遗传算法与模糊理论结合的传感网络防御算法

李 环

(东莞理工学院计算机科学与技术学院 广东 东莞 523808)

摘 要 由于无线传感器网络(WSNs)的通信链路并不如有线网络一样私密可控,在面对伪装认可攻击(FEIAs)时鲁棒性不强。针对这种情况,提出基于遗传算法与模糊理论的自适应防御(ACS)方案。针对 ACS 方案容易对 FEIAs 产生误判的问题,重点分析了安全参数选择,并对此方案进行模拟仿真。仿真结果显示提出的结合型防御算法相较于传统 ACS 方案减少了存储需求和计算复杂度,能提高检测精度,同时能在单一 WSN 模型中有效实现,实现模糊系统的自适应优化。

关键词 无线传感器网络 伪装认可攻击 模糊理论 遗传算法

中图分类号 TP391.9 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2024.07.046

WIRELESS SENSOR NETWORK DEFENSE ALGORITHM BASED ON GENETIC ALGORITHM AND FUZZY THEORY

Li Huan

(*School of Computer Science and Technology*, *Dongguan University of Technology*, *Dongguan 523808*, *Guangdong*, *China*)

Abstract The communication link of wireless sensor networks (WSNs) is not as private and controllable as wired network, and its robustness is not strong in the face of false endorsement insertion attacks (FEIAs). In view of this, this paper combines the genetic algorithm and fuzzy theory, and proposes an adaptive defense scheme (ACS). Aiming at the problem that ACS scheme was prone to misjudge FEIAs, this paper focused on the selection of safety parameters. And this scheme was simulated. The simulation results show that compared with the traditional ACS scheme, the combined defense algorithm reduces the storage requirements and computational complexity, improves the detection accuracy, and can be effectively implemented by a single WSN model, so as to realize the adaptive optimization of fuzzy system.

Keywords Wireless sensor networks Endorsement insertion attacks Fuzzy theory Genetic algorithm

0 引 言

无线传感器网络(WSN)由大量微型传感器节点组成,这些节点监视周围区域,并向基站报告该区域内目标的行为(例如车辆的外观)[1]。由于节点没有人工参与,恶意攻击者可以捕获一些节点,却不被检测到[2-3]。攻击者可以破坏节点的信息,发起各种内部攻击[4]。

这些攻击可以分为两种,虚假数据注入攻击(FDIA)与伪装认可攻击(FEIA)[5]。目前,在前一种攻击方面,学者已经进行了大量的工作[2,4-5,6-11];而关于后一种攻击方式,相关研究较少[6,12]。关于 FEIA 的应对策略可以分为两种,一种是分布式[6,12-13],另一种为集中式[14]。Lee 等[14]提出了一种集中式解决方案:自适应防御(ACS)。但 ACS 的一个主要缺点是用户需要确定许多用于检测安全攻击的安全参数。如果这些参数选择不恰当,可能会导致对事件的误报。

为了防御 WSN 中 FEIAs 攻击,本文提出了一个基于模糊理论的改进 ACS 方法。本文方法采用两种基于模糊规则的系统来检测和防御 FEIAs:一种用于检测 FEIAs,另一种用于选择对策。本文方法可以检测

收稿日期:2021-01-28。基于机器学习算法的科技成果转化精准推荐模型研究(2018KTSCX222)。**李环**,副教授,主研领域:智能计算,无线网络安全与信息技术,隐私保护。

出 FEIA,并选择最有效的防御措施。这个集中式的方案不仅能大量节约资源还具有以下优点:

1) 将遗传算法与仿真相结合,对模糊系统进行自动优化,实质上消除了 ACS 中人工参数设置的问题[15]。

2) 减少了检测 FEIAs 和选择对策输入因素,降低了存储要求和计算复杂度。

3) 不需要任何特殊的设备,如全球定位系统(GPS),就可以获得输入因子。

4) 利用模糊理论的近似计算能力,减少 FEIAs 检测过程中的误差。

1 基于模糊理论的自适应方法防御 FEIAs

本文方法是基于模糊理论改进的 ACS 方案。主要研究用来跟踪移动目标的大规模传感器节点[16]。每个节点都有与基站 s 联系的密钥来验证检测报告。在检测到目标后,节点基于密钥生成一个包含 MAC 的检测报告,此报告经过多个转发节点到达基站。基站则根据 MAC 来检验报告的完整性。

1.1 基本检测和防御程序

如图 1 所示,检测程序放在基站上。存储硬件用于临时存储由基站收集的报告。每个报告将存储一段预先防御时间 TR。基于模糊规则的系统使用存储的报告检测 FEIAs。此系统将每个目标报告的平均合法性(ARl)和 ARl 的方差(VRL)进行模糊计算来检测 FEIAs。

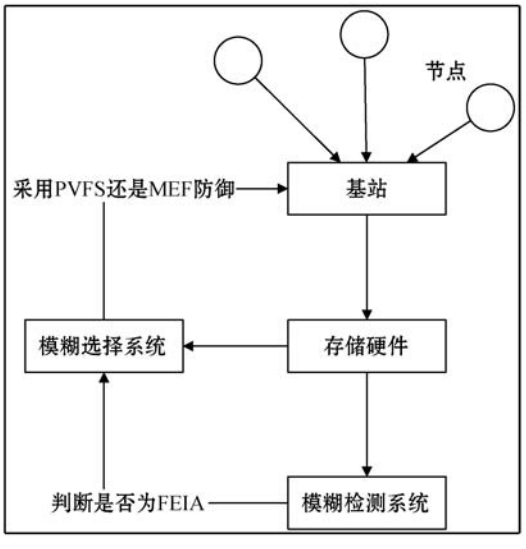

图 1 基于模糊的检测和对抗

当模糊系统检测出 FEIA 时,系统将信息立即报告给用户。这时另一个模糊系统通过使用存储的报告,考虑可靠性和资源效率两个因素,选择应对 FEIA 最有效的策略。在选择对策时,模糊计算过程使用了两个因素:报告对象的节点数(NNN)和到对象的平均距离(ADO)[14,17]。最后,通过传播控制消息激活模糊系统选择的对策。如果威胁已经消除,则可以通过传播另一条消息使对策失效。

1.2 使用模糊理论的优点

本文方法采用了两种模糊系统:一种用于检测 FEIAs,另一种用于选择对策。使用模糊系统的主要原因如下。

1) 将遗传算法与仿真相结合,可以自动优化模糊系统的隶属度函数[15]。实际上,不需要专家(例如 ACS 中的专家)来确定内部参数。

2) 模糊计算的知识可以用 if-then 规则表示[18]。所需要的基本规则可以由经验确定,如果以往的经验太少,那么可以用人工神经网络来代替模糊系统[19]。

3) 模糊系统可用于近似计算,尤其是在数据不精确的情况下。有些数据不准确现象可能是由一些节点的故障导致的[20]。

1.3 FEIA 检测的因素

在本文方法中,检测 FEIA 的模糊系统使用了两个因素:ARL 和 VRL。

ARL 是检测 FEIA 的最重要因素。当基站收到一个节点生成的报告时,它首先检查报告的合法性,对比报告内容与节点生成 MAC 的匹配度。如果它们完全匹配,报告是合法的;否则,报告被篡改。尽管节点或无线通信故障可能导致 MAC 不匹配,但大部分报告都是合法的;如果节点受到 FEIAs 攻击,大多数报告的 MAC 都将不正确。如果我们检查目标的 ARL 值较低时,则说明该节点是 FEIA 的目标。基站在 TR 内收到目标的报告,合法报告的合法性是 1.0,而非法报告的合法性是 0.0。

VRL 是检测 FEIAs 的另一个重要因素。FEIA 攻击会导致 ARL 低,但 FDIA 攻击也可以带来相同的结果。如果攻击者通过被破坏的节点发起 FDIA,那么这些节点生成的报告大都合法,但如果使用外部节点,报告则为非法。在这种情况下,攻击者的目标变为消耗节点能源资源或干扰节点间通信。与 FDIA 相比,FEIA 攻击下的 ARL 会随着目标运动变化很大,即 VRL 不为 0。因此,在检测 FEIAs 时需要考虑 VRL,以区分 FDIAs 和 FEIAs 攻击。

其他因素可以用来检测 FEIAs,但是检测过程中需要额外的存储空间。由于 if-then 规则的爆炸式增长,更多的因素在计算过程中会带来更多的计算消耗。在获得某些因素时,还需要给节点安装一些特殊设备。

1.4　对策选择的因素

经过模糊系统检测出节点已经被 FEIA 攻击后,第二个对策选择系统会根据可靠性和资源消耗选择采用 PVFS 或者 MEF 来抵御 FEIA。在这一步的计算中,我们考虑 NNN 和 ADO 两个因素。

PVFS 和 MEF 的设计主要是为了防御 FDIAs 和 FEIAs。在防御措施激活后,报告有一定的丢弃概率(RDP)。当报告在中间节点转发时,每个节点都会验证报告。如果验证失败一次(MEF)或多次(PFVS),节点就可以丢弃报告。为了防止合法的报告被丢弃,MEF 通过多条路径转发同一报告,而 PVFS 是将一条路径生成报告发送给基站。总体来说,MEF 的 RDP 比 PVFS 要好。但是多路径路由意味着 MEF 在转发过程中消耗更多的能量资源。因此,PVFS 在节能方面通常优于 MEF。

影响 RDP 的因素之一是节点个数(NNN)。如果目标很大,或者移动很快,短时间内多个节点生成和转发报告,为了节约资源 PVFS 是一个很好的选择。但如果一个目标是由很少节点报告的,为了 RDP、MEF 是一个更好的解决对策。另一个影响 RDP 的因素是报告完成的平均跳数(ADO)。与 MEF 相比,由于只有一条路径,PVFS 具有更高的能量效率。但由于经过很多节点和路径长时间的负荷,报告被丢弃的概率更大。对于路径很短的报告,经研究证明,MEF 有着与 PVFS 相当的资源效率[14]。RDP 也会受到其他因素的影响,但考虑其他因素将需要额外的空间要求和繁重的计算费用或者一些特殊的硬件。

1.5　模糊逻辑设计

模糊系统中用于检测 FEIAs 的模糊变量表示如下:

1) ARL = {VL (Very Low), L (Low), M (Medium), H (High), VH (Very High)}。

2) VRL = {DR (Decreasing Rapidly), D (Decreasing), S (Steady), I (Increasing), IR (Increasing Rapidly)}。

3) ADR (Attack Detection Result) = {FDIA? (Perhaps an FDIA), RO (Real Object), FEIA? (Perhaps an FEIA), FEIA (Probably an FEIA)}。

用来检测 FEIA 的模糊 if-then 规则如表 1 所示。表的最左侧列上的数字用作规则的标识符。规则 1 可以理解为,对于一个对象,如果 ARL 是 VL,VRL 是 DR,那么 ADR 就是 FEIA。每个输入变量(即,ARL 和 VRL)有 5 种,所以系统最多可以有 5 × 5 = 25 个规则。

表 1　模糊 if-then 规则的 FEIA 检测

Rule number	ARL	VRL	ADR
1	VL	DR	FEIA
2	VL	D	FDIA?
3	VL	S	FDIA?
4	VL	I	FDIA?
5	—	—	—
6	L	DR	FEIA
7	L	D	FEIA?
8	L	S	FDIA?
9	L	I	FDIA?
10	—	—	—
11	M	DR	FEIA
12	M	D	FEIA?
13	M	S	RO
14	M	I	FEIA?
15	M	IR	FEIA
16	—	—	—
17	H	D	RO
18	H	S	RO
19	H	I	FEIA?
20	H	IR	FEIA
21	—	—	—
22	VH	D	RO
23	VH	S	RO
24	VH	I	RO
25	VH	IR	FEIA

对于每个目标,模糊系统定期使用 ARL 和 VRL 判断目标是否被 FEIA 攻击。由于攻击者没有办法使受控制的节点随目标移动,整个系统中会有一块受 FEIA 影响的静止区域。当目标进入该区域时,会有非常多的不合法 MACs,ARL 和 VRL 较低,从而使模糊系统检测到 FEIA。在规则 1、6 和 11 的 if-clauses 中 VRL 是 DR,描述了这种情况。因此规则的 then-clauses 规则断定目标可能被 FEIA 攻击,即 ADR 是 FEIA。当它离开该区域时,ARL 和 VRL 会很高也可以使 FEIA 被检测到。规则 15、20 和 25 中描述了 VRL 是 IR 对应目标被 FEIA 攻击。如果攻击者使用外部节点进行 FDIA 攻击,那么目标的 ARL 将非常低,VRL 的变化非常小,模糊系统将警告用户节点可能被 FDIA 攻击。规则 3 的 if-clauses 规则规定了这种情况,then-clauses 规则断定节点不存在。但是本文方法不是为检测 FDIAs 而设计的。为了正确检测 FDIAs 的发生,用户

应采用其他 FDIA 检测解决方案。

系统的对策选择有 25 条模糊 if-then 规则,如表 2 所示。模糊变量的表示如下:

1) NNN = {VS (Very Small), S (Small), M (Medium), L (Large), VL(Very Large)}。

2) ADO = {VN (Very Near), N (Nera), M (medium), F (Far), VF (Very Far)}。

3) CSR (Countermeasure Selection Result) = {PVFS, PVFS? (PVFS might be better), MEF? (MEF might be better), MEF}。

表 2 模糊 if-then 规则的 FEIA 检测

Rule number	NNN	ADO	CSR
1	VS	VN	PVFS
2	VS	N	MEF
3	VS	M	MEF
4	VS	F	MEF
5	VS	VF	MEF
6	S	VN	PVFS
7	S	N	PVFS
8	S	M	MEF
9	S	F	MEF
10	S	VF	MEF
11	M	VN	PVFS
12	M	N	PVFS
13	M	M	PVFS
14	M	F	MEF
15	M	VF	MEF
16	L	VN	PVFS
17	L	N	PVFS
18	L	M	PVFS
19	L	F	PVFS
20	L	VF	MEF
21	VL	VN	PVFS
22	VL	N	PVFS
23	VL	M	PVFS
24	VL	F	PVFS
25	VL	VF	PVFS

当确定目标被 FEIA 攻击后,模糊系统利用 PVFS 或 MEF 进行对策选择,并对其有效性进行评估。模糊计算过程使用 NNN 和 ADO 作为对象。在规则 2、3、4 和 5 的 if-clauses 中,如果 NNN 是 VS,少量节点为该目标生成报告(该目标可能非常小,并且在范围内移动非常缓慢)。MEF 将是一个更好的解决对策。相反,如果 NNN 是 VL(一个非常大的或快速移动的目标),例如规则 21、22、23、24 和 25,多个节点将报告传递给基站 s,模糊系统将依据可靠性和资源消耗选择出最适合的对策。在规则 5、10、15 和 20 中目标位于离基站很远的地方(ADO 是 VF),每个报告的目标将有更大的机会被丢弃。由于在 PVFS 中,单个报告通过单一路由路径转发到基站,在可靠性方面 MEF 的性能优于 PVFS。相反在规则 1、6、11、16 和 21 中,ADO 是 VN (目标离基站 s 非常近),那么 PVFS 将是一个更好的解决方案,因为它将节省有限的资源,并可靠地向用户报告目标。

1.6 模糊隶属函数的优化

本文将仿真和遗传算法相结合,自动优化 FEIA 检测和对策选择这两个模糊系统[15]。为了优化 WSN 的模糊系统,用户只需要建立一个 WSN 的仿真模型,该模型可以使用基于 GUI 的工具获得。实际上,用户只需要确定两个因素:NNN 的最大值和 ADO。即使是没有经验的用户也可以很容易地为这两个因素选择合适的值,解决了在 ACS 中需要根据经验确定大量参数值的问题。

如图 2 所示,确定模糊隶属函数的参数表示为染色体。最初,一组染色体(一个种群)是随机产生的。每个染色体的适应度通过模拟来评估。接下来,根据模拟结果,通过选择、交叉和突变三种遗传操作对种群进行进化。重复评估进化过程,直到满足一个条件(例如,当世代数达到 400 时)。这两个模糊系统互不影响,可以分别优化。在对用于检测的模糊系统进行优化时,利用误报率来评价染色体的适应度。在选择系统的优化过程中,采用了报告传输过程的可靠性和资源效率进行评价。

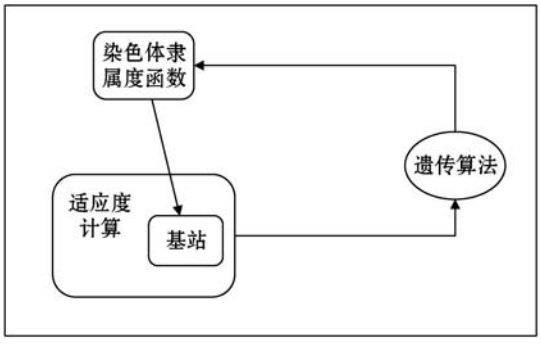

图 2 将遗传算法与仿真相结合进行模糊系统优化

1.7 策略激活和失效

当系统为 FEIA 选择了策略时,通过传播控制消息来激活策略。传播可以限制在小的区域内,例如,通过分组节点或限制传播距离,或应用于整个网络[21]。在消除威胁后,可以通过传播另一条控制消息使对策失效[22]。

2　仿真结果

为了验证本文方法的性能,我们将其与基于 ACS 的仿真进行了比较。传感器场尺寸为 1 000 × 100 m^2,单个基站位于场角。总共部署 4 500 个节点,这些节点可以探测目标并与周边 10 米内的节点进行通信。恶意攻击者可以物理捕获 1 ~ 15 个节点,然后使用它们来进行 FEIA 攻击。攻击者也可以使用外部节点进行 FDIAs 攻击。每一个目标在边界出现,以随机路线运动并消失在边界。目标的移动速度 U 米/秒(0.5, 1.0)。在检测 FEIAs 的过程中,比较了本文方法和 ACS 在两种不同设置下的合法报告误报率(FPER)和非法报告忽略率(FNER)。在策略选择过程中,采用本文方法、ACS、PVFS、MEF 分别测量向基站传输的合法报告比例和传输过程中的平均能耗。

2.1　检测性能

在图 3 和图 4 中,ACS(expert)表示参数由经验丰富用户优化的检测方法,ACS(common user)表示用户参数配置不熟练的方法,The proposed method 表示本文提出的方法。FPER 为正常报告被检测为不合法报告的比例;FNER 则指没有检测到 FEIA 的概率。图 3 显示,随着 TR 的增加,FPER 降低,这是因为基站有更多的报告来进行检测。但随之而来的是大量时间与存储能力的消耗。与 ACS 比较,本文中模糊系统的近似计算,降低了 FPER。这一点对传感器网络十分重要。本文提出方法的性能与 ACS(expert)相似,但并不需要经验丰富用户的配置。另外在 TR 特别小的时候本文的方法对 FEIAs 的检测有较大提升。在图 4 中对 FNER 的研究,有相同的结论。总的来说,本文提出的方法优于 ACS 方法,尤其是当 ACS 参数没有由经验丰富的用户配置,存储空间受限或者需要快速检测时。

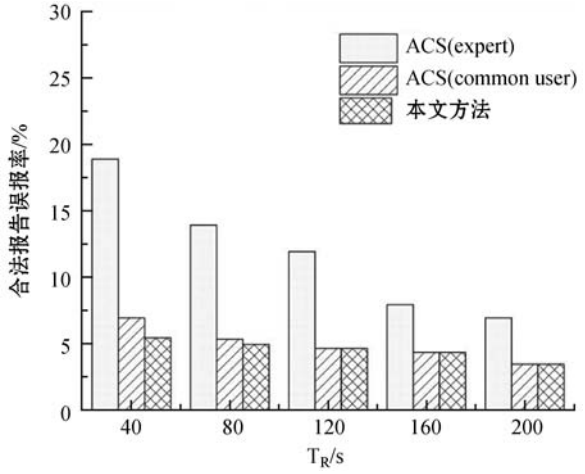

图 3　检测 FEIAs 过程中的 FPER

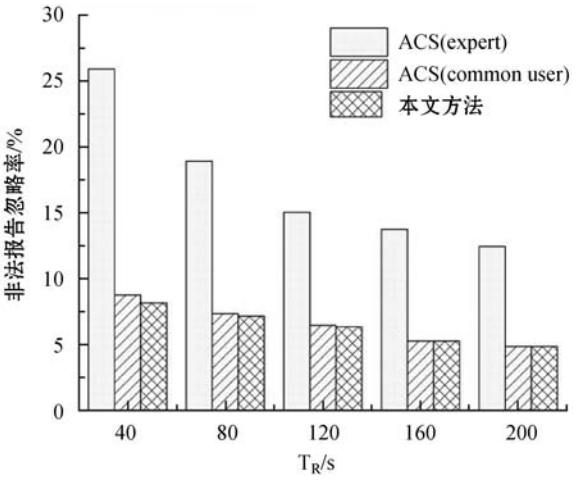

图 4　检测 FEIAs 过程中的 FNER

2.2　可靠性和资源效率

图 5 和图 6 显示了到达基站的合法报告的比例(可靠性)和每份报告到达的平均资源消耗。Case1 表示高密度 WSN 的结果,Case2 表示中密度 WSN 的结果。高密度 WSN 中一个目标平均可以被 15 个节点检测到,中密度则为 9 个。能耗水平的计算基于文献[2]中的模型。可以看出,MEF 使用了多条路径,在可靠性方面是最佳解决方案。但是无线传感器网络消耗了大量的资源。相比之下,PVFS 使用单一路径,最小化了节点验证,在资源效率方面表现最好。但是很多合法的报告并没有送到基站,这就意味着一些真实的目标并没有报告给用户。在大多数 WSN 应用中,可靠性是非常关键的。本文提出的方法和 ACS 在可靠性方面都表现得相当好,每个真实目标都报告给用户。由于 ACS 考虑了一个额外的 FEIA 干扰[23],它的能量效率略优于所提出的方法。然而,在生成报告期间,ACS 需要不相交的两条路径,很复杂而且不太容易实现。此外,ACS 在选择过程中使用了三个因素,而在本文方法中使用了两个因素,所以本文方法可以应用于更大范围的无线传感器网络。与 ACS 相比,本文方法对节点设备和基站硬件的依赖性较小。

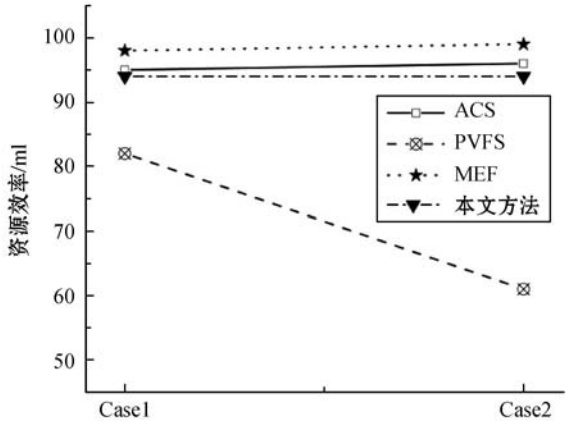

图 5　检测报告传递过程中的可靠性

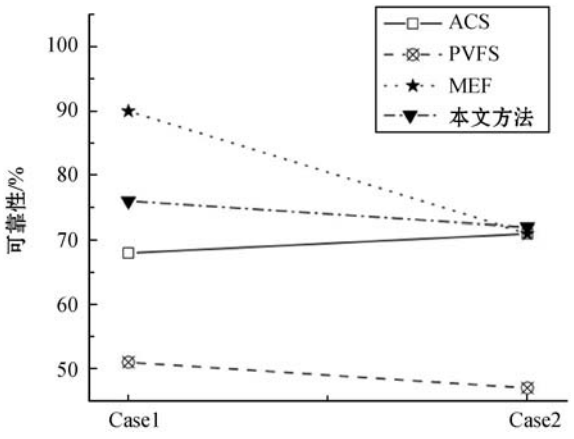

图 6 检测报告传递过程中的资源效率

3 结 语

本文提出了一种基于模糊理论的方法来自适应地防御密集无线传感器网络中的 FEIAs。该方法采用了两种模糊系统:一种用于检测 FEIAs,另一种用于选择防御 FEIA 策略。对于每个目标,前一个系统根据基站收集的数据来确定目标是否处于 FEIA 状态。在检测到 FEIA 状态后,后一个系统基于数据评估 PVFS 和 MEF 对 FEIA 的有效性。该方法的一个主要优点是可以对无线传感器网络的模糊系统进行自动优化。与 ACS 相比,基于模糊逻辑的近似计算提高了检测精度;使用的因子更少,可以减少存储需求和计算复杂度。仿真结果验证了该方法的优越性。

参 考 文 献

[1] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. A survey on sensor networks[J]. IEEE Communications Magazine,2002,40(8):102 - 114.

[2] 黄天奇,王布宏,林东. FDIA 对雷达组网系统数据融合的影响分析[J]. 火力与指挥控制,2020,45(6):67 - 72.

[3] Yu H, He J S, Liu R, et al. On the security of data collection and transmission from wireless sensor networks in the context of Internet of things[J]. International Journal of Distributed Sensor Networks,2013,2013(31):1 - 13.

[4] 宋蕾. 数据注入攻击下的信息物理系统安全控制[D]. 上海:上海交通大学,2019.

[5] Zhu S, Setia S, Jajodia S, et al. Interleaved hop-by-hop authentication against false data injection attacks in sensor networks[J]. ACM Transactions on Sensor Networks, 2007, 3(3):14.

[6] Li F, Srinivasan A, Wu J. PVFS: A probabilistic voting-based filtering scheme in wireless sensor networks[J]. International Journal of Security and Networks,2008,3:173 - 182.

[7] 赵丽莉,刘忠喜,孙国强,等. 基于非线性状态估计的虚假数据注入攻击代价分析[J]. 电力系统保护与控制,2019,47(19):38 - 45.

[8] 裔传俊. 无线传感器网络虚假数据注入攻击防御策略研究[D]. 南京:南京邮电大学,2019.

[9] 阮嘉祺. 虚假数据攻击模型及其防范策略的研究[D]. 深圳:深圳大学,2019.

[10] 李唯. 虚假数据注入攻击下信息物理系统的安全控制研究[D]. 兰州:兰州理工大学,2019.

[11] 刘珊. 信息物理系统在网络攻击下的控制问题研究[D]. 广州:华南理工大学,2019.

[12] Kim M S, Cho T H. A multipath en-route filtering method for dropping reports in sensor networks[J]. IEICE Transactions on Information and Systems,2007,90(12):2108 - 2109.

[13] Yu C M, Tsou Y T, Lu C S, et al. Constrained function-based message authentication for sensor networks[J]. IEEE Transactions on Information Forensics and Security,2011,6(2):407 - 425.

[14] Lee H Y, Cho T H. A scheme for adaptively countering application layer security attacks in wireless sensor networks[J]. IEICE Transactions on Communications,2010,93(7):1881 - 1889.

[15] Lee H Y, Cho T H. Optimized fuzzy adaptive filtering for ubiquitous sensor networks[J]. IEICE Transactions on Communications,2011,94(16):1648 - 1656.

[16] 蔡文波,张亚. 带虚假数据注入攻击识别和通信触发机制的传感器网络分布式估计[J]. 东南大学学报(自然科学版),2019,49(5):890 - 896.

[17] Krau C, Schneider M, Eckert C. Defending against false-endorsement-based DoS attacks in wireless sensor networks[C]//1st ACM Conference on Wireless Network Security, 2008:13 - 23.

[18] Takacs M. Approximate reasoning in fuzzy systems based on pseudo-analysis and Uninorm residuum[J]. Acta Polytechnica Hungarica,2004,1(2):49 - 62.

[19] Li H X, Chen C L. The equivalence between fuzzy logic systems and feedforward neural networks[J]. IEEE Transactions on Neural Networks,2000,11(2):356 - 365.

[20] Serrano N, Seraji H. Landing site selection using fuzzy rule-based reasoning[C]//IEEE International Conference on Robotics and Automation,2007:4899 - 4904.

[21] Nghiem T P, Cho T H. A fuzzy-based interleaved multi-hop authentication scheme in wireless sensor networks[J]. Journal of Parallel and Distributed Computing,2009,69(5):441 - 450.

[22] Perrig A, Szewczyk R, Tygar J D, et al. SPINS: Security protocols for sensor networks[J]. Wireless Networks,2002,8:521 - 534.

[23] Lee H Y, Cho T H, Kim H J. Fuzzy-based detection of injected false data in wireless sensor networks[J]. Communications in Computer and Information Science,2010,76:128 - 137.