面向网络攻击和隐私保护的多智能体系统分布式共识算法

徐明，张保俊，伍益明，应晨铎，郑宁
（杭州电子科技大学网络空间安全学院，浙江 杭州 310018）

摘 要：为解决网络攻击与信息窃听环境下的多智能体系统分布式共识问题，提出一种能有效抵御网络拒绝服务（DoS）攻击和保护节点状态隐私的平均共识算法。首先，结合网络化控制系统中 DoS 攻击的特性，构建与时间相关的周期性 DoS 攻击模型。其次，利用邻居节点间的信息交互给出一种基于差分隐私的分布式网络节点信息处理机制，并将其引入平均共识算法。再次，结合事件触发机制，提出一种适用于 DoS 攻击下无向通信网络的分布式共识算法，并分别对其收敛性和隐私保护性能进行了严格的数学分析。最后，通过数值仿真实验和硬件实验验证了所提算法的有效性。

关键词：多智能体系统；平均共识；拒绝服务攻击；隐私保护；网络安全
中图分类号：TP273
文献标志码：A
DOI: 10.11959/j.issn.1000–436x.2023049

Cyber attacks and privacy protection distributed consensus algorithm for multi-agent systems

XU Ming, ZHANG Baojun, WU Yiming, YING Chenduo, ZHENG Ning
School of Cyberspace, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: To solve the consensus problem of multi-agent systems in the environment of cyber attacks and information eavesdropping, an average consensus algorithm that can effectively resist network denial of service (DoS) attacks and protect nodes' state privacy was proposed. Firstly, combined with the characteristics of DoS attacks in networked control systems, a periodic DoS attack model related to time was constructed. Then a distributed network node information processing mechanism based on differential privacy was proposed by using the information interaction between neighbor nodes, and then it was introduced into the average consensus algorithm. Furthermore, combined with the event-triggered mechanism, a consensus algorithm method suitable for undirected communication networks under DoS attacks was proposed, and its convergence and privacy protection performance were analyzed strictly mathematically. Finally, the effectiveness of the proposed algorithm was verified by a numerical simulation experiment and hardware experiment.

Keywords: multi-agent system, average consensus, DoS attack, privacy protection, cyber security

0 引言

近 20 年来，多智能体系统共识课题一直备受研究者的青睐，并在诸多领域有着广泛的应用，如无人机编队控制[1]、传感器网络[2]、分布式能量管理[3]等。

多智能体系统共识算法是指在一个分布式网络中，每个智能体个体仅根据自身信息及收集到的

收稿日期：2022–11–12；**修回日期：**2023–02–08
通信作者：伍益明，ymwu@hdu.edu.cn
基金项目：国家自然科学基金资助项目（No.61803135, No.62073109）；浙江省公益技术应用研究基金资助项目（No.LGF21F020011）；浙江省教育厅科研基金资助项目（No.Y202146750）
Foundation Items: The National Natural Science Foundation of China (No.61803135, No.62073109), Zhejiang Provincial Public Welfare Research Project of China (No.LGF21F020011), The Scientific Research Fund of Zhejiang Provincial Education Department (No.Y202146750)

相邻个体信息作为算法的输入，随着时间的推移，最终所有个体的某一状态信息在该算法下达成一致[4]。经过国内外研究者的共同努力，目前多智能体系统共识算法已在理论分析上取得了丰硕的研究成果。

然而，在实际应用中，受限于单个智能体节点的构造、通信带宽以及通信方式，多智能体系统中的节点或通信链路均容易遭受网络攻击的威胁。因此，越来越多的专家学者针对网络攻击下的多智能体系统共识问题展开了研究[5-8]。现有针对多智能体系统的网络攻击的研究大致可以分为两类，即拒绝服务（DoS, denial of service）攻击[9-12]和欺骗攻击[13-15]（或称虚假数据注入攻击）。其中，DoS 攻击作为网络控制系统中一种易于实现且能对系统性能造成极大干扰的攻击方式，近年来得到了广泛研究。文献[9]在假设 DoS 攻击的持续时间和频率平均有界的前提下，设计了一种输出编码控制策略，利用有限的数据传输速率实现了系统指数收敛，并给出了系统 Lyapunov 稳定性的充分条件。文献[11]研究了 DoS 攻击下有无领导者的一致性控制问题，提出了一种基于事件触发的一致性控制协议，采用 2 种不同的组合度量，分别给出了领导者–跟随者和跟随者–跟随者情形下的控制器触发条件，并利用 Lyapunov 函数证明了系统的稳定性。该研究不仅降低了系统的更新频率，而且可以有效抵御 DoS 攻击的影响。

另一方面，在现有的共识算法执行过程中，节点会根据协议将自身的状态信息无保留地传输给周围邻居节点。值得注意的是，在一些特殊的应用场景中，节点并不希望将其真实的信息透露给其他节点。例如，在一个群体网络中，成员们采用共识算法来达成对某一主题的共同意见，但从隐私的观点来看，每名成员都不想透露自己的初始意见[16]。

基于上述考虑，越来越多的研究者将隐私保护概念引入多智能体系统共识算法中，即考虑隐私保护框架下的多智能体系统共识算法设计问题。在这类研究工作中，较典型的一类是基于密码学的算法。具有代表性的成果有文献[17]提出的密钥共享算法和文献[18]提出的同态加密共识算法等。尽管基于密码学的算法可以较好地解决节点状态值隐私保护问题，但需要节点具备强大的计算和通信能力，这对本身构造较简单的智能体节点而言，在实际应用上较难实现。因此严格来讲，该类算法并不适用于规模较大的多智能体系统。另一类是文献[19]提出的基于状态分解的算法，核心思想是通过将系统中每个节点的初始状态值随机拆分成 2 个子状态，一个子状态用于信息交互，另一个子状态则自始至终保留在内部，以此达到隐私保护的目的。该算法虽然可以实现精准的平均共识收敛，但对网络通信拓扑要求较苛刻，且对通信网络中的时延、丢包等干扰较敏感。

此外，还有一类研究工作是基于差分隐私[20]的算法，通过在节点信息传输中添加噪声的方式来避免自身状态信息被推断出来。文献[21]给出了一种基于差分隐私的平均共识通用框架，通过在系统更新过程中添加满足指数衰减且总和为零的高斯噪声，来保证节点信息的隐私，并实现系统最终状态均方意义上的平均共识。文献[22]提出了一种随机渐进型差分隐私方法。在该方法下，节点的初始状态值会以均方收敛到一个公共随机变量，而其期望值则是整个系统的初始状态平均值。进一步，文献[23]针对基于差分隐私的平均共识算法，研究了拉普拉斯噪声参数改变与隐私级别变化间的关系。尽管上述基于差分隐私的共识算法为智能体节点的隐私提供了良好的保护，但同时也给系统最终的收敛精度带来了无法避免的误差，这意味着系统无法精确地实现节点初始状态值的平均共识。

不难发现，现有的大多文献将多智能体系统网络攻击问题与隐私保护问题分开考虑，单独设计相应的共识算法，却少有文献将两者一同考虑来设计共识算法。事实上，在一些实际场景中，攻击者往往先通过前期窃听系统的内部信息，据此推算系统的脆弱点，进而发动更具有针对性的网络攻击策略。文献[24]率先将网络攻击与隐私保护问题结合起来，提出了一种安全共识算法，但系统的收敛精度会受隐私保护机制融入产生的影响，且考虑的网络攻击为固定形式的攻击模型。

基于上述分析，本文致力于设计一种兼具抵御网络 DoS 攻击和保护节点信息隐私的多智能体系统共识算法，从而补充现有共识算法的研究成果。本文主要的研究工作如下。

1) 根据 DoS 攻击在多智能体系统的攻击特性，刻画了与时间相关的周期性 DoS 攻击模型。

2) 针对网络中存在的信息窃听者，提出了一种新的适用于平均共识的分布式差分隐私算法，有效避免了系统中节点自身信息被信息窃听者推断出来。

3) 针对 DoS 攻击的影响，提出了一种分布式事件触发机制，以补偿攻击造成的信息丢失对差分隐私和平均共识算法造成的影响，确保系统最终实现零误差收敛。

1 预备知识

1.1 图论知识

通常，多智能体系统的通信拓扑可以由一个图 $G=(V,E,\boldsymbol{A})$ 来表示，其中，$V=\{1,2,\cdots,n\}$ 表示节点集，$E=V\times V$ 表示边集，$\boldsymbol{A}=[a_{ij}]$ 表示节点间的权重。如果节点 i 可以接收来自节点 j 的信息，那么满足 $(j,i)\in E$，且称节点 j 为节点 i 的邻居节点。本文考虑无向图情况。在一个无向图中，节点之间的信息是互通的，即 $(j,i)\in E\Rightarrow(i,j)\in E$。令 N_i 表示节点 i 的邻居节点集合，则有 $N_i=\{j\in V\mid(j,i)\in E,j\neq i\}$。

1.2 平均共识

假设网络中有 n 个节点，每个节点 i 的初始状态值为 $x_i(0)\in\mathbb{R}$。在常规共识算法中，每个节点根据与邻居节点交互得到的信息来更新自己下一时刻的状态值，其状态更新方程可表示为

$$\begin{cases}x_i(k+1)=x_i(k)+u_i(k)\\ u_i(k)=\sum\limits_{j\in N_i}a_{ij}\left(x_j(k)-x_i(k)\right)\end{cases}\tag{1}$$

即

$$x_i(k+1)=a_{ii}x_i(k)+\sum_{j\in N_i}a_{ij}x_j(k)\tag{2}$$

其中，a_{ii} 和 a_{ij} 分别表示节点自身和节点与邻居节点间的权重值。如果节点 i 和节点 j 之间存在通信边，则 $a_{ij}>0$，否则 $a_{ij}=0$。

根据已有的理论可知[25]，只要系统网络拓扑满足双随机矩阵，即 $\boldsymbol{A}\mathbf{1}=\mathbf{1}^{\mathrm{T}}\boldsymbol{A}=\mathbf{1}$，系统中节点在更新规则（式(2)）的作用下，最终网络中节点的值就会收敛于所有节点初始时刻的平均值，即平均共识收敛。这一过程可表示为

$$\lim_{k\to\infty}x_i(k)=\frac{\sum\limits_{l=1}^{n}x_l(0)}{n}=\bar{x}\tag{3}$$

1.3 差分隐私

差分隐私为多智能体系统隐私保护提供了一种量化评估方法。其核心思想是通过给原始数据添加噪声，保证整体的统计信息不受影响。这样即使窃听者获取了数据，仍无法获取个体的准确信息，即保护了个体隐私。本文给出 δ-邻接和 ε-差分隐私的定义。

定义 1　δ-邻接[26]。对于任意给定的 $\delta\in\mathbb{R}(\delta\geqslant 0)$，如果存在一个 $k\in\{1,2,\cdots,n\}$ 满足式(4)，那么称状态向量 $\boldsymbol{x},\boldsymbol{x}'$ 是 δ-邻接的。

$$|x_i-x_i'|\leqslant\begin{cases}\delta, & i=k\\ 0, & i\neq k\end{cases}\tag{4}$$

其中，$i\in\{1,2,\cdots,n\}$，$\boldsymbol{x},\boldsymbol{x}'\in\mathbb{R}^n$。

定义 2　ε-差分隐私[26]。对于 $\varepsilon>0$，如果式(5)成立，那么对于 δ-邻接状态向量的任意对 $\boldsymbol{x},\boldsymbol{x}'$ $(\boldsymbol{x},\boldsymbol{x}'\in\Omega\subseteq\mathbb{R}^n)$ 和任意集合 $\mathcal{O}\subseteq Ra(R)$，随机机制 R 是满足 ε-差分隐私的，其中，R 的域是 Ω，$Ra(R)$ 是机制 R 下的输出域。

$$P\left[R(\boldsymbol{x})\in\mathcal{O}\right]\leqslant\mathrm{e}^{\varepsilon}P\left[R(\boldsymbol{x}')\in\mathcal{O}\right]\tag{5}$$

其中，常数 ε 表示隐私级别，ε 的值越小，隐私保护的级别越高[27]。

1.4 网络攻击模型

DoS 攻击是现有网络攻击中常见的一种，本文讨论的 DoS 攻击为多智能体系统中各个节点的通信链路间信息传输失败的情况，即当图中 2 个节点的通信链路发生 DoS 攻击时，虽然节点之间存在链路，但是节点间的链路被攻击中断，不能进行信息交互，从而使通信资源失效，最终导致系统不能实现状态值收敛。从资源限制的角度考虑，攻击者通常需要一段休眠时间来为下一次行动积蓄能量。因此借鉴文献[11]，本文有如下假设。

假设 1　攻击者发动 DoS 攻击的能量是有限的。

基于假设 1，令 $T=\gamma k$ 为一个攻击周期。在每个周期 T 内[7]，假设 DoS 攻击开始时刻存在一个攻击序列 $\left\{\tilde{k}_l^{(i,j)}\right\}$，$l\in\mathbb{N}$。第 l 个 DoS 攻击持续时间间隔为 $\boldsymbol{\mathcal{A}}_l^{(i,j)}=m\tilde{k}_l^{(i,j)}$，其中 $m<\gamma$。因此，对于每个周期 T，通信受阻的时间间隔为 $\Xi_a(i,j)=\bigcup\boldsymbol{\mathcal{A}}_l^{(i,j)}$，那么节点可以通信的时间间隔为 $\Xi_s(i,j)=\dfrac{T}{\Xi_a(i,j)}$。

2 算法设计和分析

考虑用一个无向图 $G=(V,E,\boldsymbol{A})$ 来表示一个由 n 个节点组成的多智能体系统。图中每个节点代表一个智能体，每条边代表智能体节点之间的通信链

路。为保护节点自身信息的隐私，一种广泛采用的方法是在节点的真实状态值中添加一个随机噪声 $\theta_i(k)\in\mathbb{R}$。然后用添加噪声后的信息与邻居节点进行交互。添加噪声后的节点状态值可表示为

$$\tilde{x}_i(k)=x_i(k)+\theta_i(k), i\in V \tag{6}$$

将式(6)代入式(2)，可以得到

$$x_i(k+1)=a_{ii}\tilde{x}_i(k)+\sum_{j\in N_i}a_{ij}\tilde{x}_j(k)=$$
$$a_{ii}\left(x_i(k)+\theta_i(k)\right)+\sum_{j\in N_i}a_{ij}\left(x_j(k)+\theta_j(k)\right) \tag{7}$$

因此，本文的目标是设计一种基于差分隐私的平均共识算法，通过设计随机噪声 $\theta_i(k)$，包括噪声的分布和相关性以及相应的数据补偿机制，最后使 1) 节点与邻居节点在信息交互过程中可以保护自身初始状态值不被泄露；2) 在假设的网络攻击模型下，系统在式(7)更新规则下可以达成平均共识目标式(3)。

注 1 注意到，式(7)中的节点 i 自身也采用了添加噪声后的新状态值 $\tilde{x}_i(k)$ 进行更新，目的是配合后续去除系统中噪声的处理机制，确保系统最终可以达到精确的平均共识。

在给出具体算法之前，需要定义两类窃听者和 2 种相应的信息集。

第一类：局部好奇窃听者。局部好奇窃听者存在于多智能体网络内部，往往是内部网络中的某一个智能体节点。该节点遵循系统的控制协议，但是具有好奇性，会私自收集并存储每次迭代时从邻居节点接收到的信息，建立信息集 $I_i^1(k)$ 为

$$I_i^1(k)=\left\{\tilde{x}_j(0),\cdots,\tilde{x}_j(k)\right\} \tag{8}$$

第二类：全局好奇窃听者。全局好奇窃听者存在于多智能体网络外部。该节点掌握整个网络拓扑信息，同时通过窃听正常节点通信链路获得该节点全部的交互数据，建立信息集 $I_i^2(k)$ 为

$$I_i^2(k)=\left\{N_i,a_{ii},a_{ij},\tilde{x}_i(0),\tilde{x}_j(0),\cdots,\right.$$
$$\left.\tilde{x}_i(k),\tilde{x}_j(k)\,|\,j\in N_i\right\} \tag{9}$$

针对上述两类好奇窃听者，本文分别设计差分隐私算法来实现正常节点初始状态值的隐私保护，避免好奇节点对信息的泄露。

2.1 事件触发函数设计

为了抵御网络中的 DoS 攻击，本文使用事件触发机制。节点 i 将信息发送给节点 j 的接收器，节点 j 的接收器将收到的所有邻居节点的信息发送到缓存器，判断是否满足触发条件，若满足，则有事件触发，进入控制器进行迭代更新，然后将状态值发送到发送器；若不满足，则直接进入发送器，发送器将得到的节点状态值转发给邻居节点。事件触发机制示意如图 1 所示。

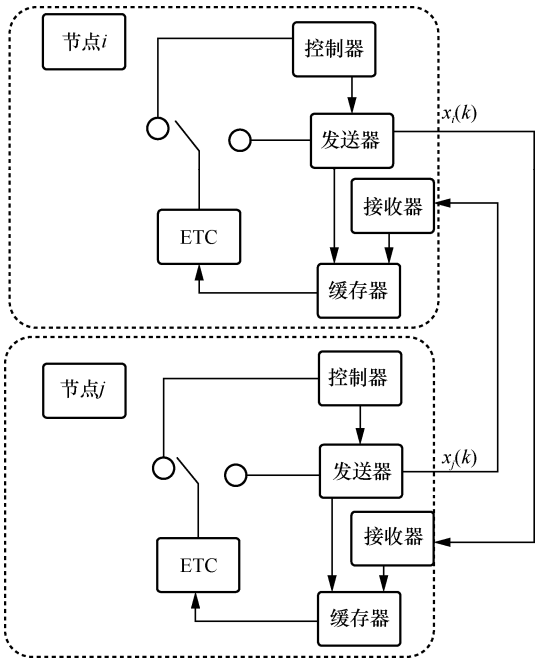

图 1 事件触发机制示意

节点 i 创建缓存空间 M_i，用来存放邻居节点传递过来的触发时刻的信息，表示为

$$M_i\left(t_k^i\right)=\left\{x_i^{j_1}\left(t_k^i\right),x_i^{j_2}\left(t_k^i\right),\cdots\right\},$$
$$k=1,2,\cdots,\ j_1,j_2\cdots\in N_i \tag{10}$$

其中，$x_i^{j_1}\left(t_k^i\right)$ 表示节点 i 的邻居节点 j_1 在触发时刻的状态值，t_k^i 表示节点 i 的触发时刻，每更新一次，则 k+1，直至达成共识。节点 i 根据自身的入度 d_i，执行触发函数

$$f\left(x_i^j,t_k^i\right)=\left|M_i\left(t_k^i\right)\right|=d_i \tag{11}$$

根据平均共识收敛条件可知，每个周期 T 内，系统中的各个节点都必须确保有一次状态更新。因此在周期 T 内，节点需要至少接收到一次每个邻居节点传递给它的信息。根据上述需求，将节点 i 在缓存空间中的触发阈值设定为邻居个数 d_i。结合触发函数式(11)，将式(7)进一步设计为

$$x_i\left(t_{k+1}^i\right)=a_{ii}\tilde{x}_i\left(t_k^i\right)+\sum_{j\in N_i}a_{ij}\tilde{x}_j\left(t_k^j\right)=$$
$$a_{ii}\left(x_i\left(t_k^i\right)+\theta_i\left(t_k^i\right)\right)+\sum_{j\in N_i}a_{ij}\left(x_j\left(t_k^j\right)+\theta_j\left(t_k^j\right)\right) \quad (12)$$

注 2　关于 Zeno 行为，即事件在有限的时间间隔内发生无数次。本文算法是基于恒定采样时间间隔的，在离散时间 k 迭代，也就是说本文算法最小时间间隔为一个步长。在最坏情况下，事件发生在每个采样时刻，即便如此，事件触发次数在有限时间间隔内仍然是有限的。因此，本文算法不会发生 Zeno 行为。

2.2　隐私保护算法设计

接下来，本文对噪声进行设计。相较于均匀噪声，拉普拉斯噪声可以实现不同程度的隐私保护，因此本文选用拉普拉斯噪声进行设计。拉普拉斯分布表示为

$$\eta_i(k)\sim \mathrm{Lap}(k,\mu,\lambda)=\frac{1}{2\lambda}\mathrm{e}^{-\frac{|k-\mu|}{\lambda}} \quad (13)$$

其中，数学期望为 μ，方差为 $2\lambda^2$。

因为局部好奇窃听者只能窃听到邻居节点发送的信息，所以在算法中添加噪声可以保护节点的真实状态值不被泄露。但对于了解整个网络拓扑和可以窃听到邻居节点状态值的全局好奇窃听者来说，仅添加噪声并不能防止节点初始状态值的泄露。为了防止好奇窃听者对隐私的泄露，本文引入了一个秘密的连续函数 $\Phi_{ij}(\cdot)$。首先，每个节点 i 选择拉普拉斯分布的随机变量 $\eta_i(0)$ 和恒定的随机序列 s_{ij}；然后，将带有噪声的状态值 $\tilde{x}_i(0)$ 和随机序列 s_{ij} 发送给邻居节点 j，每个节点计算秘密函数和随机变量，表示为

$$\tilde{\eta}_i(0)=\eta_i(0)-\sum_{j\in N_i}\left[\Phi_{ij}(s_{ij})-\Phi_{ji}(s_{ji})\right],\ \forall i\in V \quad (14)$$

最后，生成拉普拉斯分布的随机变量，并按照生成噪声的计算式生成随机噪声，表示为

$$\theta_i\left(t_k^i\right)=\begin{cases}\rho\eta_i\left(t_1^i\right)-\tilde{\eta}_i(0), k=1\\ \rho^{t_k^i}\eta_i\left(t_k^i\right)-\rho^{t_{k-1}^i}\eta_i\left(t_{k-1}^i\right), k\geqslant 2\end{cases} \quad (15)$$

每个节点将带有噪声的状态值转发给邻居节点。每个节点将收到的邻居节点的状态值放到缓存空间 M_i 中，但 M_i 中只存放与此节点时刻相同的邻居节点的状态值，若收到与邻居节点时刻相同的状态值或前一时刻、后一时刻的状态值则丢弃。然后进行触发函数的判断，若满足触发函数，则事件触发，节点进行状态更新，否则保持原来的状态值不变。基于事件触发的隐私保护平均共识算法的描述如算法 1 所示。

算法 1　基于事件触发的隐私保护平均共识算法

输入　$k=0$，$x_i(0)$，N_i，a_{ii}，a_{ij}，$\Phi_{ij}(\cdot)$，λ，ρ（$0<\rho<1$）

1) 每个节点 i 根据式(13)选择一个拉普拉斯分布的随机变量 $\eta_i(0)$，并且任意选择一个恒定的随机序列 s_{ij}；
2) 计算 $\theta_i(0)=\eta_i(0)$，$\tilde{x}_i(0)=x_i(0)+\theta_i(0)$；
3) 每个节点 i 将计算出的 $\tilde{x}_i(0)$ 和 s_{ij} 传递给邻居节点 j；
4) 节点 i 根据式(11)判断是否有事件触发；
5) 如果有事件触发，节点 i 按照式(14)计算 $\tilde{\eta}_i(0)$；
6) 节点 i 按照式(12)更新节点状态值；
7) 节点 i 根据式(13)选择一个拉普拉斯分布的随机变量 $\eta_i\left(t_k^i\right)$，$k\geqslant 1$；
8) 按照式(15)生成噪声，然后将噪声加入真实状态值中，有 $\tilde{x}_i\left(t_k^i\right)=x_i\left(t_k^i\right)+\theta_i\left(t_k^i\right)$；
9) 节点 i 将添加噪声后的状态值传递给邻居节点，每个节点接收到邻居节点的状态值后，若为自己更新时刻 t_k^i 的状态值，则接收并将其存储在 M_i 中，否则忽略；
10) 节点根据触发函数判断是否有事件触发，如果有事件触发，根据式(12)进行状态更新，返回步骤 7)进行计算；
11) 否则保持节点状态值不变，返回步骤 9)进行计算，等待节点下一次更新；
12) 输出每个节点的最终状态值 $x_i(k)$。

2.3　共识分析

定理 1　考虑 DoS 攻击下的离散时间多智能体系统式(7)，其通信网络是一个满足假设 1 的无向图，在式(11)和式(12)的作用下，系统所有节点的状态值最终可以实现精确的平均共识，即 $\lim\limits_{k\to\infty}\left(x_i(k)-\bar{x}_i(0)\right)=0$。

证明　根据文献[28]中的定理3.1可知，如果式(7)中的随机噪声满足有界、渐进收敛到 0、所有节点噪声之和为 0，就可以达到精确的平均共识。下面证明本文添加的噪声满足上述条件。

首先证明本文添加的噪声是有界的。显然，

$\theta_i(0)=\eta_i(0)$ 是有界的。因为 $\Phi_{ij}(s)$ 是连续函数，所以它的值对任意给定的 s 都是有界的，则式(14)中的 $\tilde{\eta}_i(0)$ 就是有界的。对于 $k\geqslant 1$，因为 $\eta_i\left(t_k^i\right)$ 是有界的，而式(15)中的 $\theta_i\left(t_k^i\right)$ 与其有关，所以可推出 $\theta_i\left(t_k^i\right)$ 是有界的。

接下来证明噪声收敛到 0。

$$\lim_{k\to\infty}\left|\theta_i\left(t_k^i\right)\right|\leqslant\lim_{k\to\infty}\left|\rho^{t_k^i}\eta_i\left(t_k^i\right)-\rho^{t_{k-1}^i}\eta_i\left(t_{k-1}^i\right)\right|\leqslant$$
$$\lim_{k\to\infty}\left[\rho^{t_k^i}\frac{1}{2\lambda}\mathrm{e}^{-\frac{\left|t_k^i\right|}{\lambda}}-\rho^{t_{k-1}^i}\frac{1}{2\lambda}\mathrm{e}^{-\frac{\left|t_{k-1}^i\right|}{\lambda}}\right]=0 \tag{16}$$

显然，根据式(16)可知本文添加的噪声最终收敛于 0。

最后，证明所有节点噪声之和为 0。注意到

$$\sum_{i=1}^{n}\sum_{k=0}^{\infty}\theta_i\left(t_k^i\right)=\sum_{i=1}^{n}\theta_i\left(t_0^i\right)+\sum_{i=1}^{n}\theta_i\left(t_1^i\right)+$$
$$\sum_{i=1}^{n}\sum_{k=2}^{\infty}\left(\rho^{t_k^i}\eta_i\left(t_k^i\right)-\rho^{t_{k-1}^i}\eta_i\left(t_{k-1}^i\right)\right)=$$
$$\sum_{i=1}^{n}\theta_i\left(t_0^i\right)+\sum_{i=1}^{n}\left(\rho^{t_1^i}\eta_i\left(t_1^i\right)-\tilde{\eta}_i\left(t_0^i\right)\right)+$$
$$\sum_{i=1}^{n}\left(\rho^{t_\infty^i}\eta_i\left(t_\infty^i\right)-\rho^{t_1^i}\eta_i\left(t_1^i\right)\right)=$$
$$\sum_{i=1}^{n}\eta_i\left(t_0^i\right)-\sum_{i=1}^{n}\tilde{\eta}_i\left(t_0^i\right) \tag{17}$$

显然有 $\rho^{t_\infty^i}\eta_i\left(t_\infty^i\right)=0$，再将 $\tilde{\eta}_i\left(t_0^i\right)$ 代入式(17)，得到

$$\sum_{i=1}^{n}\sum_{k=0}^{\infty}\theta_i\left(t_k^i\right)=\sum_{i=1}^{n}\eta_i\left(t_0^i\right)-\sum_{i=1}^{n}\tilde{\eta}_i\left(t_0^i\right)=\sum_{i=1}^{n}\eta_i\left(t_0^i\right)-$$
$$\sum_{i=1}^{n}\left[\eta_i\left(t_0^i\right)-\sum_{j\in N_i}\left[\Phi_{ij}\left(s_{ij}\right)-\Phi_{ji}\left(s_{ji}\right)\right]\right]=$$
$$\sum_{i=1}^{n}\sum_{j\in N_i}\left[\Phi_{ij}(s_{ij})-\Phi_{ji}(s_{ji})\right] \tag{18}$$

由于在节点 i 中使用的每对 $\Phi_{ij}(s_{ij})-\Phi_{ji}(s_{ji})$ 在节点 j 中皆存在相应的负值 $\Phi_{ji}(s_{ji})-\Phi_{ij}(s_{ij})$，因此有

$$\sum_{i=1}^{n}\sum_{j\in N_i}\left[\Phi_{ij}(s_{ij})-\Phi_{ji}(s_{ji})\right]=0 \tag{19}$$

最后，有 $\sum_{i=1}^{n}\sum_{k=0}^{\infty}\theta_i\left(t_k^i\right)=0$。

证毕。

2.4 差分隐私分析

定理 2 在本文设计的算法 1 下，所添加的噪声 $\theta_i(k)$ 可以全部用 $\eta_i(k)$ 来表示。

证明 由式(13)可知，$\eta_i(k)$ 和 $\eta_i(k-1)$ 均服从拉普拉斯分布，其中 $k\in\{1,2,\cdots,n\}$，那么 $\eta_i(k-1)$ 可以直接用 $\eta_i(k)$ 来表示，即

$$\eta_i(k-1)=\beta\eta_i(k) \tag{20}$$

那么

$$\beta=\frac{\eta_i(k-1)}{\eta_i(k)}=\frac{\frac{1}{2\lambda}\mathrm{e}^{-\frac{|k-1|}{\lambda}}}{\frac{1}{2\lambda}\mathrm{e}^{-\frac{|k|}{\lambda}}}=\mathrm{e}^{\frac{1}{\lambda}} \tag{21}$$

由此可知

$$\eta_i(k-1)=\mathrm{e}^{\frac{1}{\lambda}}\eta_i(k) \tag{22}$$

因此，系统中添加的噪声可改写为

$$\theta_i(k)=\begin{cases}\rho\eta_i(1)-\tilde{\eta}_i(0), & k=1\\ \rho^k\eta_i(k)-\rho^{k-1}\beta\eta_i(k), & k\geqslant 2\end{cases} \tag{23}$$

证毕。

定理 3 在本文设计的算法 1 下，多智能体系统式(7)能够实现 ε-差分隐私，其中隐私保护的程度为 $\varepsilon=\max\varepsilon_i$。

证明 通过定义 2 可以得到

$$\frac{P\left[\tilde{\boldsymbol{x}}\in\mathcal{O}\right]}{P\left[\tilde{\boldsymbol{x}}'\in\mathcal{O}\right]}=\prod_{k=1,\cdots,K}\frac{P\left[\tilde{x}_1(k),\tilde{x}_2(k),\cdots,\tilde{x}_n(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_1'(k),\tilde{x}_2'(k),\cdots,\tilde{x}_n'(k)\in\mathcal{O}\right]}=$$
$$\prod_{k=1,\cdots,K}\frac{P_1P_2\cdots P_n}{P_1'P_2'\cdots P_n'}=\prod_{1\leqslant i\leqslant n,k=1,\cdots,K}\frac{P\left[\tilde{x}_i(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]} \tag{24}$$

其中，$P_i=P\left[\tilde{x}_i(k)\in\mathcal{O}\right]$，$P_i'=P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]$，$i=1,2,\cdots,n$。由于状态值彼此相互独立，且基于 δ-邻接的定义，有

$$\frac{P\left[\tilde{x}(k)\in\mathcal{O}\right]}{P\left[\tilde{x}'(k)\in\mathcal{O}\right]}=\prod_{k=1,\cdots,K}\frac{P\left[\tilde{x}_i(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]} \tag{25}$$

$\tilde{x}_i(k)\in\mathcal{O}$ 的概率密度分布属于拉普拉斯分布，当 $k=1$ 时，令 $\Phi=\sum_{j\in N_i}\left[\Phi_{ij}(s_{ij})-\Phi_{ji}(s_{ji})\right]$，可以得到

$$P\left[\tilde{x}_i(k)\in\mathcal{O}\right]=P\left[x_i(k)+\theta\left(\eta_i(k)\right)\in\mathcal{O}\right]=$$
$$P\left[\rho\eta_i(1)-\eta_i(0)+\Phi\in\mathcal{O}-x_i(1)\right]=$$
$$P\left[(\rho-\beta)\eta_i(1)\in\mathcal{O}-\Phi-x_i(1)\right]=$$
$$\int_{\Psi_1}L\left((\rho-\beta)\eta_i(1),0,\lambda\right)\mathrm{d}(\rho-\beta)\eta_i(1)=$$
$$\int_{\Psi_1}L\left(H_i(1),0,\lambda\right)\mathrm{d}H_i(1) \tag{26}$$

其中，$\Psi_1=\mathcal{O}-\boldsymbol{\Phi}-x_i(1)$，$H_i(1)=(\rho-\beta)\eta_i(1)$。

通过相同的方式，可以得到

$$\begin{aligned}&P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]=P\left[x_i'(k)+\theta\left(\eta_i(k)\right)\in\mathcal{O}\right]=\\&P\left[\rho\eta_i(1)-\eta_i(0)+\boldsymbol{\Phi}\in\mathcal{O}-x_i'(1)\right]=\\&P\left[(\rho-\beta)\eta_i(1)\in\mathcal{O}-\boldsymbol{\Phi}-x_i'(1)\right]=\\&\int\limits_{\mathcal{O}-\boldsymbol{\Phi}-x_i'(1)}L\left((\rho-\beta)\eta_i(1),0,\lambda\right)\mathrm{d}(\rho-\beta)\eta_i(1)=\\&\int\limits_{\mathcal{O}-\boldsymbol{\Phi}-x_i'(1)}L\left(H_i(1),0,\lambda\right)\mathrm{d}H_i(1)\end{aligned}\tag{27}$$

由定义 1 可知，$\left|x_i(0)-x_i'(0)\right|\leqslant\delta$。根据式(4)和式(7)，可以计算

$$\begin{aligned}&\left|x_i(k+1)-x_i'(k+1)\right|=\\&\left|a_{ii}\tilde{x}_i(k)+\sum_{j\in N_i}a_{ij}\tilde{x}_j(k)-a_{ii}\tilde{x}_i'(k)-\sum_{j\in N_i}a_{ij}\tilde{x}_j'(k)\right|=\\&\left|a_{ii}\tilde{x}_i(k)-a_{ii}\tilde{x}_i'(k)\right|=a_{ii}\left|x_i(k)+\theta_i(k)-x_i'(k)-\theta_i(k)\right|=\\&a_{ii}\left|x_i(k)-x_i'(k)\right|\cdots=a_{ii}^{k+1}\left|x_i(0)-x_i'(0)\right|\leqslant a_{ii}^{k+1}\delta\end{aligned}\tag{28}$$

同时令 $x_i'(k)=x_i(k)-m_i(k)$，那么式(27)可以写为

$$\begin{aligned}&P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]=\int\limits_{\Psi_1+m_i(1)}L\left(H_i(1),0,\lambda\right)\mathrm{d}H_i(1)=\\&\int\limits_{\Psi_1}L\left(H_i(1)+m_i(1),0,\lambda\right)\mathrm{d}H_i(1)\end{aligned}\tag{29}$$

通过式(26)和式(29)，可以得到

$$\begin{aligned}&\frac{P\left[\tilde{x}_i(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]}=\frac{\int\limits_{\Psi_1}L\left(H_i(1),0,\lambda\right)\mathrm{d}H_i(1)}{\int\limits_{\Psi_1}L\left(H_i(1)+m_i(1),0,\lambda\right)\mathrm{d}H_i(1)}=\\&\frac{L\left(H_i(1),0,\lambda\right)}{L\left(H_i(1)+m_i(1),0,\lambda\right)}\leqslant\exp\left(\frac{|m_i(1)|}{\lambda}\right)=\\&\exp\left(\frac{|x_i(1)-x_i'(1)|}{\lambda}\right)\leqslant\exp\left(\frac{a_{ii}\delta}{\lambda}\right)\end{aligned}\tag{30}$$

同理，当 $k\geqslant 2$ 时

$$\begin{aligned}&P\left[\tilde{x}_i(k)\in\mathcal{O}\right]=P\left[x_i(k)+\theta\left(\eta_i(k)\right)\in\mathcal{O}\right]=\\&P\left[\rho^k\eta_i(k)-\rho^{k-1}\eta_i(k-1)\in\mathcal{O}-x_i(k)\right]=\\&P\left[(\rho^k-\beta\rho^{k-1})\eta_i(k)\in\mathcal{O}-x_i(k)\right]=\\&\int\limits_{\mathcal{O}-x_i(k)}L\left(H_i(k),0,\lambda\right)\mathrm{d}H_i(k)\end{aligned}\tag{31}$$

其中，$H_i(k)=(\rho^k-\beta\rho^{k-1})\eta_i(k)$。

通过相同的方式，可以得到

$$\begin{aligned}&P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]=P\left[x_i'(k)+\theta\left(\eta_i(k)\right)\in\mathcal{O}\right]=\\&P\left[\rho^k\eta_i(k)-\rho^{k-1}\eta_i(k-1)\in\mathcal{O}-x_i'(k)\right]=\\&P\left[(\rho^k-\beta\rho^{k-1})\eta_i(k)\in\mathcal{O}-x_i'(k)\right]=\\&\int\limits_{\mathcal{O}-x_i'(k)}L\left(H_i(k),0,\lambda\right)\mathrm{d}H_i(k)\end{aligned}\tag{32}$$

根据式(28)，式(32)可以写为

$$\begin{aligned}&P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]=\int\limits_{\mathcal{O}-x_i(k)+m_i(k)}L\left(H_i(k),0,\lambda\right)\mathrm{d}H_i(k)=\\&\int\limits_{\mathcal{O}-x_i(k)}L\left(H_i(k)+m_i(k),0,\lambda\right)\mathrm{d}H_i(k)\end{aligned}\tag{33}$$

通过式(31)和式(33)，可以得到

$$\begin{aligned}&\frac{P\left[\tilde{x}_i(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]}=\frac{\int\limits_{\mathcal{O}-x_i(k)}L\left(H_i(k),0,\lambda\right)\mathrm{d}H_i(k)}{\int\limits_{\mathcal{O}-x_i(k)}L\left(H_i(k)+m_i(k),0,\lambda\right)\mathrm{d}H_i(k)}=\\&\frac{L\left(H_i(k),0,\lambda\right)}{L\left(H_i(k)+m_i(k),0,\lambda\right)}\leqslant\exp\left(\frac{|m_i(k)|}{\lambda}\right)=\\&\exp\left(\frac{|x_i(k)-x_i'(k)|}{\lambda}\right)\leqslant\exp\left(\frac{a_{ii}^k\delta}{\lambda}\right)\end{aligned}\tag{34}$$

综上所述，由式(30)和式(34)可以得到

$$\begin{aligned}&\frac{P\left[\tilde{\boldsymbol{x}}\in\mathcal{O}\right]}{P\left[\tilde{\boldsymbol{x}}'\in\mathcal{O}\right]}=\prod_{k=1,\cdots,K}\frac{P\left[\tilde{x}_i(k)\in\mathcal{O}\right]}{P\left[\tilde{x}_i'(k)\in\mathcal{O}\right]}\leqslant\\&\prod_{k=1,\cdots,K}\exp\left(\frac{a_{ii}^k\delta}{\lambda}\right)=\mathrm{e}^{\varepsilon}\end{aligned}\tag{35}$$

由于多智能体系统中每个节点可以选择不同的 λ，即 λ_i，因此每个智能体会有不同的隐私程度 ε_i，从上述推导可以得到每个智能体的差分隐私程度为

$$\varepsilon_i=\sum_{k=1}^{\infty}\left(\frac{a_{ii}^k\delta}{\lambda_i}\right)=\frac{\delta a_{ii}\left(1-a_{ii}^k\right)}{\lambda_i(1-a_{ii})}\tag{36}$$

显然，根据定义 2，ε 越小则系统隐私保护的程度越高。因此，多智能体系统式(7)的隐私保护程度为 $\varepsilon=\max\varepsilon_i$。

证毕。

3　仿真分析

本节将通过数值仿真实验与硬件实验分别对本文算法的有效性进行验证。

3.1　数值仿真实验

在数值仿真实验中，仿真平台为一部运行 64 位

Windows 10 操作系统的 PC 机，配置为 Intel(R) Core(TM) i5-5200U CPU @ 2.20 GHz 处理器，RAM 为 8 GB。编程语言为 Python 3.8.8，编程环境为 PyCharm。根据上述数值仿真环境，本文将所提算法与文献[16,18]中的算法进行仿真比较，以验证本文算法的优越性。

本文构建了由 5 个智能体节点组成的无向图网络，其通信拓扑如图 2 所示。2 个节点之间有连接线表示节点间存在通信链路，可进行信息交互；无连接线则表示节点间无通信链路，不能传递信息。在区间[1,50]随机赋予每个节点的初始状态值为 $x_1(0)=40$，$x_2(0)=44$，$x_3(0)=31$，$x_4(0)=48$，$x_5(0)=16$。相应地，可以计算出系统初始状态值的平均值为 35.8。实验设置 $\rho=0.8$，$\lambda=\sqrt{\dfrac{1}{2}}$。根据图 2 中呈现的拓扑关系，设置其邻接矩阵为

$$\boldsymbol{A}=\frac{1}{5}\begin{bmatrix}1&2&0&0&2\\2&1&1&0&1\\0&1&2&1&1\\0&0&1&4&0\\2&1&1&0&1\end{bmatrix} \tag{37}$$

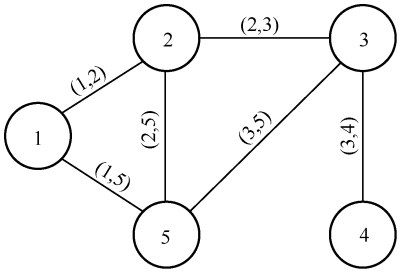

图 2 网络通信拓扑

根据上述设置条件，5 个智能体节点执行不同的共识算法进行实验，若所有节点状态值随着时间的推移最终都收敛于 35.8，则说明系统在该对应算法下实现了平均共识。此外，考虑存在好奇窃听者试图猜测节点的初始状态值，本文采用与文献[18]相同的方法，即在仿真实验中建立了一个观测器（窃听者）来推测其他节点的状态值。不妨令观测器的窃听目标为节点 1 的初始状态值，其更新规则如下。

1) 观测器的初始状态被设置为节点 1 添加噪声的初始状态，即

$$z(0)=\tilde{x}_1(0) \tag{38}$$

2) 观测器的更新是基于传输状态值与预测状态值之间的累积，即

$$z(k+1)=z(k)+\tilde{x}_1(k+1)-\left(a_{11}\tilde{x}_1(k)+\sum_{j\in N_1}a_{1j}\tilde{x}_j(k)\right) \tag{39}$$

图 3 为系统在无 DoS 攻击情形下，5 个智能体节点采用常规差分隐私算法（未加入秘密函数）下的状态轨迹。从图 3 中可以看出，在无 DoS 攻击且采用常规差分隐私算法下，5 个节点精确收敛到初始状态值的平均值 35.8，说明系统可以实现精确的平均共识，但也看到，观测器节点推测值最终收敛到节点 1 的初始状态值，说明节点 1 的信息被观测器推测出，存在隐私泄露。实验仿真结果说明，常规差分隐私算法不能对系统中节点的初始状态值进行有效保护。

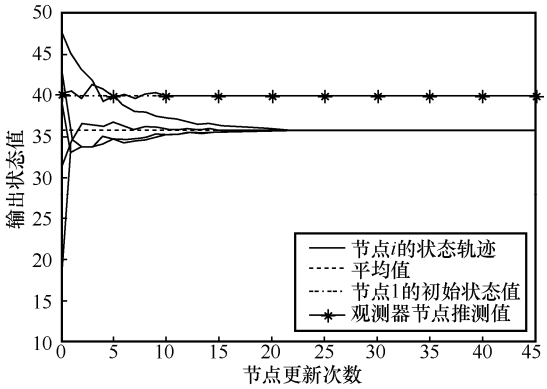

图 3 系统采用常规差分隐私算法下的状态轨迹

图 4 为系统在无 DoS 攻击下，5 个节点在加入秘密函数后的差分隐私算法下的状态轨迹。令秘密函数 $\Phi_{ij}=\dfrac{3i-2j}{40}$，从图 4 可以看出，加入秘密函数后的算法不但可以使系统准确达到所有节点初始状态值的平均值 35.8，而且可以让观测器不能推测到节点 1 的初始状态值 40，即验证了加入秘密函数后的算法可以实现对节点初始值的隐私保护。

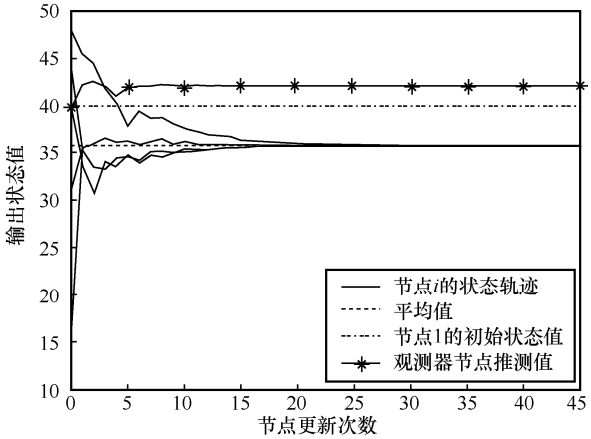

图 4 系统在加入秘密函数后的差分隐私算法下的状态轨迹

接下来，考虑网络中存在 DoS 攻击的情形。根据本文所考虑的 DoS 攻击模型，将实例中发生的 DoS 攻击对系统通信链路的影响具象化为图 5。具体地，在 $k=0$ 至 $k=1$ 时间段，链路(1,2)和(3,5)遭受 DoS 攻击，节点之间的通信受阻，即此时不能将信息传递给相邻节点，其他链路正常通信；在 $k=1$ 至 $k=2$ 时间段，链路(1,5)和(3,4)遭受 DoS 攻击，其他链路正常通信；在 $k=2$ 至 $k=3$ 时间段，链路(2,5)和(3,5)遭受 DoS 攻击，因此有一个固定的周期，没有受到攻击的链路都可以正常通信。不难验证，图 5 所示的 DoS 攻击行为满足本文的假设 1。

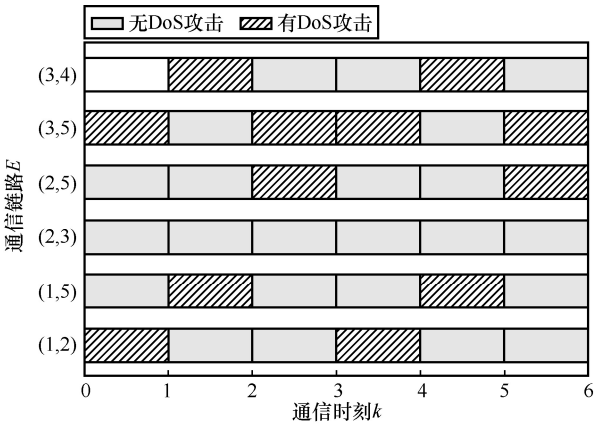

图 5　DoS 攻击对系统通信链路的影响

图 6 为系统在 DoS 攻击下，5 个节点在采用文献[16]算法下的状态轨迹。从图 6 中可以看到，虽然文献[16]算法对系统的状态值实现了隐私保护，并最终达成了共识，但距离真实平均值存在较大误差，说明文献[16]算法无法以精确的初始平均值实现收敛。

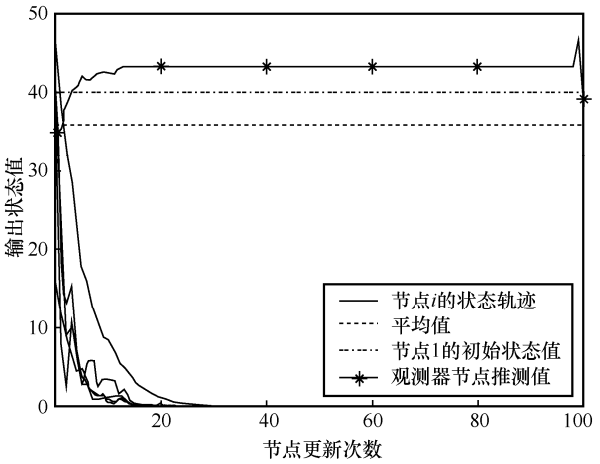

图 6　系统在 DoS 攻击下 5 个节点采用文献[16]算法下的状态轨迹

此外，为进一步验证本文算法的优越性，本文对本文算法与文献[18]算法进行了仿真对比实验。文献[18]中的节点更新规则为

$$x_i[k+1]=x_i[k]+\varepsilon\sum_{v_j\in N_i}a_{ij}^{(k)}\left(x_j[k]-x_i[k]\right) \quad (40)$$

其中，$x_i[k]$ 为节点 i 的状态值，ε 为恒定步长。为满足文献[18]中的仿真实验要求，将 ε 设为 0.5，加权邻接矩阵设置为

$$\boldsymbol{A}=\left[a_{ij}\right]=\begin{bmatrix}0 & 0.2 & 0 & 0 & 0.4\\ 0.2 & 0 & 0.3 & 0 & 0.5\\ 0 & 0.3 & 0 & 0.4 & 0.7\\ 0 & 0 & 0.4 & 0 & 0\\ 0.4 & 0.5 & 0.7 & 0 & 0\end{bmatrix} \quad (41)$$

图 7 为在 DoS 攻击下，5 个节点在文献[18]算法下的状态轨迹。从图 7 中可以看到，最终 5 个节点的状态轨迹无法达成共识，这说明网络中存在 DoS 攻击会使文献[18]共识算法失效。

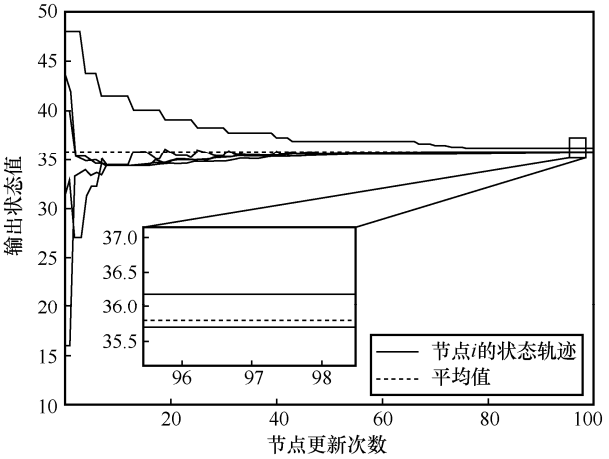

图 7　DoS 攻击下 5 个节点采用文献[18]算法下的状态轨迹

图 8 为在 DoS 攻击下，5 个节点在本文算法下的状态轨迹。从图 8 中可以看出，系统中所有节点的状态值可以实现精确的平均共识，即达到 35.8。而且在节点更新迭代的过程中，观测器节点无法准确推测到节点 1 的真实初始状态值，从而验证了本文理论，即系统采用本文算法可以在 DoS 攻击影响下实现精确的平均共识和隐私保护。图 9 为系统中各个节点的触发时刻。从图 9 中可以看出，若节点满足触发函数，则更新节点的状态值；若不满足触发函数，则保持节点状态值不变。

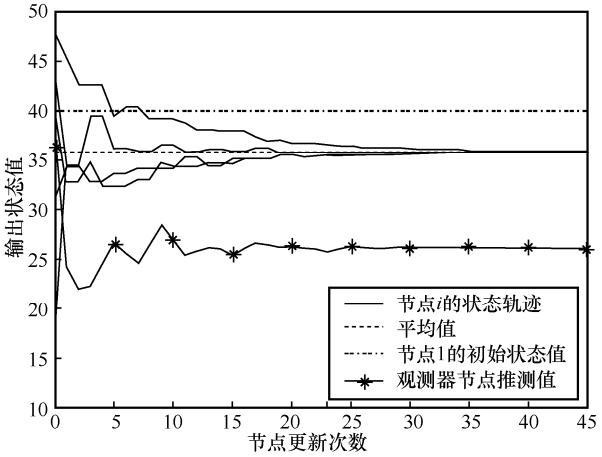

图 8　DoS 攻击下 5 个节点在本文算法下的状态轨迹

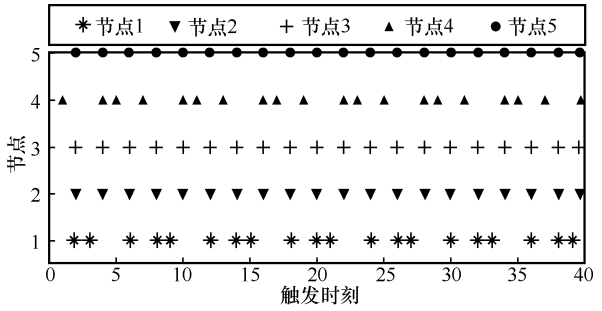

图 9　系统中 5 个节点的触发时刻

3.2　硬件实验

为了验证在真实环境中本文算法的有效性，本文在由 6 块开发板组建成的通信网络中进行了硬件实验验证。6 块开发板型号分别为一块 ESP-S3-32S、两块 ESP-12F、一块 ESP-12K、一块 ESP32-S 和一块 CORE-ESP32。前 5 块开发板作为 5 个智能体节点，最后一块 CORE-ESP32 则作为观测器节点。

在实验过程中，随机给定开发板节点 1~5 的初始值分别为 10、25、30、5 和 22，此时初始状态平均值为 18.4。节点基于 TCP 通过 Micropython 中的 socket 进行 Wi-Fi 通信。在相邻节点间的交互过程中，一个节点使用请求信息来进行发送，邻居节点使用响应信息进行响应。同样，为满足本文考虑的 DoS 攻击模型，即假设 1 条件，硬件实验中，同样考虑了图 5 所示的 DoS 攻击的影响。在程序的编写中，通过代码指令抑制节点不传送数据来模拟 DoS 攻击的发生。另外，考虑到 DoS 攻击的影响，所有节点不能同步更新状态，因此在每个节点上都使用了计数器，将计数值与将要传输节点的型号和本节点的状态值一同封装，再将信息传输出去。同时，每个节点将更新后的状态值传输至电脑端。根据电脑端收集到的所有信息，得到各开发板节点状态轨迹，如图 10 所示。硬件实验同样表明，本文算法可以在 DoS 攻击影响下实现精确的初始状态值平均共识，也可以保护节点的初始状态值不被窃听者推测出。

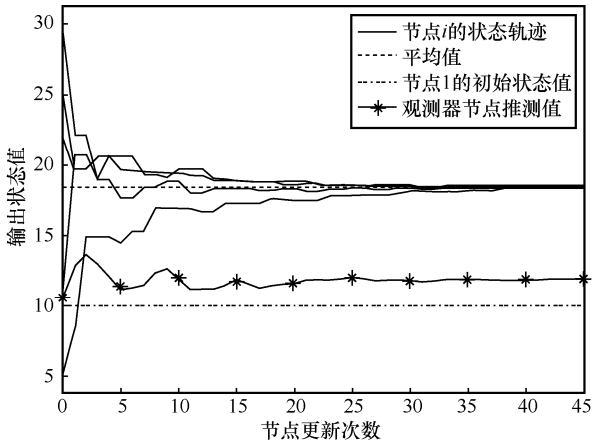

图 10　各开发板节点状态轨迹

此外，本文在上述硬件条件下将本文算法与文献[18]算法在执行效率上进行了比较。其中，文献[18]算法是具有加性同态的 Paillier 公钥密码系统。首先，节点采用文献[18]算法统计得出每个节点每一次迭代更新和信息交互所需的平均时间为 6.64 s。然后，节点采用本文算法，即基于差分隐私和秘密函数的共识算法，统计出每个节点每一次迭代更新和信息交互所需的平均时间为 0.27 s。实验结果表明，在相同的硬件环境下，虽然 2 种算法最终都可以实现精确的平均共识和隐私保护，但文献[18]算法在执行单次迭代所需花费的时间是本文算法的 24 倍以上，由此验证了本文算法在算法执行效率上相对文献[18]算法具有较强的优势。

4　结束语

本文针对实际环境中多智能体系统面临的网络攻击与信息窃听等问题，提出了一种分布式安全平均共识算法。本文算法利用差分隐私后的节点信息值作为输入，并且在整个算法的执行过程中融入事件触发机制。相比于传统基于差分隐私的算法，本文算法通过加入特定秘密函数，实现了零误差收敛精度；相比于已有的安全共识算法，本文算法结合事件触发机制，有效地满足了实际通信过程中的信息量化需求，有利于减少更新频次。同时，理论上严格分析和证明了在无 Zeno 行为的情况下该算法的收敛性和隐私性。仿真结果验证了本文算法的有效性。

参考文献：

[1] DONG X W, HU G Q. Time-varying output formation for linear multiagent systems via dynamic output feedback control[J]. IEEE Transactions on Control of Network Systems, 2017, 4(2): 236-245.

[2] QIN J H, ZHU Y D, FU W M. Distributed clustering algorithm in sensor networks via normalized information measures[J]. IEEE Transactions on Signal Processing, 2020, 68: 3266-3279.

[3] ZHAO C C, HE J P, CHENG P, et al. Consensus-based energy management in smart grid with transmission losses and directed communication[J]. IEEE Transactions on Smart Grid, 2017, 8(5): 2049-2061.

[4] REN W, BEARD R W. Consensus seeking in multiagent systems under dynamically changing interaction topologies[J]. IEEE Transactions on Automatic Control, 2005, 50(5): 655-661.

[5] ZUO Z Q, CAO X,WANG F Y J, et al. Security control of multi-agent systems under false data injection attacks[J]. Neurocomputing, 2020, 404: 240-246.

[6] DIBAJI S M, ISHII H. Resilient consensus of second-order agent networks: asynchronous update rules over robust graphs[C]// Proceedings of 2015 American Control Conference (ACC). Piscataway: IEEE Press, 2015: 1451-1456.

[7] 杨浩, 许宇航, 倪媛, 等. 网络系统的安全决策与控制: 容错博弈研究综述[J]. 控制与决策, 2022, 37(4): 769-781.
YANG H, XU Y H, NI Y, et al. Safe decision and control of network systems: a survey on fault tolerant game[J]. Control and Decision, 2022, 37(4): 769-781.

[8] GUSRIALDI A, QU Z H, SIMAAN M A. Robust design of cooperative systems against attacks[C]//Proceedings of 2014 American Control Conference. Piscataway: IEEE Press, 2014: 1456-1462.

[9] 李丽, 王夕娟. 拒绝服务攻击下领导–跟随多智能体系统的均方一致性研究[J]. 控制与决策, 2019, 34(11): 2317-2322.
LI L, WANG X J. Mean square consensus for leader-following multi-agent systems under denial-of-service attacks[J]. Control and Decision, 2019, 34(11): 2317-2322.

[10] YANG Y, LI Y F, YUE D, et al. Distributed secure consensus control with event-triggering for multiagent systems under DoS attacks[J]. IEEE Transactions on Cybernetics, 2021, 51(6): 2916-2928.

[11] XU Y, FANG M, SHI P, et al. Event-based secure consensus of mutiagent systems against DoS attacks[J]. IEEE Transactions on Cybernetics, 2020, 50(8): 3468-3476.

[12] CHANG B B, MU X W, YANG Z, et al. Event-based secure consensus of muti-agent systems under asynchronous DoS attacks[J]. Applied Mathematics and Computation, 2021, 401: 126120.

[13] RENGANATHAN V, SUMMERS T. Spoof resilient coordination for distributed multi-robot systems[C]//Proceedings of 2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS). Piscataway: IEEE Press, 2018: 135-141.

[14] WHEELER T, BHARATHI E, GIL S. Switching topology for resilient consensus using Wi-Fi signals[C]//Proceedings of 2019 International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2019: 2018-2024.

[15] WU Y M, XU M, ZHENG N, et al. Event-triggered resilient consensus for multi-agent networks under deception attacks[J]. IEEE Access, 2020, 8: 78121-78129.

[16] HE J P, CAI L, ZHAO C C, et al. Privacy-preserving average consensus: privacy analysis and algorithm design[J]. IEEE Transactions on Signal and Information Processing Over Networks, 2019, 5(1): 127-138.

[17] SHAMIR A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.

[18] RUAN M H, GAO H, WANG Y Q. Secure and privacy-preserving consensus[J]. IEEE Transactions on Automatic Control, 2019, 64(10): 4035-4049.

[19] WANG Y Q. Privacy-preserving average consensus via state decomposition[J]. IEEE Transactions on Automatic Control, 2019, 64(11): 4711-4716.

[20] DWORK C. Differential privacy[C]//Proceedings of the 33rd international Conference on Automata, Languages and Programming. New York: ACM Press, 2006: 1-12.

[21] MO Y L, MURRAY R M. Privacy preserving average consensus[J] IEEE Transactions on Automatic Control, 2017, 62(2): 753-765.

[22] WANG J M, ZHANG J F. Differentially private consensus for multi-agent systems[C]//Proceedings of 2020 39th Chinese Control Conference (CCC). Piscataway: IEEE Press, 2020: 4843-4848.

[23] GAO L, DENG S J, REN W. Differentially private consensus with an event-triggered mechanism[J]. IEEE Transactions on Control of Network Systems, 2019, 6(1): 60-71.

[24] FIORE D, RUSSO G. Resilient consensus for multi-agent systems subject to differential privacy requirements[J]. Automatica (Journal of IFAC), 2019, 106(C): 18-26.

[25] OLSHEVSKY A, TSITSIKLIS J N. Convergence speed in distributed consensus and averaging[J]. SIAM Review, 2011, 53(4): 747-772.

[26] HE J P, CAI L, GUAN X P. Differential private noise adding mechanism and its application on consensus algorithm[J]. IEEE Transactions on Signal Processing, 2020, 68: 4069-4082.

[27] HAN S, TOPCU U, PAPPAS G J. Differentially private distributed constrained optimization[J]. IEEE Transactions on Automatic Control, 2017, 62(1): 50-64.

[28] HE J P, CAI L, CHENG P, et al. Consensus-based data-privacy preserving data aggregation[J]. IEEE Transactions on Automatic Control, 2019, 64(12): 5222-5229.

[作者简介]

徐明（1970– ），男，江苏苏州人，博士，杭州电子科技大学教授、博士生导师，主要研究方向为网络信息安全、数字取证等。

张保俊（1997– ），女，河北唐山人，杭州电子科技大学硕士生，主要研究方向为隐私保护、多智能体系统共识等。

伍益明（1987– ），男，浙江慈溪人，博士，杭州电子科技大学副教授、硕士生导师，主要研究方向为多智能体系统网络安全、群体智能与智能控制、迭代学习控制等。

应晨铎（1998– ），男，浙江温州人，杭州电子科技大学硕士生，主要研究方向为多智能体系统网络安全等。

郑宁（1962– ），男，浙江杭州人，杭州电子科技大学研究员、博士生导师，主要研究方向为网络信息安全、信息管理系统等。