

## 基于博弈理论的移动自组网激励机制研究

闻英友<sup>1,2</sup>, 赵博<sup>1</sup>, 赵宏<sup>2</sup>

(1.东北大学 医学影像计算教育部重点实验室, 辽宁 沈阳 110819; 2.东软集团研究院, 辽宁 沈阳 110719)

**摘要:** 针对移动自组网中节点在报文转发过程中的表现出的自私行为, 利用博弈理论, 从静态和动态 2 个方面对其进行了完整的建模与分析。首先, 提出了一种严厉针锋相对策略, 并建立了一个无限重复报文转发博弈模型, 求得了激励一致性条件。然后, 利用演化博弈理论对节点由自私向协作转变的动态过程进行了分析, 并证明了严厉针锋相对策略的演化稳定性。仿真结果表明, 即使在自私节点比率为 1 的条件下, 只要合理选择惩罚参数, 均可有效激励自私节点的协作转发行为, 整体网络性能最多可提升 80%。

**关键词:** 移动自组网; 重复博弈; 演化博弈; 激励机制; 自私性

中图分类号: TP302

文献标识码: A

文章编号: 1000-436X(2014)04-0044-09

## Study on game-based incentive mechanism of mobile ad hoc network

WEN Ying-you<sup>1,2</sup>, ZHAO Bo<sup>1</sup>, ZHAO Hong<sup>2</sup>

(1.Ministry of Education Key Laboratory of Medical Image Computing, Northeastern University, Shenyang 110819, China;

2.Neusoft Research, Shenyang 110719, China)

**Abstract:** According to the selfishness of rational mobile ad hoc network nodes showed during the packet forwarding, the selfish behavior statically and dynamically based on game theory were analyzed and modeled. A stern tit for tat strategy(STFT) was proposed to motivate node cooperation, and an infinite repeated game model was established to analyze the node behavior. Then, an incentive-compatible condition was obtained analytically. The dynamic process of selfish node turning to cooperate using the evolutionary game theory was studied, and the evolutionary stability of STFT was proved. Simulation results show that, even if the ratio of selfish nodes is one, by setting punishment parameters reasonably, the overall network performance can be improved 80% at most.

**Key words:** mobile ad hoc network; repeated game; evolutionary game; incentive mechanism; selfishness

### 1 引言

移动自组网是由一组带有无线收发装置的移动节点组成的无须固定基础设施支持的临时性通信网络, 节点之间的通信通过无线信道、中继节点的多跳转来完成, 网络的可用性直接依赖于节点无偿与他人协作的意愿及其协作程度, 因此节点间的协作至关重要。目前的移动自组网是基于节点合作这一基本假设的, 然而这种假设在实际的网络环境中并不一定成立, 特别是当网络节点属于多个不

同组织的时候。网络节点由于受到自身处理能力、存储空间、电池容量等各种资源的限制, 其行为呈现出一定的理性化趋势: 为了追求自身利益最大化, 节点在使用网络资源的同时, 拒绝耗费自身有限的能量为他人提供服务, 这势必会影响到网络正常的路由和数据转发功能<sup>[1]</sup>。与恶意节点不同, 自私节点并不是主动、直接破坏网络的正常运行, 因为主动、直接破坏将会消耗大量的能量, 但节点自私性的影响却是不可忽视的, 研究表明, 即便存在着小部分的自私节点(10%~40%), 也将导致网络吞

收稿日期: 2013-03-19; 修回日期: 2013-07-08

基金项目: 国家自然科学基金资助项目(60903159, 61173153); 沈阳市科技计划基金资助项目(1091176-1-00); 中央高校基本科研业务费基金资助项目(N110818001, N100218001)

**Foundation Items:** The National Natural Science Foundation of China (60903159, 61173153); The Science and Technology Plan of Shenyang (1091176-1-00); The Special Fund from the Central Collegiate Basic Scientific Research Bursary (N110818001, N100218001)

吐性能的显著下降(16%~32%)<sup>[2]</sup>。因此,如何有效地激励节点的协作行为,从而保障自组网的可用性及整体网络性能,是当前移动自组网研究领域面临的重大挑战之一。

针对上述问题,在假设节点理性的基础上,提出了一种有效的惩罚策略,并利用重复博弈模型对网络节点的协作过程进行了分析,得出了激励一致性条件,在此条件下,节点将迫于惩戒策略的威慑而自愿采取合作行为。利用演化博弈理论,进一步对惩罚策略的稳定性进行了分析。

## 2 相关工作

目前,基于博弈论的移动自组网激励机制主要包括如下。Srinivasan<sup>[3]</sup>提出了一种针对不同能量等级来调整节点转发率的协作模型,从节点与网络博弈的角度分析协作发生的可能性。作者对各能级的归一化转发率(NAR, normalized acceptance rate)进行了求解,并提出了一种收敛于NAR的转发策略GTFT(generous tit-for-tat),然而,为计算NAR,必须获知所有节点的能级及其作为源或中继节点的概率分布,作者并没有明确指出如何获得这些信息;另一方面,能量等级的划分也忽略了个体的差异性和机制的公平性。Wei Yu<sup>[4]</sup>等人利用博弈论框架对自组网中的合作激励与安全问题进行了分析,作者首先建立了2个节点间的分组转发博弈模型,利用Pareto最优、子博弈完美等最优化标准对模型的纳什均衡进行了精炼,给出了最优防欺诈分组转发策略,然后将该策略扩展到多个节点的博弈场景中,提出了基于信誉的防欺诈、抗攻击的合作激励策略,不仅可有效激励节点的自私行为,同时也应对潜在的攻击行为。Komathy<sup>[5]</sup>等人利用非合作博弈理论对邻居节点的自私性进行研究,采用马科夫链分析了2个节点的分组转发博弈,提出了一种强制合作策略BNS(best neighbor strategy),并证明了该策略是演化稳定的,能够抵御自私策略的入侵。Urpi<sup>[6]</sup>等人将节点的能量等级作为参与人的类型特征,基于贝叶斯博弈对节点间的合作转发行为进行了建模,给出了相关均衡条件,并形式化地证明了:一般情况下,节点转发的分组数量不会多于其自身发送的分组数量。Jaramillo<sup>[7]</sup>提出了一种在路由发现阶段建立的合作转发博弈模型,并通过邻居发现协议来发现参与博弈的节点数,网络中的节点通过混合策略中的概率转发策略,实现了混合策略纳什均衡。

在国内研究方面,文献[8~10]综述了无线网络中因自私节点而带来的一些关键问题,特别对含有自私节点的无线环境中基于非合作博弈理论的路由机制进行了分析和研究。陆音等<sup>[11]</sup>针对自组网络节点的预期收益及其协作交互过程建立了一个重复博弈模型,改进了文献[12]中提出的孤立惩罚机制,给出了激励一致性条件,在此条件下,节点将迫于惩戒机制威慑而自愿采取合作策略,同时分析了节点对将来利益重视程度、机制参数和作弊检测效率对协作效果的影响。武汉大学的王博等<sup>[13]</sup>同样采用惩罚机制促进节点间的合作,设计了一种通用惩罚机制对节点的不合作行为进行惩罚,从而抵消了节点从不合作行为中的获益,并在此基础上建立了基于惩罚机制的激励合作转发模型,得出了激励节点之间合作的积极性条件,但是没有给出惩罚机制具体的实现方法。由于惩罚是由邻居节点实施的,节点的移动性会导致惩罚失效,因此上述基于本地监测的惩罚机制在动态适应性方面有待进一步加强。黄蕾等<sup>[14]</sup>基于邻居节点中继和生成的路由请求分组之间的统计关系,提出了一种适用于按需路由协议寻路阶段的自私行为检测和惩罚机制,并利用博弈论对激励合作机制的有效性进行了分析。

## 3 重复报文转发博弈模型

移动自组网节点在长期使用网络时,或无法预知何时退出网络时,其行为特征将不再仅由历史收益或眼前的既得利益所决定,而更取决于它对未来利益的期望,因此节点会表现得更加耐心,更具有合作的意愿。理性节点如果预见到其不合作行为将不可避免地招致惩罚并导致未来收益降低,那么它必将选择合作行为,因此可针对节点的耐心程度来设计灵活的惩罚机制,以降低自私节点未来的收益,震慑其作弊的企图,从而促成其合作。

依照重复博弈理论,将相邻节点之间的报文转发抽象成一个多次博弈的过程,将节点的耐心程度视为继续参与下一次博弈的概率,以此建立重复报文转发博弈模型,对节点的预期收益进行分析,求解其纳什均衡,最终得出激励一致性条件。

### 3.1 单阶段报文转发博弈

在没有采取任何惩罚策略的条件下,对相邻节点间的报文转发过程进行建模,首先给出博弈的假设条件。

1) 整个移动自组网  $G(V,E)$  由  $N$  个理性节点构

成,  $G$  为连通图,  $V, E$  则分别为  $G$  的节点及链路集合。

2) 节点间的通信链路是双向的, 即当且仅当节点  $u, v$  均处于彼此传输范围内时, 它们之间的通信链路  $(u, v) \in E$ 。

3) 由于相邻节点直接通信时, 不涉及到报文转发, 因此假定网络中任意两节点的通信至少经过 1 个以上的中继节点。

4) 整个系统时间由一系列离散的时隙  $t$  构成, 在任一时隙内, 单阶段报文转发博弈都会发生, 每个节点均至少有 1 个报文待发送, 且在同一时隙内的路由状态不会发生变化, 确保每一报文均能抵达目标节点。

5) 所有节点转发单个报文消耗相同的能量  $e$ , 而接收和处理报文的能耗相对较小, 可忽略不计。

根据博弈假设条件, 可定义节点间的单阶段报文转发博弈的如下。

**定义 1** 单阶段报文转发博弈的基本式为  $G = \{I, S, \mu\}$ , 其中参与人集合  $I = \{i, j\}$ ,  $i$  与  $j$  为邻居节点; 节点的策略空间为  $S = \{S_i, S_j\}$ , 效用函数为  $\mu = \{\mu_i, \mu_j\}$ 。

理性节点在报文转发过程中, 如按照协议规范进行转发, 称节点采取协作(cooperation)策略; 不进行转发则称节点采取不协作(non-cooperation)策略。因此节点的策略集合为  $S_i = S_j = \{1, 0\}$ , 其中 1 表示协作策略  $C$ , 0 表示不协作策略  $NC$ 。不妨假设当邻居节点合作时, 节点成功发送 1 个报文时的收益为  $b$ , 接收报文的收益为 0, 则可将时隙  $t$  中节点  $i$  的效用函数定义为

$$\mu_i^t = s_j b - s_i e n_i$$

类似地, 节点  $j$  的效用函数为

$$\mu_j^t = s_i b - s_j e n_j$$

其中,  $s_i \in S_i, s_j \in S_j, n_i$  和  $n_j$  分别为节点  $i$  和节点  $j$  在时隙  $t$  内转发的报文数量。结合节点的博弈策略和效用函数, 可以得出一个二维的收益矩阵, 如表 1 所示。

**表 1** 节点单阶段博弈收益矩阵

博弈节点( $i, j$ )	$C$	$NC$
$C$	$b - e n_i, b - e n_j$	$-e n_i, b$
$NC$	$b, -e n_j$	$0, 0$

当  $b > e n_k, k = i, j$  时, 不论邻居节点合作与否, 节点选择  $NC$  策略的收益最大, 因此, 策略  $NC$  为单阶段报文转发博弈的优越策略, 所有理性节点都

将采取  $NC$  策略, 此时博弈达到纳什均衡, 网络中将不存在任何的合作行为, 且所有节点的收益均为 0, 这就形成了移动自组网中的“囚徒困境”。

### 3.2 严厉针锋相对惩罚策略

导致上述协作困境的根本原因在于节点的作弊行为不会遭致相应的后续惩罚, 如果存在某种惩罚策略, 使得节点作弊获得的短期利益无法弥补因惩罚而带来的未来利益损失, 那么理性节点必将选取合作策略。

在重复博弈中, 常见的惩罚策略包括: 触发策略(trigger strategy)、有限惩罚策略(limited punishment strategy)和针锋相对策略(tit for tat strategy)。Robert Axelrod 证明了在无限重复博弈中, 参与人最好的应对策略是针锋相对策略<sup>[15]</sup>。针锋相对策略(TFT)可简述为: 在博弈双方策略可观察的条件下, 某参与者在第一个博弈阶段( $t=1$ )选择协作; 在  $t > 1$  的阶段, 选择对手在  $t-1$  阶段的策略。可以看出, TFT 策略的惩罚是单期的(只能持续一个阶段), 惩罚力度单一, 无法对自私节点形成足够的震慑。针对此问题, 提出了一种新的基于本地监测的惩罚机制: 严厉针锋相对策略(stern tit for tat strategy)。

严厉针锋相对策略, 简称 STFT, 可描述为: 在博弈双方策略可观察的条件下, 节点在第一个博弈阶段( $t=1$ )选择协作; 在  $t > 1$  的博弈阶段, 如果节点  $i$  在  $t-1$  阶段选择不协作, 则从  $t$  阶段开始所有邻居节点对节点  $i$  实施  $T$  个阶段的严厉惩罚。在惩罚期内, 所有以  $i$  为源或目标的报文将被拒绝转发, 同时节点  $i$  必须为其余节点提供无偿转发服务; 若节点  $i$  在惩罚内, 不接受惩罚而再次选择不协作, 则惩罚将变为无限期。在惩罚期结束后, 节点  $i$  的不协作行为将被遗忘, 其他节点回到合作状态。可见, 当  $T=1$  时, STFT 策略退化成为 TFT 策略。

### 3.3 重复报文转发博弈

在典型的囚徒困境特征的博弈模型中, 由于参与人的自私性和机会主义倾向, 对于一次性博弈的策略组合, {不协作, 不协作}是各方的最优策略选择, 即选择尽可能大地去享用他人的服务和资源, 同时尽可能拒绝为他人提供服务, 以达到自身利益最大化。为了打破这种困境, 引入了严厉针锋相对策略, 使得节点的作弊行为将受到后续惩罚, 从而促使节点选择不同的策略均衡, 同时, 考虑到节点无法预知博弈何时终止, 相邻节点间的多次报文转发过程不再是一系列相互独立的单阶段

报文转发博弈，而是扩展成了一个无限重复报文转发博弈。

结合上面给出的单阶段博弈模型，利用无限重复博弈理论对节点间的报文转发行为进行建模，具体如下。

**定义 2** 设  $G$  为阶段博弈，移动自组网中节点之间的长期报文转发交互过程是阶段博弈  $G$  的不断重复，并且在每次阶段博弈开始前，所有以往博弈的历史都被观察到，那么这样构成的动态博弈就称之为无限重复报文转发博弈，记作  $G(\infty, \delta)$ ， $\delta$  为贴现因子，且  $0 < \delta < 1$ 。各博弈方在  $G(\infty, \delta)$  中的收益等于各阶段收益的贴现值之和，即节点  $i$  在整个重复博弈中的预期收益可表述为

$$U_i = \mu_i(s_i^1, s_j^1) + \delta \mu_i(s_i^2, s_j^2) + \dots = \sum_{t=1}^{\infty} \delta^{t-1} \mu_i(s_i^t, s_j^t) \quad (1)$$

其中， $\mu_i(s_i^t, s_j^t)$  为节点  $i$  在时隙  $t$  中的单阶段博弈的收益，贴现因子  $\delta$  可被视为对节点协作的耐心程度： $\delta$  越大，节点越耐心，越重视长期利益；反之，节点越注重眼前既得利益。在无限重复博弈中， $\delta$  也可被解释为博弈在下一时隙继续进行的概率，一般地，临时构建的网络  $\delta$  相对较小，而运行稳定的网络  $\delta$  则相对较大。

为了便于分析，假设在稳定的路由状态下，节点  $i$  在每个时隙中所需转发报文数量的平均值为  $n$ 。考察重复报文转发博弈过程可以看出，理性节点如果在一次作弊后，不接受惩罚而选择持续作弊，则其预期收益为

$$U_i^{NC} = \mu_i(NC, C) + \delta \sum_{t=0}^{\infty} \delta^t \mu_i(NC, NC) = \mu_i(NC, C) \quad (2)$$

如果理性节点接受惩罚，那么在惩罚期结束后，节点将再次面临相同的决策场景，即重复转发博弈的每个子博弈阶段，恰好都是原博弈本身，因此，理性节点仍将再次选择不协作策略，在这种情况下，其预期收益为

$$\begin{aligned} U_i^{NC} &= \mu_i(NC, C) + \delta \sum_{t=0}^{T-1} \delta^t \mu_i(C, NC) + \\ &\quad \delta^{T+1} (\mu_i(NC, C) + \delta \sum_{t=0}^{T-1} \delta^t \mu_i(C, NC)) + \dots \\ &= \sum_{k=0}^{\infty} \delta^k \delta^{T+1} (\mu_i(NC, C) + \delta \sum_{t=0}^{T-1} \delta^t \mu_i(C, NC)) \quad (3) \end{aligned}$$

为了激励自私节点的协作行为，必须保证其在

采取持续协作策略时的预期收益不低于采取不协作策略的预期收益。根据式(1)可计算节点在持续协作时的预期收益为

$$U_i^C = \sum_{t=1}^{\infty} \delta^{t-1} \mu_i(C, C) \quad (4)$$

因此，为了保证  $U_i^C \geq U_i^{NC}$ ，即

$$\begin{cases} \sum_{t=1}^{\infty} \delta^{t-1} \mu_i(C, C) \geq \sum_{k=0}^{\infty} \delta^k \delta^{T+1} (\mu_i(NC, C) + \\ \delta \sum_{t=0}^{T-1} \delta^t \mu_i(C, NC)) \\ \sum_{t=1}^{\infty} \delta^{t-1} \mu_i(C, C) \geq \mu_i(NC, C) \end{cases} \quad (5)$$

在理性节点做出决策之前，都会对将来的预期收益进行评估，如果式(5)成立，则节点必将选择协作策略。作为系统设计者，可针对不同的  $\delta$  设置相应的  $T$  值，使式(5)条件成立，从而促进节点的协作。从式(5)中可以看出，惩罚机制是靠引入负收益来实现的，惩罚的力度取决于  $T$ 。由表 1 可知， $\mu_i(C, C) = \mu_i(NC, C) + \mu_i(C, NC)$ ，将式(5)化简可得

$$\frac{\delta - \delta^{T+1}}{1 - \delta} \geq \frac{|\mu_i(C, NC)|}{\mu_i(NC, C)} \quad (6)$$

其中， $T \geq 1$ ，式(6)即为无限重复报文转发博弈的激励一致性条件。当节点采用严厉针锋相对策略时，式(6)保证了每个单阶段子博弈的{协作，协作}策略组合为子博弈的完美纳什均衡，从而整个重复博弈也将处于纳什均衡状态，任何理性节点均无法形成足够的动机来偏离协作策略。

### 3.4 重复报文转发博弈参数分析

由式(6)可见，其右端是由节点所处网络环境以及节点自身状况所决定的，取值依赖于  $b$ 、 $e$  和  $n$ 。对于系统设计者而言，为了能够依据激励一致性条件选取适当的惩罚措施，必须首先对右式的取值进行预测和评估。一般来说， $n$  可设定为全网节点平均转发报文数量的最大值；而参数  $b$  和  $e$ ，对于不同节点，它们的取值可能是不同的，是由节点自身状态所决定的，比如 QoS 等级、当前剩余能量等，但是其比值可根据节点所处的具体应用环境进行设定。

式(6)的左端完全由贴现因子  $\delta$  和惩罚参数  $T$  构成，其中， $\delta$  由网络和应用的性质决定，而参数  $T$  代表惩罚力度的大小，在不同的应用环境中，可通过灵活设置  $T$  的取值来实现不同程度的协作激励。

令函数  $f(\delta, T) = \frac{\delta - \delta^{T+1}}{1 - \delta}$ ，由于  $0 < \delta < 1$ ，函数  $f(\delta, T)$  是  $T$  的增函数，如图 1 所示。

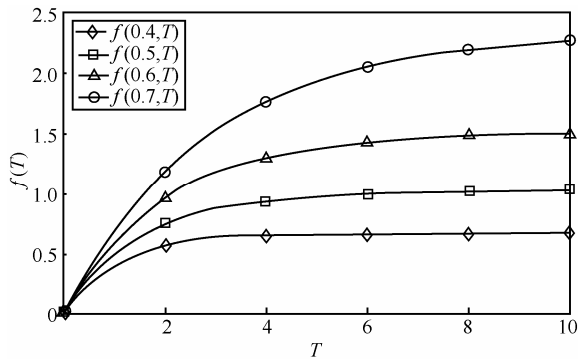


图 1  $f(\delta, T)$  函数

根据函数  $f(\delta, T)$  的性质，可对激励一致性条件做进一步分析。

**命题 1**  $T$  的增加有助于提升节点间的协作。

**证明** 对  $f$  求  $T$  的偏导可得： $\frac{\partial f}{\partial T} = \frac{\delta}{\delta - 1} \ln \delta \cdot \delta^T$ ，由于  $0 < \delta < 1$ ，所以  $\frac{\partial f}{\partial T} > 0$ ，在给定  $\delta$  和

$\frac{|\mu_i(C, NC)|}{\mu_i(NC, C)}$  的情况下， $T$  的增大将导致  $f$  变大，式 (6) 条件更容易得到满足。证毕。

**命题 1** 说明， $T$  越大，则节点预期收益中因不协作所遭受的损失就越大，那么理想节点便越倾向于选择协作策略，这与直觉相一致。

**命题 2** 当  $\delta < \delta_{\min} = \frac{|\mu_i(C, NC)|}{|\mu_i(C, NC)| + \mu_i(NC, C)}$  时，

无论  $T$  如何取值，激励一致性条件永不满足。

**证明** 对于给定的  $\delta$ ，由于  $\frac{\partial f}{\partial T} > 0$ ，因此函数

$f(\delta, T)$  的极大值为  $\lim_{T \rightarrow \infty} f(\delta, T) = \lim_{T \rightarrow \infty} \frac{\delta - \delta^{T+1}}{1 - \delta} = \frac{\delta}{1 - \delta}$ ，

当  $\frac{\delta}{1 - \delta} < \frac{|\mu_i(C, NC)|}{\mu_i(NC, C)}$  时，式 (6) 永远不成立。证毕。

**命题 2** 表明，当节点极端注重眼前既得利益时，由于交易环境过于恶劣，惩戒策略将完全失效，节点将不受任何约束的自由选择策略。然而，考虑到  $\frac{|\mu_i(C, NC)|}{\mu_i(NC, C)} = \frac{e}{b} \cdot n$ ，即使在这种最恶劣的环境下，还是可以通过限制节点最大转发的报文数量  $n$  来进行调整，以使式 (6) 成立，从而保证惩罚策略的有效性。

**命题 3** 当  $\delta > \delta_{\max} = \frac{|\mu_i(C, NC)|}{\mu_i(NC, C)}$  时，对于任何

$T$ ，激励一致条件均满足。

**证明** 对于给定的  $\delta$ ，由于  $\frac{\partial f}{\partial T} > 0$ ，且  $T \geq 1$ ，

因此函数  $f(\delta, T)$  的极小值为  $f(\delta, 1) = \delta$ ，当  $\delta > \frac{|\mu_i(C, NC)|}{\mu_i(NC, C)}$  时，对于任何  $T$  均满足式 (6) 条件。证毕。

**命题 3** 给出了一种较为长期稳定的交易场景，即理性节点由于不知何时会退出网络，它对未来的预期利益会表现得更加重视，这时，仅需 1 次惩罚便足以震慑其作弊行为。

上述关于  $\delta$  界限以及  $T$  取值设定的结论，为分析重复转发博弈模型中的节点协作程度提供了基础依据，使系统设计者能够对惩罚策略的适用性及其有效性做出迅速判断，进而做出相应的调整。

#### 4 基于演化博弈的节点协作性分析

传统的重复博弈理论得到均衡是静态的均衡，即博弈最终所能到达的稳定状态，它无法描述与人达成均衡的动态演化过程，因此，本节将从整个群体的角度出发，对节点间的协作关系进行分析，建立演化博弈模型，阐述节点由自私向协作的演化过程，并分析 STFT 策略的稳定性。

##### 4.1 演化博弈模型建立与求解

移动自组网中节点发送报文的同时，也需要转发报文。本文假定参与网络运行的实体节点可分为两类：协作节点和自私节点。双方的策略集合均为  $\{STFT, NC\}$ ，系统通过“物竞天择，适者生存”的原则自发演化，各节点根据其他节点的策略选择，通过学习、调整和模仿过程，以寻求最优策略。节点间的协作报文转发过程是在一个具有不确定性和有限理性的空间进行，同时它们之间的策略选择又是相互影响的。博弈双方的收益矩阵如表 2 所示。

策略	STFT	NC
STFT	$U(S, S), U(S, S)$	$U(S, N), U(N, S)$
NC	$U(N, S), U(S, N)$	0, 0

由于演化博弈也是基于时间序列的，因此根据收益的贴现计算方法可得

$$U(S, S) = \sum_{t=1}^{\infty} \delta^{t-1} \mu(C, C) = \frac{\mu(C, C)}{1 - \delta}$$

$$U(S, N) = \sum_{k=0}^{\infty} \delta^k \delta^{T+1} (\mu(C, NC) + \delta \sum_{t=0}^{T-1} \delta^t \mu(NC, C))$$

$$= \frac{(1-\delta)\mu(C, NC) + (\delta - \delta^{T+1})\mu(NC, C)}{(1-\delta)(1-\delta^{T+1})}$$

$$U(N, S) = \sum_{k=0}^{\infty} \delta^k \delta^{T+1} (\mu(NC, C) + \delta \sum_{t=0}^{T-1} \delta^t \mu(C, NC))$$

$$= \frac{(1-\delta)\mu(NC, C) + (\delta - \delta^{T+1})\mu(C, NC)}{(1-\delta)(1-\delta^{T+1})}$$

假设节点中协作节点所占的比例为  $\rho$ ，则自私节点所占比例为  $1-\rho$ 。由收益矩阵可计算。

协作节点的期望收益为

$$\psi_s = \rho U(S, S) + (1-\rho)U(S, N)$$

自私节点的期望收益为

$$\psi_N = \rho U(N, S) + (1-\rho)0$$

由协作节点和自私节点的期望收益，可计算整体的平均收益为

$$\bar{\psi} = \rho \psi_s + (1-\rho)\psi_N$$

$$= \rho^2 U(S, S) + \rho(1-\rho)U(S, N) + \rho(1-\rho)U(N, S)$$

可构造协作节点的复制动态方程

$$F(\rho) = \rho(\psi_s - \bar{\psi}) = \rho(1-\rho)U(N, S) \quad (7)$$

式(7)描述了演化系统中的协作群体动态。令  $F(\rho)=0$ ，若  $U(N, S) \neq 0$ ，可求得  $\rho_1=0, \rho_2=1$ ；若  $U(N, S)=0$ ，则对于任意  $\rho \in [0, 1]$  都是上述方程式的解。

#### 4.2 STFT 策略稳定性分析

由  $F(\rho)=0$  求出的均衡解并不一定都是稳定的，但根据微分方程的稳定性原理可知：稳定状态处的函数导数必须小于 0，即  $F'(\rho^*) < 0$ ， $\rho^*$  表示稳定状态，即  $\rho^*$  为演化稳定策略。下面分别对求得的均衡解进行稳定性分析。

1) 当  $U(N, S)=0$ ，任意  $\rho \in [0, 1]$  都是均衡解，且恒有  $F'(\rho)=0$ ，这说明外界的任何扰动都不能使状态发生波动，因此  $\rho \in [0, 1]$  都是动态稳定均衡解，此时复制动态方程的动态演化相位如图 2 所示。

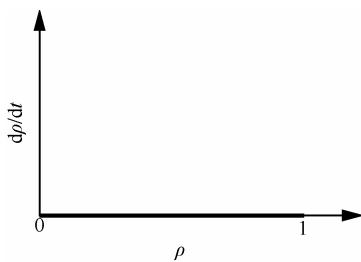


图2  $\rho \in [0, 1]$  为稳定状态

由  $U(N, S)=0$  可得， $\frac{\delta - \delta^{T+1}}{1-\delta} = \frac{|\mu(C, NC)|}{\mu(NC, C)}$ ，对

比激励一致性条件可知，此时节点选择自私行为的收益与选择合作行为的收益相等，因此节点没有足够动机改变自身行为，系统处于稳定状态。

2) 当  $U(N, S) \neq 0$ ， $\rho_1=0$  和  $\rho_2=1$  为均衡解。由  $F'(\rho) = (1-2\rho) \cdot U(N, S)$  可计算

$$F'(\rho_1=0) = \frac{(1-\delta)\mu(NC, C) + (\delta - \delta^{T+1})\mu(C, NC)}{(1-\delta)(1-\delta^{T+1})}$$

$$F'(\rho_2=1) = -\frac{(1-\delta)\mu(NC, C) + (\delta - \delta^{T+1})\mu(C, NC)}{(1-\delta)(1-\delta^{T+1})}$$

(i) 当  $\frac{\delta - \delta^{T+1}}{1-\delta} < \frac{|\mu(C, NC)|}{\mu(NC, C)}$  时，可知  $F'(\rho_1 =$

$0) < 0$ ， $F'(\rho_2=1) > 0$ ，因此  $\rho_1=0$  是唯一的演化稳定策略。这说明在开始博弈的时候，协作节点都选择协作策略，但由于惩罚机制的惩罚力度不够，导致节点选择自私策略的收益比选择协作策略时的收益要大，因此，经过无数次博弈之后，所有的节点都将选择不协作策略，网络中将不存在任何的协作行为。 $\rho_1=0$  时的复制动态方程的动态演化相位如图 3 所示。

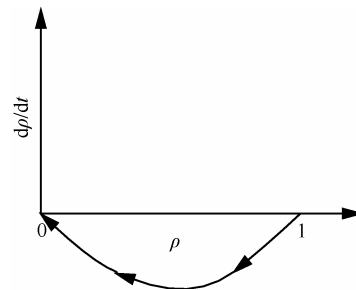


图3  $\rho = \rho_1$  为稳定状态点

(ii) 当  $\frac{\delta - \delta^{T+1}}{1-\delta} > \frac{|\mu(C, NC)|}{\mu(NC, C)}$  时，可知  $F'(\rho_1 =$

$0) > 0$ ， $F'(\rho_2=1) < 0$ ，因此  $\rho_2=1$  是唯一的演化稳定策略。这表明协作节点选择 STFT 策略时，如果自私节点选择不协作策略，则它将受到协作节点的惩罚，使得其预期收益小于选择协作行为时的收益，因此，经过无数次的博弈之后，所有节点都将选择 STFT 策略，节点间将保持稳定的合作关系。 $\rho_2=1$  时的复制动态方程的动态演化相图如图 4 所示。

综合以上分析，可以得出如下结论：当  $\frac{\delta - \delta^{T+1}}{1-\delta} > \frac{|\mu(C, NC)|}{\mu(NC, C)}$  时，STFT 策略是演化稳定策

略,可抵御不协作策略 NC 的入侵。根据演化稳定策略的性质可知,不论初期自私节点所占比例大小,经过多次的博弈后,STFT 策略都可使其演化成协作节点,从而使网络系统达到演化稳定状态,在这种状态下,系统对微小的扰动具有很好的稳健性,也就是说,即使某些节点由于偶然的错误出现偏离,复制动态仍然会使其回到稳定状态。

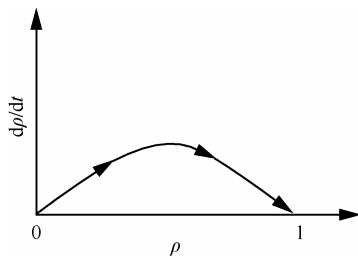


图 4  $\rho=\rho_2$  为稳定状态点

## 5 仿真测试与分析

### 5.1 仿真方法与仿真设置

为验证激励一致性条件及严厉针锋相对惩罚策略的正确性和有效性,本文在 NS-2<sup>[16]</sup>仿真平台上进行了仿真实验,采取的仿真方法如下。

1) 网络拓扑设置:利用 NS-2 中的 GT-ITM Topology Generator 生成  $N$  个节点组成的随机拓扑, $N$  默认为 20,网络直径为 8 跳。

2) 随机流量设置:节点间的业务流量由 cbngen 工具生成,随机挑选源节点以及非相邻的目标节点,生成 20 对 CBR 业务流,每个 CBR 报文大小为 512 B,每秒发送 4 个报文。任意一个节点在一轮协作时隙中均可发送 12 个报文。

3) 节点行为监测与惩罚实施:采用 Catch<sup>[17]</sup>协议对节点的转发行为进行监控,一旦发现节点的作弊行为,则通过 Catch 的 ANV 机制实现邻居节点对作弊节点的协同惩罚。

4) 路由协议采用 DSR<sup>[18]</sup>协议,源节点在报文头部中显式的设定源路由信息,以便 Catch 中的 Watchdog 进行信息统计与行为监控。

5) 节点报文转发决策:各节点依据当前转发报文数量  $n_i$ ,遵照式(6)对可行策略进行评估,进而选择转发行为(转发/丢弃)。

6) 报文完成发送与转发后,节点统计并存储报文转发信息,节点的下一阶段策略选取随本地监控结果而改变,一轮交互时隙仿真结束。

仿真的物理环境设定如下:仿真区域的大小为

800 m×800 m,节点移动速度在 1~10 m/s 之间均匀分布,节点持续运动,中间不作停留,每个节点的无线传输距离为 250 m,带宽为 2 Mbit/s,信道和无线模型为 Two-Ray Ground,链路层采用 IEEE 802.11 协议。重复博弈模型中的相关参数设置如下: $b=1$ , $e=0.048\ 63$ 。

在仿真过程中,每一类实验共进行 5 次,每次实验由 50 次随机拓扑组成,针对每一拓扑随机产生 200 次 CBR 流量,即仿真时隙为 200 个,仿真结果取 5 次仿真实验的平均值。

考虑到节点作弊行为所导致的直接后果就是分组丢失率上升,报文成功发送的比率下降,网络可用性降低,因此,在下面的仿真实验中,重点考察节点自私性、耐心程度及惩罚策略对报文投递率的影响,进而验证激励机制的有效性。这里定义报文投递率(PDR, packet delivery ratio)为:所有成功到达目标的报文与全部发送的报文的比值。

### 5.2 仿真结果与分析

#### 1) 节点自私性对协作性的影响

如果缺乏惩罚策略的震慑,节点间的交互将逐渐退化为单阶段报文转发博弈,节点趋向选择不协作,自私节点将越来越多。随着自私节点比率的增加,越来越多的传输链路将被打断,产生越来越多的路由分区,从而降低了整个网络的可用性。如图 5 所示,在没有引入惩罚机制时,原始 DSR 协议的报文投递率随自私节点比率的增加而急剧下降,最终导致整个网络瘫痪。

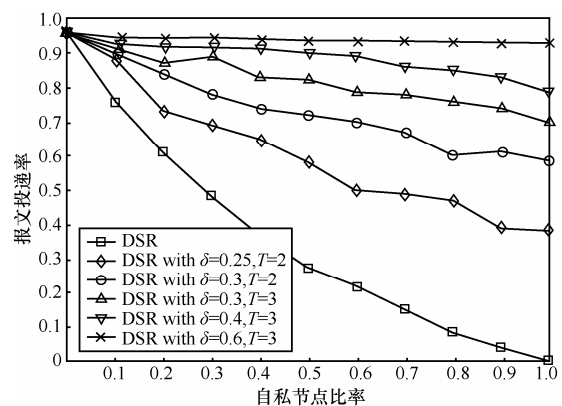


图 5 不同自私节点比率对报文投递率的影响

图 5 中也给出了惩罚机制存在时的情况,可以看出,虽然报文投递率的整体趋势都在下降,但是引入惩罚机制后,报文投递率有了明显的改善,提高幅度至少都在 10%以上。当  $\delta=0.6$  时,即便自私节点比率

高达 80%以上，只要采取合适的惩罚力度( $T=3$ )，也能确保 92%以上的报文成功抵达目的地。这说明惩罚机制确实使得原来的自私节点重新审视其转发策略，为了避免因惩罚而带来的利益损失，自私节点将不得不与其他节点合作，开始转发报文，从而提高了网络的全局协作程度，增强了网络的可用性，延长了网络的生存时间。

鉴于自私节点比率为 1 时，网络中不存在任何协作行为，此时网络环境最差，因此在后续仿真实验中，设置所有节点均为自私节点。

### 2) 节点耐心程度对协作性的影响

图 6 为  $T=2$  时的报文投递率随  $\delta$  的变化情况，从图中可以看出，投递率随  $\delta$  的增加而明显提高，整个网络的可用性随节点耐心程度的增加而提高，这与观察动态博弈中的无名氏定理相符。考察曲线  $N=20$ ，当  $b=1$ ， $e=0.04863$  时，由命题 3 可知， $\delta > 0.9177$  可保证节点均采用协作策略，图 6 证实了这一点，此时的报文投递率等于没有自私节点的情况；而当  $\delta$  接近 0 时，注意到各曲线的 PDR 并不为 0，这主要源于自私节点接受惩罚时的无偿转发服务，其贡献十分有限，最多不足 10%。

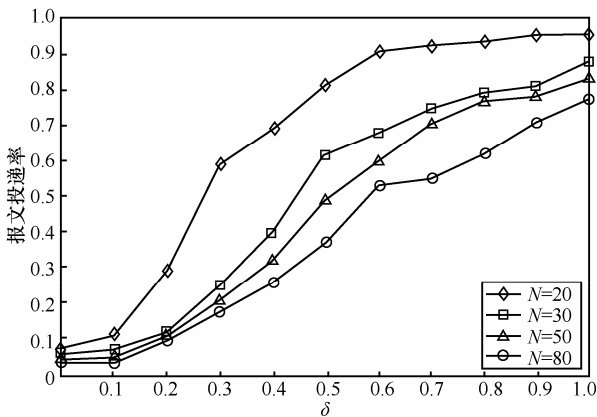


图 6 节点耐心程度对报文投递率的影响( $T=2$ )

由图 6 可见，随着  $N$  的增加，即使节点足够耐心，报文投递率仍然有所下降，当  $\delta=1$  时， $N=50$  的报文投递率与  $N=20$  的报文投递率相比也下降了 15%，此时的惩罚力度( $T=2$ )已不能有效激励自私节点合作的积极性。因为随着  $N$  的增大，每个时隙中节点需要转发的报文数量  $n$  也随之增多，而惩罚力度却相对固定，以致无法满足激励一致性条件的要求，节点趋向自私化。因此，网络中的节点数量越多，惩罚机制应越严厉。

图 7 给出了更为严厉的惩罚措施下( $T=3$ )的仿真结果，可以看出，报文投递率同样随  $\delta$  的增大而上升，

同时，随  $N$  的增大而降低。对比图 6 和图 7，网络规模相同时，惩罚力度加大，报文投递率也随之升高，而且当  $\delta$  增大时， $T=3$  对投递率的提高比  $T=2$  时更加明显，这说明从整体角度来看  $T=3$  的激励效果要好于  $T=2$  的激励效果。因此，针对不同的网络规模和不同的节点耐心程度，灵活配置合适的惩罚力度是激励节点合作的关键。

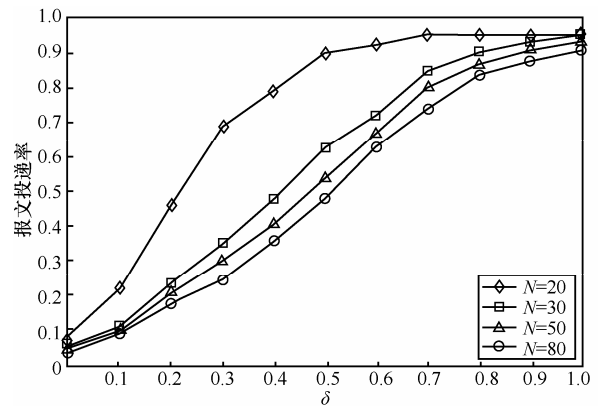


图 7 节点耐心程度对报文投递率的影响( $T=3$ )

### 3) 惩罚机制对协作性的影响

图 8 给出了惩罚力度对报文投递率的影响，同时也给出了不同节点耐心程度时的情况。从图中可以看出， $T$  的增加对节点间的协作性起到了明显的促进作用，而这一效果在节点耐心程度较低时尤为明显，这一观察与命题 1 是相符合的。另一方面，当  $T=1$  时，STFT 惩罚机制退化为 TFT 惩罚机制，可以看出，TFT 惩罚机制对节点协作性有一定的促进作用，但是当节点耐心程度较差时，其取得的效果并不理想，以  $\delta=0.3$  为例，TFT 只能将报文投递率提升到 58%，而 STFT 最多可将 PDR 提升到 80%，因此，STFT 惩罚机制更加灵活、有效，更具可扩展性。

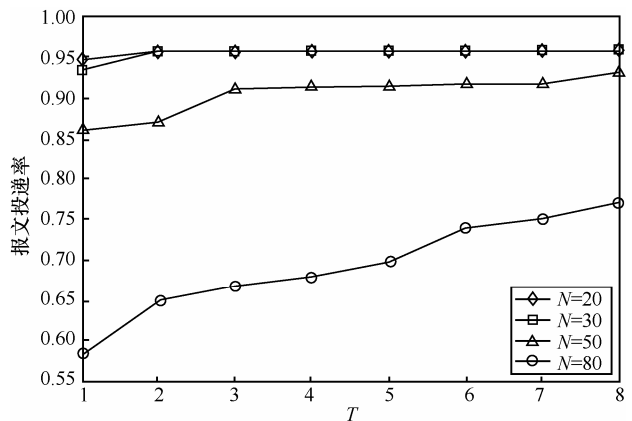


图 8 惩罚机制对报文投递率的影响

## 6 结束语

移动自组网中的节点受到自身处理能力、存储空间和电池能量等各种资源的限制,经常会表现出自私性,因此激励自私节点的合作行为成为有待解决的关键问题。本文针对节点的报文转发过程,首先建立了单阶段的报文转发博弈模型,分析了节点的自私性;提出了基于本地监测的严厉针锋相对惩罚策略,并以此为基础构建了无限重复报文转发博弈模型,求得了激励一致性条件。然后基于演化博弈理论对节点由自私向协作转变的动态过程进行了分析,并证明了严厉针锋相对策略的演化稳定性。最后通过仿真实验验证了激励机制的有效性,仿真结果表明,通过合理选择惩罚参数,可以有效激励自私节点的合作转发行为,提高网络的可用性,延长网络的生存时间,增加网络的总体收益。

### 参考文献:

- [1] URPI A, BONUCCELLI M, GIORDANO S. Modeling cooperation in mobile ad hoc networks: a formal description of selfishness[A]. Proceedings of 1st International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks[C]. Hingham, 2003. 3-5.
- [2] MARTI S, GIULI T, LAI K. Mitigating routing misbehavior in mobile ad hoc networks[A]. Proceedings of the ACM MobiCom 2000[C]. New York, 2000. 255-265.
- [3] SRINIVASAN V, NUGGEHALI P. Cooperation in wireless ad hoc networks[A]. Proceedings of the IEEE INFOCOM 2003[C]. Washington, 2003. 808-817.
- [4] YU W, LIU K J R. Game theoretic analysis of cooperation stimulation and security in autonomous mobile ad hoc networks[J]. IEEE Transactions on Mobile Computing, 2007,6(5): 507-521.
- [5] KOMATHY K, NARAYANASAMY. Best neighbor strategy to enforce cooperation among selfish nodes in wireless ad hoc networks[J]. Computer Communications, 2007, 30(18): 3721-3735.
- [6] URPI A, BONUCCELLI M A, GIORDANO A. Modeling cooperation in mobile ad hoc networks: a formal description of selfishness[J]. IEEE/ACM Workshop WiOpt, 2003, 39(1): 118-124.
- [7] JOSÉ J J, SRIKANT R. A game theory based reputation mechanism to incentivize cooperation in wireless ad hoc networks[J]. Ad Hoc Networks, 2010, 8(4): 416-429.
- [8] ZHAO L, ZHANG J, YANG K, *et al.* Using incompletely cooperative game theory in mobile ad hoc networks[A]. Proceedings of IEEE International Conference on Communications 2007[C]. Piscataway, NJ, 2007.3401-3406.
- [9] ZHAO L, ZHAN G J, ZHANG H. GDGF: game-theoretic distributed co-ordination function in WLAN[J]. Electronics Letters, 2007, 43(9): 510-511.
- [10] 汪洋, 林闯, 李泉林等. 基于非合作博弈的无线网络路由机制研究[J]. 计算机学报, 2009, 32(1): 54-68.
- WANG Y, LIN C, LI Q L. Non-cooperative game based research on routing schemes for wireless networks[J]. Chinese Journal of Computers, 2009, 32(1): 54-68.
- [11] 陆音, 石进, 谢立. 基于重复博弈的无线自组网络协作增强模型[J]. 软件学报, 2008, 19(3): 755-768.
- LU Y, SHI J, XIE L. Repeated-game modeling of cooperation enforcement in wireless ad hoc network[J]. Journal of Software, 2008, 19(3): 755-768.
- [12] BANDYOPADHYAY S. A game-theoretic analysis on the conditions of cooperation in a wireless ad hoc network[A]. Proceedings of 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks[C]. Trentino, 2005. 54-58.
- [13] 王博, 黄传河, 杨文忠等. Ad Hoc 网络中基于惩罚机制的激励合作转发模型[J]. 计算机研究与发展, 2011, 48(3): 398-406.
- WANG B, HUANG C H, YANG W Z. An incentive-cooperative forwarding model based on punishment mechanism in wireless ad hoc networks[J]. Journal of Computer Research and Development, 2011, 48(3): 398-406.
- [14] 黄蕾, 刘立祥. Ad hoc 网络寻路阶段的合作激励机制研究[J]. 计算机学报, 2008, 31(2): 262-269.
- HUANG L, LIU L X. Study on cooperation stimulation mechanism in route discovery of ad hoc networks[J]. Chinese Journal of Computers, 2008, 31(2): 262-269.
- [15] AXELROD R. The Evolution of Cooperation[M]. New York: Basic Books, 1984.
- [16] ns-2 network simulator[EB/OL]. <http://www.isi.edu/nsnam/ns/>, 2011.
- [17] MAHAJAN R, RODRIG M. Sustaining cooperation in multi-hop wireless networks[A]. Proceedings of the USENIX NSDI 2005 Symposium on Networked Systems Design & Implementation (NSDI 2005)[C]. Berkeley, 2005. 231-244.
- [18] JOHNSON D B, MALTZ D A. Dynamic Source Routing in Ad Hoc Wireless Networks[M]. Norwell: Kluwer Academic Publishers, 1996.

### 作者简介:



闻英友(1974-), 男, 辽宁沈阳人, 东北大学副教授, 主要研究方向为移动通信、网络安全。



赵博(1982-), 男, 辽宁沈阳人, 东北大学博士生, 主要研究方向为移动自组网路由与安全、P2P 路由。

赵宏(1954-), 男, 河北河间人, 东北大学教授、博士生导师, 主要研究方向为分布式多媒体信息处理、计算机网络、图像处理等。