骗产生的收益为 0, 合作产生的收益为 $2T(u-v)$, 每个以牙还牙节点收益均为 $T(u-v)$ 。

以以牙还牙的思想去设计每个节点, 其结果是大量的以牙还牙节点取代之前的系统中的合作者, 导致的结果可以通过对比两系统主要存在的博弈(1)和(7), (2)和(9)以及(4)和(10)。结论是在基本不影响其它组合的博弈情况下, 有效地迫使原来的精明欺骗者放弃以前采取的 T 轮一直欺骗, 而转为在其参与的 T 轮期间, 采取 $(T-1)$ 轮合作而只在最后一轮采用欺骗。

以以牙还牙的思想去设计每个节点, 其结果是大量的以牙还牙节点取代之前的系统中的合作者, 导致的结果可以通过对比两系统主要存在的博弈(1)和(7), (2)和(9)以及(4)和(10)。结论是在基本不影响其它组合的博弈情况下, 有效地迫使原来的精明欺骗者放弃以前采取的 T 轮一直欺骗, 而转为在其参与的 T 轮期间, 采取 $(T-1)$ 轮合作而只在最后一轮采用欺骗。

4 结束语

在一个开放、公平而且不具备强有力的惩罚机制的系统中要设计弱者战胜强者的规则是不可能的, 所以尽量削弱恶意者和合作者之间技术的差异, 使他们的收益非常接近, 而以牙还牙, 在博弈论中的特点就是在博弈中收益和对手接近, 这样就达到设计系统的目的。由于实际中并没有实现类似的系统, 因此会在后期的工作中做出更深入的研究。

参考文献:

- [1] Li Gong. Peer-to-Peer Networks in Action[J]. IEEE Internet Computing, 2002(5):40-42.
- [2] 施锦铨. 策略与博弈[M]. 上海: 上海财经大学出版社, 2005:31-41.
- [3] 王则柯. 新编博弈论平话[M]. 北京: 中信出版社, 2005.
- [4] Haveson J. Napster[EB/OL]. <http://www.napster.com>, 1998.
- [5] 王先兵, 张荣, 胡建光. Gnutella 协议的研究[J]. 计算机工程, 2001, 27(11):56-58.

All-Terminal Network Reliability[EB/OL]. 1999-08-04. www.eng.auburn.edu/aesmith/postscript/bound.pdf.

- [7] 王凌, 郑大钟. 多目标优化的一类模拟退火算法[J]. 计算机工程与应用, 2002, 38(8):4-5.

(上接第 25 页)

pute the reliability of planar cube free networks[J]. IEEE Trans Reliability, 1990, 39(12):557-563.

- [6] Konak A, Smith A E. An Improved General Upper Bound for