arXiv:2603.05023v1 [eess.SP] 5 Mar 2026

基于航迹共识的分布式多目标跟踪中的标签劫持

Helena Calatrava、Shuo Tang、Pau Closas美国马萨诸塞州波士顿东北大学，邮编 *02115*{calatrava.h、tang.shu、closas}@northeastern.edu

摘要——**视场受限的传感器网络中的分布式多目标跟踪（DMTT）通常会受到**标签不一致**问题的影响，即不同节点对同一目标的身份存在分歧。近期的航迹一致性分布式多目标跟踪（TC-DMTT）策略通过基于度量的航迹匹配，强制实现运动学和标签一致，从而缓解了这一问题。然而，这些策略在对抗条件下的表现仍未得到充分研究。本文揭示了TC-DMTT在身份层面的漏洞，并提出**标签劫持**概念：对手通过注入欺骗航迹，破坏整个网络中的目标身份。借鉴雷达中经典拖引欺骗的思想，我们形式化定义了攻击隐蔽性，并推导出一种用于构造此类攻击的基于优化的策略。一个三传感器网络案例研究展示了所提攻击对标签一致性和跟踪精度的影响，并表明该攻击能够成功实施目标冒充。总体而言，本研究凸显了重新思考分布式多目标跟踪框架中共识层鲁棒性的必要性。**

索引词—**多目标跟踪、传感器网络、对抗性攻击、数据关联。**

I. 引言

多目标跟踪（MTT）是一种统计推断问题，旨在估计动态场景中物理目标的数量、它们的轨迹，以及在某些表述中，它们的身份（标签），这些信息均根据含噪声的传感器测量 [1]得出。多目标跟踪起源于20世纪60年代，此后已成为众多感知和决策应用的基础技术 [2]。

现代多目标跟踪系统必须应对过程噪声和测量噪声之外的多种不确定性，包括随时间变化的目标数目、虚警（杂波）、漏检、目标出生与消亡以及数据关联（DA）。受 Mahler 关于有限集统计的研究 [3],启发，随机有限集（RFS）理论应运而生，成为一种数学上优雅的框架，可通过统一的贝叶斯表述 [4], [5]隐式处理这些不确定性。

多目标跟踪中的一个基本问题是信息应如何被处理

多传感器场景中，将所有原始测量值传输至中央处理器通常并不现实 [6]。为解决这一问题，分布式多目标跟踪（*DMTT*）允许每个节点执行局部跟踪，并交换摘要信息以进行融合，交换的信息可以是航迹级状态估计（通常称为航迹间融合），也可以是多目标密度。由于随机有限集框架

本研究部分获得了美国国家科学基金会资助，资助编号为 1845833、2326559 和 2530870。

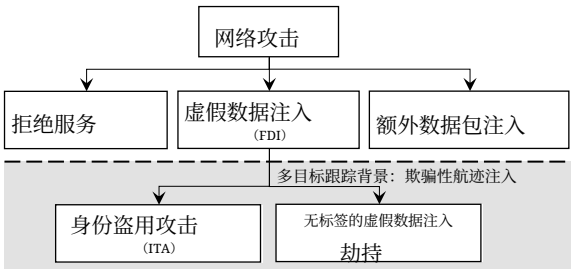


图 1：分布式传感器网络的网络攻击分类 [10]。我们专门研究多目标跟踪场景中的虚假数据注入，重点关注身份威胁。

自然地将单目标分布式估计推广至多目标情形 [7], ，各种基于随机有限集的滤波器（包括带标签形式 [8], ）已被应用于分布式场景，并且主要依赖广义协方差交集（GCI） [9]对概率密度进行对数线性几何平均。

尽管 GCI 融合方法得到了广泛应用，但仍存在两个主要局限。首先，在视场有限且不重叠的条件下，其性能会下降 [11], ，而带标签形式容易出现标签不一致，即不同节点可能为同一物理目标分配不同身份 [12]。其次，更普遍地说，与航迹间融合方法相比，密度级融合会带来显著的计算开销和通信开销 [13]。近期关于航迹一致性分布式多目标跟踪（TC-DMTT） [14], [15]的研究（见第II节）通过增强航迹间融合来应对这些挑战，利用航迹距离度量（通常取自最优子模式分配（OSPA）系列）同时确保运动学一致性和标签共识 [16]。尽管 TC-DMTT 在良性运行条件下表现有效，但其在存在对抗性威胁时的性能仍未得到充分研究。

尽管分布式架构在视场受限时能够提升可观测性，但同时也扩大了其遭受网络攻击的攻击面 [10] ，尤其是在分布式网络中（分类见图1）。本文聚焦于一种虚假数据注入（FDI）场景：攻击者将欺骗航迹[1]，(即不对应任何真实（物理）目标的航迹）注入融合过程。当对手入侵某个传感器的通信信道，并篡改传递给其他未

[1]与多目标跟踪文献中由虚警或数据关联错误造成的虚假航迹不同。

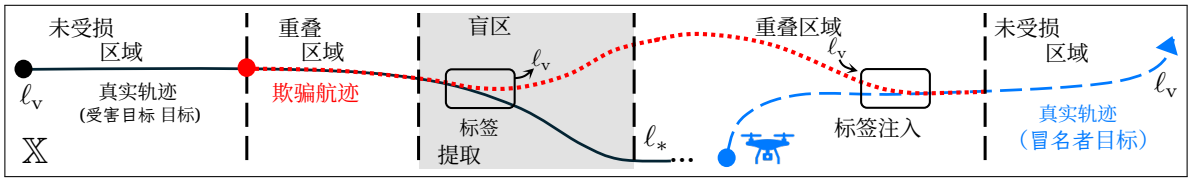


图 2：我们提出标签劫持（提取 + 注入），将其视为拉脱式雷达欺骗在数据层的类比 [17]。欺骗航迹劫持 ℓ_{v} 受害目标的身份，随后该目标便在错误标签下被跟踪 ℓ_{*}。劫持 ℓ_{v}后，欺骗航迹可以自由移动，同时仍与该身份关联；通过与另一个目标对齐， ℓ_{v} 可转移至一个冒名顶替的目标。

被攻陷节点。这使受影响的节点变得不可靠，并破坏了分布式融合过程。

只要真实目标仍能被准确跟踪，在某些应用中，仅仅存在欺骗航迹或许可以容忍。然而，当目标身份遭到篡改时，就会出现更为严重的故障模式。如图2所示，在标签劫持情况下，融合过程会继续跟踪真实目标，但使用的是错误标签，由此产生了我们所称的身份盗用攻击（*ITA*），并破坏航迹连续性。我们还进一步表明，这种威胁模型与雷达系统中的欺骗式干扰之间存在有理有据的对应关系 [17]。

通过施加航迹距离约束来强制实现标签共识，TC-DMTT架构引入了尚未得到充分探索的身份层面漏洞。本文表明，基于度量的航迹匹配阶段构成了一个关键攻击面，可被用于实施标签劫持。我们的主要贡献如下：*(i)* 我们提出航迹共识身份盗用攻击（TC-ITA），通过对航迹共识进行对抗性操纵来利用这一漏洞；*(ii)* 将TC-ITA视为经典拉脱式雷达欺骗在数据层的类比，据此定义了隐蔽性概念，并提出一种基于模型预测控制（MPC）的策略来生成此类攻击；以及*(iii)*我们在模拟的分布式跟踪场景中展示了隐蔽TC-ITA能够成功劫持身份[2]。

本文其余部分组织如下。第II节回顾TC-DMTT框架；第III 节讨论针对这些框架的对抗机会，并介绍所提出的攻击模型；第IV 节给出结果；第V 节总结本文。

II. 背景：基于航迹共识的分布式多目标跟踪

本节回顾航迹共识分布式多目标跟踪框架，我们统称为*TC-DMTT*。其分布式架构和航迹匹配度量分别概括于第II-A 节与II-C节。TC-DMTT将经典航迹间融合加以扩展，通过航迹共识解决视场有限且不重叠的网络中的标签不一致问题，如第II-B节所述。

符号说明：粗体符号表示带标签的量；大写符号表示集合值量。我们的符号体系主要沿用 [14] 中的工作。

[2]代码可在https://github.com/hcalatrava/dmtt-adversarial获取

A. 系统模型

融合以完全分布式的方式进行，无需了解网络拓扑或其他节点的视场。此外，由于融合阶段仅处理报告的轨迹估计，且不会修改本地多目标密度，因此 TC-DMTT 不受各节点所采用的具体本地多目标跟踪算法的影响。

1）网络拓扑：分布式传感器网络被建模为无向图 $G = (\mathcal{N}, \mathcal{A})$，其中 $\mathcal{N}$表示节点集合，$\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ 表示双向通信链路集合。两个节点 a，$b \in \mathcal{N}$ 若处于通信范围 [10]内，则可直接通信。节点 a 的邻居集定义为 $\mathcal{N}^{(a)} \triangleq \{b \in \mathcal{N} \mid b \neq a, \text{且 } (a, b) \in \mathcal{A}\}$。每个节点均配备视场有限、受虚警和检测不确定性影响的传感器，以及本地处理能力。

2）消息交换与全局标记：在时间 k，每个节点 a都会向其他节点传输其本地带标记的多目标状态估计$\mathbf{X}_k \subset \mathbb{X} \times \mathbb{L}$，其中每个元素$\mathbf{x}_k = (x_k, \ell_k)$由运动学状态 $x_k \in \mathbb{X}$ 和一个本地唯一标签 $\ell_k \in \mathbb{L}$组成。由于不同节点的本地标签无法相互比较，因此通过为每个标签附加发送节点的身份，得到全局标记方案，形成 $(\ell_k, a) \in \mathbb{I} = \mathbb{L} \times \mathcal{N}$。每个节点都会存储从各发送邻居节点 a 接收的带全局标签的多目标状态估计，记为$\mathbf{X}_k^{(a)} \subset \mathbb{X} \times \mathbb{I}$。

3）航迹形成：航迹是在对象被判定为存在的时间索引子集$\mathcal{D}^{(t)} \subseteq \mathcal{K}$ 上，按时间顺序排列的单目标状态估计序列。每个节点 a都会在有限时间窗口 $\mathcal{K} \triangleq \{1, \ldots, K\}$ 内，分别为自身的估计值$\mathbf{X}_k^{(a)}$ 以及从邻居节点 {接收的估计值$\mathbf{X}_k^{(b)}\}_{b \in \mathcal{N}(a)}$独立维护本地航迹记忆。对于每个源 $b \in \mathcal{N}(a) \cup \{a\}$，通过连接不同时刻具有相同全局标签的状态，构建相应的带全局标签航迹集合$\mathbf{T}^{(b)} \subset \mathbb{T} \times \mathbb{I}$。

B. 航迹共识

TC-DMTT中的航迹共识由两个阶段组成：*(1)*运动学一致性，用于对齐并融合航迹状态，以及*(2)* 标签共识，用于在网络中传播身份一致关系。

1）运动学共识：在时间 k，节点 a 可访问其自身的活动航迹集 $\mathbf{T}^{(a)}$ 以及来自其邻居的航迹集，即 {$\mathbf{T}^{(b)}\}_{b \in \mathcal{N}(a)}$。此处，活动 指已声明存在且其状态估计值可用于融合的航迹。随后，使用图3中的双节点流程进行成对运动学共识：对于每个邻居 $b \in$

$\mathcal{N}(a)$，节点 a 对航迹对 $\{\mathbf{T}^{(a)}、\mathbf{T}^{(b)}\}$执行此操作，从而生成一组中间共识结果 $\{\mathbf{X}^{(a,b,\mathrm{con})}\}_{b\in\mathcal{N}(a)}$。随后，依次合并这些中间输出，以形成节点 a处的完整（网络）运动学共识，记为$\mathbf{X}^{a,\ \mathrm{con}}$ （详见 [14]，[15] ）。

我们现在描述轨迹集 $\mathbf{T}^{(a)}$ 与$\mathbf{T}^{(b)}$ 之间的双节点（成对）运动学共识过程的细节，该过程由轨迹匹配阶段以及随后的运动学融合组成（图 3）。轨迹匹配的作用是回答以下问题：哪条航迹（位于 $\mathbf{T}^{(a)}$ ）对应于哪条航迹（位于 $\mathbf{T}^{(b)}$）？首先，使用航迹距离度量计算两个轨迹集之间的最优指派，从而得到航迹之间的一一匹配。我们在第 II-C 节回顾最常见的轨迹度量选择。然后，使用截断距离 c对所得配对进行筛选（即，仅将代价低于 c的配对视为有效匹配）。随后，对每个已匹配配对进行融合，以产生单一轨迹估计，通常通过对各个状态估计进行加权组合来实现，其中权重反映每个节点的相对置信度。未匹配航迹可根据启发式准则（例如，最小轨迹长度）予以保留 [14], [15]。

2）标签共识：在成对运动学共识阶段识别出的已匹配航迹（图3）被视为局部身份一致关系，并通过标签共识机制在网络中进行传递闭包扩展 [14]，[15]。我们省略该阶段的具体内容，因为所提出的攻击会通过受损的轨迹匹配导致身份篡改，随后这种篡改会由标签共识机制传播（见备注1）。

C. 轨迹匹配度量

运动学共识需要计算轨迹集之间的距离，因此需要采用适当的轨迹匹配度量。文献 [14], [15] 为此采用了OSPA度量的时间扩展形式，称为*OSPA-on-OSPA*，记作OSPA$^{(2)}$。下面我们定义这两种度量；关于其性质的详细讨论见 [16]。

1）*OSPA*：对于两个有限集 $X,\ Y\in\mathcal{F}(\mathbb{X})$，其中$\mathcal{F}(\mathbb{X})$是由 $\mathbb{X}$、$X=\{x^{(1)},\ldots,x^{(m)}\}$和 $Y=\{y^{(1)},\ldots,y^{(n)}\}$构成的有限集空间，$p$ 阶OSPA定义为

$$d_p^{(c)}(X,Y)=\left(\frac{1}{n}\left(\min_{\pi\in\Pi_n}\sum_{i=1}^{m}d^{(c)}(x^{(i)},y^{(\pi(i))})^p+c^p(n-m)\right)\right)^{1/p} \tag{1}$$

对于 $m\le n$，以及 $d_p^{(c)}(X,Y)\triangleq d_p^{(c)}(Y,X)$，对于 $m>n$。这里，Π_n表示所有排列的集合 $\{1,\ldots,n\}$；$d^{(c)}(x,y)=\min(c,d(x,y))$表示截断基准距离，其中$c\in(0,\infty)$为截断参数；$d(\cdot,\cdot)$是 $\mathbb{X}$上的度量，例如曼哈顿距离。注意，OSPA按如下约定定义：$d_p^{(c)}(\emptyset,\emptyset)=0$，以及当其中一个自变量为空时 $d_p^{(c)}(X,Y)=c$，由此所有超过 c 的距离均被视为同样大。

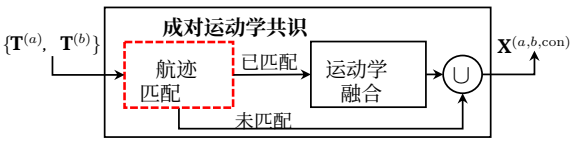


图3：两个节点之间的成对运动学共识。位于$\{\mathbf{T}^{(a)},\mathbf{T}^{(b)}\}$ 中的航迹采用航迹距离度量进行匹配（第II-C节）；匹配航迹经融合后与保留的未匹配航迹组合，形成$\mathbf{X}^{(a,b,\mathrm{con})}$。如第III节所示，仅操纵轨迹匹配阶段便足以诱发身份盗用（见注1）。

2)OSPA$^{(2)}$*:* 该度量将OSPA扩展到有限航迹集 $\mathbb{T}$的空间，即 $\mathcal{F}(\mathbb{T})$，作为时间平均OSPA距离。对于航迹 $t,u\in\mathbb{T}$，OSPA$^{(2)}$ 由下式给出

$$\tilde{d}^{(c)}(t,u)=\begin{cases}\sum_{k\in\mathcal{D}^{(t)}\cup\mathcal{D}^{(u)}}\frac{d^{(c)}(\{t(k)\},\{u(k)\})}{|\mathcal{D}^{(t)}\cup\mathcal{D}^{(u)}|}, & \text{if } \mathcal{D}^{(t)}\cup\mathcal{D}^{(u)}\neq\emptyset,\\ 0, & \text{otherwise.}\end{cases} \tag{2}$$

其中 $d^{(c)}(\{t(k)\},\{u(k)\})$是在时刻 k计算的航迹间OSPA距离，即单元素集合之间的距离，并由此强制满足 $p=1$。这里，我们用 $t(k)\in\mathbb{X}$ 表示航迹 t 在时刻 $k\in\mathcal{D}^{(t)}$的单目标状态。

III. 对抗模型

本节介绍标签劫持的对抗模型和可行性假设（第III-A节），提出TC-ITA策略（第III-B节），并提出一种基于模型预测控制的方法，用于生成隐蔽欺骗轨迹。该方法的灵感源自RGPO攻击中的接管效应（第III-C节） [17]。

该攻击目标是将受害目标的全局标签转移给受控冒充目标，使网络继续认为自己正在跟踪受害目标，而实际上跟踪的却是冒充目标。值得注意的是，如果冒充目标是由攻击者控制的智能体（例如无人机），其运动便可任意设计，从而使其在被盗身份下表现出任何期望的行为。

术语：受感染节点是指通信信道受攻击者控制的传感器节点。未受损节点是指未受对抗方控制的传感器节点；但是，攻击者可能会拦截其传输消息，或试图影响其融合输出。盲区是监视区域中不在任何未受损节点视场范围内的区域。未受损区域是指至少有一个未受损节点能够观测到的区域。

A. 对抗场景

我们考虑图4所示的对抗场景，其中攻击者通过控制一个或多个传感器节点的通信信道来攻陷这些节点。攻击者会截获被攻陷节点生成的局部航迹估计，以及其未受攻击的邻居节点传输的信息。通过汇总这些信息，攻击者便能获取目标轨迹，即使目标位于被攻陷节点的感知覆盖范围之外，只要至少有一个邻居节点观测到该目标。

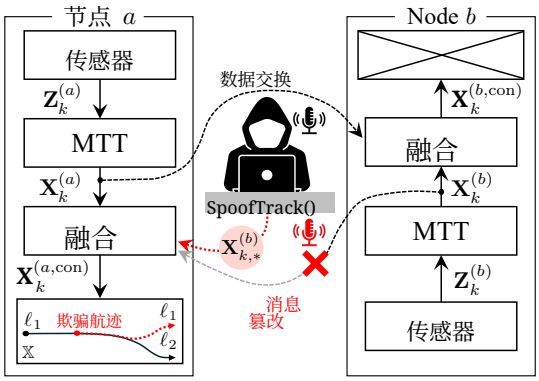


图4：TC-DMTT 中的虚假数据注入场景（可自然扩展到更大规模网络）。在时间 k，每个节点 i 根据其测量值$\mathbf{X}_k^{(i)}$ 形成局部航迹$\mathbf{Z}_k^{(i)}$。攻击者攻陷节点 b，将其名义航迹$\mathbf{X}_k^{(b)}$ 替换为对抗性航迹$\mathbf{X}_{k,*}^{(b)}$，并将其转发给未受损邻居节点，从而破坏其局部融合输出$\mathbf{X}_k^{(a,\mathrm{con})}$。

攻击者根据截获的数据构造欺骗航迹，并以其代替被攻陷节点的标称局部航迹估计进行转发，从而破坏相邻节点的局部融合过程及由此产生的一致性估计。该机制遵循与重放攻击相同的拦截–修改–再注入范式。

所提出的攻击基于以下假设：

A1）传感器节点的视场有限且部分不重叠。这使受害目标能够进入盲区，而这是进行标签提取所必需的。A2）攻击者必须攻陷其试图操纵融合过程的节点的直接邻居节点，以及其试图拦截航迹信息的节点的直接邻居节点。A3）由于节点不了解网络几何结构，未受损节点无法验证报告航迹是否由发送节点通过物理观测获得，因此会在一致性过程中将欺骗航迹视为有效航迹并予以接受。A4）攻击者能够推断受害目标何时进入和离开盲区。

该可见性推断根据假设 A4，使攻击者能够把握第 III-B 节所述标签提取与注入阶段的时机。攻击者通过拦截的通信模式以启发式方式进行推断：若连续多个时间步未收到更新报告，则表明目标已离开某个节点的视场；若仍持续收到报告，则意味着目标仍可被观测。值得注意的是，我们假定检测这些转变的延迟可以忽略不计；该延迟取决于本地 MTT 航迹管理策略，而非对网络几何结构的了解。

B. 身份盗窃攻击策略

在 TC-ITA 中，攻击者会选择网络正在跟踪、全局标签为 ℓ_{v} 的受害目标，并监控其由相邻节点和受感染节点报告的状态估计值。令 t_{v} 和 t_{i} 分别表示具有代表性的受害目标航迹和冒充航迹，

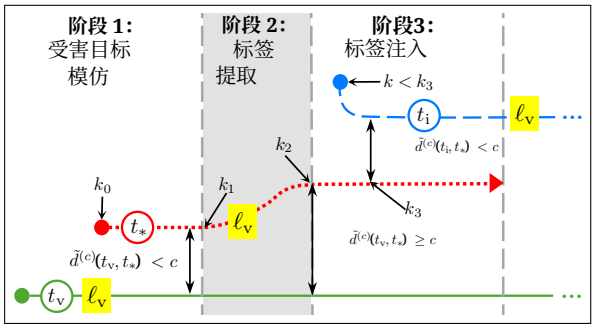


图 5：TC-ITA 三个阶段及相关符号概览（参见第 III-B 节）。白色和灰色区域分别表示未受损区域和盲区。

这些航迹由攻击者可访问的至少一个节点（受感染或未受感染）报告。当多个节点报告任一目标时，假定下文引入的时间和航迹距离约束对所有相应航迹均成立。此外，令 t_* 表示攻击者从时间 k_0开始传输的欺骗航迹。下面我们将介绍 TC-ITA 的三个阶段，如图5所示。

1）受害目标模仿： t_* 通过 ℓ_{v}，模仿 t_{v}，使其足够相似，从而依据第 II-B1 节的航迹匹配规则被选中。（参见注 2）。这要求满足匹配条件 $\tilde{d}^{(c)}(t_{\mathrm{v}}, t_*) < c$，对于某个 $k \in [k_0, k_1)$ 而言，即必须发生在受害目标进入盲区之前。在时间 k_1，攻击者检测到受害目标进入盲区（假设 A4），这会触发标签提取阶段的开始。

2）标签提取（拖引阶段）：在盲区内，即对于 $k \in (k_1, k_2)$，网络仅接收到 t_*，将其作为与标签 ℓ_{v}关联的候选航迹。由于这段时间内不存在竞争航迹， t_* 可以任意演变，同时保留被劫持标签。当受害目标于时间 k_2重新进入未受损区域时，必须满足间隔条件$\tilde{d}^{(c)}(t_{\mathrm{v}}, t_*) \geq c$，以确保确认后将 t_{v} 视为新目标，并为其分配不同于 ℓ_{v} 的标签；此后必须持续满足这一条件，以防止再次匹配。请注意，若受害目标未重新进入未受损区域，则 t_* 不会有与 t_{v}匹配的风险，因为 $\tilde{d}^{(c)}(t_{\mathrm{v}}, t_*) = \tilde{d}^{(c)}(\emptyset, t_*) = c$。值得注意的是，这一阶段与雷达跟踪中的拖引欺骗直接类似：传统的距离波门拖引会将雷达距离波门从真实目标回波处拉开，而 TC-ITA 则会拉开目标的身份标签。请注意，无论 t_* 在盲区间隔内的隐蔽性如何，都能实现标签拖引（参见第 IV 节中的硬切换变体）；隐蔽变体则进一步受到距离波门拖引接管体制的启发，通过强制实现平滑的轨迹过渡来保证运动学合理性。

3）标签注入：假设冒充目标已进入未受损区域，则必须在某个时间 $k_3 > k_2$满足匹配条件$\tilde{d}^{(c)}(t_{\mathrm{i}}, t_*) < c$，从而使 t_{i} 和 t_* 匹配并共享一个全局标签。由于航迹匹配机制会保留存活更久的航迹的标签，并假设冒充目标

比受害目标更晚进入场景，被劫持标签 ℓ_{v} 便会传播至 t_{i}。如果注入成功，那么对于 $k > k_3$，攻击者可以停止报告 t_*，此后 t_{i} 将成为与 ℓ_{v} 关联的唯一候选，并可在保留该身份的同时任意演变。

注释 1（定向航迹共识阶段）。尽管 *TC-ITA* 的目标是劫持受害目标的标签 ℓ_v，但该攻击并不操纵标签共识机制本身。相反，它完全在运动学层面运行，通过注入满足航迹匹配约束的欺骗航迹 t_*。这会导致 t_* 反复与受害者航迹 t_{v} 匹配，在盲区间发生漂移，并最终与冒名航迹 t_{i} 匹配。由于标签共识由持续的航迹匹配驱动，仅破坏运动学共识就足以诱导身份转移。

注释 2 （单步充分性）。由于 （2）中的定义采用时间平均方式，两个航迹之间只要在 c 任意单个时间瞬间处于距离范围内，就会建立有效匹配，而与窗口长度无关。这一性质被 *TC-ITA* 的模仿阶段和标签注入阶段所利用，在这些阶段中，仅需在一个时间步内保持接近，就足以触发航迹匹配。在标签提取阶段，同一性质变得至关重要：即使欺骗航迹与受害航迹短暂重新接近，也会立即重新启用匹配，因此必须保持持续间隔。

C. 面向隐蔽攻击设计的模型预测控制

TC-ITA 依赖于在图 5 所示的转换时刻满足航迹匹配约束。这可以通过硬切换策略实现，如图 6a所示：一旦受害目标的标签被提取（即，欺骗航迹已占据 ℓ_{v}），攻击者便在盲区间隔内停止传输。当冒充目标变得可见并由未受攻陷节点报告后，攻击者通过转发与冒充目标航迹匹配的轨迹恢复传输 t_{i}。尽管该策略足以诱发标签劫持，但也可能产生运动学上不合理的轨迹，从而被运动一致性检查检测到。为克服这一局限性，我们提出了 TC-ITA 的一种隐蔽型实现方式，通过实现平滑标签脱离来保持运动学一致性，如图 6d所示。该策略表现出类似于 RGPO 欺骗策略中所观察到的接管效应 [17]。

一旦该条件 $\tilde{c}^{(t)}(t_*) < \tilde{c}$ 已在第一个攻击阶段（第III-B1节）完成后，欺骗航迹 t_* 的设计必须满足三个主要目标：（*i*）平滑地向冒名航迹 t_{i}收敛；（*ii*）与重新出现的受害者航迹 t_{v}保持足够的间隔；以及（*iii*）通过对规划轨迹施加物理约束来确保隐蔽性，例如速度和加速度限制（$v_{\max}$ 和 $a_{\max}$）。为系统地满足这些目标和约束，我们将 t_* 的设计建模为约束最优控制问题，并通过模型预测控制求解，从而在轨迹生成过程中自然地考虑系统动力学、预测和物理限制。我们强调，此优化

该策略仅由攻击者使用来生成欺骗轨迹，不属于跟踪或融合过程的一部分。

下文中，我们使用简写 $t(k) \triangleq x_k$ 来将任意航迹 t 与其在时间 $k \in \mathcal{D}^{(t)}$的单目标运动学状态相关联。该状态定义为 $x_k = [p_k^\top, v_k^\top]^\top \in \mathbb{X}$，其中 p_k，$v_k \in \mathbb{R}^2$ 分别表示位置向量和速度向量。此约定适用于上一小节中引入的所有航迹，即 $t_*(k) \triangleq x_k^*$、$t_{\mathrm{i}}(k) \triangleq x_k^{\mathrm{i}}$ 和 $t_{\mathrm{v}}(k) \triangleq x_k^{\mathrm{v}}$。

欺骗状态动力学采用恒速度转移模型，并加入攻击者控制的加速度输入，表示为

$$x_k^* = Ax_{k-1}^* + Bu_{k-1}^*, \tag{3}$$

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} \frac{1}{2}\Delta t^2 & 0 \\ 0 & \frac{1}{2}\Delta t^2 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, \tag{4}$$

其中，$u_k^* \in \mathbb{R}^2$ 为攻击者控制的加速度输入，$\Delta t > 0$ 表示采样周期。

在每个时间步 k_0'，攻击者求解一个 K步有限时域最优控制问题。相应的 预测时域索引集 定义为

$$\mathcal{K}_{k_0'} \triangleq \{k_0' + \kappa : \kappa = 0, 1, \ldots, K\} = \{k_0', k_0'+1, \ldots, k_0'+K\}, \tag{5}$$

其中 $\mathcal{K}_{k_0'}^- \triangleq \mathcal{K}_{k_0'} \setminus \{k_0' + K\}$。

在初始条件 $\bar{x}_{k_0'}$ 下，我们求解控制（加速度）和预测状态序列，

$$U^* \triangleq \left[u_k^*\right]_{k \in \mathcal{K}_{k_0'}^-} \in \mathbb{R}^{2\times K}, \quad X^* \triangleq \left[x_k^*\right]_{k \in \mathcal{K}_{k_0'}} \in \mathbb{R}^{2\times(K+1)}, \tag{6}$$

根据优化问题：

$$\begin{aligned} \hat{U}^*, \hat{X}^* = \arg\min_{U^*, X^*} \; & \alpha_p \sum_{k=k_0'+1}^{k_0'+K} \gamma_p^k \|p_k^* - p_k^{\mathrm{i}}\|^2 \\ & + \alpha_v \sum_{k=k_0'+1}^{k_0'+K} \gamma_v^k \|v_k^* - v_k^{\mathrm{i}}\|^2 \\ & + \alpha_c \sum_{k=k_0'+1}^{k_0'+K} \max\big(c - \|p_k^* - p_k^v\|, 0\big)^2 \\ \text{s.t.} \quad & x_k^* = Ax_{k-1}^* + Bu_{k-1}^*, \qquad k \in \mathcal{K}_{k_0'} \setminus \{k_0'\}, \\ & x_{k_0'}^* = \bar{x}_{k_0'}, \\ & \|u_k^*\| \le a_{\max}, \qquad \forall k \in \mathcal{K}_{k_0'}^-, \\ & \|v_k^*\| \le v_{\max}, \qquad \forall k \in \mathcal{K}_{k_0'}. \end{aligned} \tag{7}$$

目标函数中的第一项推动欺骗轨迹靠近冒充目标，第二项通过惩罚速度失配来确保运动学一致性，第三项则通过与受害目标保持最小距离来体现欺骗要求。这些项由 α_p、α_v、$\alpha_c \ge 0$ 加权，以平衡跟踪、平滑性和间隔的相对重要性，而 γ_p、$\gamma_v \in (0, 1]$ 则是

算法1： `SpoofTrack()`

输入： 初始欺骗状态 $\bar{x}_0$，初始冒充者状态 x_0^{i}，初始受害者状态x_0^{v}，预测时域 K，攻击持续时间 N，OSPA阈值 c，速度和加速度界限 $v_{\max}$, $a_{\max}$。

输出： 欺骗状态序列 $\mathcal{S}$

初始化： 记录 $x_0^* \leftarrow \bar{x}_0$ 于 $\mathcal{S}$;

对于 $n \leftarrow 0$ 至 $\mathrm{N}-1$ **执行**

- $(x_n^{\mathrm{i}},\ x_n^{\mathrm{v}}) \leftarrow$ GetFromNetwork(n);
- $\hat{U}^* \leftarrow$ 求解（7），得到 $k_0' \leftarrow n$ 和参数 $\{x_n^*, x_n^{\mathrm{i}}, x_n^{\mathrm{v}}, \mathrm{K}\ c, a_{\max}, v_{\max}\}$;
- $u_n^* \leftarrow \hat{U}^*[:,1]$ 获取第一个最优控制;
- $x_{n+1}^* \leftarrow Ax_n^* + Bu_n^*$ 更新欺骗状态

返回 $\mathcal{S} = [x_0^*, \ldots, x_N^*]$

指数折扣因子，用于优先考虑近期偏差，而非预测时域中更远期的偏差。

请注意，通过调整各阶段的参考状态，可以使用 (7) 中相同的模型预测控制结构生成完整的欺骗轨迹：在模仿阶段，参考状态设为受害者状态，并使用 $\alpha_c = 0$；在拉偏阶段，参考状态设为指定的会合点；在注入阶段，参考状态设为当前的冒充者状态。

在每个时间步，通过预测系统在 K 步时域内的演化并优化相应的控制序列，在线求解 (7) 中的模型预测控制问题。在总计 N 个时间步的时长内，在线生成欺骗轨迹的策略总结于算法 1 中。在时间 k_0'，从网络中获取当前的冒充者状态和受害者状态 $x_{k_0'}^{\mathrm{i}}$ 和 $x_{k_0'}^{\mathrm{v}}$（例如，从攻击者可访问的受损或未受损节点获取）。为构造 (7) 中所需的 K 步轨迹，随后使用假定的运动模型（例如匀速）向前推演这些状态，从而得到仅用于计算模型预测控制解的辅助预测。在下一个时间步 $k_0'+1$，预测窗口向前移动，并重复该过程；初始状态再次通过 `GetFromNetwork` 从最新的网络航迹估计中刷新。

虽然 t^{i} 在（7）中的轨迹对齐项是凸的，可使用标准二次优化工具求解，但分离惩罚会引入非凸分量。将这一要求表述为软惩罚后，优化仍能保持可行，同时允许在轨迹保真度与远离真实目标之间进行可控权衡。在伪造轨迹与真实轨迹起初相近、严格规避可能暂时不可行的场景中，这一点尤为重要。在我们的仿真中，使用*CasADi* [18]求解所得的非线性规划。

IV. 仿真实验

本节评估第III-C节中介绍的硬切换与隐蔽TC-ITA变体。我们在一个代表性的三节点案例研究中说明标签劫持，并评估其对跟踪性能的影响。

A. 仿真设置

我们考虑一个三节点分布式传感器网络，在 $[-500, 2500]$ m× $[0, 1000]$ m的区域内跟踪两个物理目标：一个身份被劫持的受害目标，以及身份被转移到其上的冒充目标。传感器节点位于 $[0,0]^\top$、$[1000,0]^\top$和 $[1800,0]^\top$ m处，并从左到右依次编号为节点1–3，其中仅节点 2被攻击者受损。

我们在 [14], 的TC-DMTT框架中实现所提出的TC-ITA策略，该框架在每个节点采用局部标记多伯努利（LMB）滤波器。假设采用恒速度运动模型，过程噪声标准差为 $\sigma_v = 5$，测量噪声标准差为$\sigma_r = 2$米。每个传感器在相对角度 $[-60°, 60°]$范围内探测物体，检测概率为 $P_D = 0.98$，探测距离为 $r_D = 800$米。与 [14], 一致，（2）中的基础距离度量选为曼哈顿距离，即$d(x,\ y) = \|x - y\|_1$。

评估窗口内目标的真实轨迹 $N = 80$ 如图6a 和6d所示，其中物理目标 t_{v} 和 t_{i} 分别以绿色和蓝色显示，欺骗航迹 t_* 在两种待评估攻击中均以红色显示。图6d 中的隐蔽TC-ITA轨迹采用算法1所总结的基于MPC的策略生成，时域为 $K = 20$，权重为 $\alpha_p = 1$、$\alpha_v = \alpha_c = 0.1$，折扣因子为 $\gamma_p = \gamma_v = 0.99$，物理边界为 $v_{\max} = 30\mathrm{m/s}$ 和 $a_{\max} = 30\mathrm{m/s}^2$。

图6 给出了单次代表性实现的结果，而图7a 和7b 报告了在 $M = 100$次蒙特卡洛运行上平均得到的统计结果。

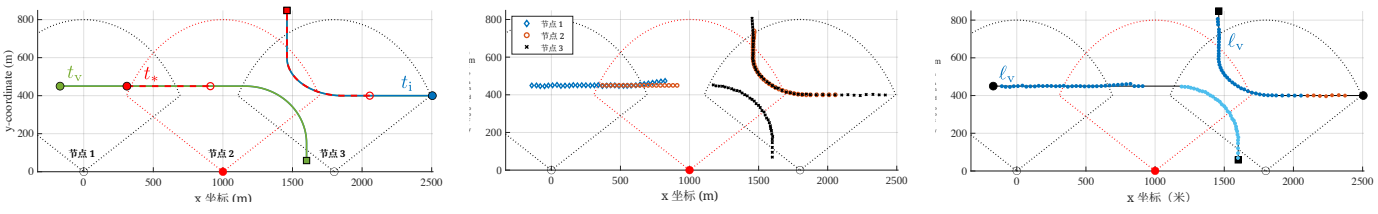
B. 标签劫持性能

我们报告节点 3的局部融合（一致性）结果，即$\hat{\mathbf{X}}_{1:N}^{(3,\mathrm{con})}$。这并不失一般性，因为在运动学一致性阶段（第II-B1节）使用了等概率权重。

如图中目标轨迹和欺骗轨迹所示，6a和 6d，硬切换攻击会产生突兀且运动学上不合理的轨迹，而基于模型预测控制的隐蔽型攻击则会生成平滑的拉偏轨迹，使其紧随冒充目标，同时与受害目标保持间隔，从而形成视觉上合理的欺骗轨迹 t_*。

图中展示了节点 3 可用的相应单独航迹流（即其自身的局部估计 $\hat{\mathbf{X}}_{1:N}^{(3)}$以及从邻居节点接收到的 $\{\hat{\mathbf{X}}_{1:N}^{(i)}\}_{i\in\mathcal{N}(3)}$）：见图 6b 和 6e。这些航迹流构成融合阶段的输入，并表明由受感染节点（节点 2，橙色）注入的欺骗航迹 t_*进入融合过程时被视为合法航迹估计，与诚实节点报告的航迹无法区分。

一致性结果如图6c 和6f所示，其中不同颜色对应不同的一致性标签分配。受害目标最初与蓝色标签相关联，而冒充者携带红色标签；两种攻击变体均成功将蓝色标签转移给冒充者。硬切换策略通过一条突变轨迹实现了这一点



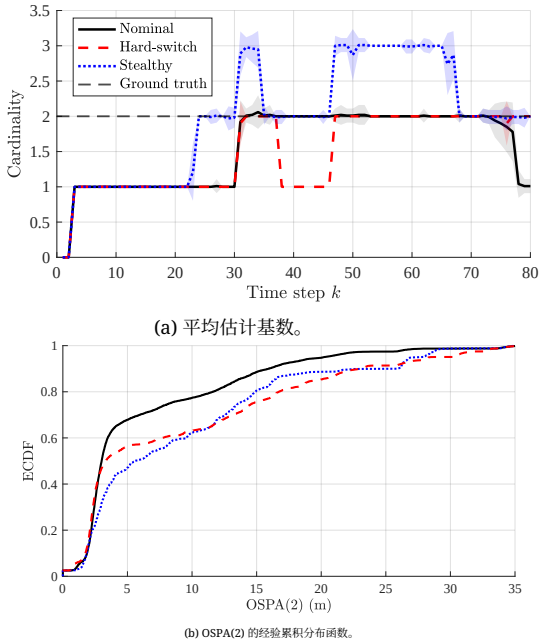
(a) 真实值（硬切换）。 (b) 个体估计（硬切换）。 (c) 共识估计（硬切换）。

(d) 真实值（隐蔽型）。 (e) 个体估计（隐蔽型）。 (f) 共识估计（隐蔽型）。

图6：硬切换（上）和隐蔽型（下）TC-ITA攻击下某次实现的跟踪结果；后者通过算法 1 生成。（左）：受害目标、冒充目标和欺骗轨迹的真实轨迹。（中）：节点 3 可用的单独航迹流，包括其自身的估计值以及从相邻节点接收的航迹流，这些构成融合阶段的输入。（右）：节点 3 的局部一致性估计值，$\hat{\mathbf{X}}_{1:N}^{(3,\mathrm{con})}$：。不同颜色表示不同的一致性标签分配；在硬切换和隐蔽型 TC-ITA 两种攻击变体中，受害目标的标签 ℓ_{v}（蓝色）均被成功转移至冒充目标。圆形标记和三角形标记分别表示每条轨迹的起点和终点。

过渡虽然在运动学上不合理，但仍未被TC-DMTT框架缓解。这表明分布式多目标跟踪框架仅依赖基于度量的匹配来实现标签共识，而忽略了运动学一致性这一额外的完整性检查。隐蔽TC-ITA进一步通过其动态上合理的轨迹规避了潜在的运动一致性检查。

基数估计结果如图 7a所示。对于硬切换攻击，主要在切换区间内出现暂时的低估，此时攻击者停止传输。相比之下，在隐蔽型 TC-ITA攻击下，基数会被高估，因为欺骗航迹 t_* 在整个攻击过程中不一定与任何真实目标重合。图 7b 展示了硬切换和隐蔽型 TC-ITA攻击以及名义场景（无攻击）下，蒙特卡洛运行中瞬时OSPA的经验累积分布函数（ECDF）。与名义场景相比，两种攻击变体均显著降低了跟踪精度，证实标签劫持也会影响目标跟踪性能。



(a) 平均估计基数。

(b) OSPA(2) 的经验累积分布函数。

图 7：在硬切换攻击、隐蔽 TC-ITA 和无攻击（名义场景）情况下，节点 3 在整个评估窗口内的跟踪精度影响。

V. 结论

本文揭示了航迹一致性分布式多目标跟踪中的身份层面漏洞。通过利用基于度量的航迹匹配阶段，所提出的航迹一致性身份盗窃攻击能够劫持标签并冒充目标。我们提出了一种硬切换攻击实现，
表明即使轨迹突变且在运动学上不合理，也足以在基于 OSPA 的匹配中触发标签劫持。随后，我们提出一种基于 MPC 的隐蔽型实现，在满足相同匹配约束的同时，确保拉偏行为平滑且符合物理规律。在有限视场网络中的仿真证实了这两种变体均能成功。这些发现凸显了在共识层实施身份感知保护的必要性。

参考文献

[1] Y. Bar-Shalom、P. K. Willett与X. Tian，跟踪与数据融合，第11卷。美国康涅狄格州斯托尔斯：YBS Publishing，2011年。[2] B.-N. Vo、M. Mallick、Y. Bar-Shalom、S. Coraluppi、R. Osborne、R. Mahler与B.-T. Vo，“多目标跟踪”，《*Wiley* 电气与电子工程百科全书》，2015年。[3] R. P. S. Mahler，统计多源多目标信息融合。美国马萨诸塞州波士顿：Artech House，2007年。[4] B.-T. Vo，多目标滤波中的随机有限集。博士论文，西澳大利亚大学，澳大利亚，2008年。[5] W. Wu、H. Sun、M. Zheng与W. Huang，基于随机有限集的目标跟踪。施普林格，2023年。

[6] ´E. L. Souza、E. F. Nakamura与R. W. Pazzi，“传感器网络中的目标跟踪：综述”，*ACM*计算调查（*CSUR*），第49卷，第2期，第1–31页，2016年。[7] R. P. Mahler，“最优/鲁棒分布式数据融合：一种统一方法”，载于信号处理、传感器融合与目标识别 *IX*，第4052卷，第128–138页，SPIE，2000年。[8] C. Fantacci、B.-N. Vo、B.-T. Vo、G. Battistelli与L. Chisci，“多传感器多目标跟踪的鲁棒融合”，*IEEE*信号处理快报，第25卷，第5期，第640–644页，2018年。[9] M. B. Hurley，“协方差交集及其推广的信息论依据”，载于第*5*届融合会议论文集，第1卷，第505–511页，2002年。[10] Y. Yu与Y. Liang，“存在恶意网络攻击的去中心化传感器网络中的安全多目标跟踪”，数字信号处理，第117卷，第103132页，2021年。[11] M. ¨Uney、J. Houssineau、E. Delande、S. J. Julier与D. E. Clark，“有限集分布的融合：逐点一致性与全局基数”，*IEEE*航空航天与电子系统汇刊，第55卷，第6期，第2759–2773页，2019年。[12] S. Li、W. Yi、R. Hoseinnezhad、G. Battistelli、B. Wang与L. Kong，“基于带标签随机有限集的鲁棒分布式融合”，*IEEE*信号处理汇刊，第66卷，第2期，第278–293页，2017年。[13] L. Gao、G. Battistelli与L. Chisci，“事件触发式分布式多目标跟踪”，*IEEE*网络信号与信息处理汇刊，第5卷，第3期，第570–584页，2019年。[14] H. Van Nguyen、H. Rezatofighi、B.-N. Vo与D. C. Ranasinghe，“有限视场传感器条件下的分布式多目标跟踪”，*IEEE*信号处理汇刊，第69卷，第5329–5344页，2021年。[15] J. K. Verma、J. K. Chhabra与V. Ranga，“移动分布式传感器网络中基于航迹共识的带标签多目标跟踪”，*IEEE*移动计算汇刊，第23卷，第6期，第7351–7362页，2023年。[16] M. Beard、B. T. Vo与B.-N. Vo，“OSPA(2)：使用OSPA度量评估多目标跟踪性能”，载于第*7*届*ICCAIS*会议论文集，第86–91页，2017年。[17] H. Calatrava、S. Tang与P. Closas，“雷达系统反欺骗干扰策略的进展：综述”，*IEEE*航空航天与电子系统杂志，第1–22页，2025年。[18] J. A. E. Andersson、J. Gillis、G. Horn、J. B. Rawlings与M. Diehl，“CasADi——面向非线性优化与最优控制的软件框架”，数学规划计算，第11卷，第1期，第1–36页，2019年。