arXiv:2301.11152v1 [eess.SY] 26 Jan 2023

非均匀时域滚动时域博弈中的多智能体系统簇形成

Yurid Nugraha、Ahmet Cetinkaya、Tomohisa Hayakawa、Hideaki Ishii 和 Quanyan Zhu

摘要——本文讨论在干扰攻击下多智能体系统的一致性与簇形成，以及防御者采取的反应式恢复行动。攻击者能够禁用网络中的部分边，旨在将智能体划分为规模更小的簇；作为应对，防御者通过提高传输功率恢复部分边。我们考虑重复博弈，并以滚动时域方式推导两位参与者相应的最优策略。攻击者和防御者计算策略的能力各不相同。这一特点通过时域长度和博弈周期的非均匀取值来体现。理论分析和仿真结果展示了时域长度与博弈周期对智能体状态的影响。

I. 引言

多智能体系统用于建模网络中多个能够进行局部决策的智能体之间的交互。由于智能体具有分布式特性，它们容易受到恶意对手发起的网络攻击 [1]。最常见的网络攻击之一是干扰攻击，对手通过发射干扰信号来阻止智能体之间的通信。关于多智能体系统一致性问题中的干扰攻击，已有研究，例如 [2], [3]。

存在对手时，遵循标准一致性协议的智能体未必总能收敛到相同状态；相反，它们可能分成若干簇。多智能体系统中的簇形成已有研究，例如 [4],，其中智能体状态更新中的权重可以取负值，表示某些智能体之间可能存在敌对关系。研究人员也采用博弈论方法，分析网络中此类对手与智能体之间的交互 [5]。

针对易受攻击的多智能体系统，研究者采用模型预测控制来应对智能体在求解带约束的特定优化问题时所面临的情况：通过考虑由时域长度所表征的未来时刻。针对在分布式模型预测控制下协同工作的智能体，一种防御机制已在 [6]中进行了研究。滚动时域概念是模型预测控制的关键组成部分，在 [7]中，该概念被应用于一种与重放攻击相关的非合作安全博弈。

本文考虑集中式攻击者与集中式防御者之间的双人博弈中的干扰攻击与防御场景。攻击者试图将智能体划分为尽可能多的簇，而防御者则力求使簇的数量尽可能少。该博弈随时间反复进行，参与者会根据滚动时域优化方法重新计算，并可能改变其策略。假设参与者具有不同的计算能力，这种差异由时域长度和博弈周期的非均匀取值表示。该问题可视为我们先前研究的扩展 [8], [9]；在 [8]，中，我们研究了参与者时域参数均匀的特殊情形，而在 [9] 中讨论了参与者在非均匀时域下的性能。

参与者以异步方式更新策略，并拥有计算能力和数据等不同资源的博弈，可用于建模现实应用。例如，涉及不同国家企业的决策无法同时进行，因为各地的工作时间不同 [10], [11]。由于时域非均匀，参与者的决策过程变得复杂，且在不同时间点执行。时域较长的参与者显然占据优势；由于我们研究的是完全信息情形，该参与者甚至可能解决对手未来将面临的优化问题。另一方面，时域较短的参与者无法再完美地观察对手计划采取的行动；这可能导致参与者浪费用于发起攻击/恢复行动的资源。网络安全背景下类似的能量分配博弈已在文献中讨论，例如 [12], [13]。

在此，我们考虑智能体簇形成背景下的能量分配博弈。我们的方法与网络效应/外部性的概念相关 [14], , 在该概念中，某个智能体在特定簇中的效用取决于属于该簇的其他智能体的数量。此类概念已用于分析社交网络和计算机网络等环境中的智能体分组，相关讨论见 [15], [16]。

本文结构安排如下。第二节介绍滚动时域博弈框架以及参与者的能耗模型。第三节详细描述具有非统一时域长度和博弈周期的博弈结构。接下来，我们将讨论关于一致性和簇

Yurid Nugraha 和 Tomohisa Hayakawa 任职于日本东京 152-8552 的东京工业大学系统与控制工程系。yurid@dsl.sc.e.titech.ac.jp, hayakawa@sc.e.titech.ac.jpAhmet Cetinkaya 任职于日本东京 135-8548 的芝浦工业大学。ah met@shibaura-it.ac.jpHideaki Ishii 任职于日本横滨 226-8502 的东京工业大学计算机科学系。ishii@c.titech.ac.jpQuanyan Zhu 任职于美国纽约州布鲁克林 11201 的纽约大学电气与计算机工程系。quanyan.zhu@nyu.edu

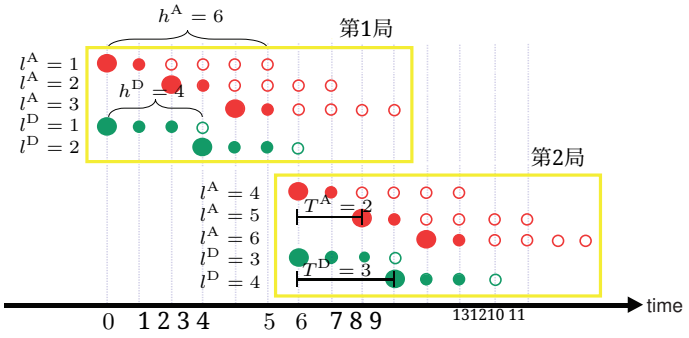


图1. 具有决策索引 l^{A} 和 l^{D}的博弈序列：攻击者时域（红色）和防御者时域（绿色），以及非均匀博弈周期。时域长度为$h^{\mathrm{A}}=6$ 和 $h^{\mathrm{D}}=4$，而博弈周期为 $T=2$ $T^{\mathrm{D}}=3$ A 和。有两个博弈以黄色矩形表示；在此示例中，每隔 lcm(TA T^{D}) $=6$ 个时间瞬间进行一局博弈。实心圆表示已实施的策略，空心圆表示未实施的博弈策略。

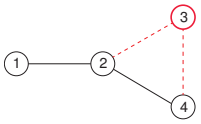


图2. 注释2.1中讨论的节点攻击示例。攻击者攻击节点/智能体3，移除边(3,2)和(3,4)。因此，智能体3与其他智能体不连通。

形成的理论结果，见第四节和第五节。随后，我们在第六节给出数值示例。最后，我们在第七节总结全文。

本文所使用的符号相当标准。我们用 $|\cdot|$ 表示集合的基数。向下取整函数和向上取整函数分别记为 $\lfloor\cdot\rfloor$ 和 $\lceil\cdot\rceil$。非负整数集合记为 $\mathbb{N}_0$。

II. 问题表述

我们研究一个由 n 个智能体组成的多智能体系统，这些智能体在离散时间下相互通信。网络拓扑由无向连通图 $\mathcal{G}=(\mathcal{V},\mathcal{E})$ 描述。该图由表示智能体的顶点集合 $\mathcal{V}$ 和表示通信链路的边集合 $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ 组成。每个智能体 i 的标量状态 x_i 在时间 $k\in\mathbb{N}_0$ 遵循一致性更新规则

$$x_i[k+1]=x_i[k]+u_i[k], \tag{1}$$

$$u_i[k]=\sum_{j\in\mathcal{N}_i[k]}a_{ij}(x_j[k]-x_i[k]), \tag{2}$$

其中，$x[0]=x_0$、$a_{ij}>0$、$\sum_{j=1,\ j\neq i}^{n}a_{ij}<1$和 $\mathcal{N}_i[k]$表示在时间 k能与智能体i 通信的智能体集合。该集合可能因攻击而发生变化。

考虑攻击者与防御者之间的双人博弈。攻击者能够通过干扰一些目标边来阻断通信，从而延迟（或完全阻止）智能体之间的一致性。这些干扰攻击通过移除$\mathcal{G}$中的边来表示。作为回应，防御者尝试分配资源以重建其中一些边，从而恢复智能体间通信。

我们考虑一名攻击者，其干扰信号有两种强度：强干扰和普通干扰。防御者只能恢复受到普通强度攻击的边；如果防御者将能量分配给受到强力攻击的边，这些边便无法重建，资源也会被浪费。同样，如果防御者将资源分配给未受攻击的边，资源也会被浪费，网络连通性不会得到任何改善。尽管近期研究 [3], [5]–[8] 在类似的多智能体系统设置中探讨了干扰，其中并未出现资源浪费这一概念。

A. 攻击-恢复序列

在我们的设置中，参与者在每个时间 $k\in\mathbb{N}_0$ 执行攻击/恢复行动。在时间 k 开始时，系统的通信拓扑表示为$\mathcal{G}$。随后，参与者分两个阶段决定攻击/恢复哪些边：攻击者先行动，防御者后行动。因此，该博弈属于斯塔克尔伯格类型。

更具体地说，在时间 k，攻击者通过使用普通干扰信号删除 $\mathcal{E}_k^{\mathrm{A}}\subseteq\mathcal{E}$，并使用强干扰信号删除 $\overline{\mathcal{E}}_k^{\mathrm{A}}\subseteq\mathcal{E}$，从而攻击 $\mathcal{G}$，其中 $\mathcal{E}_k^{\mathrm{A}}\cap\overline{\mathcal{E}}_k^{\mathrm{A}}=\emptyset$；而防御者恢复 $\mathcal{E}_k^{\mathrm{D}}\subseteq\mathcal{E}$。由于先发生攻击、随后进行恢复，网络从 $\mathcal{G}$变为 $\mathcal{G}_k^{\mathrm{A}}:=(\mathcal{V},\mathcal{E}\setminus(\mathcal{E}_k^{\mathrm{A}}\cup\overline{\mathcal{E}}_k^{\mathrm{A}}))$，并进一步变为 $\mathcal{G}_k^{\mathrm{D}}:=(\mathcal{V},\mathcal{E}\setminus(\mathcal{E}_k^{\mathrm{A}}\cup\overline{\mathcal{E}}_k^{\mathrm{A}})\cup(\mathcal{E}_k^{\mathrm{D}}\cap\mathcal{E}_k^{\mathrm{A}}))$。随后，智能体根据由此得到的图 $\mathcal{G}_k^{\mathrm{D}}$与其邻居进行通信。

在该博弈中，参与者试图根据被攻击/恢复的边（$\overline{\mathcal{E}}_k^{\mathrm{A}}$、$\mathcal{E}_k^{\mathrm{A}}$）以及 $\mathcal{E}_k^{\mathrm{D}}$选择最佳策略，以最大化各自的效用函数。这里，博弈定义在包含多个步骤的时域上。参与者以滚动时域方式进行决策，具体说明见第三节；过去某一时间获得的最优策略，可能会在参与者于未来时间重新计算策略时发生变化。图1展示了所讨论的随时间变化的序列；攻击者和防御者的时域长度，即参与者在确定策略时向未来展望的范围，分别表示为 h^{A} 和 h^{D}；而博弈周期，即参与者更新策略的频率，则分别表示为 T^{A} 和 T^{D}（稍后将进一步讨论）。由于博弈周期不统一，参与者具有相互独立的决策过程，分别由决策索引 l^{A} 和 l^{D}表示；一个博弈定义为一组决策过程，其起始时间为参与者同时更新策略的时刻，该时刻由黄色方框表示。

注释2.1：除了上述基于单条边的攻击和恢复外，我们还可以考虑一种略有不同的情形：攻击者可以攻击节点/智能体，使所有与被攻击智能体相邻的边断开，如图2所示。具体而言，攻击者现在的行动是 $\mathcal{E}_k^{\mathrm{A}}\in\mathcal{F}$ 和 $\overline{\mathcal{E}}_k^{\mathrm{A}}\in\mathcal{F}$，其中 $\mathcal{F}:=\{\emptyset,F_1,F_2,\ldots,F_n,F_1\cup F_2,\ldots,\mathcal{E}=\cup_{i\in\mathcal{V}}F_i\}$和$F_i:=\{(i,j):\ j\in\mathcal{N}_i\}$ 表示相邻边的集合

与智能体 i相邻。在这种情况下，攻击者通过攻击与节点/智能体相邻的所有边，实际上攻击了该节点/智能体。

B. 能量约束

参与者通过攻击和恢复，将能量分配给受攻击边/恢复边。这些行动受到能量资源约束的影响，而能量资源会随时间线性增加。参与者消耗的能量也与计划攻击/恢复的边数成正比。在这里，对 $\overline{\mathcal{E}}_k^{\mathrm{A}}$ 的强攻击每条边每单位时间消耗 $\overline{\beta}^{\mathrm{A}} > 0$ 能量，而对 $\mathcal{E}_k^{\mathrm{A}}$的常规攻击每条边消耗 $\beta^{\mathrm{A}} > 0$ 能量，其中 $\overline{\beta}^{\mathrm{A}} > \beta^{\mathrm{A}}$。攻击者截至时间 k 使用的总能量受以下约束:

$$\sum_{m=0}^{k} (\overline{\beta}^{\mathrm{A}} |\overline{\mathcal{E}}_m^{\mathrm{A}}| + \beta^{\mathrm{A}} |\mathcal{E}_m^{\mathrm{A}}|) \leq \kappa^{\mathrm{A}} + \rho^{\mathrm{A}} k, \tag{3}$$

其中 $\kappa^{\mathrm{A}} \geq \rho^{\mathrm{A}} > 0$, $\beta^{\mathrm{A}} > 0$。该不等式意味着，攻击者消耗的总能量不能超过由初始能量 κ^{A} 和供给率 ρ^{A} 所确定的可用能量。该能量约束限制了攻击者能够攻击的边数上限。

防御者的能量约束与此类似，表示为 (3)

$$\sum_{m=0}^{k} \beta^{\mathrm{D}} |\mathcal{E}_m^{\mathrm{D}}| \leq \kappa^{\mathrm{D}} + \rho^{\mathrm{D}} k \tag{4}$$

其中 $\kappa^{\mathrm{D}} \geq \rho^{\mathrm{D}} > 0$, $\beta^{\mathrm{D}} > 0$。需要注意的是，防御者可能会低效地分配能量，即防御者可能试图恢复未受攻击的边，或受到强干扰信号攻击的边。

对于注释2.1中讨论的节点攻击情形，攻击者的能量约束 (3) 变为 $\sum_{m=0}^{k} \overline{\beta}_{\mathcal{V}}^{\mathrm{A}} |\overline{\mathcal{V}}_m^{\mathrm{A}}| + \beta_{\mathcal{V}}^{\mathrm{A}} |\mathcal{V}_m^{\mathrm{A}}| \leq \kappa^{\mathrm{A}} + \rho^{\mathrm{A}} k$ $\overline{\beta}_{\mathcal{V}}^{\mathrm{A}} > (\ )$, $\beta_{\mathcal{V}}^{\mathrm{A}}$ $\overline{\mathcal{V}}_m^{\mathrm{A}}$ $\mathcal{V}_m^{\mathrm{A}}$ 其中能量为 , 且 和 分别表示其相邻边受到强干扰信号和普通干扰信号攻击的节点/智能体集合。注意，在这种情况下，如果一条边同时受到普通信号和强信号攻击，那么防御者就无法恢复该边。

C. 智能体聚类与状态差异

通过攻击，攻击者使图断开连通，并将智能体分隔为若干簇（即智能体集合）。我们引入一些与智能体分组/聚类相关的概念。我们将具有相同无限时间状态的每个智能体子集 $\mathcal{C} \subseteq \mathcal{V}$ 称为一个簇，即 $\lim_{k\to\infty} x_i[k] = \lim_{k\to\infty} x_j[k]$, $\forall i, j \in \mathcal{C}$。

在所考虑的博弈中，攻击者和防御者关注的是每个群组中的智能体数量。具体而言，我们采用网络效应/网络外部性的概念 [14], , 即某个集群中的智能体效用取决于属于该集群的其他智能体数量。在此博弈中，攻击者试图孤立智能体，使每个群组中的智能体数量尽可能少；而防御者则希望尽可能多的智能体处于同一群组中。接下来，我们用以下指标表示群组程度

: 图 $\mathcal{G}'$ 中的函数 $c(\cdot)$，称为智能体群组指数，其定义如下:

$$c(\mathcal{G}') := \sum_{p=1}^{\overline{n}(\mathcal{G}')} |\mathcal{V}_p'|^2 - |\mathcal{V}|^2 \quad (\leq 0). \tag{5}$$

如果 $\mathcal{G}'$ 连通，则 $c(\mathcal{G}')$的值为 0，因为此时只有一个群组（即$\overline{n}(\mathcal{G}') = 1$）。 $c(\mathcal{G}')$的值越大（越接近 0），意味着图 $\mathcal{G}'$中的连通分量越少，且每个连通分量包含的智能体越多。

在我们的问题设定中，参与者在攻击/恢复时也会考虑其行动对智能体状态的影响。例如，攻击者可能希望将状态值差异较大的智能体分隔到不同的群组中。我们将时间 k 的智能体状态差异之和 z_k 定义为

$$z_k(\overline{\mathcal{E}}_k^{\mathrm{A}}, \mathcal{E}_k^{\mathrm{A}}, \mathcal{E}_k^{\mathrm{D}}) := x^{\mathrm{T}}[k+1] L_{\mathrm{c}} x[k+1], \tag{6}$$

其中， L_{c} 是含有 n 个智能体的完全图的拉普拉斯矩阵。遭受攻击和恢复的边 $(\overline{\mathcal{E}}_k^{\mathrm{A}}, \mathcal{E}_k^{\mathrm{A}}, \mathcal{E}_k^{\mathrm{D}})$ 会影响 $x[k+1]$, 进而影响 z_k的值。请注意，即使系统遭受攻击，由于式（1）和式（2）中给出的协议， z_k 的值也不会随时间 [17] 增加。

博弈结构将在后文详细说明，如图1所示。攻击者（分别为防御者）在第 l^{A}个（分别为第 l^{D}个）决策索引处、于时间 k =开始的 $l^{\mathrm{A}}, l^{\mathrm{D}} \in \mathbb{N}$（分别为 $k = (l^{\mathrm{D}} - 1) T^{\mathrm{D}}$）的效用函数，会考虑时间范围 $h^{\mathrm{A}}, h^{\mathrm{D}} \geq 1$ 内从时间（$l^{\mathrm{A}} - 1$）T_{A}到（$l^{\mathrm{A}} - 1$）$T_{\mathrm{A}} + h^{\mathrm{A}} - 1$（分别为从（$l^{\mathrm{D}} - 1$）$T_{\mathrm{D}}$到（$l^{\mathrm{D}} - 1$）$T^{\mathrm{D}} + h^{\mathrm{D}} - 1$）的智能体群组指数 $c(\cdot)$。具体而言，攻击者在第 l^{A}个决策过程中的效用函数，以及防御者在第 l^{D}个决策过程中的效用函数为

$$U_{l^{\mathrm{A}}}^{\mathrm{A}} := \sum_{k=(l^{\mathrm{A}}-1)T^{\mathrm{A}}}^{(l^{\mathrm{A}}-1)T^{\mathrm{A}}+h^{\mathrm{A}}-1} (a z_k - b c(\mathcal{G}_k^{\mathrm{D}})), \tag{7}$$

$$U_{l^{\mathrm{D}}}^{\mathrm{D}} := \sum_{k=(l^{\mathrm{D}}-1)T^{\mathrm{D}}}^{(l^{\mathrm{D}}-1)T^{\mathrm{D}}+h^{\mathrm{D}}-1} (-a z_k + b c(\mathcal{G}_k^{\mathrm{D}})), \tag{8}$$

这些效用函数需要由参与者最大化。时域长度较长且博弈周期较短的参与者预计能更高效地利用其能量，从而随着时间推移获得更高的效用。

III. 具有非均匀滚动时域长度和博弈周期的博弈结构

我们关注的是求出迄今所述博弈的子博弈精炼均衡。为此，将该博弈划分为若干子博弈／决策点。子博弈精炼均衡必须在每个子博弈中都是均衡。通过逆向归纳法，即从最小的子博弈开始求解均衡，来得到每位参与者的最优策略。当参与者的策略产生相同效用时，就会出现平局。此时，我们假设如果参与者有足够的

能量在后续所有步骤中攻击／恢复所有边，他们就会选择攻击／恢复更多边；否则，他们会攻击／恢复较少的边。

本节中，在考虑更一般的情形之前，我们先研究一个较简单的场景：参与者采用不同的时域参数，但其博弈周期相同。随后，我们研究参与者的博弈周期也不相同的情形。第一种情形仍然体现了参与者在求解博弈时具有不同的计算能力。

A. 非均匀时域长度

在本小节中，我们说明时域长度非均匀、博弈周期均匀时的博弈结构。也就是说，我们假设 $T_{\mathrm{A}} = T^{\mathrm{D}} = T$。这意味着两位参与者在同一时间定期做出决策。因此，参与者决策过程的索引在所有时刻均相等 $l^{\mathrm{A}} = l^{\mathrm{D}} = l$。

由于滚动时域方法的特性，第 l 个决策过程中得到的策略，即被攻击的边和恢复边，仅应用于从时间 $(l-1)T$ 到 $lT-1$，其中 $T \le \min\{h^{\mathrm{A}}, h^{\mathrm{D}}\}$。参与者在第 l 个决策过程中的策略指定为 $((\overline{\mathcal{E}}^{\mathrm{A}}_{l,1}, \mathcal{E}^{\mathrm{A}}_{l,1}, \mathcal{E}^{\mathrm{D}}_{l,1}) \ldots, (\overline{\mathcal{E}}^{\mathrm{A}}_{l,h^{\mathrm{D}}}, \mathcal{E}^{\mathrm{A}}_{l,h^{\mathrm{D}}}, \mathcal{E}^{\mathrm{D}}_{l,h^{\mathrm{D}}}), (\overline{\mathcal{E}}^{\mathrm{A}}_{l,h^{\mathrm{D}}+1}, \mathcal{E}^{\mathrm{A}}_{l,h^{\mathrm{D}}+1}) \ldots, (\overline{\mathcal{E}}^{\mathrm{A}}_{l,h^{\mathrm{A}}}, \mathcal{E}^{\mathrm{A}}_{l,h^{\mathrm{A}}}))$ 若 $h^{\mathrm{A}} > h^{\mathrm{D}}$，且 $((\overline{\mathcal{E}}^{\mathrm{A}}_{l},1, \mathcal{E}^{\mathrm{A}}_{l},1, \mathcal{E}^{\mathrm{D}}_{l},1) \ldots \ldots, (\overline{\mathcal{E}}^{\mathrm{A}}_{l,h^{\mathrm{A}}}, \mathcal{E}^{\mathrm{A}}_{l,h^{\mathrm{A}}}, \mathcal{E}^{\mathrm{D}}_{l,h^{\mathrm{A}}}) \mathcal{E}^{\mathrm{D}}_{l,h^{\mathrm{A}}+1}, , \mathcal{E}^{\mathrm{D}}_{l,h^{\mathrm{D}}})$ 若 $h^{\mathrm{A}} < h^{\mathrm{D}}$，其中 $\overline{\mathcal{E}}^{\mathrm{A}}_{l,\alpha}, \mathcal{E}^{\mathrm{A}}_{l,\alpha}, \mathcal{E}^{\mathrm{D}}_{l,\alpha}$ 表示在具有 $\alpha \in \mathbb{N}$ 的第 l 个决策过程的第 α 步所采取的策略。请注意，若 $h^{\mathrm{A}} > h^{\mathrm{D}}$，则只有攻击者在第 h^{D} 步之后制定其策略。同理，若 $h^{\mathrm{A}} < h^{\mathrm{D}}$，则只有防御者在第 h^{A} 步之后制定其策略。$h^{\mathrm{A}} = h^{\mathrm{D}}$ 的情况也可类似处理，其策略为 $((\overline{\mathcal{E}}^{\mathrm{A}}_{l},1, \mathcal{E}^{\mathrm{A}}_{l},1, \mathcal{E}^{\mathrm{D}}_{l},1) \ldots, (\overline{\mathcal{E}}^{\mathrm{A}}_{l,h^{\mathrm{A}}}, \mathcal{E}^{\mathrm{A}}_{l,h^{\mathrm{A}}}, \mathcal{E}^{\mathrm{D}}_{l,h^{\mathrm{A}}}))$。

在这些得到的策略中，仅应用从第 1 步到第 T 步的策略。由于我们考虑的是完全信息设定，双方参与者都知道 h^{A} 和 h^{D} 的值。

现在举例说明在 $h^{\mathrm{A}} = 3$ 和 $h^{\mathrm{D}} = 2$ 的情况下如何得到最优边。在决策索引 l 处，按时间逆向求解得到的参与者最优策略如下：

- Step 3:

$$\mathcal{E}^{\mathrm{D}*}_{l,3}(\overline{\mathcal{E}}^{\mathrm{A}}_{l,3}, \mathcal{E}^{\mathrm{A}}_{l,3}) \in \arg\max_{\mathcal{E}^{\mathrm{D}}_{l,3}} -U^{\mathrm{A}}_{l,3}, \tag{9}$$

$$(\overline{\mathcal{E}}^{\mathrm{A}*}_{l,3}(\mathcal{E}^{\mathrm{D}}_{l,2}), \mathcal{E}^{\mathrm{A}*}_{l,3}(\mathcal{E}^{\mathrm{D}}_{l,2})) \in \arg\max_{(\overline{\mathcal{E}}^{\mathrm{A}}_{l,3}, \mathcal{E}^{\mathrm{A}}_{l,3})} U^{\mathrm{A}}_{l,3}(\mathcal{E}^{\mathrm{D}*}_{l,3}), \tag{10}$$

- Step 2:

$$\mathcal{E}^{\mathrm{D}\prime}_{l,2}(\overline{\mathcal{E}}^{\mathrm{A}}_{l,2}, \mathcal{E}^{\mathrm{A}}_{l,2}) \in \arg\max_{\mathcal{E}^{\mathrm{D}}_{l,2}} U^{\mathrm{D}}_{l,2}, \tag{11}$$

$$(\overline{\mathcal{E}}^{\mathrm{A}\prime}_{l,2}(\mathcal{E}^{\mathrm{D}}_{l,1}), \mathcal{E}^{\mathrm{A}\prime}_{l,2}(\mathcal{E}^{\mathrm{D}}_{l,1})) \in \arg\max_{(\overline{\mathcal{E}}^{\mathrm{A}}_{l,2}, \mathcal{E}^{\mathrm{A}}_{l,2})} -U^{\mathrm{D}}_{l,2}(\mathcal{E}^{\mathrm{D}\prime}_{l,2}), \tag{12}$$

$$(\overline{\mathcal{E}}^{\mathrm{A}*}_{l,2}(\mathcal{E}^{\mathrm{D}}_{l,1}), \mathcal{E}^{\mathrm{A}*}_{l,2}(\mathcal{E}^{\mathrm{D}}_{l,1})) \in \arg\max_{(\overline{\mathcal{E}}^{\mathrm{A}}_{l,2}, \mathcal{E}^{\mathrm{A}}_{l,2})} U^{\mathrm{A}}_{l,2}(\mathcal{E}^{\mathrm{D}\prime}_{l,2}), \tag{13}$$

- 步骤 1:

$$\mathcal{E}^{\mathrm{D}\prime}_{l,1}(\overline{\mathcal{E}}^{\mathrm{A}}_{l,1}, \mathcal{E}^{\mathrm{A}}_{l,1}) \in \arg\max_{\mathcal{E}^{\mathrm{D}}_{l,1}} U^{\mathrm{D}}_{l}(\overline{\mathcal{E}}^{\mathrm{A}'}_{l,2}, \mathcal{E}^{\mathrm{A}\prime}_{l,2}), \tag{14}$$

$$(\overline{\mathcal{E}}^{\mathrm{A}*}_{l,1}, \mathcal{E}^{\mathrm{A}*}_{l,1}) \in \arg\max_{(\overline{\mathcal{E}}^{\mathrm{A}}_{l,1}, \mathcal{E}^{\mathrm{A}}_{l,1})} U^{\mathrm{A}}_{l}(\mathcal{E}^{\mathrm{D}\prime}_{l,1}), \tag{15}$$

其中，$U_{\mathrm{A}l,\ \alpha} := \sum_{k=\ l-1}^{(l-1)T+h^{\mathrm{A}}-1} {}_{()T+\alpha-1} = az_k - bc(\mathcal{G}^{\mathrm{D}}_k)$（分别地，$U_{\mathrm{D}l,\ \alpha} := \sum_{k=\ l-1}^{(l-1)T+h^{\mathrm{D}}-1} {}_{()T+\alpha-1} = -az_k + bc(\mathcal{G}^{\mathrm{D}}_k)$）定义为 $U_{\mathrm{A}l}$（分别地，$U_{\mathrm{D}l}$）的一部分，由第 α 步至第 h^{A} 步（分别地，第 h^{D} 步）计算得出，属于第 l 个决策过程。

这些优化问题在每个博弈周期 T 中，均从第 l 个决策过程的第（$\max\{h^{\mathrm{A}}, h^{\mathrm{D}}\}$）步开始，按逆序求解。注意，要找到（$\overline{\mathcal{E}}^{\mathrm{A}*}_{l\,1}, , \mathcal{E}^{\mathrm{A}*}_{l\,1}, $），需要预先得到（$\mathcal{E}^{\mathrm{D}*}_{l\,1}$, ($\overline{\mathcal{E}}^{\mathrm{A}}_{l}$, 1, $\mathcal{E}^{\mathrm{A}}_{l}$,1)）。同样，要找到（$\mathcal{E}^{\mathrm{D}*}_{l\,1}$, ($\overline{\mathcal{E}}^{\mathrm{A}}_{l}$,1, $\mathcal{E}^{\mathrm{A}}_{l}$,1)），需要先得到（$\overline{\mathcal{E}}^{\mathrm{A}*}_{l\,2}$, ($\mathcal{E}^{\mathrm{D}}_{l}$,1) $\mathcal{E}^{\mathrm{A}*}_{l\,2}$, ($\mathcal{E}^{\mathrm{D}}_{l}$,1)），依此类推。此外，注意虽然 $\mathcal{E}^{\mathrm{D}*}_{l\,3}$, 不属于防御者的策略，但攻击者仍需要它来获得（$\overline{\mathcal{E}}^{\mathrm{A}*}_{l\,3}, , \mathcal{E}^{\mathrm{A}*}_{l\,3}, $）。因此，超出由其时域长度 h^{D} 所界定的防御者能力范围时，我们假设攻击者采用一种策略，模拟时域更长的防御者的最佳响应，即效用函数 $-U^{\mathrm{A}}_{l}$ 的一部分。本文中，我们将（$\overline{\mathcal{E}}^{\mathrm{A}*}_{l\,\alpha}, , \mathcal{E}^{\mathrm{A}*}_{l\,\alpha}, , \mathcal{E}^{\mathrm{D}*}_{l\,\alpha}, $）表示为时域长度较长的参与者所对应的最优策略，将（$\overline{\mathcal{E}}^{\mathrm{A}\prime}_{l\,\alpha}, , \mathcal{E}^{\mathrm{A}\prime}_{l\,\alpha}, , \mathcal{E}^{\mathrm{D}\prime}_{l\,\alpha}, $）表示为时域长度较短的参与者所对应的最优策略。

在步骤 $\alpha \le h^{\mathrm{D}} (< h^{\mathrm{A}})$ 中，防御者假定攻击者的最优边（例如式 (12) 中的边）是基于防御者的效用函数确定的，而该函数仅包含 h^{D} 步骤。根据攻击者的判断，防御者的最优策略（例如式 (14) 中的策略）基于防御者对攻击者最优策略的认知，即 ($\overline{\mathcal{E}}^{\mathrm{A}\prime}_{l\,2}$, , $\mathcal{E}^{\mathrm{A}\prime}_{l\,2,}$)，因为防御者无法预见攻击者在 h^{D} 之后的策略。对于攻击者而言，由于它也能够计算出防御者的最优策略（这是因为其 h^{A} 更长），因此攻击者在索引为 $\alpha \le h^{\mathrm{D}}$ 的步骤中的策略（例如式 (13) 和 (15) 中的策略）基于 $\mathcal{E}^{\mathrm{D}\prime}_{l\,\alpha,}$。

在这种斯塔克尔伯格博弈设定中，防御者的策略空间取决于同一步骤中攻击者的策略，也就是说，防御者只能恢复遭到攻击的边。因此，防御者可能无法完美地执行其策略。具体而言，当攻击者改变自身策略时，防御者可能无法恢复其中一些已分配能量的 $\mathcal{E}^{\mathrm{D}}_{k}$。在这种情况下，防御者只会在能够恢复的边上执行策略。然而，如上所述，恢复未遭攻击的边仍会消耗能量，这一点是合理的。这对于后文讨论一致性和聚类十分重要。

该示例中参与者的决策过程如图 3 的博弈树所示，其中蓝线表示 $h^{\mathrm{A}} = h^{\mathrm{D}} = 3$ 的均衡路径，即参与者通过逆向归纳得出的策略。绿线表示 $h^{\mathrm{A}} = h^{\mathrm{D}} = 2$ 的均衡路径，洋红线表示 $h^{\mathrm{A}} = 3, h^{\mathrm{D}} = 2$ 的均衡路径。在步骤 2 中，攻击者

g 替换

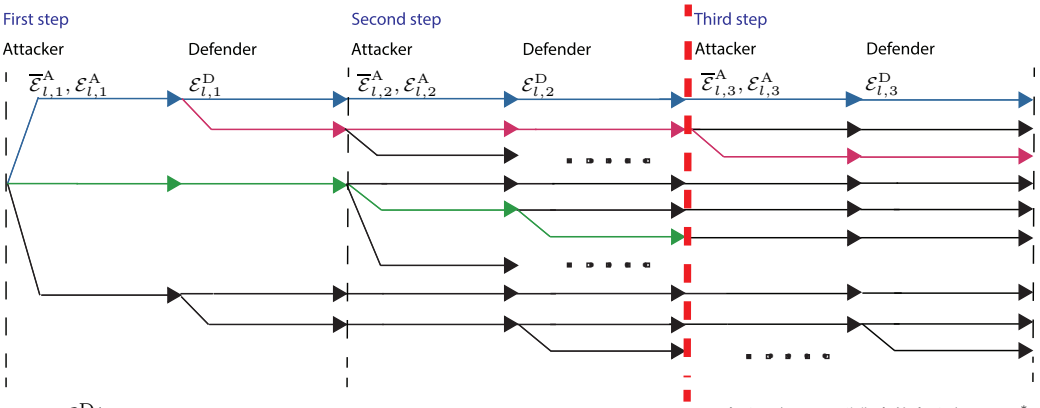


图 3. 不同 h^{A} 和 h^{D}下的扩展式博弈。竖直虚线表示博弈的不同步骤，而红色虚线表示不同玩家时域长度的边界。超出此限制的优化仅由时域较长的玩家（本例中为攻击者）完成。

假设 $\mathcal{E}_{l}^{\mathrm{D}*},{}_{2}$ 来自对 $h^{\mathrm{A}}=3$ 的效用。 $h^{\mathrm{A}}<h^{\mathrm{D}}$ 的情况也可以类似地描述。

B. 非均匀博弈周期

在本小节中，我们将讨论扩展到攻击者和防御者分别采用非均匀博弈周期 T_{A} 和 T^{D} 的情况。相应的决策索引为 l^{A} 和 l^{D}，分别包含 α^{A} 和 α^{D} 个时间步。由于我们考虑完全信息情形，双方参与者都知道这些周期 T_{A} 和 T^{D}。为确保双方参与者在任意 k 都能获得各自的策略，我们设定 $T_{\mathrm{A}}\leq h^{\mathrm{A}}$ 和 $T^{\mathrm{D}}\leq h^{\mathrm{D}}$。

非均匀博弈周期下的博弈如图1所示。黄色矩形表示一个博弈中的决策过程集合，该过程遵循特定模式。每经过 T_{A} 和 T^{D} 的最小公倍数，双方参与者便进行一次博弈，即同时更新各自的策略，记作 $\mathrm{lcm}(T_{\mathrm{A}}\,T_{\mathrm{D}})$；在图1中，每经过 2 个时间步便进行一次博弈。采用这种表述方式，预期参与者在博弈周期较短时会有更好的表现。

从图1可以看出，参与者可能不会同时决定策略。例如，在时间 $k=2$，只有攻击者更新其策略，而防御者由于 T^{D} 较长而不更新。由于 T_{A} 和 T_{D} 为双方参与者所知，在 $k=2$，攻击者会根据防御者此前在 $k=0$时确定的策略来决定自己的策略。此外，由于图1中的 $h^{\mathrm{A}}=6$，此处具备向前计算三个时间步能力的攻击者，只能预见防御者第二个决策过程向前的三个时间步。

由于非均匀博弈周期使参与者在不同时间决定策略，我们使用不同的决策索引 l^{A} 和 l^{D}，分别表示攻击者和防御者在时间 $(l^{\mathrm{A}}-1)T_{\mathrm{A}}$ 和 $(l^{\mathrm{D}}-1)T^{\mathrm{D}}$ 进行的决策过程；在这些过程中，参与者将效用函数 (7) 和 (8) 最大化。请注意，参与者的这些索引取值不同，也可能对应同一时间步；例如，在图1中，$l^{\mathrm{A}}=2,\alpha^{\mathrm{A}}=1$ 和 $l^{\mathrm{D}}=1,\alpha^{\mathrm{D}}=3$ 都对应于 $k=2$。

由于参与者在不同时间决定各自的策略，因此每个时间对应的优化问题也不同。例如，在图1所示的情形中，攻击者在时间 $k=2$ 的最优策略如下（受篇幅限制，仅展示部分步骤）：

- 步骤6（$k=7$，防御者的步骤1）：$\mathcal{E}_{\mathrm{D}3\,2,}^{*}$ $(\overline{\mathcal{E}}_{\mathrm{A}\,2,\ 6,}$ $\mathcal{E}_{\mathrm{A}2,\ 6})\in\arg\max_{\mathcal{E}_{3,2}^{\mathrm{D}}} -U_{\mathrm{A}2,6},$ (16)

$$(\overline{\mathcal{E}}_{2,6}^{\mathrm{A}*}(\mathcal{E}_{2,4}^{\mathrm{D}}),\mathcal{E}_{2,6}^{\mathrm{A}*}(\mathcal{E}_{2,4}^{\mathrm{D}}))\in\arg\max_{(\overline{\mathcal{E}}_{2,6}^{\mathrm{A}},\mathcal{E}_{2,6}^{\mathrm{A}})} U_{2,6}^{\mathrm{A}}(\mathcal{E}_{3,2}^{\mathrm{D}*}),\tag{17}$$

- Step 5 ($k=6$, Step 4 for defender):

$$\mathcal{E}_{2,4}^{\mathrm{D}\prime}(\overline{\mathcal{E}}_{2,5}^{\mathrm{A}},\mathcal{E}_{2,5}^{\mathrm{A}})\in\arg\max_{\mathcal{E}_{2,4}^{\mathrm{D}}} U_{2,4}^{\mathrm{D}},\tag{18}$$

$$(\overline{\mathcal{E}}_{2,5}^{\mathrm{A}*}(\mathcal{E}_{2,3}^{\mathrm{D}}),\mathcal{E}_{2,5}^{\mathrm{A}*}(\mathcal{E}_{2,3}^{\mathrm{D}}))\in\arg\max_{(\overline{\mathcal{E}}_{2,5}^{\mathrm{A}},\mathcal{E}_{2,5}^{\mathrm{A}})} U_{2,5}^{\mathrm{A}}(\mathcal{E}_{2,4}^{\mathrm{D}\prime}),\tag{19}$$

$$\vdots$$

- 步骤 2（$k=3$，防御者的步骤 1）：

$$\mathcal{E}_{2,1}^{\mathrm{D}\prime}(\overline{\mathcal{E}}_{2,2}^{\mathrm{A}},\mathcal{E}_{2,2}^{\mathrm{A}})\in\arg\max_{\mathcal{E}_{2,1}^{\mathrm{D}}} U_{2,1}^{\mathrm{D}}(\overline{\mathcal{E}}_{2,3}^{\mathrm{A}\prime},\mathcal{E}_{2,3}^{\mathrm{A}\prime}),\tag{20}$$

$$(\overline{\mathcal{E}}_{2,2}^{\mathrm{A}*}(\mathcal{E}_{2}^{\mathrm{D}}),\mathcal{E}_{2,2}^{\mathrm{A}*}(\mathcal{E}_{2}^{\mathrm{D}}))\in\arg\max_{(\overline{\mathcal{E}}_{2,2}^{\mathrm{A}},\mathcal{E}_{2,2}^{\mathrm{A}})} U_{2,2}^{\mathrm{A}}(\mathcal{E}_{2,1}^{\mathrm{D}\prime}),\tag{21}$$

- Step 1 ($k=2$):

$$(\overline{\mathcal{E}}_{2,1}^{\mathrm{A}*},\mathcal{E}_{2,1}^{\mathrm{A}*})\in\arg\max_{(\overline{\mathcal{E}}_{2,1}^{\mathrm{A}},\mathcal{E}_{2,1}^{\mathrm{A}})} U_{2}^{\mathrm{A}}(\mathcal{E}_{2}^{\mathrm{D}}).\tag{22}$$

攻击者无法计算超过 $h^{\mathrm{A}}=6$ 个时间步的未来，因此在上文的 (16) 式中，攻击者将使用自己的效用函数 $U_{l^{\mathrm{A}}}^{\mathrm{A}}$ 来估计防御者在 $l^{\mathrm{D}}=3$ 时的最优边。根据 k（模 T_{D}）$\neq 0$，防御者不会做出新的决策，因此会改为采用前一时间获得的策略，例如在 $k=0$ 时获得的 $\mathcal{E}_{2}^{\mathrm{D}}$。因此，博弈周期较短的参与者（本例中为攻击者）可以通过改变策略获益；例如，在上述情形中，攻击者可以通过改变 $\mathcal{E}_{2}^{\mathrm{A}}$ 获益，从而避免防御者在 $\mathcal{E}_{2}^{\mathrm{D}}$ 进行恢复，而该恢复已被设定，无法更改。在此博弈中，假设时域长度较长的参与者能够正确回忆起对手与自身策略同时确定的策略，即在 $i\mathrm{lcm}(T^{\mathrm{A}},T_{\mathrm{D}})$ $i\in\mathbb{N}_{0}$ 时确定的策略。例如，本例中的攻击者知道 $\mathcal{E}_{2}^{\mathrm{D}}$，因为它是在时间 $k=0$ 决定的，也就是与攻击者的决策时间相同。只要较短时域参与者的整个时域落在较长时域参与者的同一时域内，较长时域参与者也能知道该较短时域参与者过去确定、但并非同时确定的策略。

IV. 共识分析

现在，我们考察博弈结构和参与者的能量约束对一致性的影响。我们注意到，先前的工作 [9] 研究了非均匀时域下参与者的表现，但没有讨论一致性和簇形成，而且效用的形式也有所不同。

我们首先研究防御者的最优策略。

引理 4.1：存在一个无限序列 $\bar{l}^{\mathrm{D}} = \{\bar{l}_1^{\mathrm{D}}, \bar{l}_2^{\mathrm{D}}, \ldots\}$，由防御者的决策索引构成，其中 $\bar{l}_{i+1}^{\mathrm{D}} > \bar{l}_i^{\mathrm{D}}$ 且 $\bar{l}_i^{\mathrm{D}} \in \mathbb{N}$，使得在第 $\bar{l}_i^{\mathrm{D}}$ 个决策过程中，防御者在第一步的最优策略是恢复 $\mathcal{E}^{\mathrm{D}}_{\bar{l}_i^{\mathrm{D}},1} \neq \emptyset$，只要 $\mathcal{E}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}},1} \neq \emptyset$。

证明：注意，如果防御者在博弈的第一步没有从非零常规攻击中恢复，最坏的情况是智能体状态最终会收敛到不同的值。因此，攻击者需要持续攻击连接状态不同的智能体的边，使它们与其他簇分离。假设在索引为 $\bar{l}_i^{\mathrm{D}}$ 的博弈中，智能体被分成若干簇。在这里，我们可以验证以下关系始终成立：

$$z_{\bar{l}_i^{\mathrm{D}},1}(\cdot, \tilde{\mathcal{E}}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}},1}, \emptyset) \geq z_{\bar{l}_i^{\mathrm{D}},1}(\cdot, \tilde{\mathcal{E}}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}},1}, \mathcal{E}^{\mathrm{D}}_{\bar{l}_i^{\mathrm{D}},1}), \tag{23}$$

其中 $\tilde{\mathcal{E}}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}}}$ 是分隔状态不同的智能体的边。在这种情况下，攻击者需要攻击 $\tilde{\mathcal{E}}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}}}$，以阻止智能体达成一致性。请注意，(23) 是 $z_k(\emptyset, \mathcal{E}, \emptyset) \geq z_k(\emptyset, \mathcal{E}, \mathcal{E}_k^{\mathrm{D}})$ 的一个更具体形式，其中攻击所有边总能使 z_k 达到最大值。

根据 (23)，防御者总能从恢复非零数量的边中获益，因为如果防御者不恢复任何边，则 $z_{\bar{l}_i^{\mathrm{D}},2} = z_{\bar{l}_i^{\mathrm{D}},1}$，这会带来最低的效用值。由于如果防御者不进行恢复，$c(\mathcal{G}^{\mathrm{D}}_{\bar{l}_i^{\mathrm{D}},1})$ 也会达到最低值，因此在第 $\bar{l}_i^{\mathrm{D}}$ 场博弈的第一步，防御者恢复非零数量边时的效用总是优于不恢复任何边的情况。由于防御者在每个时间点都会持续获得 ρ^{D} 数量的能量，因此在后续编号为 $\bar{l}_{i+1}^{\mathrm{D}}$、$\bar{l}_{i+2}^{\mathrm{D}}$ 等的博弈中，防御者采取的这一行动也相同。 ■

以下两个命题给出了必要条件：智能体若要在无限长的时间内分散为多个簇且不达成一致性，必须满足这些条件。

命题4.2：一致性不发生的必要条件是 $\rho^{\mathrm{A}}/\beta^{\mathrm{A}} \geq \lambda$，其中 λ 是 $\mathcal{G}$ 的连通度。

证明：我们注意到，如果防御者没有进行任何恢复（$\mathcal{E}_k^{\mathrm{D}} = \emptyset$），那么攻击者在任意时间 k 都必须使用常规信号攻击至少 λ 条边，才能使 $\mathcal{G}_k^{\mathrm{D}}$ 不连通。如果攻击者始终使用普通干扰信号攻击 λ 条边，则能量约束 (3) 变为 $(\beta^{\mathrm{A}}\lambda - \rho^{\mathrm{A}})k \leq \kappa^{\mathrm{A}}$。因此，对于所有 k，条件 $\rho^{\mathrm{A}}/\beta^{\mathrm{A}} \geq \lambda$ 都必须成立。 ■

命题 4.3：如果满足以下两个条件之一，则不发生一致性的必要条件是 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq \lambda$

(a) $b = 0$, $h^{\mathrm{D}} \geq h^{\mathrm{A}}$ and $\mathrm{lcm}(T^{\mathrm{A}}, T^{\mathrm{D}}) = T^{\mathrm{A}}$; or
(b) $b = 0$ and $T^{\mathrm{D}} = 1$.

证明：我们采用逆否命题进行证明；具体而言，我们证明在满足指定条件时，一致性总会发生 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} < \lambda$。

我们首先假设攻击者试图始终强力攻击 λ 条边，以使图 $\mathcal{G}_k^{\mathrm{D}}$ 不连通。根据 (3)，攻击者在时间 k 的能量约束变为 $(\overline{\beta}^{\mathrm{A}}\lambda - \rho^{\mathrm{A}})k \leq \kappa^{\mathrm{A}}$。如果 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} < \lambda$，那么当 k 较大时，该不等式不成立，因为左侧会变为正值，而 κ^{A} 是有限的。因此，如果 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} < \lambda$，攻击者就无法始终强力攻击 λ 条边，只能改用普通干扰信号攻击来使图不连通。由此，根据引理 4.1，存在一个时间区间，在此期间防御者始终进行恢复，即 $\mathcal{E}^{\mathrm{D}}_{\bar{l}_i^{\mathrm{D}},1} \neq \emptyset$，$i = 1, 2, \ldots$，在给定 $\mathcal{E}^{\mathrm{A}}_{\bar{l}_i^{\mathrm{D}},1} \neq \emptyset$ 的情况下是最优的。请注意，由于该策略用于博弈的第一步，因此始终会采用该策略。

由（8）中的效用函数可知，给定 $b = 0$ 时，智能体彼此越接近，防御者获得的效用越高。这意味着，当需要恢复的边数非零时（在上述索引为 $\bar{l}_i^{\mathrm{D}}$ 的博弈的第一步），防御者会恢复连接距离较远的智能体的边。具体而言，对于决策索引序列 $[\bar{l}_i^{\mathrm{D}}, \bar{l}_{i+1}^{\mathrm{D}}]$，存在一个时间步，使得 $U_b^{\mathrm{D}}(\mathcal{E}^{\mathrm{D}}_{l^{\mathrm{D}},1} = \mathcal{E}_1) \geq U_b^{\mathrm{D}}(\mathcal{E}_2)$，其中 $\mathcal{E}_1$ 和 $\mathcal{E}_2$ 分别表示连接状态距离较远和较近的智能体的边集。由于智能体通过一致性协议（1）进行通信后，其状态会逐渐接近，因此，如果恢复边所连接的智能体的状态 $\mathcal{E}_k^{\mathrm{D}}$ 变得足够接近，防御者就会选择恢复不同的边。

对于情形 (a)，在 $h^{\mathrm{D}} \geq h^{\mathrm{A}}$ 且 $\mathrm{lcm}(T_{\mathrm{A}}\, T_{\mathrm{D}}) = T^{\mathrm{A}}$ 的条件下，可以保证防御者不会因恢复边而浪费能量，因为更长的时域长度使防御者能够准确预测攻击者的行动。博弈周期 $\mathrm{lcm}(T_{\mathrm{A}}\, T_{\mathrm{D}}) = T^{\mathrm{A}}$ 意味着攻击者绝不会单独更新其决策，即攻击者更新时防御者也会更新，从而阻止攻击者单方面改变策略以规避防御者计划的恢复操作。另一方面，对于情形 (b)，拥有 $T^{\mathrm{D}} = 1$ 的防御者能够完美观察攻击者的行动，因此可以完全避免浪费能量的可能性。

因此，对于情况（a）和（b），如果 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} < \lambda$，那么存在依赖于 i 的 $j \in \mathbb{N}$，使得在决策索引 $[\bar{l}_i^{\mathrm{D}}, \bar{l}_{i+j(i)}^{\mathrm{D}}]$ 的范围内，由每个图的边所组成的图的并，即包含各图边的并的图 $(\mathcal{V}, \bigcup((\mathcal{E} \setminus (\overline{\mathcal{E}}_k^{\mathrm{A}} \cup \mathcal{E}_k^{\mathrm{A}})) \cup \mathcal{E}_k^{\mathrm{D}}))$，对于所有 i 都成为连通图。这些区间 $[(\bar{l}_i^{\mathrm{D}} - 1)T^{\mathrm{D}}, (\bar{l}_{i+j(i)}^{\mathrm{D}} - 1)T^{\mathrm{D}}]$，$i = 1, 2, \ldots$ 会无限次出现，因为防御者的能量上界会随时间不断增加。

文献 [18] 表明，采用协议（1）时，若有界时间区间内各图的并图为连通图，则智能体可在时变图中实现一致性。这意味着，若 $(\mathcal{V}, \bigcup((\mathcal{E} \setminus (\overline{\mathcal{E}}_k^{\mathrm{A}} \cup \mathcal{E}_k^{\mathrm{A}})) \cup \mathcal{E}_k^{\mathrm{D}}))$ 在所有 i 上均于 $[\bar{l}_i^{\mathrm{D}}, \bar{l}_{i+j(i)}^{\mathrm{D}}]$ 连通，则可实现一致性。

■

下一个结果给出了彻底阻止一致性、使所有智能体彼此分离的条件。该结果表明，攻击者必须能够在所有时间对所有边发动强攻击。

命题 4.4：所有智能体在无限时间内无法实现一致性的一个充分条件是 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq |\mathcal{E}|$。

证明：由于 $z_{l^{\mathrm{A}},\alpha^{\mathrm{A}}}(\mathcal{E},\emptyset,\emptyset) \geq z_{l^{\mathrm{A}},\alpha^{\mathrm{A}}}(\overline{\mathcal{E}}^{\mathrm{A}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}},\mathcal{E}^{\mathrm{A}}_{l^{\mathrm{A}}\ \alpha^{\mathrm{A}}}\ \mathcal{E}^{\mathrm{D}}_{l^{\mathrm{A}}\ \alpha^{\mathrm{A}}} \quad c\ \mathcal{V}\ \emptyset\ \geq c\ \mathcal{V}\ \mathcal{E}\setminus\ \overline{\mathcal{E}}^{\mathrm{A}}_{l^{\mathrm{A}}\ \alpha^{\mathrm{A}}}\cup\mathcal{E}^{\mathrm{A}}_{l^{\mathrm{A}}\ \alpha^{\mathrm{A}}}\cup,,,)$ 且 $((,))((,(,,)(\mathcal{E}^{\mathrm{D}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}}\cap\mathcal{E}^{\mathrm{A}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}})))$ 成立，因此当攻击者攻击所有边时，函数 $U\mathrm{A}_{l^{\mathrm{A}}}$ 取最大值，即 $\overline{\mathcal{E}}^{\mathrm{A}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}} = \mathcal{E}$ 或 $\mathcal{E}^{\mathrm{A}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}} = \mathcal{E}$。在 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq |\mathcal{E}|$ 条件下，攻击者可以在任意时间 k 使用强干扰信号攻击 $\mathcal{G}$ 的所有边。因此，攻击者将在第 l^{A} 个决策过程的第 α^{A} 步强力攻击 $\mathcal{E}$，即 $\overline{\mathcal{E}}^{\mathrm{A}*}_{l^{\mathrm{A}}}$，$_{\alpha^{\mathrm{A}}} = \mathcal{E}$，这也会阻止防御者恢复任何边，即 $\mathcal{E}^{\mathrm{D}}_{l^{\mathrm{A}},\alpha^{\mathrm{A}}} = \emptyset$，对所有 $l^{\mathrm{A}},\alpha^{\mathrm{A}}$ 均如此。因此，攻击者将始终强力攻击 $\mathcal{E}$，使智能体分成 n 个簇。

■

注释 4.5：到目前为止，我们已在假设恢复失败（即 $\mathcal{E}^{\mathrm{D}}_k \setminus \mathcal{E}^{\mathrm{A}}_k$）仍会消耗能量，并按照式 (4) 表述的基础上，得到了必要条件和一个充分条件。对于防御者不会因恢复失败而损失能量的情形，其条件可以用更直观的形式表示。在这种情况下，我们假设防御者的能量消耗满足 $\sum_{m=0}^{k}\beta^{\mathrm{D}}|\mathcal{E}^{\mathrm{D}}_m \cap \mathcal{E}^{\mathrm{A}}_m| \leq \kappa^{\mathrm{D}} + \rho^{\mathrm{D}}k$。那么，防止一致性的必要条件为：若 $b=0$，则为 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq \lambda$；否则为 $\rho^{\mathrm{A}}/\beta^{\mathrm{A}} \geq \lambda$。充分条件可轻易得到，为 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq |\mathcal{E}|$。

注释 4.5 中必要条件和充分条件所用的术语，即 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq \lambda$、

$\rho^{\mathrm{A}}/\beta^{\mathrm{A}} \geq \lambda$

和 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq |\mathcal{E}|$，与上文命题 4.2 和 4.4 中的相同，因为正如这些命题的证明中所讨论的，这些结果的条件是根据攻击者的能力而非防御者的能力推导得出的。对于 $b=0$ 的情形，与命题 4.3 不同的是，在无能量浪费情形下，防御者的时域参数不再影响获得更严格必要条件的要求。这意味着，在能量约束 (4) 下，防御者会变得更弱，因为防止一致性的必要条件变得不那么严格。

注4.6：对于注释2.1中所描述的节点攻击情形，命题4.2和4.3中防止一致性所需的必要条件 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq \lambda$ 和$\rho^{\mathrm{A}}/\beta^{\mathrm{A}} \geq \lambda$ 分别改为 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}_{\mathcal{V}} \geq 1$ 和 $\rho^{\mathrm{A}}/\beta^{\mathrm{A}}_{\mathcal{V}} \geq 1$，因为攻击者只需隔离一个智能体即可阻止一致性。

V. 聚类分析

本节中，我们推导关于智能体在无限时间内形成的簇数量的一些结果。由上文的命题 4.4 可知，如果 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} \geq |\mathcal{E}|$，攻击者就可以通过在所有时间强力攻击所有边，使簇的数量达到 n。因此，对于下述结果，我们考虑以下情形

其中 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}} < |\mathcal{E}|$。

命题 5.1：定义一个向量 $\Theta \in \mathbb{R}^{|\mathcal{E}|}$，其元素为$\Theta_i := \max_{|\mathcal{E}^{\mathrm{A}}|=i}\overline{n}((\mathcal{V},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}))$，其中 $\overline{n}((\mathcal{V},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}))$是图$(\mathcal{V},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}})$的群组数量。那么无限时间时形成的簇数量有如下上界：

- $\Theta_{\lfloor\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}\rfloor}$ if either
  (a) $b=0$, $h^{\mathrm{D}} \geq h^{\mathrm{A}}$ and $\mathrm{lcm}(T^{\mathrm{A}},T^{\mathrm{D}}) = T^{\mathrm{A}}$; or
  (b) $b=0$ and $T^{\mathrm{D}}=1$,
- $\Theta_{\min\{|\mathcal{E}|,\lfloor\rho^{\mathrm{A}}/\beta^{\mathrm{A}}\rfloor\}}$ otherwise.

证明：向量 Θ 的第 i个元素由给定受攻击边数量为 i时形成的群组最大数量 $\overline{n}((\mathcal{V},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}))$构成。由于需要持续攻击某些边才能将智能体划分为簇，因此无限时间时形成的簇数量绝不会超过在相同数量的强攻击边下任意时间 k 的群组最大数量。

证明的其余部分可由命题 4.3 得出。对于其中（a）和（b）所述的情形，$\lfloor\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}\rfloor$ 是在所有时间都能被强力攻击的边数上限。如果情形（a）或（b）的条件不成立，则边数上限为 $\lfloor\rho^{\mathrm{A}}/\beta^{\mathrm{A}}\rfloor$。因此，给定已知的图拓扑 $\mathcal{G}$ 及其 $|\mathcal{E}|$ 条边，我们可以推知，取决于时域长度和博弈周期 h^{A}、h^{D}、T^{A} 和T^{D}的取值，$\Theta_{\lfloor\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}\rfloor}$ 或 $\Theta_{\min\{|\mathcal{E}|,\lfloor\rho^{\mathrm{A}}/\beta^{\mathrm{A}}\rfloor\}}$ 中的值给出了无限时间下簇的最大数量。

■

我们通过将图2中的图的未受攻击版本视为 $\mathcal{G}$，来说明命题 5.1 的结果。在该图中，向量 Θ 为 $\Theta = [2,2,3,4]^{\mathrm{T}}$。假设 $\beta^{\mathrm{A}} = 1$、$\overline{\beta}^{\mathrm{A}} = 2$、$\rho^{\mathrm{A}} = 3.5$。在此情况下，如果 $b=0$和 h^{A}、h^{D}、T^{A}、T_{D}s满足命题 5.1 中情形（a）或（b）的条件，那么簇的最大数量为$\Theta_1 = 2$；否则，簇的最大数量为$\Theta_3 = 3$。当时域长度和博弈周期的取值满足命题 5.1 中情形（a）或（b）的条件时，可知对每个 i，$\Theta_{i+1} \geq \Theta_i$ 都成立。这是因为在条件（a）或（b）下，防御者更强，因此攻击次数增加时，攻击者可能无法形成更多的簇。

注释 5.2：对于注释2.1中描述的节点攻击情形，命题5.1中簇的上界里的向量 Θ 变为元素为 $\Theta_{\mathcal{V},i}$ 的 $\Theta_{\mathcal{V}} \in \mathbb{R}^{|\mathcal{V}|} := \max_{|\mathcal{V}^{\mathrm{A}}|=i}\overline{n}_{\mathcal{V}}((\mathcal{V}\setminus\mathcal{V}^{\mathrm{A}},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}_{\mathcal{V}}))$，其中 $\overline{n}_{\mathcal{V}}((\mathcal{V}\setminus\mathcal{V}^{\mathrm{A}},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}_{\mathcal{V}}))$ 是图 $(\mathcal{V}\setminus\mathcal{V}^{\mathrm{A}},\mathcal{E}\setminus\mathcal{E}^{\mathrm{A}}_{\mathcal{V}})$ 的群组数量，$\mathcal{E}^{\mathrm{A}}_{\mathcal{V}}$是与智能体 $\mathcal{V}^{\mathrm{A}}$ 相邻的所有边的集合。假设 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}_{\mathcal{V}} < |\mathcal{V}|$，则项 $\Theta_{\lfloor\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}\rfloor}$和 $\Theta_{\lfloor\rho^{\mathrm{A}}/\beta^{\mathrm{A}}\rfloor}$分别变为 $\Theta_{\mathcal{V},\lfloor\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}_{\mathcal{V}}\rfloor}$ 和 $\Theta_{\mathcal{V},\lfloor\rho^{\mathrm{A}}/\beta^{\mathrm{A}}_{\mathcal{V}}\rfloor}$。

VI. 数值示例

本节通过数值仿真展示时域长度和博弈周期对一致速度及簇数量的影响。具体而言，我们证明，即使攻击者的能量参数相对较低，只要防御者时域参数足够小，一致性仍可能受到阻止，甚至

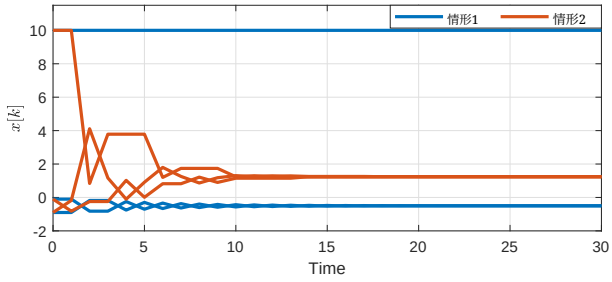


图4. 防御者在情形1和情形2中的智能体状态

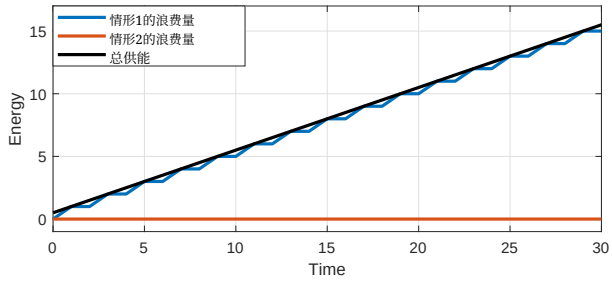


图5. 情形1和情形2中防御者浪费的能量和总供能，即 $\kappa^{\mathrm{D}}+\rho^{\mathrm{D}}k$

与命题4.2和4.3中的条件有关。

考虑一个由三个智能体（$|\mathcal{E}|=2$、$\lambda=1$）组成的简单路径图，其参数如下：

- Case 1: $h^{\mathrm{A}}=3$, $h^{\mathrm{D}}=2$, $T^{\mathrm{A}}=1$, $T^{\mathrm{D}}=2$,
- Case 2: $h^{\mathrm{A}}=h^{\mathrm{D}}=T^{\mathrm{A}}=T^{\mathrm{D}}=2$,

在两种情形中均有 $b=0$、$\kappa^{\mathrm{D}}=\rho^{\mathrm{D}}=0.5$、$\kappa^{\mathrm{A}}=\rho^{\mathrm{A}}=1.5$、$\overline{\beta}^{\mathrm{A}}=2$、$\beta^{\mathrm{A}}=\beta^{\mathrm{D}}=1$ 。注意，由于 $\rho^{\mathrm{A}}/\beta^{\mathrm{A}}<\lambda=1$，攻击者无法在所有时间都对边实施强攻击，以使图保持不连通。因此，为了阻止一致性，攻击者需要不断改变其策略，使恢复无法成功。

在情形1中，攻击者比防御者更强，因为凭借 $h^{\mathrm{A}}>h^{\mathrm{D}}$ 和 $T^{\mathrm{A}}<T_{\mathrm{D}}$，攻击者可以前瞻更长时间，并更频繁地更新策略。因此，在这种情形下，攻击者可以通过取消计划中的攻击，或改用强攻击，来避免在 $\mathcal{E}_k^{\mathrm{D}}$ 上进行恢复。在情形2中，两位参与者的时域参数完全相同，这意味着防御者从不浪费能量，因为攻击者无法单方面改变其策略。

图4展示了智能体状态的演变：在情形1中，智能体被划分为两个簇；与之相反，在情形2中，它们收敛到相同状态。这是因为在情形2中，防御者试图将资源分配给通常未遭受攻击的边，结果耗尽了所有能量，如图5所示。请注意，情形2中的时域长度和博弈周期满足命题 4.3中的要求，从而使必要条件更加严格，即 $\rho^{\mathrm{A}}/\overline{\beta}^{\mathrm{A}}\geq\lambda$ 而非of $\rho^{\mathrm{A}}/\beta^{\mathrm{A}}\geq\lambda$。另一方面，情形1中的时域参数不满足这些要求，因此更容易阻止一致性。

VII. 结论

我们构建了一个用于研究弹性多智能体系统簇形成的双人博弈。参与者会考虑自身行动对未来通信拓扑和智能体状态的影响，并根据滚动时域方法调整策略。我们推导出了智能体形成簇的条件，并讨论了时域参数对可能形成的簇数量及一致性的影响。一般而言，攻击者除了需要有足够的能量发动攻击，还需要较长的时域长度和较短的博弈周期，才能阻止一致性。

参考文献

[1] H. Sandberg、S. Amin 和 K. H. Johansson，“网络化控制系统中的信息物理安全特刊”，IEEE Control Syst. Mag.，第 35 卷，第 20–23 页，2015 年。[2] D. Senejohnny、P. Tesi 和 C. De Persis，“一种用于自触发网络协调的抗干扰算法”，IEEE Trans. Control Netw. Syst.，第 5 卷，第 981–990 页，2018 年。[3] Y. Nugraha、A. Cetinkaya、T. Hayakawa、H. Ishii 和 Q. Zhu，“应用于多智能体一致性的动态弹性网络博弈”，IEEE Trans. Control Netw. Syst.，第 8 卷，259，2021 年 246 页。–。[4] C. Altafini，“具有对抗性交互的网络上的一致性问题”，IEEE Trans. Autom. Control，第 58 卷，第 4 期，第 935–946 页，2013 年。[5] Y. Li、D. E. Quevedo、S. Dey 和 L. Shi，“远程状态估计中的基于信干噪比的拒绝服务攻击：一种博弈论方法”，IEEE Trans. Control Netw. Syst.，第 4 卷，第 632–642 页，2017 年。[6] J. M. Maestre、P. Velarde、H. Ishii 和 R. R. Negenborn，“针对基于拉格朗日的分布式模型预测控制中漏洞的情景防御机制”，Control Eng. Practice，第 114 卷，文章编号 104879，2021 年。[7] M. Zhu 和 S. Martinez，“重放攻击下弹性网络化控制系统的性能分析”，IEEE Trans. Autom. Control，第 59 卷，第 3 期，第 804–808 页，2014 年。[8] Y. Nugraha、A. Cetinkaya、T. Hayakawa、H. Ishii 和 Q. Zhu，“弹性多智能体系统簇形成的滚动时域博弈”，载于 IEEE 决策与控制会议论文集，2021 年，第 4829–4934 页。[9] ——，“非一致时域弹性网络的滚动时域博弈”，European Journal of Control，文章编号 100693，2022 年。[10] R. Lagunoff 和 A. Matsui，“重复协调博弈中的异步选择”，Econometrica，第 65 卷，第 1467–1477 页，1997 年。[11] K. Yoon，“异步重复博弈的民间定理”，Econometrica，第 69 卷，第 191–200 页，2001 年。[12] H. Mo 和 G. Sansavini，“针对不确定攻击，在信息物理系统中动态分配防御资源以最大限度减少未满足需求”，IEEE Trans. Reliability，第 66 卷，1265，2017 年 1253 页。–。[13] M. Abdallah、P. Naghizadeh、A. R. Hota、T. Cason、S. Bagchi 和 S. Sundaram，“以攻击图建模的相互依存系统中的行为与博弈论安全投资”，IEEE Trans. Control Netw. Syst.，第 7 卷，第 1585–1596 页，2020 年。[14] M. L. Katz 和 C. Shapiro，“系统竞争与网络效应”，Journal of Economic Perspective，第 8 卷，第 93–115 页，1994 年。[15] Y. Li、C. A. Courcoubetis、L. Duan 和 R. Weber，“具有网络外部性的点对点共享最优定价”，IEEE/ACM Trans. Netw.，第 29 卷，第 1 期，第 148–161 页，2021 年。[16] X. Gong、L. Duan、X. Chen 和 J. Zhang，“无线网络中社交网络效应与拥塞效应相遇时：数据使用均衡与最优定价”，IEEE J. Sel. Areas Commun.，第 35 卷，第 2 期，第 449–462 页，2017 年。[17] F. Bullo，《网络系统讲义》。Kindle Direct，2019 年。[18] W. Ren 和 R. W. Beard，“动态变化的交互拓扑下多智能体系统的一致性寻求”，IEEE Trans. Autom. Control，第 50 卷，第 5 期，第 655–661 页，2005 年。