浙江工业大学

硕士学位论文

基于主动注入随机噪声的分布式安全融合估计

作者姓名	严新豪
指导教师	陈博　教授
学科专业	控制科学与工程
学位类型	工学硕士
培养类别	全日制学术型硕士
所在学院	信息工程学院

提交日期：　　2023 年 6 月

Distributed Secure Fusion Estimation Based on Actively Injected Random Noises

Dissertation Submitted to

Zhejiang University of Technology

in partial fulfillment of the requirement

for the degree of

Master of Engineering

by

Xinhao YAN

Dissertation Supervisor: Prof. Bo CHEN

Jun., 2023

浙江工业大学学位论文原创性声明

本人郑重声明：所提交的学位论文是本人在导师的指导下，独立进行研究工作所取得的研究成果。除文中已经加以标注引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不含为获得浙江工业大学或其它教育机构的学位证书而使用过的材料。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人承担本声明的法律责任。

作者签名：尹新豪　　日期：2023年 5 月

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权浙江工业大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于：　1、保密□，在一年解密后适用本授权书。

2、保密□，在二年解密后适用本授权书。

3、保密□，在三年解密后适用本授权书。

4、不保密☑。

（请在以上相应方框内打“√”）

作者签名：尹新豪　　日期：2023年 5 月

导师签名：陈博　　日期：2023年 5 月

中图分类号　TP301　　学校代码　10337

UDC　007　　密级　公开

研究生类别　全日制学术型硕士研究生

浙江工业大学

工学硕士学位论文

基于主动注入随机噪声的分布式安全融合估计

Distributed Secure Fusion Estimation Based on Actively Injected Random Noises

作者姓名　严新豪　　第一导师　陈博 教　授

学位类型　工学硕士　　第二导师

学科专业　控制科学与工程　　培养单位　信息工程学院

研究方向　安全融合估计

答辩日期：　2023　年　05　月　14　日

基于主动注入随机噪声的分布式安全融合估计

摘　要

多传感器信息融合是把多个传感器的冗余信息或者互补信息基于某种融合准则进行组合以获得被检测对象的一致性描述，而分布式融合估计作为其中一个重要的分支有着重要的应用价值，因为它具有较高的鲁棒性和扩展性。随着通讯技术的快速发展，通信网络已经成为多传感器融合估计系统传输信息的主要手段，从而构成网络化多传感器融合估计系统（Networked Multi-sensor Fusion Estimation Systems, NMFESs），这样的系统有易维护和布线简单等众多优点，受到了工业界和学术界的广泛关注，包括无人自主系统、智能制造、智能电网等众多领域。

然而，由于网络的开放性和匿名性，网络攻击已成为影响 NMFESs 性能的一个无法避免的问题。特别地，窃听攻击是一类典型的网络攻击方式，它虽然不会直接对传输信道或数据包进行破坏，但是会将关键数据泄露出去，从而对系统安全造成多方面的威胁。为了降低窃听攻击对 NMFESs 的影响，如何保护 NMFESs 中传输数据的隐私成为一个亟待解决的基础性问题。传统的密码学和物理层方法虽然能有效保护隐私，但是存在计算量和能量消耗大等问题，而采用随机噪声加密方式则无需复杂计算和能量消耗。然而，噪声的注入也会对 NMFESs 中合法用户的估计性能等方面产生消极影响。为此，本文针对分布式融合估计系统遭受恶意窃听的问题，研究了主动注入随机噪声的方法，并设计高效的补偿策略以提升噪声机制影响下合法用户的估计性能。本文的主要工作如下：

1. 针对 NMFESs 中局部数据被窃听的问题，通过分析加权融合矩阵的结构，提出了一种面向局部估计的随机噪声污染策略。该策略基于零空间的设计方案，采用主动注入随机噪声的方式使窃听者进行加权融合后损失部分局部信息，而合法用户则采用一步预测在线补偿由于噪声注入产生的估计误差，并根据预测的协方差重新融合以获得性能较好的融合估计器。最后，通过移动机器人仿真系统验证了所提方法的有效性。

2．针对 NMFESs 的隐私保护问题，通过在多个融合估计的统计信息上主动注入高斯白噪声，提出了差分隐私下的安全融合估计方法。该方法基于加权融合矩阵的逆设计了传感器端的主动噪声注入方式，在实现差分隐私的同时使得窃听者进行线性加权融合也无法获取准确的状态估计值，而合法用户则根据主动注入噪声的统计特性重新设计融合准则，从而降低所提随机噪声机制误差带来的估计

性能损失。最后，通过目标跟踪仿真系统验证了所提方法的有效性。

3．针对 NMFESs 遭受窃听攻击的问题，结合线性加权融合准则中加权融合矩阵的零空间与差分隐私中的高斯机制提出了基于两步序贯噪声主动注入的分布式安全融合估计方法。该方法利用差分隐私机制设计注入噪声的协方差下界，并且通过构建差分的融合方式使得窃听者由于噪声的累加导致其估计误差协方差发散。与此同时，合法用户根据主动注入噪声的协方差重新设计融合准则，进而基于零空间的噪声设计方案抵消部分噪声影响，从而获得稳定且性能较优的融合估计器。最后，通过航空发动机模型验证了所提方法的有效性。

关键词：分布式融合估计，窃听攻击，随机噪声，零空间，差分隐私

DISTRIBUTED SECURE FUSION ESTIMATION BASED ON ACTIVELY INJECTED RANDOM NOISES

ABSTRACT

Multi-sensor information fusion is a technique that combines redundant or complementary information of multiple sensors based on some fusion criteria to obtain a consistent description of the detected object. As an important branch, distributed fusion estimation has significant application value, because it has high robustness and scalability. With the rapid development of communication technology, communication networks have become the main tools of transmitting information in multi-sensor fusion estimation systems, thus forming networked multi-sensor fusion estimation systems (NMFESs). These systems have many advantages, such as easy maintenance and simple wiring. They have received wide attention from industry and academia, including unmanned autonomous systems, smart manufacturing, smart grids and many other fields.

However, due to the openness and anonymity of the network, cyber-attack has become an unavoidable problem that affects the performance of NMFESs. In particular, eavesdropping attack is a typical type of cyber-attack. Although it does not directly damage the transmission channel or data packet, it leaks key data and poses multiple threats to system security. To reduce the impact of eavesdropping attacks on NMFESs, how to preserve the privacy of transmitted data in NMFESs has become an urgent basic problem to be solved. Although traditional cryptography and physical layer methods can effectively protect privacy, there are some problems such as large amount of computation and energy consumption, while adopting random noise encryption method does not need complicated calculation and energy consumption. Nonetheless, injecting noise will have a negative impact on the estimation performance of legitimate user in NMFESs. Therefore, focusing on the problem of malicious eavesdropping on distributed fusion estimation systems, this paper studies the method of actively injecting random noises and further designs effective compensation strategies to improve the estimation performance of legitimate user under the influence of noise mechanisms. The main work of this paper is given as follows:

1. In response to the issue of local data being eavesdropped in NMFESs, a random noise contamination strategy is proposed with regard to local estimates by analyzing the structure of weighting fusion matrix. Based on the design of null space, this strategy adopts actively injected random noises to make the eavesdropper lose some local information after weighting fusion. Moreover, legitimate user applies one-step prediction to online compensate the estimation error generated by the noise contamination. Then, it re-fuses local data based on the prediction covariances to obtain a better-performing fusion estimator. Finally, the effectiveness of the proposed method is verified by the mobile robot simulation systems.

2. To address the privacy protection issue in NMFESs, a secure fusion estimation method with differential privacy is proposed by actively injecting white Gaussian noises on the statistical information of multiple fusion estimates. Based on the inverse design of weighting fusion matrix, the active noise injection method at sensor side is designed. In this case, difference privacy is realized and the accurate state estimates cannot also be obtained with linear weighting fusion by the eavesdropper. Moreover, legitimate user redesigns the fusion criterion based on the statistical characteristics of the actively injected noises to reduce the estimation performance loss generated by the proposed random noise mechanism. Finally, the effectiveness of the proposed method is verified by the target tracking simulation systems.

3. To address the issue of eavesdropping attacks in NMFESs, combining the null space of weighting fusion matrix in the linear weighting fusion criterion and the Gaussian mechanism in the differential privacy, a distributed secure fusion estimation method based on two step sequential noise active injection is proposed. This method applies the differentially private mechanism to design the lower bound of the covariances of actively injected random noises, and constructs a differential fusion method to make the estimation error covariance of eavesdropper diverge due to the accumulation of noises. At the same time, legitimate user redesigns the fusion criterion based on the covariance of actively injected noises. Then, it counteracts certain noises impact based on the noise design scheme of null space, thus obtaining a stable and better-performing fusion estimator. Finally, the effectiveness of the proposed method is verified through an aircraft engine model.

KEY WORDS: distributed fusion estimation, eavesdropping attack, random noises, null space, differential privacy

目 录

插图清单

符号说明

R^n	——	n维实向量的集合
$\mathrm{R}^{m\times n}$	——	$m\times n$维实矩阵的集合
Z	——	整数集合
$\mathrm{col}\{a_1,\cdots,a_n\}$	——	以$a_1,\cdots,a_n$为元素的列向量
$\mathrm{diag}\{A_1,\cdots A_n\}$	——	以$A_i(i=1,\cdots,n)$为对角元素的分块对角矩阵
I	——	具有适当维数的单位矩阵
$\mathrm{E}(\bullet)$	——	数学期望函数
$\mathrm{P}(\bullet)$	——	概率函数
min	——	所有取值中的最小值
$\mathrm{rank}(\bullet)$	——	矩阵的秩
$\sigma_{\max}(\bullet)$	——	最大奇异值
$\rho_{\max}(\bullet)$	——	最大特征值
$\mathrm{Tr}\{\bullet\}$	——	矩阵的迹
$X>0$	——	矩阵X为正定矩阵
$X<0$	——	矩阵X为负定矩阵
X^{T}	——	矩阵X的转置
X^{-1}	——	矩阵X的逆
$\|\bullet\|$	——	向量或矩阵的范数
$\triangleq$	——	定义为
sup	——	上确界
iff	——	当且仅当
$\lim_{t\to\infty}$	——	时间t趋向无穷时的极限

第一章 绪论

1.1 研究背景及意义

多传感器信息融合是一门随着传感器、计算机和通信等相关技术发展而诞生的交叉学科。早在 20 世纪 70 年代，信息融合就被美国国防部用于声呐研究，后来被广泛应用于国防军事与民用科技等众多方向，包括目标跟踪、机器人系统和工业互联网等[1-5]。多传感器信息融合技术的原理是综合多个传感器在时间或者空间上冗余或者互补的信息，并按照某种特定的准则进行组合以获得某个系统过程或者参数的一致性描述[1]。状态估计是信息融合的一种典型应用，相比较简单的单传感器状态估计，多传感器融合估计由于能利用多个传感器的观测信息往往拥有更好的性能；同时，通过多传感器融合的方式，利用成本较低的传感器组合也能达到较高的估计精度。因此，基于多传感器的融合估计方法获得了广泛应用[3, 4]。

传统的多传感器融合估计系统采用专线连接，但是随着信息技术的发展，其局限性逐渐产生且日益明显。比如，由于系统规模不断增大，导致系统的布线复杂度变得越来越高，并且传感器种类和布置数量增加也使得系统的成本大大增加[6, 7]。幸运的是，随着通信技术的快速发展，尤其是无线传感网络技术的兴起，使得网络通信逐渐取代传统通信方式，被广泛地应用于传感器与传感器之间、传感器与融合中心之间的连接通信，这就构成了网络化的多传感器融合估计系统（Networked Multi-Sensor Fusion Estimation Systems，NMFESs）[8-10]。由于通信网络的引入，网络化多传感器融合系统拥有众多优点[9]：（1）多个功能模块集成到一个设备中且模块之间通过网络互联，利于共享数据和资源且能有效减少布线复杂度；（2）系统可以按需部署和协作，使得可扩展性高且分布式融合结构更易于实现。因此，NMFESs 是多传感器信息融合领域一个非常值得研究的方向，且已经得到了广泛的关注和应用[11]。具体地，无人机[12, 13]、智能制造过程[14, 15]、智能电网[16]等都是网络化多传感器融合估计的经典应用场景。图 1-1 展示了一种简单的无人机（Unmanned Aerial Vehicle，UAV）系统，其中，布置在无人机上和地面上等多种传感器通过各自的传输方式将其感知或者简易处理后的数据传输至无人机或者地面控制站（Ground Control Station, GCS），然后通过信息筛选和预处理，并借助特定的信息融合方案，进而计算出性能更优的状态估计值，从而可用于无人机自身定位或监视目标的跟踪等。

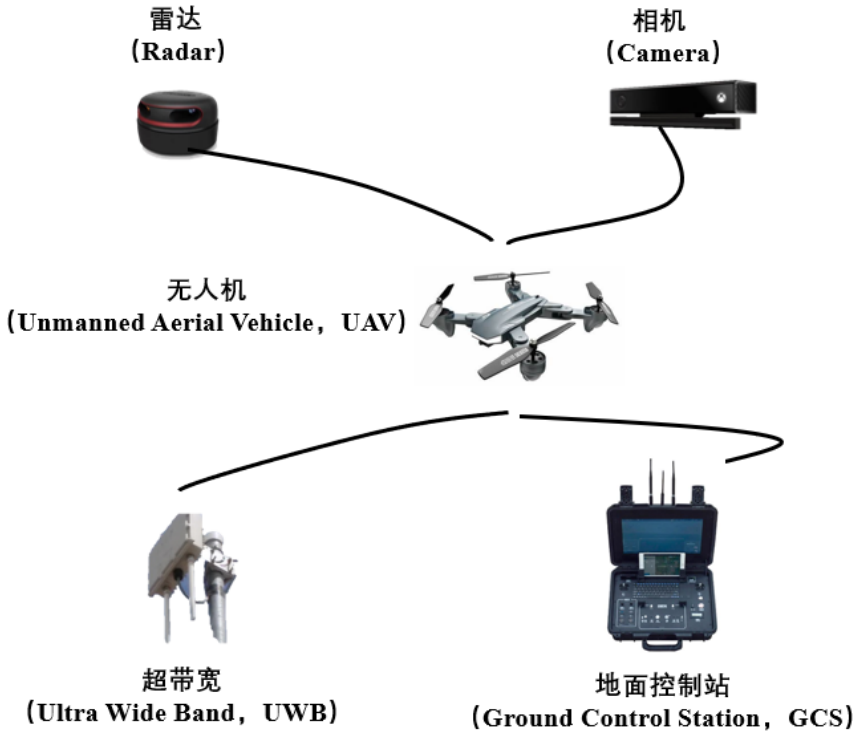

图 1-1　一种简单的无人机系统

Figure 1-1.　A simple unmanned aerial vehicle system

然而，通信网络的开放性和匿名性等性质可能会导致 NMFESs 有遭受网络攻击的风险[17]，比如虚假数据注入（False Data Injection, FDI）攻击[18]、拒绝服务（Denial of Service, DoS）攻击[19]、重放攻击[20]和窃听攻击[21]等。在某些情况下，这些攻击不但会使系统性能受损，还有可能会造成巨大的社会危害。比如，当智能电网中的传输数据被篡改，控制终端可能会发送错误信号，使得电网中的电压或电流过大，轻则破坏电路，严重的则可能会造成人员伤亡。特别地，在这些攻击方式中，通信信道窃听攻击是一种较为特殊的攻击方式。与 DoS、FDI 和重放等直接对信道或数据包进行攻击的方式不同，窃听攻击一般不对通信数据造成任何影响，仅仅通过实时监听信道中传输的信息来获取或推测其需要的相关数据。但是，窃听引起信息泄露的后果可能更加严重，不仅会带来经济损失和声誉损害等诸多风险，泄露的数据还可能被其他攻击者进一步利用造成更严重的打击，甚至可能导致系统崩溃瘫痪。因此，如何主动地保护系统隐私以对抗潜在的窃听攻击，以及在窃听发生后如何减少窃听所造成的损失是网络化融合系统的一个重要研究课题。

传统的隐私保护方法主要与密码学[22, 23]等相关研究有关，其在发送端利用公钥等加密方式将原始数据转换为密文，并且在接收端根据密钥等特定协议进行相应解密操作以获取所需信息。由于解密的难易程度与加密复杂度有关，所以性能

较好的密码学算法一般需要较大的计算量。然而，大部分传感器一般直接根据相应物理原理设计感知元件，不嵌入计算单元，或者仅配备性能一般的主控芯片用于简单的算法实现、控制以及通信。因此，在大多数情况下，NMFESs 传感器端的计算性能十分有限，无法支撑大计算量的加密方式。同时，如果窃听者拥有性能足够强大的设备，则传输的加密数据可能被完全解密，仍会导致系统关键信息的隐私泄露。

为了解决上述实际应用问题，关于数据扰动的策略被广泛地讨论，比如随机噪声扰动等[24]。由于随机数的生成无需复杂的计算过程，并且仅通过叠加或相乘等简单的运算就可以产生一定效果，因此，基于随机噪声的相关方法可以在有效避免大规模加解密计算的同时实现较高的隐私水平。

1.2 国内外研究现状

1.2.1 多传感器信息融合研究现状

一般情况下，单一的传感器只能提供有限的信息量，无法提供全面的数据，这可能会导致数据不准确或不完整。此外，其测量结果容易受到环境的影响，例如温度、湿度、光线等因素的影响，复杂的环境可能会导致数据误差增加，从而影响估计的准确性。为解决基于单一传感器状态估计存在的问题，及满足日益增长的高精度和高鲁棒性等需求，基于多传感器的融合估计被提出，并成为信息融合理论的一个重要研究领域[2, 4]。

信息融合估计可追溯至 20 世纪 70 年代，在假设各个状态估计相互独立的前提下，R. Singer 等[25]首先讨论了多传感器的融合估计问题，标志着多传感器融合估计系统研究的开始。后来，逐渐涌现出众多传感器融合方法，经典的有 D. Willner 等[26]提出的多传感器集中式 Kalman 融合估计算法和 X. Li 等[27]提出的线性无偏估计方法。随后，邓自立和孙书利等[4, 28, 29]考虑了局部估计误差的相关性，基于加权最小二乘进一步给出按矩阵、对角矩阵和标量加权三种线性最小方差意义下的最优融合估计方法。

多传感器信息融合框架主要分为两种：集中式融合框架[26, 30-32]和分布式融合框架[28, 33, 34]。在集中式融合估计框架中，每个传感器发送原始观测值至融合中心，融合中心将有效的观测数据进行增广形成一个高维观测矩阵，并以此高维观测信息使用传统的单传感器方法进行估计，比如 Kalman 滤波[4, 9, 20]和有界递归优化算法[34]等。另一方面，在分布式融合框架中，传感器需要先根据自身有限维数的初始量测进行局部状态估计，并将该估计值发送给融合中心，接着，融合中心再基于特定的融合准则对收到的局部状态估计进行融合，经典的融合准则有按矩阵加

权融合准则和协方差交叉融合准则等[4]。

虽然集中式融合估计性能更高，但是分布式融合估计框架所需的计算量更小，并且拥有更好的鲁棒性和更高的容错率[9]。以融合估计系统遭受虚假数据注入攻击的情况为例，在集中式融合估计框架中，如果局部观测值在传输过程中被攻击者篡改，那么该离群值的影响会由估计器的迭代计算而逐渐扩大，最后甚至可能使估计误差发散。不同的是，在分布式融合框架中，如果传输的局部状态估计遭受攻击，分布式融合估计只会在受攻击时刻产生一定偏差，不会影响到其他时刻的融合估计值。这是因为估计器真正的迭代只发生在传感器端局部估计的计算中，并且分布式融合估计值只是对应时刻所有局部估计值的线性加权组合，不构成迭代计算。因此，基于分布式融合估计框架设计隐私保护策略更具实际意义。

1.2.2 数据加密方法研究现状

对抗窃听攻击的隐私保护方法有很多种，主要分为以下几大类：匿名化、密码学和数据扰动等。匿名化方法有 k-匿名[35]等；密码学方法主要有哈希函数、安全多方计算[36]、同态加密[37]、差分隐私[35, 38, 39]等；数据扰动主要采用随机噪声[24]，同时，差分隐私也是一种经典的数据扰动方法。

匿名化针对个人与存储数据联系起来的标识符，通过消除或加密的方式保护私人或敏感信息[35]。但是，匿名化加密方法并不能完全消除用户的身份信息，用户的信息还是可能被攻击者所获取。这是因为如果攻击者获取了足够多的数据信息，就可能通过分析数据的特征进而采用排除法等策略从侧面推断用户的身份信息甚至敏感信息，从而威胁隐私安全。

密码学和信息论也是传统的隐私保护方法[22, 23]。在信息论中，一般使用信息熵表示随机变量的不确定性，用互信息来表示两个变量之间的关系。比如，E. Erdemir 等[40]将窃听者获取的数据和真实的状态数据的互信息作为隐私保护程度的度量，并利用深度强化学习的方法优化该指标。在密码学中，一般用特定规则对数据进行编码，改变数据传输的数值，使窃听者在无正确解码规则的情况下无法解译出原始数据。同态加密[37]为其中一种经典方法，被应用于众多领域，比如线性二次高斯（Linear Quadratic Gaussian, LQG）控制[41]和二次优化[42]等。其允许数据处理操作在加密状态下进行，而无需解密即可得到所需结果。同态加密可以分为三种类型：加法同态、乘法同态和完全同态加密。加法和乘法同态分别允许对加密的数值进行加法和乘法操作，完全同态加密则允许对加密的数值进行任意的计算操作，包括加法、乘法和其他复杂计算。第一个同态加密算法由 C. Gentry 于 2009 年提出[43]，该方案是基于理想格构造的，并且依赖于稀疏子集和问题的假设。

注意到，上述传统密码学方法一般需要较大的计算量，比如同态加密中需要大量的指数函数计算[41, 42]，而 NMFESs 的处理器性能一般有限，除了简单的控制和通信外，过量的计算负担可能会影响系统的正常运行。为了适应计算性能有限情况，可以采用更为简单的数据扰动方法。差分隐私是其中一种计算量较小的隐私保护方法，其由 C. Dwork 于 2006 年提出[38, 39]并被持续深入研究。该策略首先被应用于数据库领域，通过对数据集合的统计信息注入随机噪声的方式保护大量数据中某些特定敏感数据的隐私。在差分隐私机制作用下，窃听者对两个统计数据进行差分而获取的单个用户数据是不准确的，从而达到保护隐私的目标。同时，差分隐私机制所注入的噪声一般对整体数据影响较小，这使得差分隐私能在某些特定用户的隐私得到有效保护的同时保留其他有效数据的统计信息。因此，差分隐私拥有严谨的数学模型以及优越的隐私保护性能被广泛应用于各个领域，包括机器学习[44]，博弈论[45]，控制论[46]等。

此外，除了直接对传输数据的数值进行改变，还可以在传输数据的调制信号上叠加复噪声进行干扰。一种经典的信道复噪声注入方法由 S. Goel 于 2008 年提出[47]，其主要思想为在调制信号后叠加复噪声波形从而影响信噪比。在窃听者接收系统中，传输的信号将受到极大的噪声干扰。由于接收到信号的信噪比较小，导致窃听者大概率无法解调出准确的数据，进而实现完美的隐私保护，即窃听者获取不到任何在合法系统中通信的数据。同时，信道中所叠加的噪声被设计为依赖于合法用户信道增益的零空间，使得合法用户接收时可以免除噪声干扰从而正常解调。注意到，该方法存在一个较为严重问题，就是复信号噪声的产生需要消耗系统较多额外的能量。然而，许多传感器，的供电方式为电池之类的可移动能源，比如锂电池和铅蓄电池，其储存的能量与体积、材料、重量等众多因素都有关，并且出于轻便等原因考虑，电池的能量往往比较有限。在这些情况下，若使用该种信道加密策略会影响传感器的使用寿命，因为随机复噪声波形的产生加快了能量消耗。并且，噪声的设计需要已知合法用户信道增益，如果是类似于图 1-1 中无人机系统[12, 13]的快速移动目标，即使利用信道估计[48]等手段，也难以发挥该方法的有效性。

1.2.3 状态估计系统加密方法研究现状

在网络化估计系统中，大量的数据会在各级通信网络中传输，比如传感器的观测数据和预处理后的状态估计。这些数据可以反应出系统的相关信息，如果被恶意的窃听者获取，将会造成严重的隐私泄露。因此，如何实现隐私保护是网络化估计系统实际应用中的一个关键问题。

复噪声注入方法因为在隐私保护方面的性能极高，所以被引入了融合估计系

统用于保护分布式融合估计的隐私[49, 50]，并给出了合理的假设解释了该方法的可行性。在实现隐私保护的同时，文献[49]讨论了带宽受限的情况，采用降维的策略减少了带宽资源的消耗，而文献[50]针对能量受限系统，在有限时域内优化了估计性能与能量损耗，提出了能量约束下最优的噪声调度策略。但是，如何在能量受限的情况下利用该策略实现高效的隐私仍是一个重要问题。针对大多数 NMFESs，相比较有额外消耗的复噪声方法，不叠加复噪声而直接对调制前的数据进行处理是一种更节能的方法。

作为密码学的经典方法，同态加密也被考虑用于保护估计系统的隐私。Z. Zhang 等[51]利用部分同态加密实现了单传感器下的安全估计，讨论了加密机制导致的量化误差问题，并给出了充分稳定条件。S. Emad 等[52]则考虑了多传感器融合估计系统的同态加密机制，利用公钥对估计器增益与传感器信息进行加密，并在合法用户端根据加法同态性利用密钥解密出原始的融合估计值。此外，该文献分别针对集中式和分布式两种多传感器融合结构提出了相应的协议，并证明了所提算法满足计算不可分性。然而，NMFESs 的计算能力一般不强，在众多实际应用中，无法有效发挥复杂加解密算法的性能。

因此，无额外能量消耗且计算量较小的数据扰动策略被广泛地研究。差分隐私作为经典且有效的扰动方法，在滤波与估计领域也获得了许多关注。最初，J. Le Ny 等[53]建立了基于差分隐私的估计系统模型，完善了状态空间模型相关的差分隐私概念和定义，并提出了基于差分隐私的 Kalman 滤波。随后，多种扩展滤波方法也被进一步讨论，比如集合 Kalman[54]，无迹 Kalman[55]和 MIMO 滤波[56]等。差分隐私机制主要存在的问题是噪声的扰动会影响所有接收数据的第三方，包括合法用户和窃听者。在这种情况下，合法用户收到的数据也是经过干扰的不准确信息，而如何减少噪声对合法用户的影响，即如何平衡合法估计的性能损失与隐私保护强度是一个重要的研究方向。此外，如何将差分隐私从数据库领域引入到其他研究领域也是一个研究热点。

除了差分隐私，状态隐私编码也是隐私保护估计领域中一种有效的方法，其由 A. Tsiamis 等[57]于 2018 年提出，而后进一步完善理论体系，并讨论了不稳定系统的情况[58]。该方法利用之前时刻的状态信息对当前状态进行修改，从而达到加密的目的。具体地，其通过构造某个时刻窃听者丢包事件使其丢失初始迭代数据，并利用系统的状态转移矩阵进行关联，导致窃听者的误差由于状态转移矩阵不断叠加。特别地，当系统不稳定时，由于状态转移矩阵的范数大于 1，窃听者均方误差随着该范数累乘而最终趋向无穷。与简单的白噪声加密机制不同，该设计方案是一种利用信息物理系统的物理层数据进行加密的典型方法。但是，当系统稳定时，其有效性也难以保证。

此外，还有一大类对抗窃听攻击的研究专注于调度策略的设计。在该框架中，窃听者与合法用户的信道均有一定概率丢包[59-61]，通过动态规划等方法优化目标函数。A. S. Leong 等[59]在最小化合法用户的估计误差协方差的同时最大化窃听者的估计误差协方差，L. Wang 等[61]进一步讨论了传感器能量受限的问题。

一般情况下，网络中并没有窃听者与合法用户之分，所有接收数据的设备都是不可信第三方。在这种情况下，需要同时讨论系统的隐私性能和数据实用性[40, 62]。E. Erdemir 等[40]限定了两条轨迹距离在一定范围内，即保证扰动后的数据精度有一定实用性，进而通过最小化互信息来保护隐私，而 E. Nekouei 等[62]则限定了隐私度量的最小值，之后最小化了估计的平均损失获取最优的加密策略。

值得注意的是，上述各种方法大多考虑单传感器观测的情况[53, 57-59]。相比较之下，多传感器融合估计系统有用较高的估计精度和鲁棒性，因此，保护其隐私安全亟待深入研究。在一般的多传感器融合结构中，存在一个融合中心用于收集局部的信息。因为各个传感器和融合中心都有频繁的通信，并且每个传感器发送的局部信息都可能被用于最终的融合估计，所以一般情况下不能只在某个局部传感器上加密，需要考虑所有传感器的信息安全[49, 50, 52]。此外，单一地利用传统策略进行加密可能无法完全发挥其有效性，如何根据信息物理系统的特性设计隐私保护策略将会有更为显著的效果。比如，利用通信信道增益设计零空间[47]以及利用状态转移矩阵放大误差[58]等。

1.3 本文研究工作

随着通信网络在多传感器信息融合领域的普及，NMFESs 的隐私保护研究也逐渐成为该领域的热点。然而，已有的隐私保护方案大多基于传统加解密策略，大多数成果只是将一些经典算法引入融合估计领域，鲜有根据融合估计系统的特性导出的有效机制。并且，许多经典算法可能不适用于实时估计领域。由于分布式融合估计框架的鲁棒性和高容错性，本文选择其作为主要讨论对象。为此，本文主要进行以下几个方面的研究。

1．针对 NMFESs 中传感器发送数据被窃听的问题，提出了一种随机噪声污染策略以保护分布式融合估计的隐私。该方法结合加权融合矩阵构造的零空间以及随机高斯白噪声等构造了一种污染向量，将其用于替代真实局部估计值的某些分量，使得窃听者依据局部估计误差协方差进行线性加权后的估计性能受到损害，并通过最大化窃听者估计误差协方差求取了最优的分量污染方式。同时，合法用户的融合中心采用了一步预估的补偿方式，即用误差相对较大但方差有界的预测值代替被污染的无用数据，从而避免估计误差协方差受剧烈影响，并根据预估的

协方差重新求得线性最小方差意义下的补偿融合矩阵。

2．针对多传感器融合估计系统的隐私保护问题，将差分隐私的概念与机制引入 NMFESs，通过在多个融合估计的平均值上添加高斯白噪声的方式实现差分隐私。进一步地，在传感器端设计了局部随机噪声注入结构，利用融合矩阵的逆将局部噪声机制转化为与融合估计上添加白噪声类似的结构，并给出所注入噪声方差的下界实现了与其相等的差分隐私水平。在这种情况下，窃听者不仅无法通过差分的方式获取局部数据，也不能利用线性加权融合的方式估计出准确的系统状态。同时，合法用户根据所注入噪声的协方差更新扰动后的局部估计误差协方差，并通过重新计算最小方差意义下的加权融合准则获取性能较高的融合估计值。

3．针对 NMFESs 遭受窃听攻击的问题，结合加权融合矩阵的零空间与差分隐私中的噪声机制构建了两步序贯噪声注入的加密方法，使得窃听者进行差分时由于噪声的累加引起局部估计误差协方差发散，进而导致其稳态加权融合估计的误差协方差发散。同时，合法用户根据所注入噪声的统计特性重新计算线性最小方差意义下的融合准则，并基于零空间的加密设计方式抵消了部分噪声的影响，获得了稳定且性能较优的分布式融合估计器。此外，结合了合法用户与窃听者双方的性能指标，提出了一种新的隐私保护度量方式，引入分数和对数函数对该度量进行放缩，并通过对该性能指标进行分段使得系统隐私性能的结果更加直观，方便了定性分析。

本文共有五章，全文组织结构如图 1-2 所示。

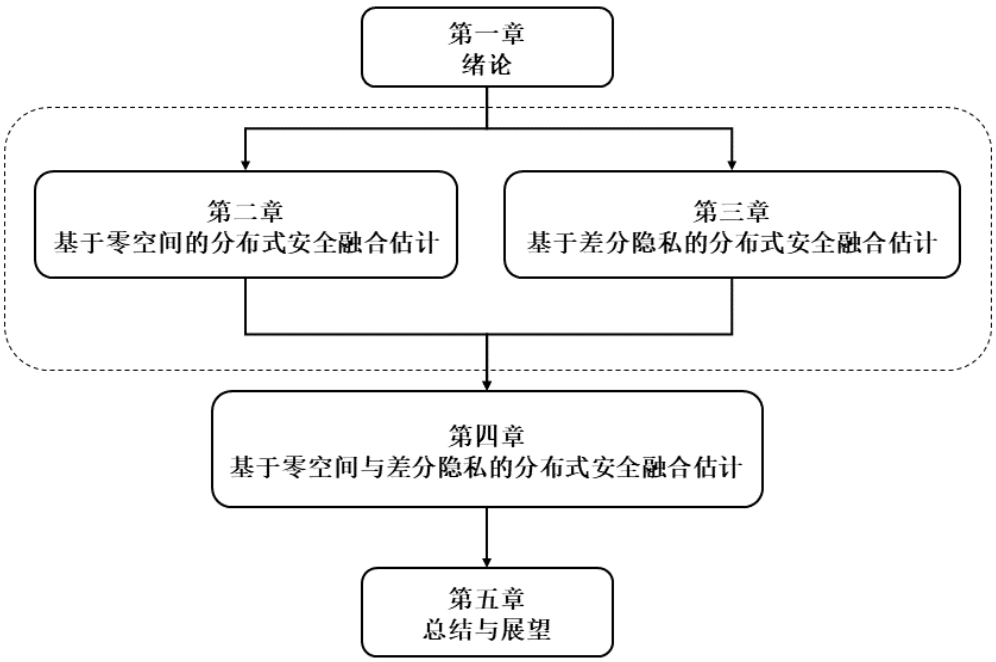

图 1-2　全文组织结构

Figure 1-2.　Structure of overall organization

论文章节的具体安排如下：第一章为绪论，包括研究背景及意义、相关概念以及研究现状；第二章首先介绍了基于 Kalman 滤波的线性加权分布式融合估计，然后给出了利用零空间以及随机噪声保护分布式融合估计隐私的研究结果；第三章首先介绍了基于稳态 Kalman 滤波的线性加权融合方式，接着给出了使用随机高斯白噪声实现差分隐私的研究结果；第四章结合零空间以及差分隐私机制，给出了针对稳定系统的分布式安全融合估计方法的研究结果；第五章总结了本文的研究工作，指出了在所提分布式安全融合估计方法中存在的问题和限制，并讨论了进一步的研究方向。

第二章 基于零空间的分布式安全融合估计

2.1 引言

由于网络的开放性和匿名性等特点，NMFESs 中的通信有遭受窃听攻击的风险。根据第一章分析，在信道中插入复噪声的方法[47]能使窃听者无法成功解码从而丢失传输信息。但是，该方法对通信信道的要求较高，并且，由于 NMFESs 能量一般有限，使用该方法所需的额外能量消耗会造成 NMFESs 传感器等使用寿命降低等问题。注意到，即使存在利用量化[49]、降维[19]、事件触发[63]、最优调度[50]等辅助工具减少加密消耗的可能性，该方法的能耗也比直接修改传输数据值的方法高。密码学[22, 23]是一类经典、有效且无需额外能量的隐私保护方法，然而其中大量的复杂计算可能占用 NMFESs 的计算资源，使得其正常工作受到严重影响，从而无法实现实时的状态估计；或者，在满足正常运行的条件下，系统没有足够的资源有效实现加解密机制。比如，同态加密[51, 52]虽然能实现计算不可分性，但是其中的指数计算等过程需要较大的计算量，也会在一定程度影响 NMFESs 的运行。

Kalman 滤波[53]是状态估计中一种经典的无偏估计方法，至今已经得到了广泛的应用，并且衍生出许多扩展形式[43, 44]。同时，根据第一章分析，分布式融合估计框架拥有高鲁棒性和高扩展性等众多优点。因此，本章首先介绍了基于时变 Kalman 滤波的线性加权分布式融合估计结构。为了提升算法的实用性，本章假设窃听者能力较强，即其已经知晓该融合准则，并且可以通过局部信息的融合获取高精度的状态估计值。然后，为了避免 NMFESs 中能量和计算量的过多消耗，本章考虑采用随机实噪声注入的方法实现隐私。具体地，本章利用线性加权融合准则设计了一种基于零空间的噪声注入加密策略，称之为随机噪声污染策略，使得窃听者采用线性最小方差融合时丢失部分局部估计的信息，导致融合估计精度下降从而无法获取准确的系统状态信息。并且，在合法用户端采用了一步预测的方式减小了所提加密策略带来的估计性能损失，根据相应统计特性重新计算融合准则从而获得性能较好的融合估计。最后，本章通过一个移动机器人仿真实例证明了所提方法的有效性。

2.2 系统建模与问题描述

2.2.1 基于时变 Kalman 滤波的线性加权分布式融合估计

考虑如下离散时间非线性时变状态空间模型：

$$\begin{cases} x(t+1) = f(x(t)) + B(t)w(t) \\ y_i(t) = h_i(x(t)) + v_i(t) \ (i = 1, \ldots, L) \end{cases} \tag{2-1}$$

其中 $x(t) \in \mathrm{R}^n$ 为系统状态， $y_i(t) \in \mathrm{R}^{m_i}$ 是传感器 i 的观测输出， $B(t) \in \mathrm{R}^{n \times n_w}$ 是具有合适维数的时变矩阵，L 为传感器数量。$w(t) \in \mathrm{R}^{n_w}$ 为系统的过程噪声，$v_i(t) \in \mathrm{R}^{m_i}$ 为传感器的观测噪声，并对噪声的统计特性作下列假设。

假设 2.1： $w(t)$ 与 $v_i(t)$ 是均值为 0 的高斯白噪声，并满足如下条件：

$$\mathrm{E}\{[w^{\mathrm{T}}(t) \quad v_i^{\mathrm{T}}(t)]^{\mathrm{T}}[w^{\mathrm{T}}(t_1) \quad v_j^{\mathrm{T}}(t_1)]\} = \delta(t, t_1)\mathrm{diag}\{Q, \delta(i, j)R_i\} \tag{2-2}$$

其中 Q 为噪声 $w(t)$ 的协方差，R_i 为噪声 $v_i(t)$ 的协方差，$\delta(i, j)$ 为克罗内克函数，即

$$\delta(i, j) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \tag{2-3}$$

在分布式融合估计框架中，每个传感器需要利用其自身观测进行局部状态估计。基于传感器的观测序列 $\{y_i(1), \cdots, y_i(t)\}$，非线性系统(2-1)的局部估计可用如下扩展 Kalman 滤波（Extended Kalman Filter, EKF）计算[64]：

$$\begin{cases} \varepsilon_i(t) = y_i(t) - h_i(f(\hat{x}_i(t-1))) \\ \hat{x}_i(t) = A_{J_i}(t-1)\hat{x}_i(t-1) + K_i^N(t)\varepsilon_i(t) \\ P_{ii}(t) = (I - K_i^N(t)C_{J_i}(t))P_{ii}^z(t) \\ P_{ii}^z(t) = A_{J_i}(t-1)P_{ii}(t-1)A_{J_i}^{\mathrm{T}}(t-1) + Q_B(t) \\ K_i^N(t) = P_{ii}^z(t)C_{J_i}^{\mathrm{T}}(t)(C_{J_i}(t)P_{ii}^z(t)C_{J_i}^{\mathrm{T}}(t) + R_i)^{-1} \end{cases} \tag{2-4}$$

其中 $Q_B(t) = B(t-1)QB^{\mathrm{T}}(t-1)$， $P_{ii}(t)$ 表示传感器 i 局部估计误差协方差矩阵，$P_{ii}^z(t)$ 表示预估误差协方差矩阵， $K_i^N(t)$ 为时变扩展 Kalman 增益， $\varepsilon_i(t)$ 代表新息序列，$\hat{x}_i(t)$ 代表局部状态估计（Local State Estimate, LSE），且

$$\begin{cases} A_{J_i}(t-1) = \left.\dfrac{\partial f(x(t))}{\partial x(t)}\right|_{x(t)=\hat{x}_i(t)} \\ C_{J_i}(t-1) = \left.\dfrac{\partial h_i(x(t))}{\partial x(t)}\right|_{x(t)=f(\hat{x}_i(t))} \end{cases} \tag{2-5}$$

然后，基于上述扩展 Kalman 滤波器，传感器 i 与传感器 j 之间的估计误差互协方差矩阵 $P_{ij}(t)$ 可由下式递归计算：

$$\begin{aligned} P_{ij}(t) &= (I - K_i^N(t)C_{J_i}(t))(A_{J_i}(t-1)P_{ij}(t-1)A_{J_i}^{\mathrm{T}}(t-1) \\ &\quad + B(t-1)QB^{\mathrm{T}}(t-1))(I - K_j^N(t)C_{J_j}(t))^{\mathrm{T}} \end{aligned} \tag{2-6}$$

通常情况下，非线性系统(2-1)可以简化为如下线性系统：

$$\begin{cases} x(t+1) = A(t)x(t) + B(t)w(t) \\ y_i(t) = C_i(t)x(t) + v_i(t) \ (i = 1, \ldots, L) \end{cases} \tag{2-7}$$

其中 $A(t) \in \mathrm{R}^{n \times n}$ 和 $C_i(t) \in \mathrm{R}^{m_i \times n}$ 是具有合适维数的时变矩阵，并且假设求上述离散时间的线性时变状态空间模型(2-7)满足如下一致完全可观可控条件。

假设 2.2：系统(2-4)一致完全可控并且一致完全可观，即存在标量 $\rho_i (i = 1,2,3,4) > 0$ 和正整数 $N > 0$ 使得当 $t > N$ 时如下不等式成立：

$$\begin{cases} \rho_1 I \le \sum_{j=t-N+1}^{t} \Pi(t, j)B(j)QB^{\mathrm{T}}(j)\Pi^{\mathrm{T}}(t, j) \le \rho_2 I \\ \rho_3 I \le \sum_{j=t-N+1}^{t} \Pi^{\mathrm{T}}(j,t)C_i^{\mathrm{T}}(j)R_i^{-1}C_i(j)\Pi(j,t) \le \rho_4 I \end{cases} \tag{2-8}$$

其中

$$\begin{cases} \Pi(j, j) = I \\ \Pi(t, j) = \prod_{l=1}^{t-j} A(t-l)(t > j) \\ \Pi(j,t) = \Pi^{-1}(t, j)(t < j) \end{cases} \tag{2-9}$$

类似于非线性系统下的扩展 Kalman 滤波，可以利用如下线性系统(2-7)的经典 Kalman 滤波器（Kalman Filter, KF）[65]计算局部估计 LSE：

$$\begin{cases} \varepsilon_i(t) = y_i(t) - C_i(t)A(t-1)\hat{x}_i(t-1) \\ \hat{x}_i(t) = A(t-1)\hat{x}_i(t-1) + K_i(t)\varepsilon_i(t) \\ P_{ii}(t) = (I - K_i(t)C_i(t))P_{ii}^z(t) \\ P_{ii}^z(t) = A(t-1)P_{ii}(t-1)A^{\mathrm{T}}(t-1) + Q_B(t) \\ K_i(t) = P_{ii}^z(t)C_i^{\mathrm{T}}(t)(C_i(t)P_{ii}^z(t)C_i^{\mathrm{T}}(t) + R_i)^{-1} \end{cases} \tag{2-10}$$

为了方便对非线性和线性系统进行统一分析，上式仍然用 $P_{ii}(t)$ 表示局部估计误差协方差，$P_{ii}^z(t)$ 表示预估误差协方差，$\hat{x}_i(t)$ 表示局部状态估计，而用 $K_i(t)$ 代表线性系统下的时变 Kalman 增益。同时，根据上述经典 Kalman 滤波方法(2-10)，估计误

差互协方差矩阵 $P_{ij}(t)$ 为：

$$\begin{aligned} P_{ij}(t) = (I - K_i(t)C_i(t))(A(t-1)P_{ij}(t-1)A^{\mathrm{T}}(t-1) \\ +B(t-1)QB^{\mathrm{T}}(t-1))(I - K_j(t)C_j(t))^{\mathrm{T}} \end{aligned} \tag{2-11}$$

在分布式多传感器融合估计结构中，每个传感器根据局部观测计算得到状态估计后，需要将其相应的实时局部估计 LSE 传输至融合中心，即上述所有经过 Kalman 滤波计算的信号 $\hat{x}_i(t)$ 都将被发送给融合中心。接着，融合中心在收到所有局部信息之后，采用线性加权求和的方式计算分布式融合估计，即：

$$\hat{x}(t) = \sum_{i=1}^{L} W_i(t)\hat{x}_i(t) \tag{2-12}$$

其中 $W(t) = [W_1(t) \quad W_2(t) \quad \cdots \quad W_L(t)] \in \mathrm{R}^{n \times nL}$ 为权重矩阵。根据线性最小方差（Linear Minimum Variance, LMV）准则，最优的权重矩阵可由如下引理计算。

引理 2.1[4]：线性加权融合方式(2-12)中，线性最小方差意义下的最优权重矩阵 $W(t)$ 由下式计算：

$$W(t) = (e^T P^{-1}(t)e)^{-1} e^T P^{-1}(t) \tag{2-13}$$

其中 $e = [I_n \quad \cdots \quad I_n]^{\mathrm{T}} \in \mathrm{R}^{nL \times n}$ 为 $nL \times n$ 的矩阵， $P(t) = (P_{ij}(t)) \in \mathrm{R}^{nL \times nL}(i, j = 1, 2, \cdots, L)$ 为 $nL \times nL$ 维的矩阵。此时，该分布式融合估计的误差协方差为 $P_f(t) = (e^T P^{-1}(t)e)^{-1}$ 。

2.2.2 问题描述与分析

在本章讨论的分布式融合估计系统中，由于窃听者的存在，需要对所有传感器发送的局部状态估计进行加密。根据之前分析，随机噪声方法拥有计算量低和无需额外能量消耗等众多优点。因此，本章基于主动注入随机噪声的方式提出了一种随机噪声污染的加密策略，并将加密处理之后的数据称为主动污染后的局部估计（Actively Contaminated Local Estimates, ACLEs），符号表示为 $\hat{x}_i^r(t)$，该策略的具体设计在后续中给出。在本章中，窃听者可以持续监听传感器与融合中心的传输信道，即窃听者可以获取实时的 $\hat{x}_i^r(t)$ 。根据加权融合准则(2-12)，窃听者的分布式融合估计为：

$$\hat{x}_e(t) = \sum_{i=1}^{L} W_i(t)\hat{x}_i^r(t) \tag{2-14}$$

而在合法用户端，为了减小加密机制带来的性能损失以获取较高精度的融合估计器，需要基于接收到的 ACLEs 采用相应的补偿方式，以获取补偿局部估计（Compensated Local Estimates, CLEs），符号表示为 $\hat{x}_i^c(t)$ 。然后，利用精度相对较高的 CLEs 进行加权融合估计，合法用户最终得到如下分布式融合估计（Distributed

Fusion Estimate, DFE）：

$$\hat{x}_f(t)=\sum_{i=1}^{L}W_{fi}(t)\hat{x}_i^c(t) \tag{2-15}$$

图 2-1 展示了基于随机噪声污染策略的分布式融合估计框架。

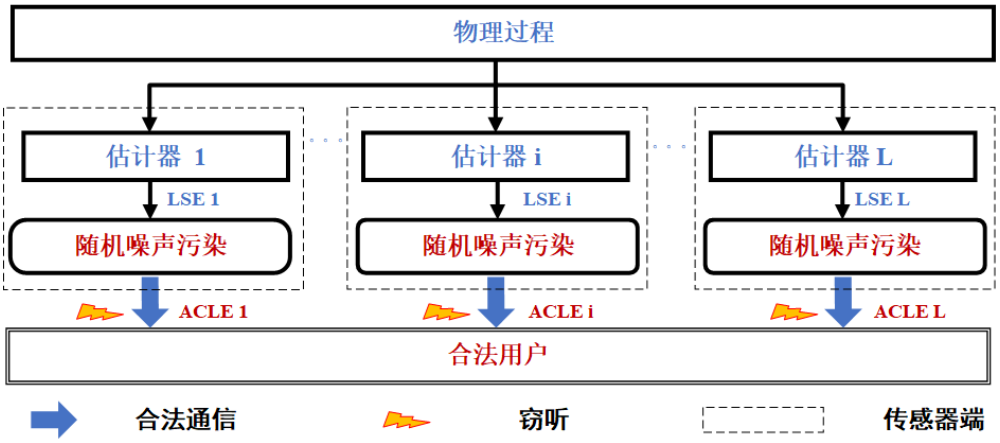

图 2-1　基于随机噪声污染策略的分布式安全融合估计框架

Figure 2-1.　The structure of distributed secure fusion estimation based on the actively contamination strategy

根据上述窃听者与合法用户估计器的分析，为了保护系统的隐私，本章需要解决的问题主要为如下两点：（1）如何基于随机噪声设计有效的加密策略使窃听者的融合估计性能下降；（2）如何设计有效的补偿策略使合法用户的融合估计器稳定且性能较优。

注释 2.1：在分布式估计框架中，每个子系统利用自身观测以及邻居信息估计其自身状态。而在本章考虑的分布式融合估计框架中，存在一个融合中心用于收集所有传感器的信息，其基于特定的融合准则对所有局部信息进行融合以提升估计性能。此外，在该框架中，由于传感器将局部估计发送至融合中心后采用线性加权的准则进行融合，因此，该框架也具有较高的鲁棒性。

2.3 基于零空间的分布式安全融合估计器

2.3.1 基于零空间的随机噪声污染策略

在网络化多传感器融合估计系统中，传感器与融合中心的通信数据容易被窃听者窃取造成隐私泄露。更具体地，在分布式融合估计结构中，窃听者首先窃听传感器发送的局部估计，然后采用最优的加权融合方式对这些数据进行处理。进行融合的理由有两个：（1）仅仅根据协方差等统计信息就可以获取精度较高的状态估计，且无需较大的计算量；（2）合法用户也对局部信息进行了融合，与合法

用户操作的一致性有利于获取更准确的系统信息。在这种情况下，为了对窃听者获取的信息造成干扰，本章提出了一种随机噪声污染策略对局部信息进行加密处理，即：将部分局部估计用随机噪声替换，从而使窃听者损失相应分量的信息。局部传感器的具体操作如下：

$$\hat{x}_i^r(t) = H_i(t)\hat{x}_i(t) + (I - H_i(t))\phi_i(t) \tag{2-16}$$

其中 $\hat{x}_i^r(t)$ 表示被噪声主动污染后的局部估计 ACLE，污染向量 $\phi_i(t)$ 与状态估计同等维数，用于替换相应的局部估计分量，分量选择矩阵 $H_i(t)(i=1,\cdots,L)$ 则用于确定哪些分量被污染，该矩阵为二值对角矩阵，即 $H_i(t) \triangleq \mathrm{diag}\{\gamma_{i1}(\mathrm{t}),\cdots,\gamma_{in}(\mathrm{t})\}$，其中 $\gamma_{ij}(\mathrm{t}) \in \{0,1\}$ 表示第 j 个对角元素。同时，本章假设遭受污染的状态估计分量的数量不变，即 $\sum_{j=1}^{n}\gamma_{ij}(\mathrm{t}) = \bar{s}_0$，其中 $\bar{s}_0$ 为常值。此时，分量选择矩阵所有可能的值将构成如下集合：

$$\hat{H}_F(t) = \{H_F^1(t),\cdots,H_F^k(t),\cdots,H_F^{\bar{k}}(t)\} \tag{2-17}$$

其中 $\bar{k} = C_{nL}^{\bar{s}_0}$ 表示集合中元素的数量，k 为索引，而 $H_F^k(t)$ 表示索引 k 对应的二值对角矩阵。因为上述集合包含了受污染数量固定情况下所有可能的分量选择矩阵，所以此时的最优分量选择矩阵必然从上述集合中选取。

在上述所提加密机制的框架下，主动污染向量 $\phi_i(t)$ 的设计尤为重要。为了能使窃听者估计时损失有效信息，本章考虑采用与融合准则相关的零空间来设计污染向量，具体地，零空间 $\Psi(t)$ 满足下列方程：

$$\bar{W}(t)\Psi(t) = 0 \tag{2-18}$$

其中

$$\begin{cases} \bar{W}(t) = W(t)\bar{H}_F(t) \\ \bar{H}_F(t) \triangleq \mathrm{diag}\{I - H_1(t),\cdots,I - H_L(t)\} \\ \Psi(t) \in \mathrm{R}^{nL\times(nL-r(t))}, r(t) = \mathrm{rank}\{\bar{W}(t)\} \end{cases} \tag{2-19}$$

然后，结合该零空间 $\Psi(t)$ 以及随机高斯白噪声 $a_F(t)$ 构成如下污染向量：

$$\phi(t) \triangleq \mathrm{col}\{\phi_1(t),\cdots,\phi_L(t)\} = \Psi(t)a_F(t) \tag{2-20}$$

其中 $a_F(t) \triangleq \mathrm{col}\{a_f(t),\cdots,a_f(t)\} \in \mathrm{R}^{nL-r(t)}$，而 $a_f(t)$ 是均值为 0 方差为 Q_a 的高斯白噪声。这里，添加高斯噪声的目的是为了保证传输数据有足够的随机性，防止窃听者检测出异常。随着协方差的收敛，权重矩阵等也将收敛至固定区间，进而导致零空间波动逐渐减小。再加上量化和编码等一些必要的压缩，某些污染向量的数值波动可能变得很小甚至不变，导致窃听者有很大可能觉察到相应稳定数据的异常，提升其故障检测的能力，从而影响隐私保护机制的性能。因此，本文采用一

种简单的随机扰动方式，即在零空间上叠加高斯白噪声，使得在不改变其隐私保护性能的情况下保证传输数据的随机性。

在上述加密机制作用下，当窃听者根据线性最小方差准则进行按矩阵加权融合时，特定分量的求和结果在零空间的作用下变为 0，从而导致分布式融合估计的准确性大打折扣。

2.3.2 线性系统下的安全融合估计器设计

虽然上述扰动策略可以有效降低窃听者融合估计的精度，但是合法用户的融合估计性能也同样地会受影响，原因是合法用户接收到的局部传感器信息也是扰动后的局部估计，并且也只能利用这些估计进行加权融合。因此，如何减少融合中心的估计性能损失成为一个至关重要的问题。为此，本章考虑在融合中心基于分量选择矩阵采用一步预估的补偿方式减少所提加密带来的负面影响。最优分量选择矩阵的设计以及具体的补偿方式由下列定理给出。

定理 2.1：利用所提随机噪声策略(2-16)保护线性系统状态估计隐私时，最大化窃听者融合估计误差协方差的最优分量选择矩阵为：

$$H_F^L(t)=\arg\max_{H_F^k(t)}\left\{\left\|H_F^k(t)-H_F^m(t)\right\|_2\right\} \tag{2-21}$$

其中

$$\begin{cases} H_F^m(t)=\Lambda_F(t)(\Lambda_F(t)+P(t))^{-1} \\ \Lambda(t)=\mathrm{E}\{x(t)x^{\mathrm{T}}(t)\}=A(t-1)\Lambda(t-1)A^{\mathrm{T}}(t-1)+B(t-1)QB^{\mathrm{T}}(t-1) \\ \Lambda_F(t)=\mathrm{E}\{x_F(t)x_F^{\mathrm{T}}(t)\},\ x_F(t)\triangleq\mathrm{col}\{x(t),\cdots,x(t)\}\in\mathrm{R}^{nL} \end{cases} \tag{2-22}$$

同时，合法用户融合中心采用如下一步预测进行补偿可得稳定的融合估计：

$$\hat{x}_i^c(t)=H_i(t)\hat{x}_i^r(t)+(I-H_i(t))A(t-1)\hat{x}_f(t-1) \tag{2-23}$$

证明：为了获得最优的分量选择矩阵，需要尽可能破坏窃听者的估计性能，即：最大化窃听者估计误差协方差。然而，根据之前分析，可能的分量矩阵的数量有限。因此，最优分量选择矩阵的设计需要分为两步。

第一步，需要计算出窃听者估计误差方差的最小值。根据式(2-12)以及系统状态方程(2-1)，可以得到受污染的局部估计误差表达式：

$$\tilde{x}_i^r(t)=H_i(t)\tilde{x}_i(t)+(I-H_i(t))(x(t)-\phi_i(t)) \tag{2-24}$$

进一步地，根据线性加权融合准则(2-10)，窃听者的融合估计误差可写为：

$$\tilde{x}_e(t)=W(t)H_F(t)\tilde{x}_F(t)+\bar{W}(t)x_F(t) \tag{2-25}$$

从上式可以直观地看出，窃听者的融合估计由于损失了部分状态信息，性能将受

到严重影响，且该性能损失与实际系统状态的大小有关。若系统不稳定，即系统状态值发散的情况，则窃听者的融合估计误差必然也发散。但是，本章考虑的随机噪声污染策略旨在保护任意估计系统的隐私，即实际系统稳定性未知。为此，当系统稳定时，就需要最大化窃听者估计误差协方差以寻求最优的隐私保护性能。注意到，当系统不稳定时，该最大化的目标仍适用，虽然发散的最终目的不变，但是发散的速率将被最大化，也有利于安全融合估计。

为了求得最小方差意义下的变量，首先给出窃听者的融合估计误差协方差如下：

$$P_e(t)=W(t)(H_F(t)P(t)H_F(t)+\bar{H}_F(t)\Lambda_F(t)\bar{H}_F(t))W^{\mathrm{T}}(t) \tag{2-26}$$

对上述协方差的迹求关于选择矩阵的偏导可得：

$$\frac{\partial \mathrm{Tr}P_e(t)}{\partial H_F(t)}=W^{\mathrm{T}}(t)W(t)(2H_F(t)P(t)-2\Lambda_F(t)+2H_F(t)\Lambda_F(t)) \tag{2-27}$$

当令该偏导数等于 0 时，可以得到最小方差意义下的分量选择矩阵，即 $H_F^m(t)=\Lambda_F(t)(\Lambda_F(t)+P(t))^{-1}$。

因为第一步得到的是最小值，并且对应的分量选择矩阵大概率不是二值对角矩阵，所以需要进一步的操作。为了获取最大化窃听者方差的结果，优化的第二步将从相应的有限集合(2-13)中选择距离该最小值对应变量最远的元素。因此，本章通过最大化 2-范数得到所需最值，如定理 2.1 中式(2-17)所示。

此外，根据之前的分析，该随机噪声污染策略同时也会降低合法用户融合中心的估计性能。为了削弱该负面影响，本章借用文献[9]中降维补偿的思想，采用一步预估代替被污染的局部数据。为了计算补偿后最优的权重矩阵，需要重新计算估计误差协方差矩阵。

首先，根据补偿局部估计与系统状态得出其误差表达式：

$$\tilde{x}_i^c(t)=H_i(t)\tilde{x}_i(t)+(I-H_i(t))(A(t-1)\tilde{x}_f(t-1)+B(t-1)w(t-1)) \tag{2-28}$$

$\tilde{x}_f(t-1)$ 为融合估计(2-11)的误差，可由下式计算：

$$\begin{aligned}\tilde{x}_f(t-1)=&W_f(t-1)H_F(t-1)\tilde{x}_F(t-1)+W_f(t-1)\bar{H}_F(t-1)\\&\times(A(t-2)\tilde{x}_f(t-2)+B(t-2)w(t-2))\end{aligned} \tag{2-29}$$

其中 $\tilde{x}_F^{\mathrm{T}}(t-1)\triangleq[x(t-1)-\hat{x}_1(t-1)\quad\cdots\quad x(t-1)-\hat{x}_L(t-1)]$。然后，补偿局部估计误差(2-23)的互协方差为：

$$\begin{aligned}\Sigma_{ij}(t)=&H_i(t)P_{ij}(t)H_i(t)\\&+H_i(t)(\Phi_i^{\mathrm{T}}(t)A^{\mathrm{T}}(t-1)+(I-K_i(t)C_i(t))B(t-1)QB^{\mathrm{T}}(t-1))(I-H_j(t))\\&+(I-H_j(t))(A(t-1)\Phi_j(t)+B(t-1)QB^{\mathrm{T}}(t-1)(I-K_j(t)C_j(t))^{\mathrm{T}})H_j(t)\\&+(I-H_i(t))(A(t-1)P_f(t-1)A^{\mathrm{T}}(t-1)+B(t-1)QB^{\mathrm{T}}(t-1))(I-H_j(t))\end{aligned} \tag{2-30}$$

其中

$$\begin{cases}\Phi_i(t)=(W_f(t-1)\bar{H}(t)A(t-2)\Phi_i(t-1)+W_f(t-1)H_F(t)\hat{P}_i(t-1)\\ \qquad +W_f(t-1)\bar{H}_F(t)B(t-2)QB^{\mathrm{T}}(t-2)(I-K_i(t-1)C_i(t-1))^{\mathrm{T}})\\ \qquad \times(A(t-1)-K_i(t)C_i(t)A(t-1))^{\mathrm{T}}\\ \hat{P}_i(t)\triangleq \mathrm{col}\{P_{1i}^{\mathrm{T}}(\mathrm{t}),\cdots,P_{Li}^{\mathrm{T}}(\mathrm{t})\}\end{cases} \tag{2-31}$$

根据引理 2.1，可得式(2-15)中的权重矩阵为：

$$W_f(t)=(e^T\Sigma_f^{-1}(t)e)^{-1}e^T\Sigma_f^{-1}(t) \tag{2-32}$$

其中 $W_f(t)=[W_{f1}(t)\quad W_{f2}(t)\quad \cdots\quad W_{fL}(t)]$，$\Sigma_f(t)=(\Sigma_{ij}(t))\in \mathrm{R}^{nL\times nL}(i,j=1,2,\cdots,L)$。并且，相应的融合估计误差协方差为：

$$P_f(t)=(e^T\Sigma_f^{-1}(t)e)^{-1} \tag{2-33}$$

证毕。

值得注意的是，无论传感器端如何设置各个估计误差协方差的初值，本章的随机噪声污染策略对窃听者造成的影响将趋于相同，即最终的最优分量选择矩阵接近一致。为证明该性质，首先根据局部 Kalman 滤波器写出相应估计误差系统：

$$\tilde{x}_i(t+1)=\Omega_i(t)\tilde{x}_i(t)+\delta_i(t) \tag{2-34}$$

其中

$$\begin{cases}\Omega_i(t)=(I-K_i(t)C_i(t))A(t)\\ \delta_i(t)=(I-K_i(t)C_i(t))B(t)w(t)-K_i(t)v_i(t)\end{cases} \tag{2-35}$$

根据上述误差的表达式，相应的协方差系统可表示为：

$$P_{ij}(t)=\Gamma_i(t,t_0)P_{ij}(t_0)\Gamma_j^{\mathrm{T}}(t,t_0)+\hat{\Delta}_{ij}(t,t_0) \tag{2-36}$$

其中

$$\begin{cases}\Gamma_i(t,t_0)=\prod_{l=0}^{t-t_0-1}\Omega_i(t_0+l)\ (\Gamma_i(t_0,t_0)=I)\\ \Delta_{ij}(t)\triangleq \mathrm{E}\{\delta_i(\mathrm{t})\delta_j^{\mathrm{T}}(\mathrm{t})\}\\ \hat{\Delta}_{ij}(t,t_0)=\sum_{l=0}^{t-t_0-1}\Gamma_i(t,t_0+l+1)\Delta_{ij}(t+1)\Gamma_j^{\mathrm{T}}(t,t_0+l+1)\end{cases} \tag{2-37}$$

根据假设 2.2 中系统一致可观可控的条件，经典 Kalman 滤波是一致渐进稳定的，即上述协方差系统(2-36)满足下列条件[4]：

$$\begin{cases}\left\|\Gamma_i(t,t_0)\right\|\le c_1e^{-c_2(t-t_0)}\to 0\ \left(t\to\infty,c_1,c_2>0\right)\\ \left\|\Gamma_j(t,t_0)\right\|\le c_3e^{-c_4(t-t_0)}\to 0\ \left(t\to\infty,c_3,c_4>0\right)\end{cases} \tag{2-38}$$

然后，任意假设两个初始协方差分别为 $P_{ij}^1(t_0)$ 和 $P_{ij}^2(t_0)$，根据上式可得：

$$\left\|P_{ij}^{1}(t)-P_{ij}^{2}(t)\right\| \le c_1 c_3 e^{-(c_2+c_4)(t-t_0)} \left\|P_{ij}^{1}(t_0)-P_{ij}^{2}(t_0)\right\| \to 0 \ \left(t\to\infty, c_1, c_2, c_3, c_4 > 0\right) \tag{2-39}$$

即无论初始值如何调整，Kalman 滤波终会收敛，这使得分量选择矩阵最终接近相同，从而保证随机噪声污染策略最终的效果一致。

注释 2.2：从随机噪声污染策略的构成可以看出，其中所有矩阵，包括估计误差协方差和状态方差，都是迭代求解的，不需要知道系统的实时信息。因此，可以通过提前或者利用系统空闲时间来进行加密策略的计算，以进一步降低该方法的计算量，保证加密策略与状态估计的实时性。此外，由于加密策略的相关信息不通过传感器与融合中心的信道进行实时传输，因此，该策略不会被窃听者获取，即窃听者无法根据分量选择矩阵进行相应补偿。

2.3.3 非线性系统下的安全融合估计器设计

实际场景中，系统的动态一般为非线性，因此，将所提随机噪声污染策略嵌入至非线性系统可以获得更为广泛的应用。但是，与线性系统的分析不同，非线性系统由于其状态协方差无法迭代，导致 $\Lambda_F(t)$ 无法计算。为了解决该问题，本章提出了如下定理。

定理 2.2：利用所提随机噪声污染策略(2-16)保护非线性系统状态估计隐私时，最大化窃听者融合估计误差协方差的最优分量选择矩阵为

$$H_F^N(t) = \arg \max_{H_F^k(t)} \mathrm{Tr}\{H_F^k(t) P(t) H_F^k(t)\} \tag{2-40}$$

并且，合法用户融合中心采用如下一步预测进行补偿：

$$\hat{x}_i^c(t) = H_i(t)\hat{x}_i^r(t) + (I - H_i(t)) f(\hat{x}_f(t-1)) \tag{2-41}$$

证明：为了尽可能防止窃听者获取准确的系统状态估计值，同样地，需要最大化窃听者估计误差协方差(2-26)。但是，非线性系统下的 $\Lambda_F(t)$ 无法计算，因此，本章考虑仅最大化协方差前半部分。由于分量选择矩阵的数量有限，从相应的有限集合中容易选出最优的分量选择矩阵如式(2-40)所示。

此外，合法用户采用式(2-41)中非线性一步预测，获得的补偿局部估计误差协方差与线性系统类似：

$$\begin{aligned}\Sigma_{ij}(t) = {} & H_i(t)P_{ij}(t)H_i(t) + H_i(t)(\Phi_i^{\mathrm{T}}(t)A_{J_i}^{\mathrm{T}}(t-1) + (I - K_i^N(t)C_{J_i}(t)) \\ & \times B(t-1)QB^{\mathrm{T}}(t-1))(I - H_j(t)) + (I - H_j(t))(A_{J_i}(t-1)\Phi_j(t) \\ & + B(t-1)QB^{\mathrm{T}}(t-1)(I - K_j^N(t)C_{J_j}(t))^{\mathrm{T}})H_j(t) + (I - H_i(t)) \\ & \times (A_{J_i}(t-1)P_f(t-1)A_{J_i}^{\mathrm{T}}(t-1) + B(t-1)QB^{\mathrm{T}}(t-1))(I - H_j(t))\end{aligned} \tag{2-42}$$

其中

$$\begin{aligned}\Phi_i(t)=&(W_f(t-1)\bar{H}(t)A_{J_i}(t-2)\Phi_i(t-1)+W_f(t-1)H_F(t)\hat{P}_i(t-1)\\&+W_f(t-1)\bar{H}_F(t)B(t-2)QB^{\mathrm{T}}(t-2)(I-K_i^N(t-1)C_{J_i}(t-1))^{\mathrm{T}})\\&\times(A_{J_i}(t-1)-K_i^N(t)C_{J_i}(t)A_{J_i}(t-1))^{\mathrm{T}}\end{aligned}\tag{2-43}$$

根据上述协方差计算方式，合法用户基于各个估计的协方差，并根据引理 2.1，可同样计算出补偿后的最优权重矩阵为(2-32)。证毕。

2.4 示例

2.4.1 线性系统仿真实例

为了证明本章所提方法的有效性，本节考虑采用一个移动机器人目标跟踪系统作为仿真实例。本节假设机器人作匀速直线运动，并且其状态转移为线性。令 $s(t)$ 和 $\dot{s}(t)$ 分别表示移动机器人的位置和速度，将状态向量设为 $x(t)=\mathrm{col}\{s(t),\dot{s}(t)\}$，根据匀速直线运动模型[2]，此时，式(2-7)中的系统矩阵可表达为如下：

$$A(t)=\begin{bmatrix}1 & f_s(t)\\0 & 1\end{bmatrix},B(t)=\begin{bmatrix}\dfrac{f_s^2(t)}{2}\\f_s(t)\end{bmatrix}\tag{2-44}$$

同时，本节考虑布置两个传感器对该机器人进行观测，式(2-7)中对应的两个观测矩阵分别为：

$$C_1(t)=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},C_2(t)=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\tag{2-45}$$

机器人的移动以及传感器的观测可能会存在扰动，因此，根据经验将系统噪声和观测噪声的方差设置为：

$$Q=1,R_1(t)=\begin{bmatrix}0.9 & 0\\0 & 0.5\end{bmatrix},R_2(t)=\begin{bmatrix}0.7 & 0\\0 & 0.8\end{bmatrix}\tag{2-46}$$

同时，设置插入的随机噪声方差为 $Q_a=1$。

根据本章所提方法，利用随机噪声污染策略对需要发送的局部 Kalman 估计值进行加密，并在合法用户端利用线性系统下的一步预测方式进行补偿。图 2-2 绘制了系统的位置与时间关系图，包括目标、合法用户以及窃听者的轨迹。从该图中可以看出，窃听者由于传感器端的随机噪声加密机制，其获取的目标位置信息与机器人真实轨迹差距较大，体现了所提随机噪声污染策略对窃听者融合估计的影响。此外，合法用户的融合中心基于补偿策略计算的分布式融合估计能准确地跟上机器人的位置状态，还体现了一步预测补偿方式的可行性。

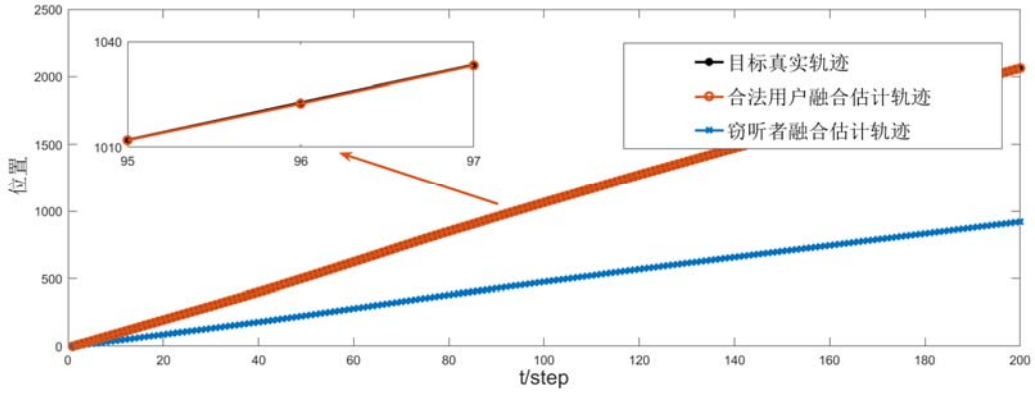

图 2-2　线性系统下目标、合法用户以及窃听者的轨迹

Figure 2-2.　The trajectories of target, legitimate user and eavesdropper for stable systems

为了进一步分析，接下来对系统各个稳定的估计作比较。由于各种随机噪声的存在，本章采用经过 100 步 Monte-Carlo 方法模拟得到的均方误差（Mean Square Error, MSE）来衡量估计的性能。图 2-3 展示了补偿局部估计、融合估计以及最优的 MSE。从该图中可以看出，补偿融合估计的 MSE 有界且明显小于两个局部估计的 MSE，验证了补偿融合估计器的稳定性，即补偿策略的有效性，也说明了线性加权的融合方法能够有效利用多个传感器的信息提升性能。但是，由于仅采用预测进行补偿，真正的局部估计分量仍未知，使得该补偿融合估计的 MSE 不可避免地大于不受污染情况下最优的 MSE。

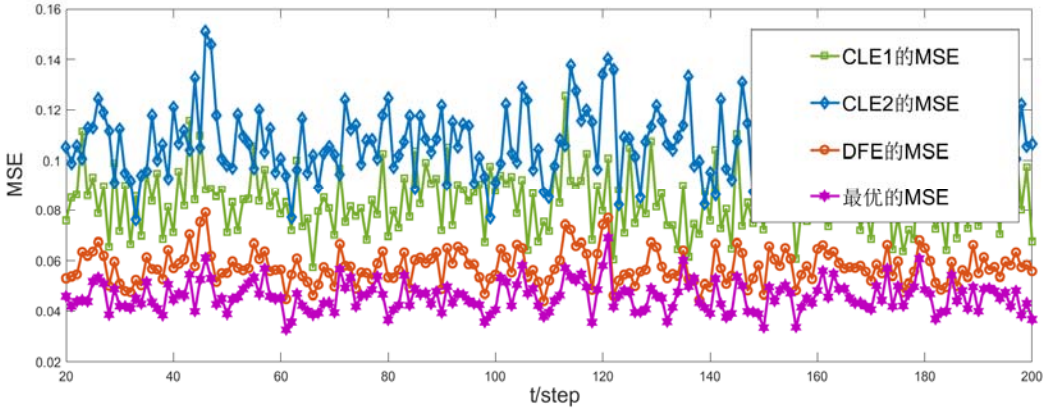

图 2-3　线性系统下补偿局部估计、融合估计以及最优的 MSE

Figure 2-3.　The MSEs of CLEs, DFE and optimal estimate for stable systems

接着，图 2-4 绘制了窃听者的估计误差协方差，并与文献[49]所提的复噪声注入方法进行了对比。从下图中可以看出，在所提随机实噪声污染策略的干扰下，窃听者融合估计的 MSE 趋向无穷，验证了所提加密方法的有效性。当文献[49]中加密机制的复噪声能量为 1 时，窃听信道的信噪比仍然比较大，导致窃听者仍有

较大概率可以成功解调，所以其 MSE 可以保证有界；而当复噪声能量足够大时，比如 5.6097，窃听者的 MSE 才可以以指数级别发散，可见该方法需要较大的能量消耗，也体现了所提随机噪声方法在能量方面的优势。

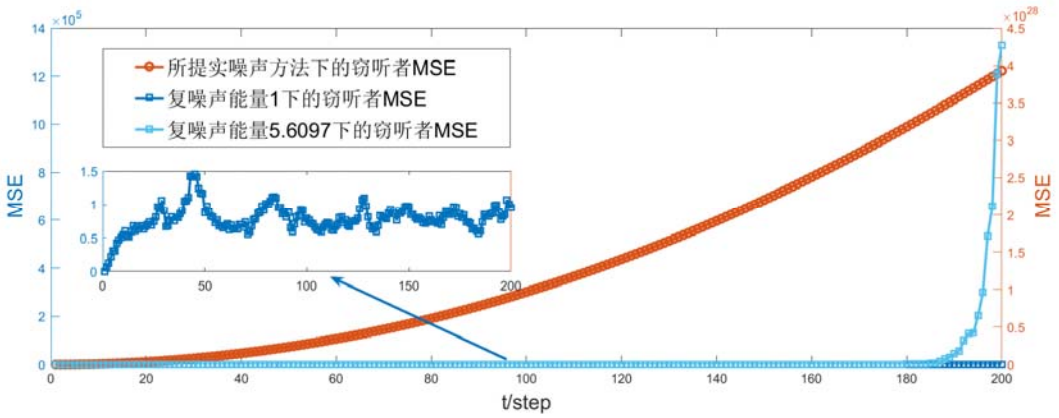

图 2-4　线性系统下窃听者的 MSE 对比

Figure 2-4.　The Comparison of MSEs of eavesdroppers for stable systems

2.4.2 非线性系统仿真实例

本节同样考虑采用一个机器人仿真模型以验证非线性系统情况下所提加密策略的有效性。设置机器人以非线性的动态运行，其状态转移方式如下[32]：

$$\begin{cases} S_x(t+1) = S_x(t) + T_0\zeta_v(t)\cos(\Theta(t)) \\ S_y(t+1) = S_y(t) + T_0\zeta_v(t)\sin(\Theta(t)) \\ \Theta(t+1) = \Theta(t) + T_0\zeta_w(t) \end{cases} \tag{2-47}$$

其中 $T_0 = 1$ 为采样时间，$(S_x(t), S_y(t))$ 为机器人的中心笛卡尔坐标，$\Theta(t)$ 为方向角，$\zeta_v(t)$ 和 $\zeta_w(t)$ 分别为线速度和角速度的控制信号。当构建状态向量为 $x(t) \triangleq \mathrm{col}\{S_x(t), S_y(t), \Theta(t)\}$ 之后，根据上述动态(2-48)可以得出非线性系统(2-1)中的状态转移函数为：

$$f(x(t)) = \begin{bmatrix} S_x(t) + T_0\zeta_v(t)\cos(\Theta(t)) \\ S_y(t) + T_0\zeta_v(t)\sin(\Theta(t)) \\ \Theta(t) + T_0\zeta_w(t) \end{bmatrix} \tag{2-48}$$

同时，本节考虑布置四个固定的地标 (S_{x_i}, S_{y_i}) $(i = 1,2,3,4)$ 用于实时监测机器人的动向，其主要观测机器人与地标之间的距离 $d_i(t)$ 和角度 $\theta_i(t)$，相应的观测值可表达为：

$$\begin{cases} d_i(t) = \sqrt{(S_{x_i} - S_x(t))^2 + (S_{y_i} - S_y(t))^2} \\ \theta_i(t) = \Theta(t) - \arctan\left(\dfrac{S_{y_i} - S_y(t)}{S_{x_i} - S_x(t)}\right) \end{cases} \tag{2-49}$$

考虑两个地标为一组，相应的观测方程按如下构造：

$$\begin{cases} y_1(t)=\begin{bmatrix} h_1(x(t)) \\ h_2(x(t)) \end{bmatrix}+v_1(t) \\ y_2(t)=\begin{bmatrix} h_3(x(t)) \\ h_4(x(t)) \end{bmatrix}+v_2(t) \\ h_i(x(t))=\begin{bmatrix} d_i(t) \\ \theta_4(t) \end{bmatrix}\ (i=1,2,3,4) \end{cases} \tag{2-50}$$

为了使用扩展 Kalman 滤波，需要对上述非线性函数进行泰勒级数展开，包括非线性状态转移和非线性观测。受限，定义如下变量：

$$\begin{cases} \tilde{S}_{x_i}(t)=S_{x_i}-S_x(t) \\ \tilde{S}_{y_i}(t)=S_{y_i}-S_y(t) \\ \tilde{S}_{xy_i}(t)=\tilde{S}_{x_i}^2(t)+\tilde{S}_{y_i}^2(t) \end{cases} \tag{2-51}$$

然后通过线性化可以得到式(2-5)中矩阵为：

$$\begin{cases} A_{J_i}(t)=\begin{bmatrix} 1 & 0 & -T_0\zeta_v(t)\sin(\Theta(t)) \\ 0 & 1 & T_0\zeta_v(t)\cos(\Theta(t)) \\ 0 & 0 & 1 \end{bmatrix}\Bigg|_{x(t)=x^*} \\ D_{J_i}(t)=\begin{bmatrix} -\dfrac{\tilde{S}_{x_i}(t)}{\sqrt{\tilde{S}_{xy_i}(t)}} & -\dfrac{\tilde{S}_{y_i}(t)}{\sqrt{\tilde{S}_{xy_i}(t)}} & 0 \\ -\dfrac{\tilde{S}_{y_i}(t)}{\sqrt{\tilde{S}_{xy_i}(t)}} & \dfrac{\tilde{S}_{x_i}(t)}{\sqrt{\tilde{S}_{xy_i}(t)}} & 1 \end{bmatrix}\Bigg|_{x(t)=x^*} \\ C_{J_1}(t)=\begin{bmatrix} D_{J_1}(t) \\ D_{J_2}(t) \end{bmatrix}, C_{J_2}(t)=\begin{bmatrix} D_{J_3}(t) \\ D_{J_4}(t) \end{bmatrix} \end{cases} \tag{2-52}$$

接着，将非线性的噪声加密策略与补偿策略应用至上述系统，得到了如下结果。首先，在图 2-5 中绘制了非线性系统下的机器人位置跟踪结果。从图中可以看出，窃听者获取的状态轨迹存在较大偏差，说明即使是针对非线性系统设计的随机噪声污染策略，也可以有效地防止窃听者进行跟踪。同时，合法用户的融合估计也能较好地跟上目标，说明合法用户可以利用一步预测实现有效补偿，从而较为准确地跟上机器人的真实轨迹。但是，合法用户的估计也存在一定偏移，这是线性化误差与噪声扰动共同作用的结果。

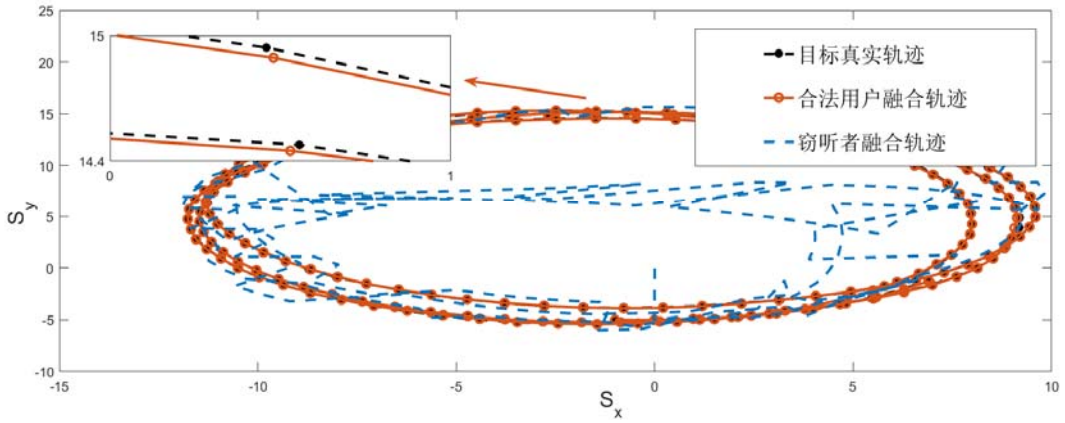

图 2-5　非线性系统下目标、合法用户以及窃听者的轨迹

Figure 2-5.　The trajectories of target, legitimate user and eavesdropper for unstable systems

同样地，本节也用 MSE 来表示估计的精度。图 2-6 中展示了合法用户中各个估计的 MSE，并在图 2-7 绘制了窃听者融合估计的 MSE。从图 2-6 中可以看出，合法用户所有补偿估计的 MSE 均有界，且补偿融合估计的 MSE 均小于补偿局部估计 CLE 的 MSE，同时，融合估计的 MSE 仅略大于最优估计的 MSE，说明了非线性一步预测补偿方式的可行性，也体现了非线性补偿融合方式的有效性。相比较图 2-6 中合法用户的 MSE 量级，图 2-7 中窃听者的 MSE 明显属于大量级，其数值范围为合法用户的数千倍，说明了随机噪声污染策略用于保护非线性系统的隐私也较为有效。

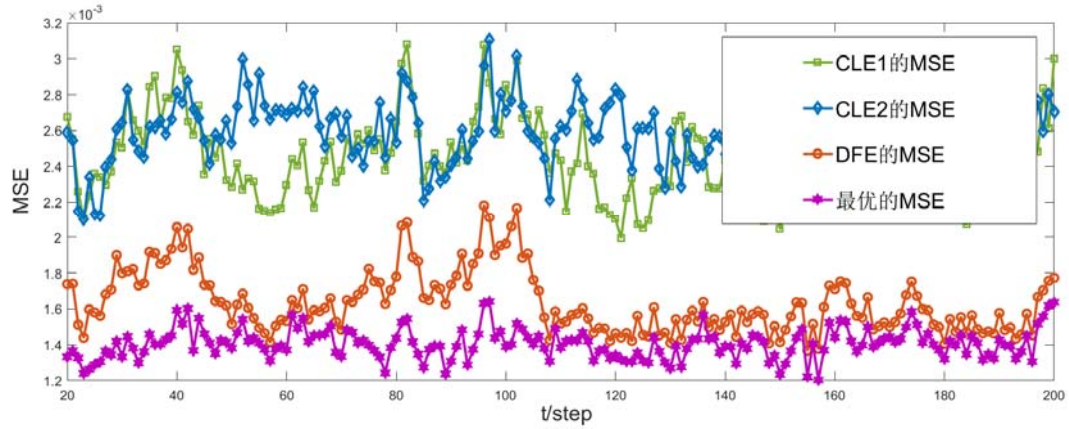

图 2-6　非线性系统下补偿局部估计、融合估计以及最优的 MSE

Figure 2-6.　The MSEs of CLEs, DFE and optimal estimate for unstable systems

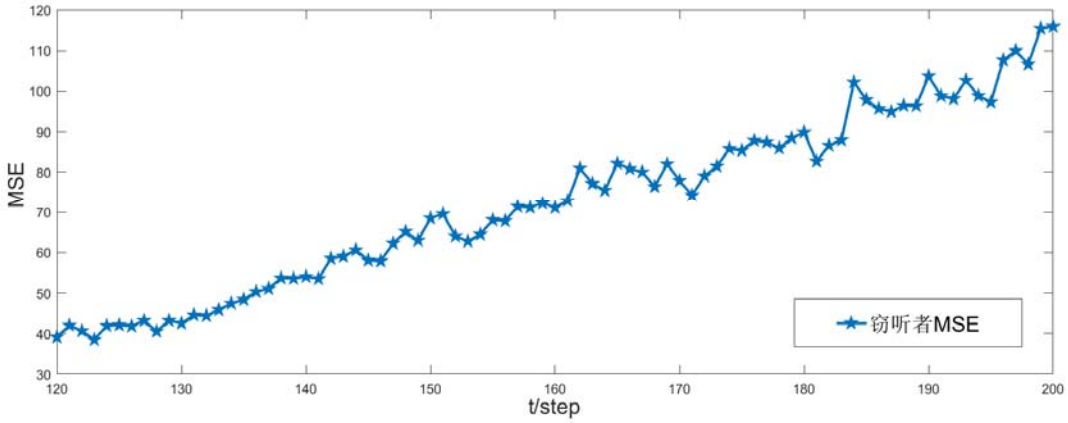

图 2-7　非线性系统下窃听者的 MSE

Figure 2-7.　The MSE of eavesdroppers for unstable systems

2.5 小结

本章针对 NMFESs 中传感器与融合中心的通信被窃听的问题，基于线性加权融合准则的零空间提出了一种随机噪声污染策略对传感器发送的局部估计进行加密，并设计了该框架中最大化窃听者估计误差协方差的分量选择矩阵。同时，针对加密机制在合法用户端产生副作用的问题，采用了一步预测的方法对相应丢失分量进行补偿，计算了线性最小方差意义下的补偿融合准则，并证明了补偿分布式融合估计的稳定性。最后，通过一个移动机器人跟踪的仿真实例证明了所提方法的有效性。

第三章 基于差分隐私的分布式安全融合估计

3.1 引言

根据第二章的分析，所注入的噪声实际上仅仅用于保证污染向量的随机性，防止传输数据的时不变性导致窃听者检测出异常。但是，噪声的具体形式，比如分布、方差和上界等，都没有固定的设计准则。差分隐私[38, 39]是设计噪声分布的一种方法，其原本被用于保护数据库中某些特定数据的隐私，并且同样无需复杂的计算。由于差分隐私严谨的数学模型和出色的性能，其被广泛应用于各个领域。文献[56]将差分隐私引入估计与滤波系统，但是该方法仅仅考虑了单传感器的情况，并且所注入噪声对合法用户性能也有较大的负面影响。在本章考虑的多传感器融合系统中，每个局部传感器都需要将数据发送给融合中心，而融合中心需要进一步将融合后的数据发送给合法用户。在这种情况下，窃听者不仅可以通过第二章类似的线性加权融合方式计算融合估计，还可以通过窃听融合中心发布的数据直接获取融合估计相关信息。因此，如何防御上述双重窃听攻击是本章考虑的一个主要问题。

为了减小噪声带来的复杂度影响，本章针对线性时不变状态空间模型，采用基于 Kalman 滤波的稳态加权融合估计框架[28]。此时，将差分隐私的相关概念和方法引入其中，可以保证随机噪声的协方差为常值，即可以直接在系统初始化时产生随机噪声序列，而不用根据实时变化的系统状态重新产生随机数。进一步地，本章在融合中心构建统计数据的发布方式。以平均值为例，融合中心在其中注入高斯白噪声以实现差分隐私，从而保护其发布数据的隐私安全。然后，本章提出了多传感器融合估计系统特有的局部扰动方式，即仅在传感器端注入随机噪声，通过利用融合矩阵的逆设计所注入的噪声增益，就可以在实现与融合中心注入噪声相等差分隐私水平的同时，使窃听者也无法通过线性加权融合来推测准确的系统状态。接着，根据所注入噪声的统计特性对融合估计进行补偿，即通过计算最小方差意义下的权重矩阵减少噪声对融合中心的估计性能的影响，提高噪声干扰下合法用户端的融合估计精度。最后，通过一个目标跟踪仿真实例证明了所提方法的有效性。

3.2 系统建模与问题描述

3.2.1 基于 Kalman 滤波的稳态分布式融合估计

本章考虑如下离散时不变的状态空间模型：

$$\begin{cases} x(t+1)=Ax(t)+w(t) \\ y_i(t)=C_i x(t)+v_i(t)(i=1,\ldots,L) \end{cases} \tag{3-1}$$

其中 $x(t)\in \mathrm{R}^n$ 为系统状态，$y_i(t)\in \mathrm{R}^{m_i}$ 是传感器 i 的观测输出，A 和 C_i 分别是状态转移矩阵和观测矩阵，且均时不变。$w(t)$ 为系统噪声，$v_i(t)$ 为观测噪声，且 $w(t)$ 和 $v_i(t)$ 均为独立同分布的高斯白噪声，并满足如下条件：

$$\mathrm{E}\{[w^{\mathrm{T}}(t)\quad v_i^{\mathrm{T}}(t)]^{\mathrm{T}}[w^{\mathrm{T}}(t_1)\quad v_j^{\mathrm{T}}(t_1)]\}=\delta(t,t_1)\mathrm{diag}\{Q_w,\delta(i,j)Q_{v_i}\} \tag{3-2}$$

其中 Q_w 为噪声 $w(t)$ 的方差，Q_{v_i} 为噪声 $v_i(t)$ 的方差。同时，本章假设矩阵对 $(A,\sqrt{Q_w})$ 完全可控并且矩阵对 $(A,C_i)(\forall i)$ 完全可观，即：

$$\begin{cases} \mathrm{rank}([\sqrt{Q_w},A\sqrt{Q_w},\cdots,A^{n-1}\sqrt{Q_w}])=n \\ \mathrm{rank}(\mathrm{col}\{C_i,C_iA,\cdots,C_iA^{n-1}]\})=n \end{cases} \tag{3-3}$$

与第二章类似，给出上述线性时不变系统(3-1)的局部 Kalman 估计器[65]：

$$\begin{cases} P_{ii}^z(t+1)=AP_{ii}(t)A^T+Q_w \\ K_i(t+1)=P_{ii}^z(t+1)C_i^T(C_iP_{ii}^z(t+1)C_i^T+Q_{v_i})^{-1} \\ P_{ii}(t+1)=(I-K_i(t+1)C_i)P_{ii}^z(t+1) \\ \varepsilon_i(t+1)=y_i(t+1)-C_iA\hat{x}_i(t) \\ \hat{x}_i(t+1)=A\hat{x}_i(t)+K_i(t+1)\varepsilon_i(t+1) \end{cases} \tag{3-4}$$

其中 $\hat{x}_i(t)$ 为局部状态估计 LSE，$K_i(t)$ 为 Kalman 滤波增益。为了对局部估计进行融合以获取最终的一致性估计，与第二章相同，本章也采用线性加权的方式计算分布式融合估计，即：

$$\hat{x}_f(t)=\sum_{i=1}^{L}W_i(t)\hat{x}_i(t) \tag{3-5}$$

根据引理 2.1 的结果，上式中线性最小方差意义下按矩阵加权的最优权重矩阵 $W(t)=[W_1(t)\quad W_2(t)\quad\cdots\quad W_L(t)]$ 可由下式计算：

$$W(t)=(e^TP^{-1}(t)e)^{-1}e^TP^{-1}(t) \tag{3-6}$$

其中 $P(t)=(P_{ij}(t))\in \mathrm{R}^{nL\times nL}(i,j=1,2,\cdots,L)$。

针对本章讨论的时不变系统，在可观可控条件(3-3)下，上述Kalman滤波器(3-4)可以收敛，并且收敛速度极快，一般数个时间步长就可以达到稳态[4, 28]。因此，为了尽可能减少计算量以保证估计的实时性，也为了腾出空余时间保证加密算法的

正常运行，本章采用稳态 Kalman 滤波，即直接假设 Kalman 滤波在算法迭代开始之前已经达到稳态，而在迭代过程中估计器参数均为相应稳态值。具体地，稳态 Kalman 增益的计算如下[28]：

$$K_i = P_{ii}^z C_i^T (C_i P_{ii}^z C_i^T + Q_{v_i})^{-1} \tag{3-7}$$

其中稳态预测误差协方差 P_{ii}^z 满足如下 Riccati 方程：

$$P_{ii}^z = AP_{ii}^z A^T + Q_w - AP_{ii}^z C_i^{\mathrm{T}} (C_i P_{ii}^z C_i^{\mathrm{T}} + Q_{v_i}) C_i P_{ii}^z A^T \tag{3-8}$$

同时，稳态估计误差协方差 P_{ii} 满足如下 Lyapunov 方程：

$$P_{ii} = \Psi_i P_{ii} \Psi_i^T + (I_n - K_i C_i) Q_w (I_n - K_i C_i)^T - K_i Q_{v_i} K_i^T \tag{3-9}$$

其中 $\Psi_i = (I_n - K_i C_i) A$。由于上述稳态局部 Kalman 滤波的设计，所有估计误差协方差均为时不变矩阵，这使得加权融合准则也变为相应稳态形式，即加权融合估计(3-5)变为如下形式：

$$\hat{x}_f(t) = \sum_{i=1}^{L} W_i \hat{x}_i(t) \tag{3-10}$$

其中稳态权重矩阵 $W = [W_1 \quad W_2 \quad \cdots \quad W_L]$ 为

$$W = (e^T P^{-1} e)^{-1} e^T P^{-1} \tag{3-11}$$

需要注意的是，本章最终采用的是基于 Kalman 滤波的稳态分布式融合估计框架，前面给出的时变估计器仅为得出相应稳态估计器做铺垫。

3.2.2 问题描述与分析

与第二章的分布式融合框架稍有不同，在本章讨论的结构中，融合中心在进行加权融合之后会继续将所需相关数据发送至合法用户。在第二章的结构中，如果融合中心需要进一步传递消息，其将会直接发布相应时刻的瞬时融合估计。这种情况下，一旦窃听者通过网络窃取了该发布数据，将会严重地危险系统隐私安全，因为精度极高的融合估计值将被直接被窃听者获取，系统精确的状态信息会被泄露。为了防止窃听者直接获取瞬时估计值，本章考虑将多个融合估计的统计信息用于融合中心的信息发送，比如求和与求平均等，并将其称之为公开发布估计（Publicly Released Estimate, PRE）。具体地，本章考虑对所需时刻融合估计的与其他时刻融合估计求平均值，即：

$$\hat{x}_u(t) = \frac{1}{\kappa} \sum_{k \in \Phi(t)} \hat{x}_f(k) \tag{3-12}$$

其中 $\kappa \in Z > 1$ 是时间索引集合 $\Phi(t)$ 中元素的数量，该集合的构成取决于合法用户的实际需求。同时，由于稳态融合方式的设计，集合中元素的变动并不影响估计误差协方差的分析。

在这种情况下，即使窃听者通过对融合中心发布的信息进行窃听从而获取传输数据(3-12)，也无法单从这一个数据中分析出瞬时的融合估计值。但是，窃听者仍然有可能通过对两个特定的邻居数据进行差分等操作获取实时系统状态信息。比如，窃听者虽然无法直接从1到t时刻所有融合估计的平均值中推测出t时刻的瞬时估计值，但是当其获取1到$t-1$时刻的平均值后，就可以对这两个值进行相减和放缩，从而得到t时刻准确的估计信息。

为了防止上述窃听方式，本章考虑进一步采用差分隐私机制防止隐私被泄露给该类窃听者。具体地，需要在公开发布估计(3-12)中额外注入随机噪声，进而得到如下扰动公开发布估计（Perturbed Publicly Released Estimate, PPRE）用于融合中心的数据传输：

$$\hat{x}_u^p(t) = M(\hat{x}_u(t)) = \hat{x}_u(t) + a(t) \tag{3-13}$$

其中$a(t) \in \mathrm{R}^n$是方差为$Q_a^2 I$的高斯白噪声。基于该噪声扰动机制，最终需要满足的差分隐私性能指标如下[53]：

$$\mathrm{P}(M(\hat{x}_u(t)) \in \mathrm{O}) \le e^{\varepsilon}\mathrm{P}(M(\hat{x}_u^{'}(t)) \in \mathrm{O}) + \delta, \forall \mathrm{O} \subseteq \mathrm{M} \tag{3-14}$$

其中$\mathrm{M} \triangleq \mathrm{Range}(M)$表示机制$M$的观测域，$\varepsilon$和$\delta$为差分隐私参数。该指标通过相关参数限制了两个邻居的概率分布的差异，即两个概率分布的差距与参数ε和δ相关联，这些参数的数值越小，则两个概率分布越接近，说明越难区分，即隐私保护性能越好。当参数ε和δ都不为0，式(3-14)被称为(ε,δ)-差分隐私，一般可由高斯噪声[53]实现，此外，当ε为0时，其可被称为δ-差分隐私，一般可由均匀噪声[66]实现，而δ为0时则称为ε-差分隐私，一般可由Laplace机制[66, 67]实现。为了统一噪声的分布以及方便估计误差协方差的计算，本章考虑采用高斯白噪声实现(ε,δ)-差分隐私。

注意到，上述噪声叠加形式(3-13)可借用现有高斯隐私机制[53]实现差分隐私，前提是需要计算敏感度，即求得$V_{\hat{x}_u}(t) = \hat{x}_u(t) - \hat{x}_u^{'}(t)$的上界，其中$\hat{x}_u^{'}(t)$为$\hat{x}_u(t)$的邻居估计。然而，由于估计器设计的多样化以及估计过程中的不确定性，上述敏感度一般无法直接给出。因此，根据文献[53]的假设，本章给出如下关于系统状态的邻居关系假设。

假设 3.1：$x(t)$与$x'(t)$互为邻居关系，即

$$\begin{aligned} &\mathrm{Adj}(x(t), x'(t)): \\ &\text{iff for time } t: \left\| Hx(t) - Hx'(t) \right\|_2 \le \zeta \\ &\text{and } (I-H)x(t) = (I-H)x'(t) \end{aligned} \tag{3-15}$$

其中$H \in \mathrm{R}^{n\times n}$是一个二值对角矩阵，用于指代特定分量是否存在邻居关系。

然而，在本章所考虑的结构中，仅考虑差分隐私机制是不够的，因为窃听者也可能同时窃听传感器发送的局部数据。与第二章相类似，窃听者可以基于加权融合准则计算高精度的分布式融合估计值，从而造成系统隐私的泄露。为了对抗该双重窃听方式，本章考虑直接在传感器端注入随机噪声，构成局部扰动状态估计（Perturbed Local State Estimate, PLSE）后再进行发送。在该加密结构下，窃听者获取的局部与融合数据都是被噪声干扰过的。但是，由于在传感器端额外噪声的注入，融合中心的融合估计性能势必受到不利影响。为了减少噪声带来的估计性能损失，融合中心需要利用所知信息尽可能计算出性能较优的补偿分布式融合估计（Compensated Distributed Fusion Estimate, CDFE）。在上述局部和融合双重窃听者下，基于局部噪声扰动的分布式融合估计框架如图 3-1 所示。

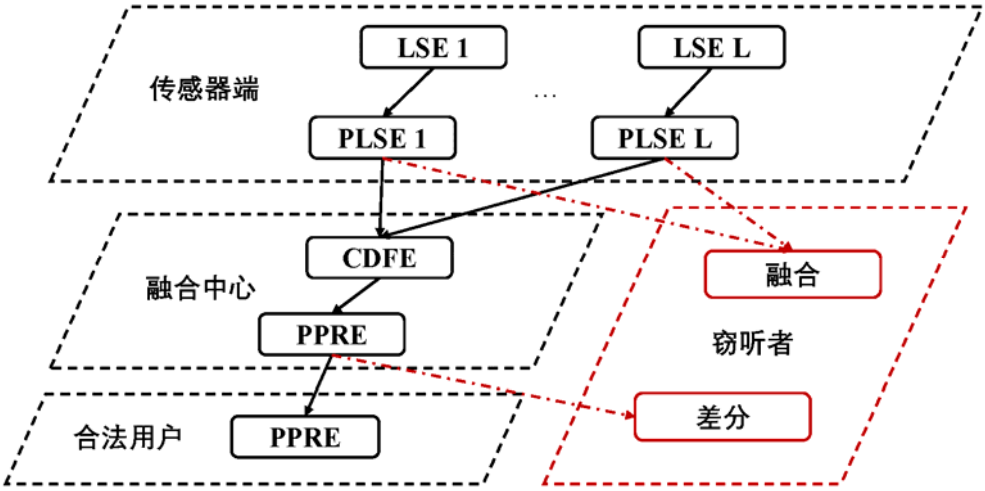

图 3-1　基于局部注入噪声扰动的分布式安全融合估计框架

Figure 3-1.　The structure of distributed secure fusion estimation based on the local noise perturbation

最终，根据上述分析，本章主要待解决的问题主要有如下三个：（1）如何基于假设 3.1 和文献[53]的机制设计主动注入噪声的协方差以满足差分隐私的条件。（2）如何在传感器端设计有效的局部噪声扰动策略，在实现同等差分隐私水平的同时保护基于线性加权的融合估计的安全。（3）如何根据系统已有信息设计有效的补偿策略，以减少噪声对合法用户融合估计性能的损害。

3.3 基于差分隐私的分布式安全融合估计器

3.3.1 融合中心差分隐私的实现

由于本章结果主要参考文献[53]中的高斯机制，下面给出该机制作为引理 3.1。

引理 3.1[53]：当机制 $M=q+w$ 中的高斯白噪声 w 的方差 $\omega^2 I$ 满足如下条件时，该机制满足 (ε,δ) -差分隐私：

$$\omega \geq \Delta^2 \Gamma^2(\varepsilon,\delta) \tag{3-16}$$

其中 $\Gamma(\varepsilon,\delta)=(Q^{-1}(\delta)+\sqrt{(Q^{-1}(\delta))^2+2\varepsilon})/2\varepsilon$，$Q(x)=\left(\int_x^{\infty} e^{-\frac{t^2}{2}}dt\right)/\sqrt{2\pi}$，$\Delta$ 为敏感度。

根据问题分析与描述，首先，需要在融合中心端实现差分隐私以保护公开发布估计，即基于假设 3.1 设计机制(3-13)中的噪声发布使之满足差分隐私条件。因此，根据引理 3.1 中的高斯机制，所注入的噪声方差可由下面定理设计。

定理 3.1：当随机噪声机制(3-13)中噪声 $a(t)$ 协方差 $Q_a^2 I$ 满足如下条件时，该机制满足 (ε,δ) -差分隐私：

$$Q_a \geq \Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(WK_dCH)\lambda^2 \tag{3-17}$$

其中 $K_d \triangleq \mathrm{diag}(K_1,\cdots,K_L)$，$C \triangleq \mathrm{col}(C_1,\cdots,C_L)$。

证明：为了得到高斯机制(3-16)中所需的敏感度，需要根据假设 3.1 以及融合估计的形式逐步计算。首先，根据(3-1)中的观测方程与定义的状态邻居关系，容易得到：

$$V_{y_i}(t)=C_i V_x(t) \tag{3-18}$$

类似地，根据稳态 Kalman 滤波(3-7)以及加权融合估计(3-10)的结构，可得如下邻居关系：

$$V_{\hat{x}_f}(t)=WK_dCV_x(t) \tag{3-19}$$

由于假设 3.1 中的 H 是二值对角矩阵，并且用于指代向量分量是否为邻居，可得 $HV_x(t)=V_x(t)$ [53]。因此，上式可重写为：

$$V_{\hat{x}_f}(t)=WK_dCHV_x(t) \tag{3-20}$$

最后，根据柯西-施瓦兹不等式以及假设 3.1 所给出的上界，可得如下结果：

$$\begin{aligned}\left\|\hat{x}_u(t)-\hat{x}_u^{'}(t)\right\|_2 &= \left\|V_{\hat{x}_f}(t)\right\|_2 = \left\|WK_dCHV_x(t)\right\|_2 \\ &\leq \left\|WK_dCH\right\|_2 \left\|V_x(t)\right\|_2 \leq \sigma_{\max}(WK_dCH)\lambda\end{aligned} \tag{3-21}$$

从上述不等式中可以直接得到敏感度 $\Delta(\hat{x}_u(t))=\sup_{\mathrm{Adj}(x(t),x'(t))}\left\|\hat{x}_u(t)-\hat{x}_u^{'}(t)\right\|_2$ 为：

$$\Delta(\hat{x}_u(t))=\sigma_{\max}(WK_dCH)\lambda \tag{3-22}$$

因此，根据引理 3.1 的结果，直接得出所注入噪声的方差条件如定理 3.1 中式(3-17)

所示。证毕。

3.3.2 局部扰动机制设计

根据前述分析，除了直接对融合中心发布的信息进行窃听之外，窃听者很有可能同时窃取传感器发送的局部数据并用于对系统状态值的估计。也就是说，窃听者有两种获取融合估计的方式：除了对融合估计的统计信息进行差分之外，窃听者也会对局部估计值进行加权融合，间接获取瞬时分布式融合估计值。因此，本章中传感器需先对局部估计进行处理后再将其发送至融合中心，以保证较高的隐私性。具体地，在传感器端就对局部估计上注入随机噪声，得到如下用于传输的扰动局部状态估计 PLSE：

$$\hat{x}_i^p(t) = \hat{x}_i(t) + a_i(t) \tag{3-23}$$

其中 $a_i(t) = G_i\xi_i(t)$，$\xi_i(t)$ 为高斯白噪声并且 $G_i = W_i^{-1}$。局部扰动结构的示意图如下。

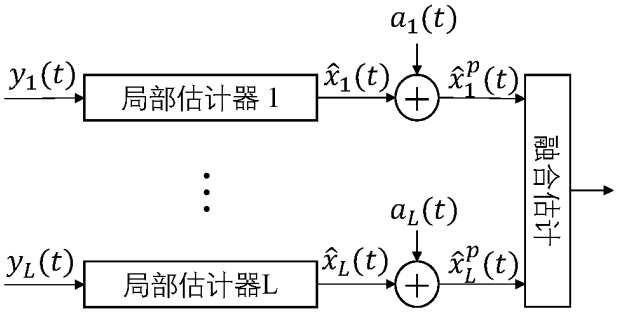

图 3-2　局部扰动结构

Figure 3-2.　The framework of local perturbation

在该噪声机制的影响下，窃听者基于局部估计误差协方差按照矩阵加权融合准则求得的融合估计为：

$$\hat{x}_f^p(t) = W\hat{x}_d^p(t) \tag{3-24}$$

其中 $\hat{x}_d^p(t) = \mathrm{col}\{\hat{x}_1^p(t), \cdots, \hat{x}_L^p(t)\}$。

根据上述扰动方式，合法用户的估计器也不可避免地会受到相同影响。为了减少所提局部噪声机制带来的副作用，合法用户根据噪声统计特性计算补偿融合估计，即：

$$\hat{x}_f^c(t) = W^C\hat{x}_d^p(t) \tag{3-25}$$

其中 W^C 为补偿权重矩阵。

为了求得线性最小方差意义下的权重矩阵，需要先知道相应的局部估计误差协方差和互协方差。由于式(3-23)中的噪声 $\xi_i(t)$ 是白噪声，故局部状态估计之间的误差互协方差的计算不受其影响，即 $P_{ij}^c = P_{ij}(i \neq j)$，而局部估计的误差协方差则需

要重新计算。根据 $\xi_i(t)$ 已知的方差 $Q_{\xi_i}I$，易得补偿局部估计误差协方差为 $P_{ii}^c = P_{ii} + G_i Q_{\xi_i} G_i^{\mathrm{T}}$。然后，直接根据引理 2.1 中的线性加权融合准则，计算出式(3-25)中线性最小方差意义的补偿权重矩阵如下：

$$W^c = (e^T (P^c)^{-1} e)^{-1} e^T (P^c)^{-1} \tag{3-26}$$

其中 $P^c \triangleq (P_{ij}^c) \in \mathrm{R}^{nL\times nL}(i,j=1,2,\cdots,L)$。

最后，为了实现差分隐私的目标，式(3-22)中噪声 $\xi_i(t)$ 的方差 $Q_{\xi_i}I$ 由下面定理设计。

定理 3.2：当噪声协方差 Q_{ξ_i} 满足如下条件时，局部噪声扰动机制(3-22)满足 (ε,δ)-差分隐私：

$$\sum_{i=1}^{L} Q_{\xi_i} \geq \Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(WK_dCH)\lambda^2 / \kappa \tag{3-27}$$

证明：差分隐私机制的目的是为了保护融合估计，因此，需要先给出融合估计的噪声注入形式。根据局部扰动方式(3-23)以及原始加权融合结构(3-10)，扰动后的分布式融合估计可表示为：

$$\hat{x}_f^p(t) = \hat{x}_f(t) + \sum_{i=1}^{L} \xi_i(t) \tag{3-28}$$

然后，根据合法用户求平均的结构(3-12)，可得如下类似于式(3-13)最终的随机噪声注入形式：

$$M(\hat{x}_u(t)) = \hat{x}_u(t) + \sum_{k\in\Phi(t)} \sum_{i=1}^{L} \xi_i(k) \tag{3-29}$$

因此，根据定理 3.1 的结果(3-16)，可以直接得出式(3-29)中噪声需要满足的条件：

$$\sum_{k\in\Phi(t)} \sum_{i=1}^{L} Q_{\xi_i} \geq \Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(WK_dCH)\lambda^2 \tag{3-30}$$

根据之前给出的假设条件，集合 $\Phi(t)$ 内元素的数量固定，所以 $\sum_{k\in\Phi(t)} \sum_{i=1}^{L} Q_{\xi_i} = \kappa \sum_{i=1}^{L} Q_{\xi_i}$。然后，在式(3-29)两边都除以 κ，即可得到定理 3.2 中的不等式结果(3-26)。证毕。

注释 3.1：注意到本章没有直接在式(3-23)中采用简单的白噪声去实现差分隐私，而是在白噪声前面引入相应乘子 G_i，该设计的原因如下：（1）引入该乘子方便计算和推导，因为在该乘子作用下，融合中心最终发布的平均值的结果上只叠加了白噪声，如式(3-29)所示，而根据本章之前给出的定理 3.1，可以轻易得出噪声方差需要满足的充分条件；（2）如果单一地使用白噪声，后续过程中需要求取

权重矩阵W的左伪逆。然而，一般情况下，L个传感器中至少有2个传感器有效，否则没有融合的意义。在这种情况下，权重矩阵W一般为为行满秩，因此，该矩阵没有左伪逆而只有右伪逆，会导致无法进一步求解的情况。

注释 3.2：传感器端噪声的分配一般比较灵活，可以根据实际系统的情况具体设计。比如，当多个传感器的精度和算力等性能类似时，一般考虑每个传感器均匀分配噪声：

$$Q_{\xi_i} \geq \Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(W\bar{K}_d CH)\lambda^2 / L\kappa,\ \forall\ \mathrm{i} \tag{3-31}$$

此外，也可以将噪声的产生只赋予某个或某些性能较强或遭受窃听可能性较大的传感器，即：

$$Q_{\xi_{i_0}} \geq \Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(W\bar{K}_d CH)\lambda^2 / L,\ \exists\ i_0 \tag{3-32}$$

3.4 示例

由于通信网络的开放性，各种固定设施、移动车辆和个人位置等众多信息在网络上传播。由于这些数据大多数都非常敏感，尤其是位置信息，泄露给恶意的窃听者会造成严重后果。因此，本章考虑用一个移动目标跟踪系统来对所提算法进行仿真。假设一个目标在二维平面上移动，则系统的曲线动态模型可表达为[2]：

$$\begin{cases} \dot{s}_x(t) = v(t)\cos\theta(t) \\ \dot{s}_y(t) = v(t)\sin\theta(t) \\ \dot{v}(t) = \alpha_{\tan}(t) \\ \dot{\theta}(t) = \dfrac{\alpha_n(t)}{v(t)} \end{cases} \tag{3-33}$$

其中$\left(s_x(t), s_y(t)\right)$为目标的中心笛卡尔坐标，$v(t)$表示速度，$\theta(t)$表示方向角，$\alpha_{\tan}(t)$和$\alpha_n(t)$分别代表切向加速度和法向加速度。当$\alpha_{\tan}(t)=0$且$\alpha_n(t)$为常数时，该模型可简化为匀速转弯模型，此时速度和角速度都变为常数，即$\gamma = v(t)$，$\phi = \theta(t)$。此时，定义状态向量为$x(t) = \mathrm{col}\{s_x(t), \dot{s}_x(t), s_y(t), \dot{s}_y(t)\}$，则连续时间下的状态微分方程为：

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\phi \\ 0 & 0 & 0 & 1 \\ 0 & \phi & 0 & 0 \end{bmatrix} x(t) + w(t) \tag{3-33}$$

通过相应离散化方法，可得如下离散状态转移矩阵[2]：

$$A=\begin{bmatrix}1 & (\sin\phi T_s)/\phi & 0 & -(1-\cos\phi T_s)/\phi \\ 0 & \cos\phi T_s & 0 & -\sin\phi T_s \\ 0 & (1-\cos\phi T_s)/\phi & 1 & (\sin\phi T_s)/\phi \\ 0 & \sin\phi T_s & 0 & \cos\phi T_s\end{bmatrix} \tag{3-34}$$

其中 T_s 为采样时间。同时，本节采用两个传感器对该移动目标进行观测，相应的观测矩阵分别设计如下：

$$C_1=\begin{bmatrix}0.8 & 0.5 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 \\ 0.4 & 0 & 0.2 & 0\end{bmatrix}, C_2=\begin{bmatrix}0 & 0 & 0.7 & 0.3 \\ 0.6 & 0.2 & 0 & 0 \\ 0 & 0.3 & 0 & 0.8\end{bmatrix} \tag{3-35}$$

然后，系统噪声和观测噪声的方差初始化如下：

$$\begin{cases}Q_{v_1}=\mathrm{diag}\{0.01,0.04,0.02\} \\ Q_{v_2}=\mathrm{diag}\{0.03,0.02,0.01\} \\ Q_{w}=\mathrm{diag}\{0.04,0.01,0.04,0.01\}\end{cases} \tag{3-36}$$

接着，将本章所提噪声注入算法应用到上述系统进行仿真，得到如下结果。首先，图 3-3 展示了所提分布式融合框架中各估计的性能。与第二章类似，本章也用均方误差 MSE 去衡量状态估计的性能。从该图可以看出，在传感器端噪声的扰动下，窃听者计算的扰动分布式融合估计 PDFE 虽然相对局部估计 PLSE 提升了精度，但还是远低于补偿融合估计 CDFE 的精度，说明了窃听者各个估计的精度都较低，即无法准确获取系统的状态信息。而补偿分布式融合估计 CDFE 的性能最好，且明显优于扰动情况下的各种估计，并且，补偿融合估计的 MSE 仅仅略高于无扰动情况下最优的分布式融合估计 DFE，证明了基于噪声方差的补偿方法的有效性。

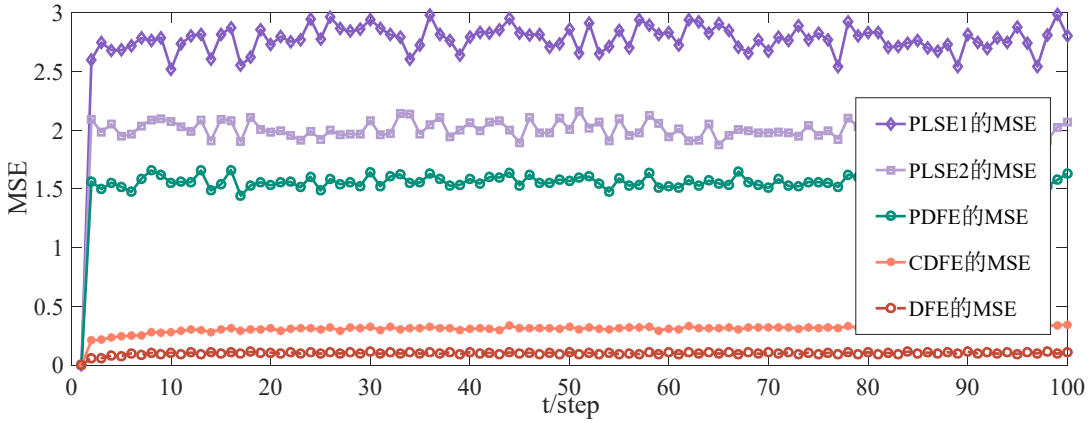

图 3-3　所提分布式融合框架中各估计的均方误差

Figure 3-3.　The MSEs of estimates in the proposed distributed fusion structure

图 3-4 描述了窃听者扰动融合估计 PDFE 与合法用户补偿融合估计 DCFE 的

MSE 与参数κ的关系。从图中可以看出，该参数越大，融合中心求平均值的集合中元素越多，即每个局部估计分配到的噪声越少，从而降低了窃听者的 MSE，削弱了融合方式的隐私保护程度。但是，可以看出该参数对合法用户的补偿融合估计性能影响较小，体现了补偿方法能比较较稳定地减少噪声影响。

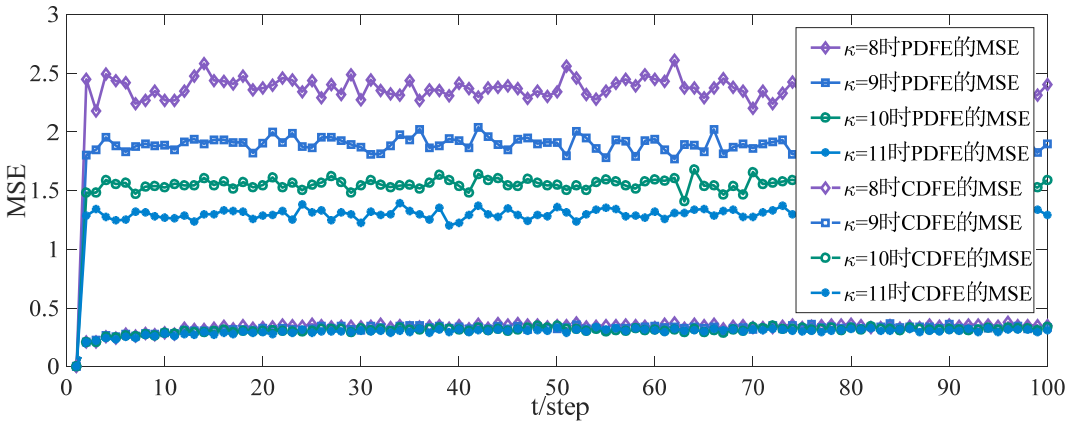

图 3-4　均方误差与参数κ的关系

Figure 3-4.　The relationship between MSE and κ

同样地，融合估计性能也会随着差分隐私相关参数的变化而变化。图 3-5 给出了窃听者与合法用户的均方误差 MSE 与差分隐私参数ε和δ的关系。从该图中可以看出，差分隐私的参数值越小，所需噪声方法越大，则窃听者估计的 MSE 越大，隐私保护程度越高，与理论分析结果相同。同时，合法用户估计器的 MSE 的变化受差分隐私参数的影响也较小，也说明了补偿方式稳定的性能。进一步地，可以根据该图给出的具体关系模糊设计实际系统中的某些参数。比如，当需要保证窃听者估计 MSE 大于 1 的时候，就可以从图中挑选出合适的差分隐私参数，大致为$\varepsilon=\log(2)$，$\delta=0.1$。通过简单粗略的选择，可以在实现差分隐私的同时，尽可能减少噪声对合法用户的影响。

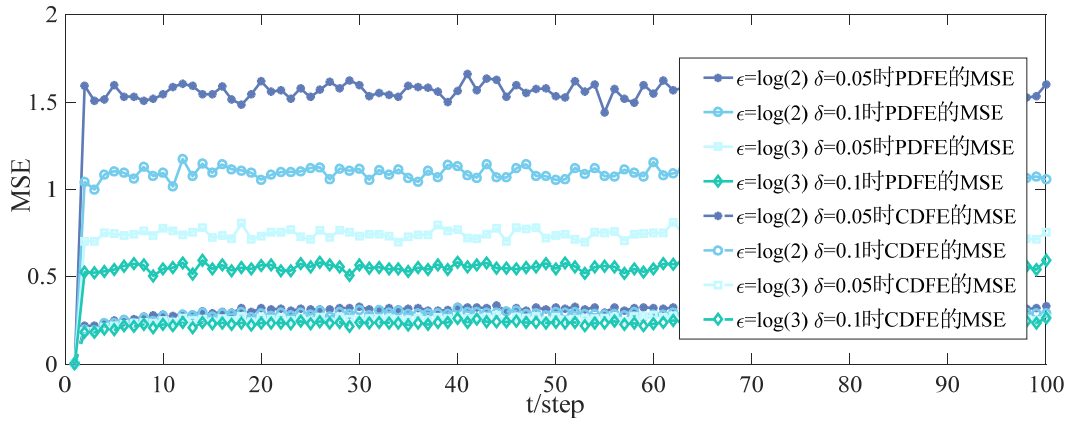

图 3-5　均方误差与差分隐私参数ε和δ的关系

Figure 3-5.　The relationship between MSE and differential privacy parameters ε and δ

此外，图 3-6 给出了某个时刻两个邻居 PPREs 的概率分布函数（Probability Distribution Functions, PDFs），并于图 3-7 中绘制了相应的实际采样值。从图 3-6 可以看出，两个邻居概率分布较为接近，这是因为叠加了特定方差的高斯白噪声，并且与差分隐私等参数有关。同时，图 3-7 绘制了图 3-6 中概率分布的 100 次真实采样，即经过噪声注入后可能的位置变化。其中，红色三角代表真实位置，而蓝色圆圈则代表经过随机噪声注入后的位置。从该图中可以看出噪声注入机制所造成的影响，即随机产生偏离真实状态的数据，从而使得窃听者差分后的结果偏离原始值。

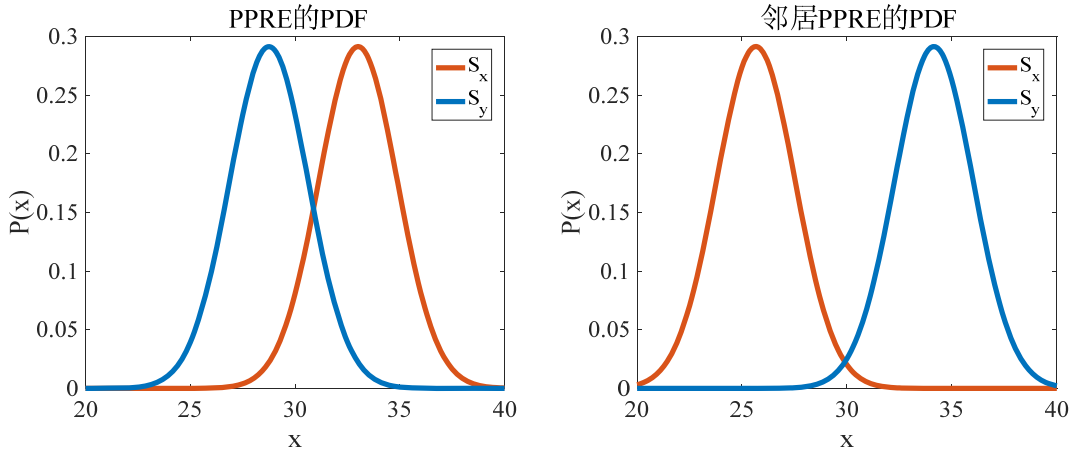

图 3-6　邻居 PPREs 的概率分布函数

Figure 3-6.　The PDFs of adjacent PPREs

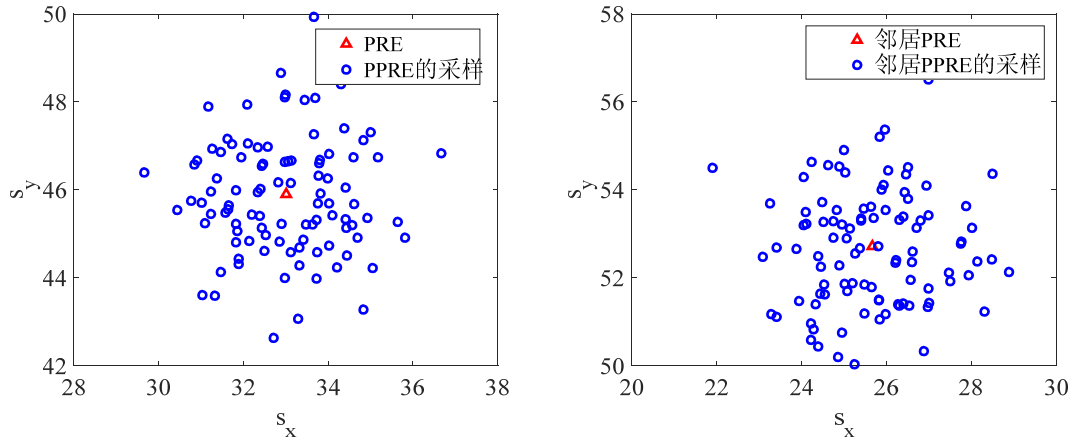

图 3-7　邻居 PPREs 的实际采样

Figure 3-7.　The samples of adjacent PPREs

3.5 小结

本章针对 NMFESs 的隐私保护问题，将差分隐私引入稳态分布式融合估计系统，采用随机噪声注入的方式设计了高斯机制保护了融合估计的隐私安全。然后，通过将噪声的注入分散至传感器端，在保证相等差分隐私水平的同时，使得窃听者通过线性加权获得的融合估计值产生较大误差。而合法用户端则根据所注入噪声的统计特性，设计了最小方差意义下的补偿融合策略，有效地减少了所注入噪声带来的负面影响。最后，通过一个机动目标跟踪的仿真实例验证了所提方法的有效性。

第四章 基于零空间与差分隐私的分布式安全融合估计

4.1 引言

根据前两章的结果可以看出，零空间和差分隐私都是设计主动注入噪声较为有效的方法，能在特定背景假设下达到较高的隐私水平。但是，第三章的差分隐私机制只能一定程度上损害窃听者的瞬时融合估计性能而无法使其发散，因为所注入噪声仅仅影响线性加权过程而非估计器迭代过程。而第二章的随机噪声污染策略虽然在系统不稳定的情况下使窃听者的分布式融合估计发散，但是针对稳定系统设计的策略无法到达类似较好的性能，因为窃听者估计误差协方差的发散与系统的状态大小有关。此外，状态隐私编码也是一种新颖且有效的加密策略，然而其也存在一定局限性。虽然文献[58]针对不稳定系统证明了状态隐私编码的有效性，但是在文献[57]讨论的稳定系统中，该方法仍然不能使窃听者估计器发散，这是因为状态隐私编码策略的效果与系统的状态转移矩阵有关，其需要使窃听者丢失某些时刻的数据包产生初始误差，进而利用状态转移矩阵的迭代放大该误差。当系统不稳定时，状态转移矩阵范数大于 1，经过累乘可以趋向无穷从而造成相应方差的发散，而当系统稳定时，状态转移累乘到最后趋于 0，反而削弱了噪声的影响。

为了保护任意稳定性系统的隐私，本章结合零空间与差分隐私的方法设计了两步序贯噪声注入的加密方法。该方法使得窃听者的估计误差协方差由于随机噪声的累加而发散，从而保证其高效的隐私性能与系统的稳定性无关，因为噪声的设计不依赖于状态转移矩阵和系统状态向量。进一步地，通过设计噪声协方差使相应数据满足差分隐私条件，从而使得窃听者利用差分也无法获取准确的瞬时局部估计。同时，通过基于加权融合准则的零空间设计，合法用户可以利用线性加权的融合方式直接抵消部分噪声，并且可以根据另外一部分噪声的统计特性重新计算最小方差意义下的最优融合矩阵，与第二章类似地减少相应噪声影响。此外，本章还同时结合了合法用户与窃听者的性能提出了一个新的隐私度量，结合了对数函数以方便对实际系统进行定性分析。最后，通过一个航空发动机的仿真实例验证了所提出方法的有效性。

4.2 系统建模与问题描述

4.2.1 系统及估计器模型

本章考虑如下离散时不变状态空间模型：

$$\begin{cases} x(t+1) = Ax(t) + Bw(t) \\ y_i(t) = C_i x(t) + v_i(t) (i = 1, \ldots, L) \end{cases} \tag{4-1}$$

其中，$x(t) \in \mathrm{R}^n$ 和 $y_i(t) \in \mathrm{R}^{m_i}$ 分别表示系统状态和传感器 i 的观测输出，L 为传感器数量。A，B 和 C_i 是具有合适维数的时不变矩阵。$w(t)$ 为系统噪声，$v_i(t)$ 为观测噪声，且 $w(t)$ 与 $v_i(t)$ 均为高斯白噪声，满足如下条件：

$$\mathrm{E}\{[w^{\mathrm{T}}(t) \quad v_i^{\mathrm{T}}(t)]^{\mathrm{T}}[w^{\mathrm{T}}(t_1) \quad v_j^{\mathrm{T}}(t_1)]\} = \delta(t, t_1)\mathrm{diag}\{Q_w, \delta(i, j)Q_{v_i}\} \tag{4-2}$$

与第三章类似，假设矩阵对 (A, B) 完全可控并且矩阵对 (A, C_i) $(\forall i)$ 完全可观，即：

$$\begin{cases} \mathrm{rank}([B, AB, \cdots, A^{n-1}B]) = n \\ \mathrm{rank}(\mathrm{col}\{C_i, C_i A, \cdots, C_i A^{n-1}]\}) = n \end{cases} \tag{4-3}$$

根据第三章给出的结果，本章设计如下的稳态 Kalman 估计器[28]：

$$\hat{x}_i(t) = \bar{\Psi}_i \hat{x}_i(t-1) + \bar{K}_i y_i(t) \tag{4-4}$$

其中 $\bar{\Psi}_i \triangleq (I - \bar{K}_i C_i)A$，$\hat{x}_i(t)$ 为局部状态估计 LSE，并且稳态 Kalman 增益 $\bar{K}_i$ 由下式计算：

$$\begin{cases} \bar{P}_{ii}^z = A\bar{P}_{ii}A^T + BQ_w B^T \\ \bar{K}_i = \bar{P}_{ii}^z C_i^T (C_i \bar{P}_{ii}^z C_i^T + Q_{v_i})^{-1} \\ \bar{P}_{ii} = (I - \bar{K}_i C_i)\bar{P}_{ii}^z \end{cases} \tag{4-5}$$

同时，根据引理 2.1，可得如下基于线性加权的最优稳态分布式融合估计：

$$\hat{x}_f(t) = \sum_{i=1}^{L} \bar{W}_i \hat{x}_i(t) \tag{4-6}$$

最优权重矩阵 $\bar{W}(t) = [\bar{W}_1(t) \quad \bar{W}_2(t) \quad \cdots \quad \bar{W}_L(t)]$ 的计算如下：

$$\bar{W}(t) = (e^T \bar{P}^{-1} e)^{-1} e^T \bar{P}^{-1} \tag{4-7}$$

其中 $\bar{P} = (\bar{P}_{ij}) \in \mathrm{R}^{nL \times nL} (i, j = 1, 2, \cdots, L)$。

4.2.2 问题描述与分析

为了保证传输数据的隐私，传感器利用随机机制 M 对局部状态估计进行处理得到 $\hat{x}_i^s(t) = M(\hat{x}_i^m(t))$，并将其称之为扰动局部发布估计（Perturbed Local Released Estimate, PLRE），其中 $\hat{x}_i^m(t) \triangleq \mathrm{col}\{\hat{x}_i(t - t_q), \cdots, \hat{x}_i(t)\}$ 为局部估计的增广向量，可以通过矩阵操作获取相应统计信息。与第三章类似，定义其邻居向量为

$\left[\hat{x}_i^s(t)\right]' = M\left(\left[\hat{x}_i^m(t)\right]'\right)$，其中$\left[\hat{x}_i^m(t)\right]' \triangleq \mathrm{col}\left\{\hat{x}_i(t-t_q),\cdots,\left[\hat{x}_i(t)\right]'\right\}$。如果没有隐私保护机制，窃听者可能通过对$\hat{x}_i^s(t)$和$\left[\hat{x}_i^s(t)\right]'$进行差分间接获取瞬时数据。因此，与第三章类似，机制M需满足如下差分隐私条件：

$$\mathrm{P}(\hat{x}_i^s(t)=\bar{x}_i) \le e^{\varepsilon}\,\mathrm{P}([\hat{x}_i^s(t)]'=\bar{x}_i)+\delta, \forall \bar{x}_i \in \mathrm{M} \tag{4-8}$$

其中M=range(M)代表该机制的观测域。同时，系统状态需满足如下假设[53]。

假设 4.1： $x^m(t) \triangleq \mathrm{col}\{x(t-t_q),\cdots,x(t)\}$与$\left[x^m(t)\right]' \triangleq \mathrm{col}\{\left[x(t-t_q)\right]',\cdots,\left[x(t)\right]'\}$互为邻居关系，即：

$$\begin{aligned}&\mathrm{Adj}_{\varsigma}(x(t),x'(t)):\\&\text{iff for time } t:\left\|Hx(t)-Hx'(t)\right\|_2 \le \zeta\\&\text{while } (I-H)x(t)=(I-H)x'(t);\\&\text{for time } \mu=t-t_q,\cdots,t-1: x(\mu)=\left[x(\mu)\right]'\end{aligned} \tag{4-9}$$

其中$H \in \mathrm{R}^{n\times n}$是一个二值对角矩阵，用于指代特定分量是否存在邻居关系。

基于上述假设，可直接得到$V(x(t)) \triangleq Hx(t)-Hx'(t)$的上界$\left\|V(x(t))\right\|_2$为$\zeta$。在本章中，窃听者首先对收到的局部数据进行差分以获得窃听局部状态估计（Wiretapped Local State Estimate, WLSE），然后，根据系统统计特性进行线性最小方差加权融合以获得窃听分布式融合估计（Wiretapped Distributed Fusion Estimate, WDFE）。同时，为了减少加密带来的误差，合法用户根据已有信息对损失的数据进行补偿，计算补偿分布式融合估计 CDFE，以保证其估计器的稳定。本章所考虑的基于差分隐私的分布式融合估计框架如图 4-1 所示。

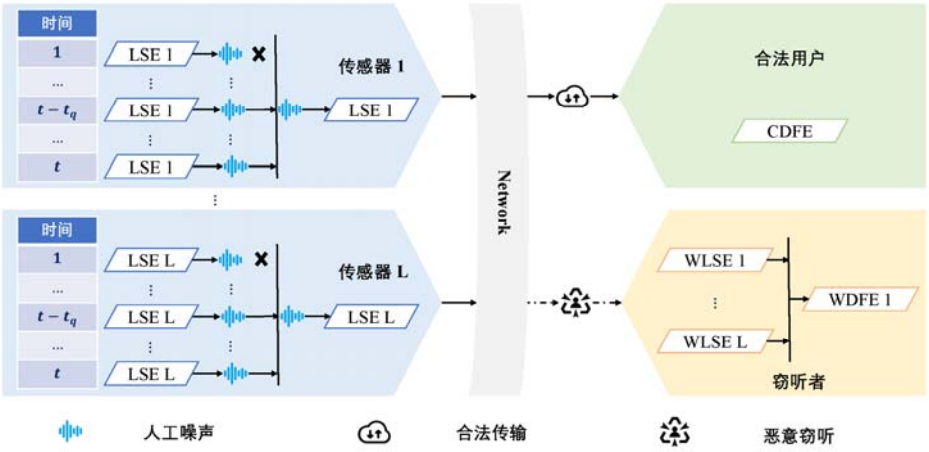

图 4-1　基于差分隐私的分布式融合估计框架

Figure 4-1.　The structure of distributed fusion estimation based on differential privacy

此外，为了对系统的隐私性能指标进行定性分析，本文提出了如下保密性指标（Confidentiality Index, CI）：

$$CI(t) \triangleq \log_2 \frac{\mathrm{Tr}\{P_f^w(t)\}}{\mathrm{Tr}\{P_f^c(t)\}} \tag{4-10}$$

其中 $P_f^w(t)$ 是窃听者融合估计 WDFE 的误差协方差，$P_f^c(t)$ 是合法用户补偿融合估计 CDFE 的误差协方差。一般情况下，仅用窃听者的协方差就可以大致度量隐私性能，因为多数加密方法不会对合法用户造成影响，比如 A. Tsiamis 等[57, 58]考虑的合法用户，其可以利用系统转移矩阵完全解密出局部信息。然而，仍有一部分加密机制可能导致合法用户性能受损，尤其是一些随机机制，比如 J. Le Ny 等[53]考虑的差分隐私滤波以及 A. S. Leong 等[59]考虑的随机传输信道。虽然 A. S. Leong 等[59]结合了合法用户与窃听者的估计误差协方差，但是其形式仅可用于定量分析，不易直观定性。因此，本章提出如式(4-9)所示的度量，利用分数与对数函数，使得该指标的结果易于定性分析。

注意到，本章考虑的系统虽然是一般的线性时不变系统，但是没有明确规定其稳定性，也就是说包括了稳定和不稳定系统。需要指出的是，当系统稳定时，本章考虑的窃听者不会觉察到其估计的异常，因为在窃听者的视角里，其融合估计是最优的，且估计误差协方差为稳态。同时，窃听者不知道系统稳定性，无法简单地通过估计值来判断其误差协方差是否发散。因此，根据上述假设与分析，窃听者无法使用一步预测或者其他稳定的估计方式进行在线补偿从而减少其实时估计的性能损失。此外，本章假设窃听者也无法获取事先设定好的加解密策略，包括所注入随机噪声的统计特性。

最后，本章需要解决的主要问题可以总结为如下两方面：（1）如何设计加密策略使得所提性能指标最大化，即实现高水平的隐私；（2）如何设计随机机制 M 使其满足差分隐私，使窃听者无法通过差分获取精确数据。

注释 4.1：注意到，仅根据窃听者的相关指标来表示整个系统的隐私性能是不靠谱的。在众多的隐私保护方法中，合法用户的性能指标同样也会收到加密机制的影响，比如利用随机噪声的差分隐私机制[53]以及随机丢包的信道假设[59]。在这些情况下，仅凭窃听者估计误差协方差较大无法确定系统的隐私水平，因为合法用户的估计性能也可能较差，导致窃听者获取的估计与合法用户所差无几。因此，合理的性能指标应该同时包含合法用户与窃听者双方的性能指标，比如本章所提出的度量方式(4-10)。

注释 4.2：本章在性能指标式(4-9)中引入以 2 为底的对数，主要有如下两个目

的：（1）作为一个性能的度量指标，其值应该为非负数，故利用对数函数保证其非负性。（2）当窃听者与合法用户性能相同时，该指标为 0，当窃听者均方误差为合法用户两倍时，该指标为 1，这些数字的直观性可以用于系统性能的定性分析，即：能更快地判断系统的大致状态。

4.3 基于零空间与差分隐私的分布式安全融合估计器

4.3.1 结合零空间与差分隐私的加密策略

根据前两章的讨论，零空间可以有效减少噪声机制带来的影响，但是单一地在分布式结构上注入噪声无法有效保护隐私。为了解决这些问题，本章提出了如下两步序贯噪声加密方法。

步骤 1：在局部估计上叠加噪声得到扰动局部估计：

$$\hat{x}_i^p(t)=\hat{x}_i(t)+R_i\xi_i(t) \tag{4-11}$$

其中 $\xi_i(t)$ 为高斯白噪声， R_i 为噪声增益矩阵。

步骤 2：在扰动局部估计上叠加噪声得到扰动局部发送估计：

$$\hat{x}_i^s(t)=\begin{cases}(t_q-t)\hat{x}_i^p(0)+\sum\limits_{\tau=1}^{t}\hat{x}_i^p(\tau)+N_i\alpha(t) & 0<t\le t_q\\ \sum\limits_{\tau=t-t_q}^{t}\hat{x}_i^p(\tau)+N_i\alpha(t) & t>t_q\end{cases} \tag{4-12}$$

其中 $N\triangleq\mathrm{col}\{N_1,\cdots,N_L\}$ 为融合矩阵的零空间，即满足如下条件：

$$\bar{W}^cN=0,N^{\mathrm{T}}N=I \tag{4-13}$$

在第一步中，扰动局部估计的误差协方差与互协方差可计算得 $\bar{P}_{ii}^p=\bar{P}_{ii}+R_iQ_{\xi_i}R_i^{\mathrm{T}}$ 并且 $\bar{P}_{ij}^p=\bar{P}_{ij}(i\neq j)$。根据之前的问题分析，所注入随机噪声需要实现差分隐私。因此，基于引理 3.1，上述加密策略中噪声的方差可按如下设计。

定理 4.1：当注入噪声方差满足如下条件时，机制(4-10)与(4-11)满足 (ε,δ)-差分隐私：

$$Q_{\phi_i}\ge\Gamma^2(\varepsilon,\delta)\sigma_{\max}^2(\bar{K}_iC_iH)\zeta^2I \tag{4-14}$$

其中 $Q_{\phi_i}=t_qR_iQ_{\xi_i}R_i^{\mathrm{T}}+N_iQ_\alpha N_i^{\mathrm{T}}$。

证明：由于差分隐私机制用于保护局部估计，故需要先构造基于局部估计的噪声叠加形式。根据机制(4-11)与(4-12)，最终的噪声叠加可表达为：

$$\hat{x}_i^s(t)=\sum_{\tau=t-t_q}^{t}\hat{x}_i(\tau)+\phi_i(t) \tag{4-15}$$

其中 $\phi_i(t)=\sum_{\tau=t-t_q}^{t}R_i\xi_i(\tau)+N_i\alpha(t)$，其协方差如定理所示。与第三章定理 3.1 推导类似，根据稳态 Kalman 滤波的结构，可得如下关系：

$$V(\hat{x}_i(t))=\bar{K}_iC_iHV(x(t)) \tag{4-16}$$

然后，利用柯西-施瓦兹不等式，可得 $\|V(\hat{x}_i(t))\|_2\le\|\bar{K}_iC_iH\|_2\|V(x(t))\|_2$。根据假设 4.1 中给出的上界，可得敏感度 $\Delta(V(\hat{x}_i(t)))\triangleq\sup_{\mathrm{Adj}(x(t),x'(t))}\|V(\hat{x}_i(t))\|_2$ 为

$$\Delta(V(\hat{x}_i(t)))\triangleq\sigma_{\max}(\bar{K}_iC_iH)\zeta \tag{4-17}$$

同样地，根据引理 3.1，可得定理中的给出噪声方差条件。证毕。

4.3.2 基于零空间与差分隐私的安全融合估计器设计

在设计好加密策略之后，本章的最终目的是最大化隐私保护性能，即：在保证合法用户估计误差协方差有界的同时，使窃听者估计误差协方差发散。首先，需要分析合法用户与窃听者的估计方式。

一方面，对于合法用户来说，由于加密机制中零空间的设计，其可以直接通过加权融合抵消部分噪声，同时，由于求和形式的发送设计，合法用户在融合之后需要利用差分的方式获取瞬时 CDFE，即：

$$\hat{x}_f^c(t)=\sum_{i=1}^{L}\bar{W}_i^c(\hat{x}_i^s(t)-\hat{x}_i^s(t-1))+\hat{x}_f^c(t-t_q) \tag{4-18}$$

根据引理 2.1，结合所注入噪声的协方差信息，线性最小方差意义下的补偿加权融合矩阵 $\bar{W}^c(t)=[\bar{W}_1^c(t)\quad\bar{W}_2^c(t)\quad\cdots\quad\bar{W}_L^c(t)]$ 为

$$\bar{W}^c(t)=(e^{\mathrm{T}}(\bar{P}^p)^{-1}e)^{-1}e^{\mathrm{T}}(\bar{P}^p)^{-1} \tag{4-19}$$

其中 $\bar{P}^p(t)=(\bar{P}_{ij}^p(t))\in R^{nL\times nL}(i,j=1,2,\cdots,L)$。

另一方面，为了尽可能获取系统信息，窃听者首先对局部数据进行差分，获取 WLSE：

$$\hat{x}_i^w(t)=\hat{x}_i^s(t)-\hat{x}_i^s(t-1)+\hat{x}_i^w(t-t_q) \tag{4-20}$$

不同于合法用户的是，窃听者不知道加密机制，包括所注入噪声的方差，所以其仅能使用式(4-6)中的原始权重矩阵进行融合，获取如下 WDFE：

$$\hat{x}_f^w(t)=\sum_{i=1}^{L}\overline{W}_i\hat{x}_i^w(t) \tag{4-21}$$

下面给出定理用于证明所提加密方法的有效性。

定理 4.2：针对本章分布式融合估计系统，基于所提加密策略(4-10)与(4-11)，所提隐私保护指标(4-9)将趋向无穷，即：

$$CI(t)\to\infty,\text{as } t\to\infty \tag{4-22}$$

证明：定理 4.2 中结论的证明需要分为两部分：（1）证明合法用户的性能指标的有界性；（2）证明窃听者的性能指标的发散性。首先，利用补偿融合估计 CDFE 的稳定性证明合法用户性能的有界性。根据引理 2.1 中的加权融合准则，融合后的协方差不会高于任意一个局部协方差，即：$\overline{P}_f^c\le\overline{P}_{ii}^p(\forall i)$。同时，利用矩阵迹的运算，易得 $\mathrm{Tr}\{\overline{P}_f^c\}\le\min\limits_{i=1,\cdots L}\mathrm{Tr}\{\overline{P}_{ii}+R_iQ_{\xi_i}R_i^{\mathrm{T}}\}$。该上界中所有矩阵均为时不变矩阵，因此，合法用户的性能有界得证。

接着，需要证明窃听者融合估计的协方差发散。根据式(4-19)中的 WLSE，窃听者的估计可更具体地表达为 $\hat{x}_i^w(t)=\hat{x}_i^w(t-t_q)+\hat{x}_i^p(t)+N_i\alpha(t-1)-\hat{x}_i^p(t-t_q)$。结合系统状态，该估计的误差：

$$\tilde{x}_i^w(t)=\begin{cases}\tilde{x}_i^p(t)+N_i\alpha(t),\ \text{for time } 0<t\le t_q\\ \tilde{x}_i^w(t-t_q)+\theta_i(t),\ \text{for time } t>t_q\end{cases} \tag{4-23}$$

其中 $\theta_i(t)=N_i\alpha(t)-N_i\alpha(t-1)+\tilde{x}_i^p(t)-\tilde{x}_i^p(t-t_q)$。当 $t>t_q$ 时，上式误差方差的迭代形式如下：

$$\overline{P}_{ii}^w(t)=\overline{P}_{ii}^w(t\bmod t_q)+\lfloor t/t_q\rfloor\overline{P}_{\theta_{ii}} \tag{4-24}$$

其中 $\overline{P}_{\theta_{ii}}=E\{\theta_i(t)\theta_i^{\mathrm{T}}(t)\}$ 为

$$\overline{P}_{\theta_{ii}}=(2I-\overline{\Psi}_i^{t_q})\overline{P}_{ii}-\overline{P}_{ii}[\overline{\Psi}_i^{t_q}]^{\mathrm{T}}+2N_iQ_\alpha N_i^{\mathrm{T}}+2R_iQ_{\xi_i}R_i^{\mathrm{T}} \tag{4-25}$$

结合上述噪声方差与系统方差等，窃听者局部估计误差方差最终可表示为

$$\overline{P}_{ii}^w(t)=\begin{cases}\overline{P}_{ii}+R_iQ_{\xi_i}R_i^{\mathrm{T}}+N_iQ_\alpha N_i^{\mathrm{T}},\ \text{for time } 0<t\le t_q\\ (2\lfloor t/t_q\rfloor+1)(R_iQ_{\xi_i}R_i^{\mathrm{T}}+N_iQ_\alpha N_i^{\mathrm{T}})\\ +(\lfloor t/t_q\rfloor(2I-\overline{\Psi}_i^{t_q})+I)\overline{P}_{ii}-\lfloor t/t_q\rfloor\overline{P}_{ii}[\overline{\Psi}_i^{t_q}]^{\mathrm{T}},\ \text{for time } t>t_q\end{cases} \tag{4-26}$$

从上式可以看出，只有时间 t 是变量，其他矩阵均为时不变矩阵，因此，窃听者局部估计误差协方差随时间趋向无穷而发散。与上面步骤相似地，窃听者互协方差为

$$\bar{P}_{ij}^{w}(t)=\begin{cases}\bar{P}_{ij}+N_iQ_\alpha N_j^{\mathrm{T}}, \text{ for time } 0<t\le t_q\\ (2\lfloor t/t_q\rfloor+1)N_iQ_\alpha N_i^{\mathrm{T}}+(\lfloor t/t_q\rfloor(2I-\bar{\Psi}_i^{t_q})+I)\bar{P}_{ij}\\ -\lfloor t/t_q\rfloor\bar{P}_{ij}[\bar{\Psi}_j^{t_q}]^{\mathrm{T}}, \text{ for time } t>t_q\end{cases} \tag{4-27}$$

同样地，上式仅有时间一个变量，故 $\bar{P}_{ij}^{w}(t)$ 也随着时间的增长而发散。

然后，根据窃听者的融合准则(4-20)，计算相应的融合估计误差为 $\tilde{x}_f^w(t)=\sum_{i=1}^{L}\bar{W}_i\tilde{x}_i^w(t)$，由此可得其协方差形式如下：

$$P_f^w(t)=\sum_{i=1}^{L}\sum_{j=1}^{L}\bar{W}_iP_{ij}^w(t)\bar{W}_j^{\mathrm{T}} \tag{4-28}$$

基于之前的分析，所有局部协方差与互协方差均发散，而所有权重矩阵都是元素为常数的稳态矩阵，因此，上述融合估计误差协方差也随时间增长而趋向无穷，即窃听者性能指标发散。结合前面证得的合法用户的性能指标有界，所提出的度量方式(4-9)如定理中所示发散。证毕。

注释 4.3：根据引理 2.1 中线性加权的融合方式可知，融合过程不构成迭代，这就使得直接对局部估计注入随机噪声无法使融合估计发散。因此，本章构建求和的发送形式(4-11)与(4-12)，使得数据接收端需要利用差分的方式解密，从而将额外注入的噪声引入迭代过程。实际上，只要局部的求和过程构成迭代形式，所提方法都可以发挥其有效性。为了选取一种典型的方式解释所提方法，也为了方便读者的理解，本章采用的是固定长度的求和方式，类似于常见的滑动平均形式。

注释 4.4：值得注意的是，窃听者融合估计的发散与局部和融合估计的计算顺序无关。本章所讨论的先差分后融合结构是为了保证窃听者能够更全面地获取信息。为了与本章中的合法用户端保持一致，窃听者同样可以利用差分直接求得分布式融合估计：

$$\hat{x}_f^w(t)=\sum_{i=1}^{L}\bar{W}_i(\hat{x}_i^s(t)-\hat{x}_i^s(t-1))+\hat{x}_f^w(t-t_q) \tag{4-29}$$

通过对表达式中项的拆分和重组，不难看出，上述估计形式与本章中先差分后融合的结果完全相等。但是，这种情况下，窃听者没有试图去获取精度较低局部估计值。

注释 4.5：尽管窃听者误差协方差的方差仅与时间有关，即与噪声的方差无关，仍然需要设计差分隐私机制。因为某些窃听者可能不进行实时跟踪，而是仅获取某个时刻的状态估计值。在这种情况下，窃听者只对特定的两个邻居发布估计进

行差分处理。因此，需要借用差分隐私相关方法以保证单步估计的隐私。需要指出的是，差分隐私机制与窃听者的发散并不冲突，是相辅相成的两种操作。具体地，在所提两步序贯噪声加密方法中，随机噪声的引入可直接使得窃听者迭代的估计误差协方差发散，而差分隐私机制则通过设计噪声的具体方差用于保护单步隐私。

4.4 示例

本章考虑采用一个航空发动机模型验证所提算法的有效性，且该系统为稳定系统，更能够突出本章工作在系统稳定情况下的贡献。首先，定义发动机的状态向量为 $x(t)=\mathrm{col}\{s_x(t),s_y(t),s_z(t)\}$，其中 $\left(s_x(t),s_y(t)\right)$ 表示发动机的平面位置，$s_z(t)$ 表示发动机的高度。根据其动态模型，离散时间的状态转移矩阵为[55, 56]：

$$A=\begin{bmatrix}0.87 & 0.00 & 0.20\\ 0.03 & 0.98 & -0.03\\ 0.03 & 0.00 & 0.80\end{bmatrix},\ B=\begin{bmatrix}0.1 & 0.0\\ 0.5 & 0.0\\ -0.2 & 0.0\end{bmatrix} \tag{4-30}$$

同时，考虑布置两个传感器对该发动机进行观测，量测矩阵分别为：

$$C_1=\begin{bmatrix}1 & 1 & 0\end{bmatrix},C_2=\begin{bmatrix}0 & 1 & -1\end{bmatrix} \tag{4-31}$$

然后，设置如下系统噪声方差：$Q_w=\mathrm{diag}\{0.04,0.01\},Q_{v_1}=0.06,Q_{v_2}=0.02$。根据上述系统矩阵，可得如下矩阵的秩：

$$\begin{cases}\mathrm{rank}([B\quad AB\quad A^2B])=3\\ \mathrm{rank}(\mathrm{col}\{C_i,C_iA,C_iA^2\})=3\ (i=1,2)\end{cases} \tag{4-32}$$

可见该系统满足可观可控条件，即 Kalman 滤波可以收敛至稳态。然后，根据本章给出的稳态 Kalman 融合估计方式，可得如下稳态参数，包括稳态局部 Kalman 增益以及稳态融合权重矩阵：

$$\begin{cases}\bar{K}_1=\begin{bmatrix}0.03\\ 0.32\\ -0.10\end{bmatrix},\bar{K}_2=\begin{bmatrix}0.06\\ 0.44\\ -0.16\end{bmatrix}\\ \bar{W}=\begin{bmatrix}0.19 & 0.00 & 0.00 & 0.81 & -0.00 & -0.00\\ 0.16 & 0.21 & 0.12 & -0.16 & 0.79 & -0.12\\ -0.03 & -0.01 & 0.16 & 0.03 & 0.01 & 0.84\end{bmatrix}\end{cases} \tag{4-33}$$

接着，将本章所提两步序贯噪声加密算法应用至上述系统，得到以下结果。图 4-2 给出了各种策略下合法用户 MSEs 的对比，包括状态隐私编码[43, 44]和第二章提出的随机噪声污染策略。从该图中可以看出，状态隐私编码下合法用户的性能

最优，因为其可以使用状态转移矩阵完全解密出局部数据。随机噪声污染策略下的性能最差，因为其丢失了部分分量的数据，虽然使用一步预估保证了稳定性，但是性能的损失较大。本章所提方法下的精度稍逊于最优融合估计，因为具体的噪声值无法实时获取，只能利用统计信息减少其影响，但是从结果可以看出这种负面影响较小。

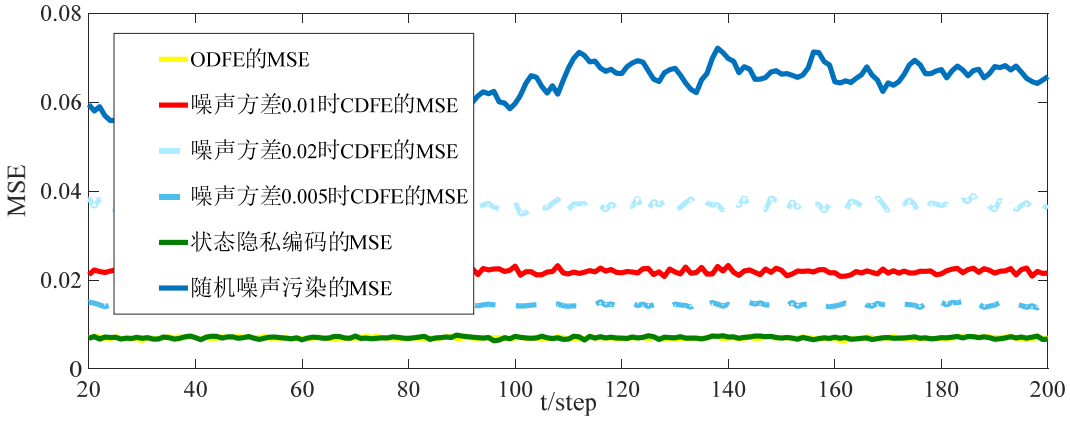

图 4-2　各种策略下合法用户 MSEs 的对比

Figure 4-2.　The comparison of MSEs in legitimate user under various strategies

图 4-3 展示了各种策略下窃听者 MSEs 的对比。从图中可以看出，状态隐私编码和随机噪声污染策略都无法在系统稳定的情况下使窃听者发散，因为状态隐私编码方法中稳定系统的状态转移矩阵叠加无法发散，而随机噪声污染策略中稳定系统的状态收敛导致最终协方差的收敛。但是，在本章所提的噪声加密策略下，窃听者协方差可以发散，证明了所提方法的优越性，并且发散速度随着噪声方差的增加而加快。

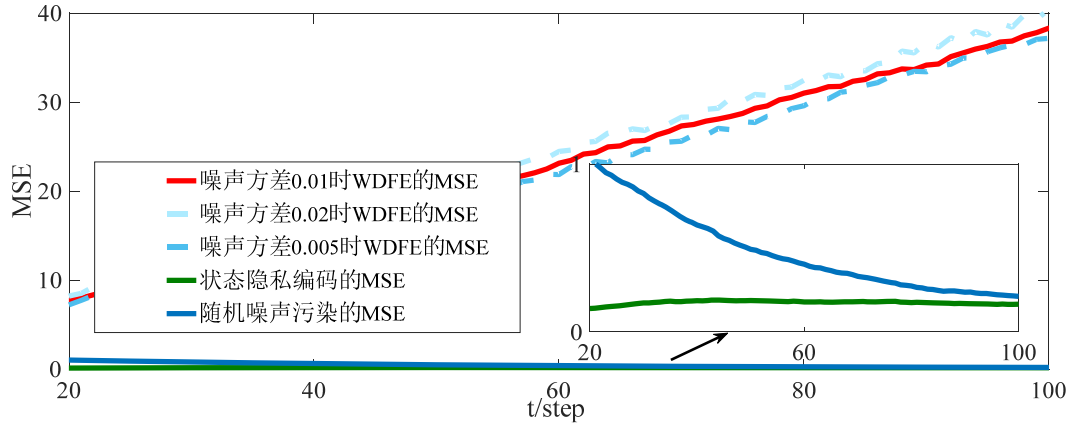

图 4-3　各种策略下窃听者 MSEs 的对比

Figure 4-3.　The comparison of MSEs in eavesdropper under various strategies

此外，可以从图 4-2 中看出估计的 MSE 随着噪声 $\xi_i(t)$ 的方差增大而增加，并

且图 4-3 也说明了窃听者的方差与所注入噪声方差有关。因此，一个值得进一步研究的方向为系统各项性能的平衡。比如，当需要合法用户性能较好时，可以在保证融合估计 MSE 小于某个阈值的时候，通过最大化窃听者融合估计误差协方差的方式计算最优的噪声协方差。

进一步地，本节结合合法用户与窃听者双方的 MSEs，利用所提出的度量方式(4-9)对系统隐私性能进行评估。同时，为了方便系统性能的定性分析，本章提出了如下保密性等级（Confidentiality Ranks, CRs）：

$$CR(t)=\begin{cases}\text{低, if } 0\le CI(t)<1;\\ \text{相对低, if } 1\le CI(t)<2;\\ \text{中等, if } 2\le CI(t)<4;\\ \text{相对高, if } 4\le CI(t)<8;\\ \text{高, otherwise}\end{cases} \tag{4-34}$$

根据保密性指数的定义，计算得随机噪声污染策略下的保密性指数为 1.23，处于相对低的水平；状态隐私编码下的保密性指数为 3.66，处于中等的水平；而所提序贯噪声加密方法下的保密性指数为无穷大，处于最高水平。根据该指标的计算结果，可以看出所提方法的有效性和优越性。

最后，图 4-4 给出了邻居 PLREs 各个分量的概率分布函数。与第三章结果类似，该图从宏观层面展示了差分隐私机制下两个概率分布之间的关系。并且，根据图 4-3 以及图 4-4 可知，本章所提算法同时实现了两种隐私目标，即差分隐私与窃听者估计误差协方差的发散，体现了本章算法的优越性。

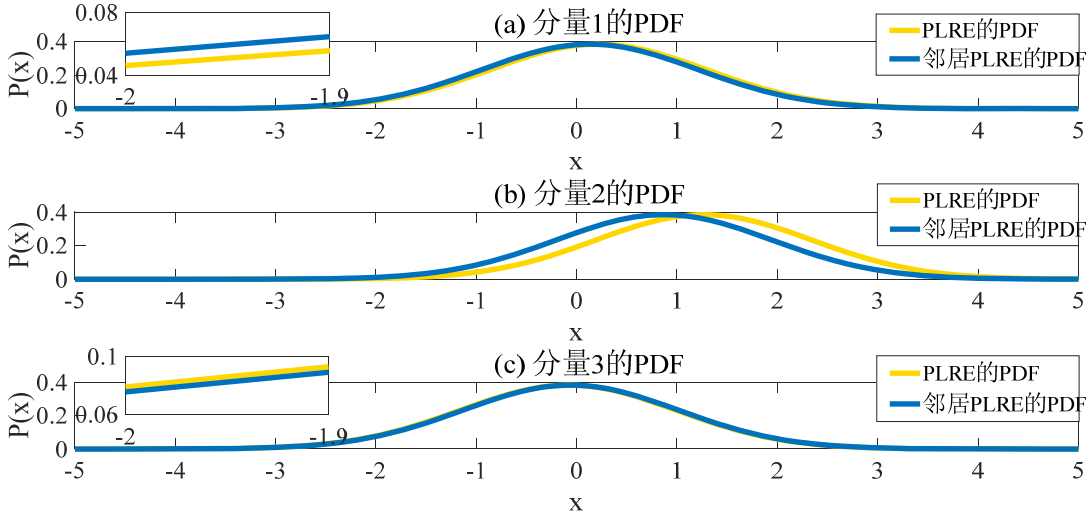

图 4-4　邻居 PLREs 的概率分布函数

Figure 4-4.　The PDFs of adjacent PLREs

4.5 小结

本章针对分布式融合估计遭受窃听攻击的问题，提出了一种两步序贯注入噪声的加密方法，其保证了窃听者无法基于局部估计误差协方差利用线性加权融合来获取准确的系统状态，也使得窃听者通过差分获得的单步瞬时局部数据受到干扰。同时，结合所注入噪声的统计特性，融合中心以最小方差为目的重新计算融合矩阵，进而基于零空间的加密设计方案获取了稳定且性能较优的补偿分布式融合估计值。此外，本章结合了合法用户与窃听者双方的估计误差协方差提出了一种新的系统隐私性能度量方式，并利用分数与对数函数对该指标进行了定性分析。最后，通过一个稳定的航空发动机实例验证了所提出方法的有效性和优越性。

第五章 总结与展望

5.1 总结

由于通信网络的引入，NMFESs 存在众多优点，比如易于维护、布线复杂度低、可扩展性强等。然而，通信网络的开放性和匿名性等也导致了 NMFESs 中存在恶意的窃听者，其通过窃取通信信道中传输的敏感数据对系统隐私等众多方面的安全性造成巨大威胁。分布式融合估计由于其较高的精度和鲁棒性，保护其数据不被窃听是 NMFESs 中一个重要方向。在分布式安全融合估计的研究中，需要考虑如下几个重要问题：（1）大多数已有的加解密方法需要消耗系统正常运行以外较多的资源，比如传统的密码学方法和物理层信道噪声法等，将其应用至 NMFESs 可能会导致其计算资源和能量的快速消耗，从而影响系统正常工作甚至使得系统寿命大大缩短；（2）许多隐私保护机制会对合法用户的正常运行产生负面影响，比如差分隐私等，这些方法所注入的随机噪声会同时影响合法用户，降低 NMFESs 原有的估计精度。为此，根据现有的研究成果，本文针对上述两个问题展开研究，主要工作总结如下：

1. 针对 NMFESs 中传输数据被窃听的问题，基于零空间和高斯白噪声设计了无需消耗额外能量的随机噪声污染策略，使得窃听者通过线性加权融合后损失特定局部估计分量的信息，导致其融合估计性能受损。同时，利用一步预测对合法用户端收到的局部估计进行补偿，根据相应协方差计算了最小方差意义下的补偿融合矩阵，使得合法用户的融合估计性能较优。

2. 针对 NMFESs 的隐私保护问题，利用差分隐私机制保护了融合估计的隐私，并设计了局部噪声注入的加密方式，在实现相等差分隐私水平的同时，使得窃听者对局部估计的加权融合性能也受到损害。同时，利用所注入噪声的统计特性采用最小方差意义下的再融合对合法用户的融合估计性能进行了补偿。

3. 针对 NMFESs 遭受窃听攻击的问题，设计了两步序贯噪声注入的加密方法，使窃听者的估计器在融合时由于噪声的叠加而发散。同时，基于零空间的特性以及噪声协方差设计了补偿融合结构，使合法用户获取了稳定且性能较优的融合估计值。此外，提出了一种新的隐私度量方式，实现了对系统隐私性能的定性分析。

本文设计的各种加密策略都无需额外的能量消耗且计算量极小，基于特定系统信息仅利用随机数的产生与叠加就可以实现较高水平的隐私。特别地，第四章

所提方法可以适用于任意系统，不用考虑系统的稳定性，即使在系统稳定的情况下也能使窃听者估计误差协方差发散。

5.2 展望

虽然本文在特定背景下基于主动注入的随机噪声设计了分布式隐私保护融合估计方法，但是所提的策略仍存在一些问题亟待解决。具体地，可以从以下几个方面作进一步研究。

1．虽然本文所提方法利用了大量的补偿策略以减少加密机制带来的负面影响，但是却无法达到最优估计的目标，因为合法用户仅知道噪声协方差而无法确定每个时刻随机噪声的具体值。如何进一步改进策略或者设计新的加解密方案使得合法用户的融合估计性能损失尽可能减小是一个重要的问题。

2．注意到本文所提各种方法下噪声对合法用户与窃听者的估计性能均有影响，而本文并未分析两者的关系等。因此，可以进一步讨论最优的权衡方式，在给出某个指标界限的情况下，最优化另一个指标。尤其是窃听者估计误差协方差无法发散的情况下，必须权衡合法用户的估计与隐私性能。

3．本文对系统中的窃听方式给出了某些特定的假设，实际上，存在一些特殊的情况，即合法系统无法保证窃听者窃取数据的具体方式，从而意味着传感器端需要平等对待所有第三方的融合中心。在这种情况下，无法设计加解密策略，而同样需要综合考虑隐私与估计性能，利用优化的方法进行权衡。

4．本文仅将密码学中差分隐私的概念引入了分布式融合估计系统，其它密码学方法也可以考虑用于保护状态估计的隐私。比如，当系统计算性能较为强大时，一般可以采用多方安全计算和同态加密等方法，尤其是线性加权的融合过程，其简单的乘加结构有利于许多方法的嵌入。

5．本文考虑的系统多数为线性定常系统，尤其是差分隐私机制，而实际中大多数的系统都为非线性时变系统，如何根据非线性函数或者根据线性化的误差设计有效的加解密策略是一个值得研究的课题，并且，如何在参数时变的情况下设计有效的策略也是一个难点。

参考文献

[1] Hall D L, Llinas J. An introduction to multi-sensor data fusion[J]. Proceedings of The IEEE, 1987, 85(1): 6-23.

[2] Bar-Shalom Y, Li X R, T. Kirubarajan. Estimation with applications to tracking and navigation[M]. John Wilely and Sons, Inc, 2001.

[3] 韩崇昭, 朱洪艳, 段战胜. 多源信息融合[M]. 北京: 清华大学出版社, 2006.

[4] 邓自立. 信息融合滤波理论及其应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 2007.

[5] Khaleghi B, Khamis A, Karray F O, Razzavi S N. Multisensor data fusion: A review of the state-of-the art[J]. Information Fusion, 2013, 14(1): 28-44.

[6] Xia Y, Fu M, Liu G. Analysis and synthesis of networked control systems[M]. New York: Springer, 2011.

[7] Zhu Y, Zhou J, Shen X, Song E, Luo Y. Networked multisensor decision and estimation fusion: Based on advanced mathematical methods[M]. CRC Press, 2012.

[8] 潘泉, 王增福, 梁彦, 杨峰, 刘准钆. 信息融合理论的基本方法与进展（II）[J]. 控制理论与应用, 2012, 29(10): 1233-1244.

[9] 陈博. 网络化多传感器信息融合估计算法研究[D]. 浙江工业大学, 2013.

[10] Chen B, Yu L, Zhang W A. Networked multi-sensor fusion estimation with delays, packet losses and missing measurements [C]. 12th International Conference on Control Automation Robotics & Vision, Guangzhou, China, 2012: 695-700.

[11] Sun S L, Lin H, Ma J, Li X Y. Multi-sensor distributed fusion estimation with applications in networked systems: A review paper[J]. Information Fusion, 2017, 38: 122-134.

[12] Huang H, Savkin A V, Ni W. Decentralized navigation of a UAV team for collaborative covert eavesdropping on a group of mobile ground nodes[J]. IEEE Transactions on Automation Science and Engineering, 2022, 19(4): 3932-3941.

[13] Huang H, Savkin A V, Huang C. Decentralized autonomous navigation of a UAV network for road traffic monitoring[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(4): 2558-2564.

[14] Shi J. In-process quality improvement: Concepts, methodologies, and applications[J]. IISE Transactions, 2023, 55(1), 2-21.

[15] Du J, Zhang X, Qu W. Knowledge-infused process monitoring for quality improvement in solar cell manufacturing processes[J]. Journal of Quality Technology, 2021, 54(5), 561-572.

[16] Xu Q, Dragicevic T, Xie L, Blaabjerg F. Artificial intelligence-based control design for reliable virtual synchronous generators[J]. IEEE Transactions on Power Electronics, 2021, 36(8) 9453-9464.

[17] Ding D, Han Q, Xiang Y, Ge X, Zhang X. A survey on security control and attack detection for industrial cyber-physical systems[J]. Neurocomputing, 2018, 275: 1674-1683.

[18] Xu W, Jaimoukha I M, Teng F. Robust moving target defence against false data injection attacks in power grids[J]. IEEE Transactions on Information Forensics and Security, 2023, 18: 29-40.

[19] Chen B, Ho D W C, Zhang W A, Yu L. Distributed dimensionality reduction fusion estimation

for cyber-physical systems under DoS attacks[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019, 49(2): 455-468.

[20] Chen B, Ho D W C, Hu G, Yu L. Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks[J]. IEEE Transactions on Cybernetics, 2018, 48(6): 1862-1876.

[21] Topal O A, Kurt G K. Whispering lies to the eavesdropper: Physical layer spoofing against eavesdropping attacks[J]. IEEE Systems Journal, 2023, 17(1): 1581-1590.

[22] Goldreich O. Foundations of cryptography: Volume 1, basic tools[M]. Cambridge, U.K.: Cambridge University Press, 2003.

[23] Goldreich O. Foundations of cryptography: Volume 2, basic applications[M]. Cambridge, U.K.: Cambridge University Press, 2004.

[24] Kailkhura B, Nadendla V S S, Varshney P K. Distributed inference in the presence of eavesdroppers: a survey[J]. IEEE Communications Magazine, 2015, 53(6): 40-46.

[25] Singer R, Kanyuck A J. Computer control of multiple site track correlation [J]. Automation, 1971, 7: 455-464.

[26] Willner D, Chang C B, Dunn K P. Kalman filter algorithm for a multisensor system[C]. IEEE Conferences on Decision and Control including the 15th Symposium on Adaptive Processes, 1976, 15: 570-574.

[27] Li X R, Zhu Y, Wang J, Han C Z. Optimal linear estimation fusion I: Unified fusion rules [J]. IEEE Transactions on Information Theory, 2003, 49(9): 2192-2208.

[28] Deng Z, Gao Y, Mao L, Li Y, Hao G. New approach to information fusion steady-state Kalman filtering[J]. Automatica, 2005, 41(10): 1695-1707.

[29] Sun S, Deng Z. Multi-sensor optimal information fusion Kalman filter[J]. Automatica, 2004, 40(6): 1017-1023.

[30] Shen Q, Liu J, Zhou X, Qin W, Wang L, Wang Q. Centralized fusion methods for multi-sensor system with bounded disturbances[J]. IEEE Access, 2019, 7: 141612-141626.

[31] Zhang Y, Chen B, Yu L. Fusion estimation under binary sensors[J]. Automatica, 2020, 115: 108861.

[32] Chen B, Hu G. Nonlinear state estimation under bounded noises[J]. Automatica, 2018, 98: 159-168.

[33] Chen B, Ho D W C, Hu G, Yu L. Delay-dependent distributed Kalman fusion estimation with dimensionality reduction in cyber-physical systems[J]. IEEE Transactions on Cybernetics, 2022, 52(12): 13557-13571.

[34] Chen B, Hu G, Ho D W C, Yu L. A new approach to linear/nonlinear distributed fusion estimation problem[J]. IEEE Transactions on Automatic Control, 2019, 64(3): 1301-1308.

[35] Zhao P; Jiang H; Lui J C S; Wang C; Zeng F; Xiao F; Li Z. P3-LOC: A privacy-preserving paradigm-driven framework for indoor localization[J]. IEEE/ACM Transactions on Networking, 2018, 26(6): 2856-2869.

[36] Yao A C. Protocols for secure computations[C]. 23rd Annual Symposium on Foundations of Computer Science, Chicago, IL, USA, 1982: 160-164.

[37] Rivest R L, Adleman L, Dertouzos M L. On data banks and privacy homomorphisms[J]. Foundations of Secure Computation, 1978, 4(1): 169-180.

[38] Dwork C, Kenthapadi K, McSherry F, Mironov I, Naor M. Our data, ourselves: Privacy via distributed noise generation[C]. Advances in Cryptology, 2006: 486-503.

[39] Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis[C]. Third Theory of Cryptography Conference, 2006: 265-284.
[40] Erdemir E, Dragotti P L, Gündüz D. Privacy-aware time-series data sharing with deep reinforcement learning[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 389-401.
[41] Alexandru A B, Pappas G J. Encrypted LQG using labeled homomorphic encryption[C]. The 10th ACM/IEEE International Conference on Cyber-Physical Systems, New York, NY, USA, 2019, 129–140.
[42] Alexandru A B, Gatsis K, Shoukry Y, Seshia S A, Tabuada P, Pappas G J. Cloud-based quadratic optimization with partially homomorphic encryption[J]. IEEE Transactions on Automatic Control, 2021, 66(5): 2357-2364.
[43] Gentry C. Fully homomorphic encryption using ideal lattices[C]. The Forty-First Annual ACM Symposium on Theory of Computing, New York, NY, USA, 2009, 169–178.
[44] Abadi M, Chu A, Goodfellow I, McMahan H B, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy[C]. The 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016, 308–318.
[45] Ye M, Hu G, Xie L, Xu S. Differentially private distributed Nash equilibrium seeking for aggregative games[J]. IEEE Transactions on Automatic Control, 2022, 67(5): 2451-2458.
[46] Kawano Y, Cao M. Design of privacy-preserving dynamic controllers[J]. IEEE Transactions on Automatic Control, 2020, 65(9): 3863-3878.
[47] Goel S, Negi R. Guaranteeing secrecy using artificial noise[J]. IEEE Transactions on Wireless Communications, 2008, 7(6): 2180-2189.
[48] Luo H, Yu X, Zhang Z, Gan C. Channel estimation for 5G mm wave communications systems: a survey[J]. Telecommunication Engineering, 2021, 61(2): 254-262.
[49] Xu D, Chen B, Yu L, Zhang W. Secure dimensionality reduction fusion estimation against eavesdroppers in cyber-physical systems[J]. ISA Transactions, 2020, 104: 154-161.
[50] Xu D, Chen B, Zhang Y, Yu L. Distributed anti-eavesdropping fusion estimation under energy constraints[J]. IEEE Transactions on Automatic Control, 2023, doi: 10.1109/TAC.2023.3250094.
[51] Zhang Z, Cheng P, Wu J, Chen J. Secure state estimation using hybrid homomorphic encryption scheme[J]. IEEE Transactions on Control Systems Technology, 2021, 29(4): 1704-1720.
[52] Emad S, Alanwar A, Alkabani Y, El–Kharashi M W, Sandberg H, Johansson K H. Privacy guarantees for cloud-based state estimation using partially homomorphic encryption[C]. 2022 European Control Conference (ECC), London, United Kingdom, 2022: 98-105.
[53] Ny J Le, Pappas G J. Differentially private filtering[J]. IEEE Transactions on Automatic Control, 2014, 59(2): 341-354.
[54] Andre H, Ny J Le. A differentially private ensemble Kalman filter for road traffic estimation[C]. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017: 6409-6413.
[55] Wang J, Zhu R, Liu S. A differentially private unscented Kalman filter for streaming data in IoT[J]. IEEE Access, 2018, 6: 6487-6495.
[56] Ny J Le, Mohammady M. Differentially private MIMO filtering for event streams[J]. IEEE Transactions on Automatic Control, 2018, 63(1): 145-157.
[57] Tsiamis A, Gatsis K, Pappas G J. State-secrecy codes for stable systems[C]. 2018 Annual

American Control Conference (ACC), 2018: 171-177.

[58] Tsiamis A, Gatsis K, Pappas G J. State-secrecy codes for networked linear systems[J]. IEEE Transactions on Automatic Control, 2020, 65(5): 2001-2015.

[59] Leong A S, Quevedo D E, Dolz D, Dey S. Transmission scheduling for remote state estimation over packet dropping links in the presence of an eavesdropper[J]. IEEE Transactions on Automatic Control, 2019, 64(9): 3732-3739.

[60] Huang L, Ding K, Leong A S, Quevedo D E, Shi L. Encryption scheduling for remote state estimation under an operation constraint[J]. Automatica, 2021, 127: 109537.

[61] Wang L, Cao X, Zhang H, Sun C, Zheng W X. Transmission scheduling for privacy-optimal encryption against eavesdropping attacks on remote state estimation[J]. Automatica, 2022, 137: 110145.

[62] Nekouei E, Sandberg H, Skoglund M, Johansson K H. Optimal privacy-aware estimation[J]. IEEE Transactions on Automatic Control, 2022, 67(5): 2253-2266.

[63] Li C, Wang Z, Song W, Zhao S, Wang J, Shan J. Resilient unscented Kalman filtering fusion with dynamic event-triggered scheme: Applications to multiple unmanned aerial vehicles[J]. IEEE Transactions on Control Systems Technology, 2023, 31(1): 370-381.

[64] Reif K, Gunther S, Yaz E, Unbehauen R. Stochastic stability of the discrete-time extended Kalman filter[J]. IEEE Transactions on Automatic Control, 1999, 44(4): 714–728.

[65] Kalman R E. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering, 1960, 82: 35–45.

[66] Geng Q, Viswanath P. Optimal noise adding mechanisms for approximate differential privacy[J]. IEEE Transactions on Information Theory, 2016, 62(2): 952-969.

[67] Du M, Wang K, Xia Z, Zhang Y. Differential privacy preserving of training model in wireless big data with edge computing[J]. IEEE Transactions on Big Data, 2020, 6(2): 283-295.

[68] Eustace R W, Woodyatt B A, Merrington G L, and A. Runacres. Fault signatures obtained from fault implant tests on an F404 engine[J]. ASME Trans. J. Engine, Gas Turbines, Power, 1994, 116(1): 178-183.

[69] Chen B, Hu G, Zhang W, Yu L. Distributed mixed $H_2/H\infty$ fusion estimation with limited communication capacity[J]. IEEE Transactions on Automatic Control, 2016, 61(3): 805-810.

致 谢

时光荏苒，岁月如梭，接近四年的硕士生涯将告一段落。从大四推免开始，我便跟随导师学习相关知识。从通信到控制，几乎是一窍不通，即使是参加过各种自动化专业竞赛，也对纯理论的研究感到陌生。因此，能有今天这样丰硕的成果，三分靠努力，七分得益于他人的支持和帮助。

首先，非常感谢我的导师陈博教授，他给予了我非常大的帮助。在我刚入门的时候，他就不厌其烦地帮我解答困惑。即使是现在看来都是非常基础的问题，当时基础薄弱的我也难以理解，然而，他仍非常细心地教导我，带我一步一步地分析理解。后来，在我基础打好之后，他就把很多很好的创新点都交给我去做，不仅跟我一起推导数学公式的细节，还对我论文进行认真修改。在他的帮助下，我才能有如此多的成果。

其次，非常感谢学院里俞立教授、张文安教授等老师在各个方面的帮助，感谢同课题组的沈英老师、孙哲老师、王浙明教授的共同指导。感谢我的师兄章宇晨、胡中尧、许大星、翁品迪、翁世清、王如生、杨望卓、胡明南、项秉铜、李同祥、周京以及同实验室其他的师兄弟们在这三年里的相互学习。还要感谢我的室友李琛玮、王颖、金学成、许建华在生活上对我的照顾，让研究生这段时间成为我最充实快乐的大学时光。

此外，感谢香港理工大学的刘亮教授给予我的交流访学机会以及其在多输入多输出系统研究上的帮助，感谢香港城市大学 Ehsan Nekouei 教授在信息论与强化学习等方法上对我的指导，感谢瑞典皇家理工学院徐倩雯教授、香港科技大学杜娟教授和香港理工大学黄海龙教授给予的交流面试机会。

最后，感谢我的家人在此期间对我的支持，让我没有后顾之忧。感谢所有评阅和答辩的专家在百忙之中所提出的专业性意见。

谨以此文献给所有帮助我、关心我、支持我的人。

作者简介

1 作者简历

1998 年 04 月出生于浙江江山。

2016 年 09 月——2020 年 06 月，浙江工业大学信息工程学院通信工程专业学习，获得工学学士学位。

2020 年 10 月——至今，浙江工业大学信息工程学院控制科学与工程专业学习，攻读工学硕士学位。

2022 年 07 月——2022 年 10 月，香港理工大学电子及资讯工程学系联合培养。

2 攻读硕士学位期间发表的学术论文

以第一作者发表的期刊论文

[1] **Yan X**, Chen B, Zhang Y, Yu L. Distributed encryption fusion estimation against full eavesdropping[J]. **Automatica**, 2023, 153: 111025.（SCI，中科院 2 区 top，对应论文第四章）

[2] **Yan X**, Chen B, Zhang Y, Yu L. Guaranteeing differential privacy in distributed fusion estimation[J]. **IEEE Transactions on Aerospace and Electronic Systems**, 2022, doi: 10.1109/TAES.2022.3219799.（SCI，中科院 2 区 top，对应论文第三章）

[3] **Yan X**, Zhang Y, Xu D, Chen B. Distributed confidentiality fusion estimation against eavesdroppers[J]. **IEEE Transactions on Aerospace and Electronic Systems**, 2022, 58(4): 3633-3642.（SCI，中科院 2 区 top，对应论文第二章）

[4] **Yan X**, Chen B, Hu Z. Distributed estimation for interconnected dynamic systems under binary sensors[J]. **IEEE Sensors Journal**, 2022, 22(13): 13153-13161.（SCI，中科院 2 区 top）

[5] **Yan X**, Chen B, Qiu X. Distributed dimensionality reduction fusion Kalman filtering with quantized innovations[J]. **Circuits, Systems, and Signal Processing**, 2021, 40: 5234-5247.（SCI，中科院 3 区）

[6] **Yan X**, Zhuo S, Wu Y, Chen B. Distributed privacy-preserving fusion estimation using homomorphic encryption[J]. **Journal of Beijing Institute of Technology**, 2022, 31(6): 551-558.（EI）

非第一作者发表的期刊论文

[7] Xu D, **Yan X**, Chen B, Yu L. Energy-constrained confidentiality fusion estimation against eavesdroppers[J]. **IEEE Transactions on Circuits and Systems II: Express Briefs**, 2022, 69(2): 624-628.（SCI，中科院 2 区）

以第一作者发表的学术会议论文

[8] **Yan X**, Chen B, Teng Y, Ge L. Distributed estimation for discrete sequential systems under binary sensors[C]. **IEEE 30th International Symposium on Industrial Electronics**, 2021: 01-06, Kyoto, Japan.（EI）

3 参与的科研项目及获奖情况

[1] 信息物理系统中基于分布式融合的网络攻击检测与安全状态估计，国家自然科学基金面上项目，编号 61973277，2020.

[2] 网络化融合系统的攻击检测与安全状态估计理论研究，浙江省自然科学基金杰出青年项目，编号 LR20F030004，2020.

学位论文数据集

<table>
<tr><td>密　级*</td><td>中图分类号*</td><td>UDC*</td><td>论文资助</td></tr>
<tr><td>公开</td><td>TP301</td><td>007</td><td></td></tr>
<tr><td>学位授予单位名称*</td><td>学位授予单位代码*</td><td>学位类型*</td><td>学位级别*</td></tr>
<tr><td>浙江工业大学</td><td>10337</td><td>工学硕士</td><td>硕士</td></tr>
<tr><td>论文题名*</td><td colspan="3">基于主动注入随机噪声的分布式安全融合估计</td></tr>
<tr><td>关键词*</td><td colspan="2">分布式融合估计，窃听攻击，随机噪声，零空间，差分隐私</td><td>论文语种*</td></tr>
<tr><td>并列题名</td><td colspan="2"></td><td>中文</td></tr>
<tr><td>作者姓名*</td><td>严新豪</td><td>学　号*</td><td>2112003034</td></tr>
<tr><td>培养单位名称*</td><td>培养单位代码*</td><td>培养单位地址*</td><td>邮政编码*</td></tr>
<tr><td>浙江工业大学信息工程学院</td><td>10337</td><td>杭州市潮王路 18 号</td><td>310014</td></tr>
<tr><td>学科专业*</td><td>研究方向*</td><td>学　制*</td><td>学位授予年*</td></tr>
<tr><td>控制科学与工程</td><td>安全融合估计</td><td>3</td><td>2023</td></tr>
<tr><td>论文提交日期*</td><td colspan="3">2023 年 6 月</td></tr>
<tr><td>导师姓名*</td><td>陈博</td><td>职　称*</td><td>教授</td></tr>
<tr><td>评阅人</td><td>答辩委员会主席*</td><td colspan="2">答辩委员会成员</td></tr>
<tr><td></td><td>邓瑞龙</td><td colspan="2">史秀纺、董山玲</td></tr>
<tr><td colspan="4">电子版论文提交格式：文本（ √ ）图像（　）视频（　）音频（　）多媒体（　）其他（　）</td></tr>
<tr><td>电子版论文出版（发布）者</td><td>电子版论文出版（发布）地</td><td colspan="2">版权声明</td></tr>
<tr><td></td><td></td><td colspan="2"></td></tr>
<tr><td>论文总页数*</td><td colspan="3">59</td></tr>
<tr><td colspan="4">注：共 33 项，其中带*为必填数据，为 25 项。</td></tr>
</table>