引用格式:范若楠,张晓荣,王鑫,等. 多 Agent 系统安全保护机制研究[J]. 信息对抗技术,2022, 1(3):46-56. [FAN Ruonan, ZHANG Xiaorong, WANG Xin, et al. Research on multi-agent system security protection mechanism[J]. Information Countermeasure Technology, 2022, 1(3):46-56. (in Chinese)]

多 Agent 系统安全保护机制研究

范若楠[1*],张晓荣[2],王 鑫[2],李 斌[1],张 俊[1]

(1. 中国船舶重工集团公司第七〇五研究所,陕西西安,710000;
2. 西安交通大学计算机科学与技术学院,陕西西安,710049)

摘 要 针对现有多 Agent 系统易受到恶意攻击、难以验证代理身份及代理间难以建立信任机制等安全问题,围绕多 Agent 分层系统架构,提出一种采用基于数字签名和非对称加密的多 Agent 通信加密方法,有效保证了通信环节的系统安全,防范了 Agent 代理身份被冒用。基于动态博弈对抗的思路,在多 Agent 体系评估的基础上,利用贪婪算法和遗传算法优化多 Agent 系统安全选择策略,提高了防御者在攻防博弈中获胜概率并缩短了攻防博弈轮次,保证了通信环节的安全防御,防止 Agent 代理信息被攻击者窃取。

关键词 多 Agent 系统;加密通信;数字签名;贪婪算法;遗传算法

中图分类号 TP 399∶TP 392 **文献标志码** A **文章编号** 2097-163X(2022)03-0046-11

DOI 10.12399/j.issn.2097-163x.2022.03.004

Research on multi-agent system security protection mechanism

FAN Ruonan[1*], ZHANG Xiaorong[2], WANG Xin[2], LI Bin[1], ZHANG Jun[1]

(1. The 705th Research Institute of China Shipbuilding Industry Corporation, Xi'an 710000, China;
2. School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract Aiming at the security problems such as the existing multi-agent system is vulnerable to malicious attacks, it is difficult to verify the identity of the agent and to establish a trust mechanism between agents, this paper focuses on the multi-agent layered system architecture, and proposes a multi-agent encryption based on asymmetric encryption algorithm and digital signature. The communication method can effectively ensure the system security of the communication and prevent the agent identity from being used fraudulently. Based on the dynamic game ideas in multiple agent system assessment, using the greedy algorithm and genetic algorithm to optimize the multi-agent system security selection strategy, we effectively increase the probability of winning and shorten the rounds of gaming, ensure the security of the communication, and prevent the agent information from being eavesdropped.

Keywords multi-agent system; encrypted communication; digital signature; greedy algorithm; genetic algorithm

收稿日期:2022-07-28 **修回日期**:2022-10-08

通信作者:范若楠,E-mail:fannyforfun84@126.com

作者简介:范若楠(1984—),女,高级工程师,研究方向为作战仿真; 张晓荣(1999—),女,硕士研究生,研究方向为网络测量; 王鑫(1998—),男,硕士研究生,研究方向为网络测量; 李斌(1983—),男,研究员,研究方向为作战仿真与指挥控制;张俊(1984—),男,高级工程师,研究方向为鱼雷仿真

0　引言

复杂系统，又称复合系统，是由许多部件组成的系统。这些部件之间存在着复杂的相互作用，因此系统的整体特性、功能难以从其组成部件的特性、功能中推理得出。复杂系统表现出的复杂性、偶然性和不可重复性使其难以通过传统的解析法、数值分析法或其他形式化方法进行描述[1-3]。基于多 Agent（人工智能体、可智能处理和执行相关操作）系统的建模方法，利用不同的 Agent 模拟复杂系统的组成部件，利用各个 Agent 间的交互模拟复杂系统各组成部件之间的联系与相互作用，从而能比较充分地刻画出复杂系统的微观细节和宏观上的“涌现”行为[4]。

目前多 Agent 系统的安全性面临着极大的挑战：一是使用从邻近主体或环境中获得的信息、知识用于自身决策的行为使得 Agent 易受到恶意攻击；二是由于缺乏可信任的权威中心，验证代理的身份和在代理之间建立信任机制比较困难，移动代理可能会受到恶意代理的不良影响[5-8]。

本文从 Agent 加密通信角度出发，为多 Agent 系统的安全保护提供了一种新的方式及未来的发展思路，并对多 Agent 体系进行了系统的评估。在基于动态博弈对抗的思路上提出了多 Agent 体系优化模型并进行了相关攻防实验。

1　相关工作

近年来，多 Agent 系统拥有自身独特的能力，例如自主性、便携性、流动性、智能性等[9-10]，因此多 Agent 系统在开放的分布式环境中被开发，并被广泛应用于与日常生活及工业生产相关的领域[11]。在开放领域中，因为多 Agent 系统的多个 Agent 的逻辑和物理位置呈现分散状态，所以每个 Agent 需要通过预先制定的通信协议，与其他 Agent 通过网络进行数据传输，以此协调彼此的任务，并协同实现总体目标。在此过程中，开放的环境带来了诸多安全隐患，意外的物理故障和对单个 Agent 的网络攻击可能会非常快速地通过信息交互传递给多 Agent 系统中的其他 Agent[10]。早在 1998 年，Greenberg 等[12]就已经归纳了几种可能针对开放环境中移动 Agent 的攻击方法。Wang 等[13]则提出了针对移动 Agent 的几种攻击路径和后果。

从多 Agent 系统环节的角度，多 Agent 系统的安全性应分为分布式调控中的安全问题与单个智能体中的安全问题。从多 Agent 系统安全性需求考虑，其安全性问题可以分为信息来源安全、授权机制安全、通信安全、主机保护安全、智能体保护安全以及共享资源安全[14]。在多 Agent 系统中，当网络通信受到攻击者的间歇性攻击时，可能会导致多 Agent 系统之间传输效率的丧失和通信中断，针对这个安全问题，Jin 等[15]开发了一种分布式自适应共识控制策略，以确保非线性多智能体系统在正常网络情况下的有界共识。对于多 Agent 的攻击数据传输和通信协议过程[16]，可能会发生 2 种类型的安全威胁：一种是探索性攻击（exploratory attack），即攻击者不会干扰通信，但会尝试提取可用信息，在这种情况下，传输的敏感信息可能泄露；另一种是诱发性攻击（causative attack），此时攻击者可能会尝试拦截、修改、删除甚至替换数据包。

2015 年，Basheer 等[17]通过智能体的可信度建立了 AgentOpCo（agent opinion confidence）模型，这种通用模型通过检测多 Agent 系统中智能体的置信度来保证系统安全。2020 年，Zuo 等[18]提出了一种安全机制来保证多 Agent 系统在受到的虚假数据注入攻击时系统的安全共识。Li 等[19]针对多 Agent 系统在受到的虚假数据注入攻击问题，采用了一种具有状态相关阈值的事件触发控制机制解决该问题。Wang 等[20]提出一种新颖的事件驱动机制针对多 Agent 系统受到拒绝服务攻击时及时更改为参数不确定性的事件触发方法。Sethi 等[21]提出一个基于强化学习的多 Agent 智能入侵检测系统。Naresh 等[22]基于 Quantum Diffie-Hellman 协议提出一种新的 Quantum GKA（QGKA）技术来保证多 Agent 系统之间共享密钥的安全性。

2006 年，Huynh 等[23]提出了有关多 Agent 系统中智能体的信任与声誉认证模型，为子系统验证提供了框架。2016 年，Zikratov 等[24]设计了动态网络下的信任模型，由于加入了时间驱动的信任级别认证，提升了多 Agent 系统在开放网络下的安全性；Majd 等[25]等则通过设置传递性、符合性、相似性与可靠性 4 个信任组件，提出了名为 TtSSR 的多 Agent 安全信任模型。

为了更好地进行分析，表 1 列出了本文的主要术语及其定义。其中，综合收益是综合探测收益、综合攻击收益之和，综合收益能够更好地对多 Agent 系统体系从探测模块、攻击模块进行综合评估。

表 1 关键术语定义

Tab. 1 Definition of key terms

符号	术语	定义
G_p	探测收益	探测收益，其中考虑敌方探测干扰我方的而造成损失情况下的我方探测收益为综合探测收益，不考虑敌方探测干扰对我方造成损失所产生的探测收益为简单探测收益
G_a	攻击收益	攻击收益，其中考虑敌方攻击干扰对我方造成损失所产生的攻击收益为综合攻击收益，不考虑敌方攻击干扰对我方而造成损失情况下我方的攻击收益为简单攻击收益
G_o	探测-攻击整体收益	综合探测收益与综合攻击收益之和
c_p	探测成本	对敌方进行探测所需的成本
g_p	主动探测收益	主动探测敌方的收益
d_p	干扰探测收益	对敌方进行干扰所产生的探测收益
i_p	诱导探测收益	对敌方进行诱导所产生的探测收益
a_p	对抗探测收益	对敌方进行打击所产生的探测收益
c_a	攻击成本	敌方进行打击的成本
g_a	攻击收益	对敌方进行攻击的收益

2 多 Agent 分层系统架构

随着人类知识体系的发展，在万物互联时代，人类正面临着大量具有大规模跨域分布、高密度交互关联、多时间尺度演变等特点的复杂系统问题[1]。与以往的复杂系统相比，以网络信息战、战场推演为代表的现代化复杂系统具有规模大、通信频率高、变化快等特征，这对仿真系统的通信能力、计算能力和决策速度都提出了极高要求[25]。而采用多 Agent 分层的系统架构，利用该系统架构的分布式结构特征以及智能化基本单元能够较好地适应对当前复杂系统的仿真要求[26-28]。

多 Agent 分层系统采用的分布式结构如图 1 所示。不同于传统的集中式系统构架，分布式系统构架将决策任务从决策大脑分发给不同的决策单元，由决策单元或 Agent 执行对应的决策任务，充分利用系统资源。相较于集中式系统，分布式系统构架具有并行计算能力强、通信效率高、容错率高、系统健壮性好等优点。

并行计算能力强则能快速调整分配方案，不受中央节点的算力、带宽限制，可充分发挥各 Agent 的计算能力。决策大脑可以将一些简单的、容错性高的任务分发给系统中的决策单元和 Agent 进行，充分发挥决策单元和 Agent 的计算能力，并保证决策大脑可以去处理重要的、复杂的任务。决策单元乃至各个 Agent 的决策活动都必须受决策大脑控制，并依据决策大脑指挥进行调度，保证系统整体可控。

通信效率高可以避免因通信拥塞而影响算法效率，不受中央节点的可用带宽限制，从而提高通信质量。将决策任务分发给决策单元和 Agent，利用决策单元和 Agent 直接计算结果并应用于自身，实时性好。尤其对于瞬息万变的战场，毫秒级别的网络延迟可能导致错过最佳攻击、防御和探测时间，而直接在 Agent 端进行计算和决策可以保证动作的及时性，同时可减轻决策大脑的压力。

容错率高则不会因为中央节点故障而导致系统崩溃；系统健壮性好体现在即使注册中心宕机，一段时间内各 Agent、决策单元还是能够调用决策大脑提供的决策指令。

在图 1 所示的多智能体分层结构的系统中，多个 Agent 之间能够相互通信，交换彼此的探测、攻击、防御等信息。每个 Agent 至少连接一个决策单元，帮助 Agent 制定其探测、攻击和防御策略，分发决策大脑和指挥人员的作战指挥信息。同时，决策大脑与决策单元直接保持通信，分发总指挥中心的指挥信息和策略。决策单元之间也可将自身的决策任务分发给其他决策单元，由其他决策单元帮助其完成决策工作。

为了防止决策大脑受到打击，必须保证其有多个备份。同时，通过整体备份和 AOF（append only file，以日志的形式记录服务器所处理的每一个写、删除操作）方式动态地存储决策大脑的所有指令，保证决策大脑备份和决策大脑信息一

致，一旦决策大脑出现被打击、断电、崩溃等紧急情况，有选举策略自动选举出最高效的决策大脑备份以替代决策大脑工作，保证系统整体运行平稳。

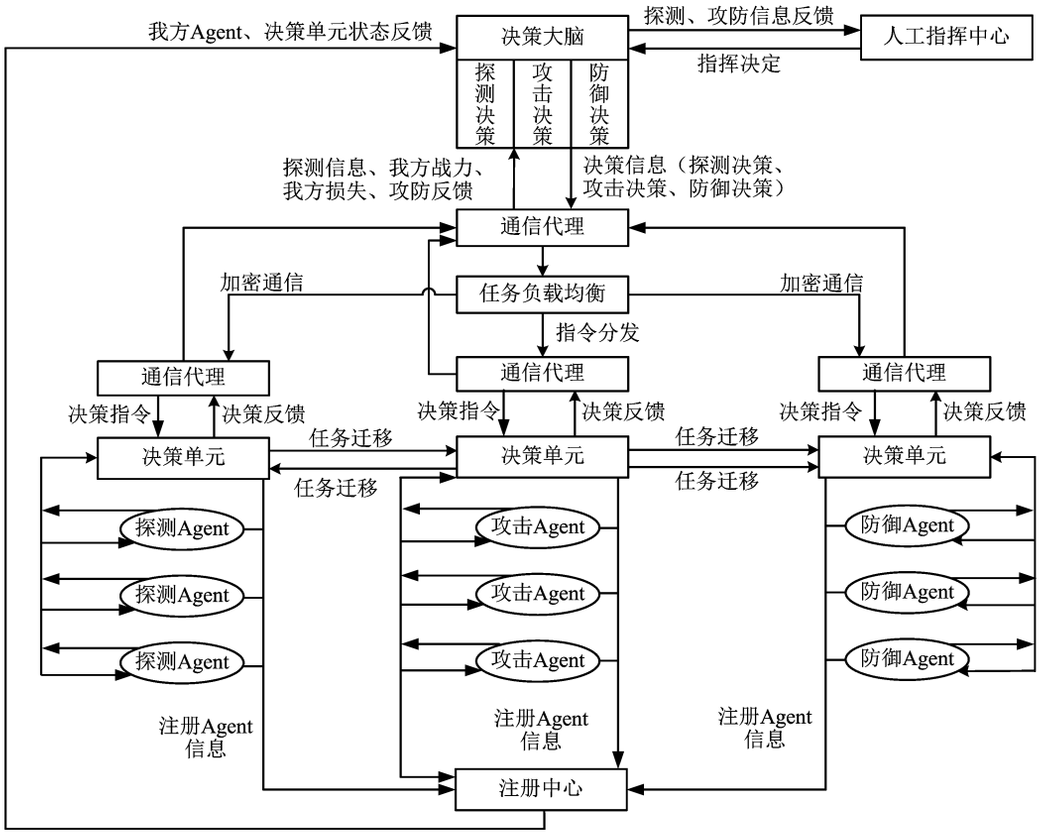

图 1　基于多智能体分层结构的作战体系系统决策 Agent 构架

Fig. 1　Combat system based on multi-intelligent body layer structure Agent framework

3　多 Agent 系统安全机制

多 Agent 系统安全机制能够从机密性、真实性、可用性 3 个方面来保证系统安全。针对多 Agent 分层系统架构，提出了一种基于数字签名和非对称加密算法的多 Agent 系统安全通信加密方法，能够在多 Agent 系统通信双方进行通信时保证数据传输的机密性。同时，通过发送方使用自己的私钥对摘要进行加密，生成数字签名来保证多 Agent 系统通信双方身份的真实性。该方法还引入了消息队列单元来保证多 Agent 系统安全通信时能够顶住增长的访问压力，使得关键组件不会因为超出负荷的请求而完全崩溃，进一步加强了多 Agent 系统在安全层面上的可用性。

对比 Chae 等[11]将认证证书和通信密钥组合在一起提出的 Agent 相互认证方法以及 Naresh 等[23]基于 Quantum Diffie-Hellman 协议提出 Quantum GKA(QGKA)技术来保证多 Agent 系统之间共享密钥安全性方法，本文提出的基于数字签名和非对称加密算法的多 Agent 系统安全通信加密方法能够实现对 Agent 身份的互相认证，具有其自身的优势和适用场景，具体过程为对于要传输的消息原文使用消息摘要算法 MD5(message digest algorithm5)、安全散列算法 SHA256(secure hash algorithm)生成消息摘要，发送方使用自己的私钥对摘要进行加密，生成数字签名。同时发送方对于要传输的消息原文使用非对称加密算法进行加密，发送方通过事先获取接收方公开的公钥信息，使用接收方的公钥信息对消息原文进行加密为密文。本文所提出的 Agent 认证方法与其他方法的具体对比情况见表 2 所列，在攻防对抗环境下，多 Agent 系统中有些轻量级 Agent 节点本身的算力不够，并不支持非对称加密。因此在需要认证轻量级 Agent 节点时，可以用重量级 Agent 节点对其进行认证，重量级节点经过长期观察，根据长期行为模式和观察结果形成历史数据，存储到区块链中，防止数

据被篡改，保证日后认证可追溯。

表 2 3 种方法对比结果

Tab. 2 Comparison results of the three methods

对比属性	本文方法	Chae 方法	Naresh 方法
适用场景	攻防场景	稳定场景	流量攻击场景
所需信息	密钥	证书、密钥	共享密钥
资源消耗	消耗少	消耗多	消耗多
安全性	不易被窃听	易被窃听	不易被窃听

为了保证 Agent 之间、Agent 与决策单元之间、决策单元与决策大脑之间以及决策大脑之间的通信安全，首先，需要保证数据安全，因此不能直接采用明文进行数据传输；同时，通信双方需经过权威认证，防止其通信过程遭受劫持、监听、篡改、伪造通信数据信息等攻击行为。图 2 为基于多智能体分层结构的系统通信方式。

如图 2 所示，对于要传输的消息原文使用消息摘要算法（MD5）、安全散列算法（SHA256）生成消息摘要，发送方使用自己的私钥对摘要进行加密，生成数字签名。同时发送方对于要传输的消息原文使用非对称加密算法进行加密，发送方通过事先获取接收方公开的公钥信息，使用接收方的公钥信息将消息原文加密为密文。只有持有对应私钥的接收方才能够使用私钥对密文信息进行解密，还原为明文信息。为了保证通信双方的信息完整，同时避免消息量较大时系统中的消息出现丢失现象，图 2 中的通信系统还引入了消息队列单元以保证消息传输的剩余力和安全性。

消息队列的中间插入了一个隐含的、基于数据的接口层，通信双方的处理过程都要实现这一接口，该接口允许独立地修改或扩展通信双方的处理过程。消息队列的持久化、削峰、解耦、异步的特性加强了系统的可用性。由于消息队列能解耦数据的处理过程，所以能够更加简单地增加消息入队和被处理的速度。同时不需要额外调整参数以及更改代码，只需简单地增加处理过程即可实现。

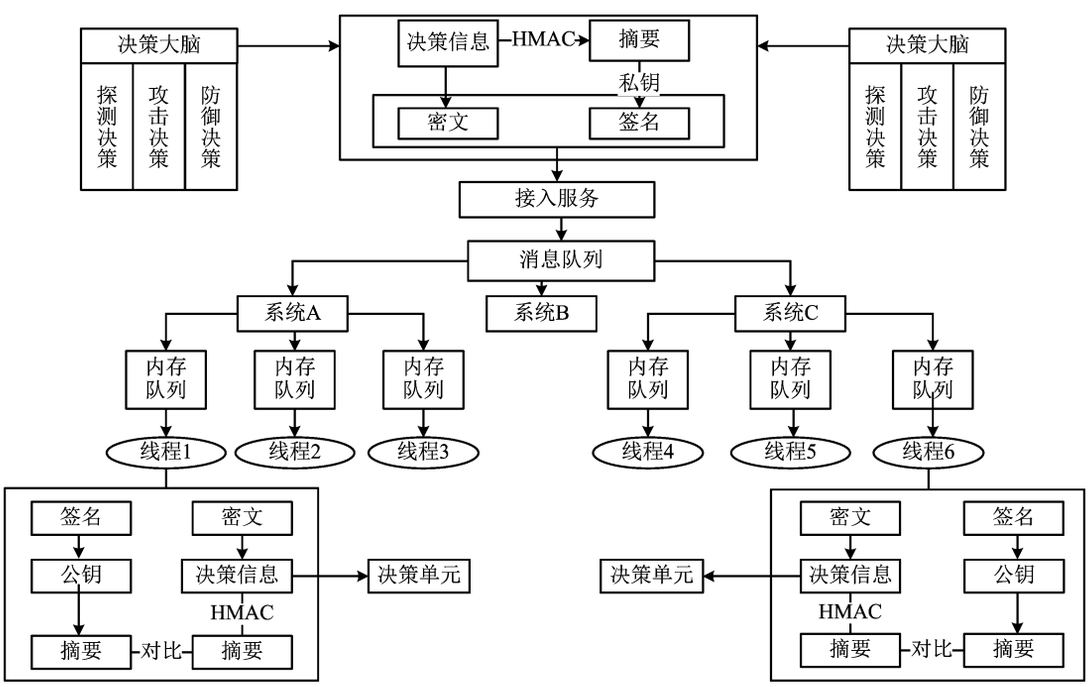

图 2 基于多智能体分层结构的作战体系系统通信方式

Fig. 2 Communication mode of combat system based on multi-agent layered structure

使用消息队列能够使关键组件顶住增长的访问压力，不会因为超出负荷的请求而完全崩溃。当体系部分组件失效时，也不会影响到整个系统的处理过程。大多数情况下，处理 Agent 数据的顺序很重要，因此对消息队列进行排序可以保证数据按照特定的顺序被处理。消息队列还提供异步处理机制，允许 Agent 将消息放入队列等待处理。

数字签名与消息一同传输给接收方，接收方使用发送方的公钥对数字签名解密还原摘要，对得到的解密后的原文进行哈希消息认证码（Hash-based message authentication code，

HMAC)计算，得出消息摘要并比较，可以保证消息的剩余力和可用性。

发送方用自身私钥生成数字签名，接收方用发送方公钥对数字签名进行解密，该过程可以验证消息是否由真实的发送方发出。根据两份消息摘要的比对结果，判断消息在传输的过程中是否被第三方恶意篡改，从而实现了 Agent 之间、Agent 与决策单元、决策单元之间、决策单元与决策大脑之间的通信加密的传输过程。

4　多 Agent 系统体系评估

多类型复杂战场的情况下，对多 Agent 系统体系评估需要从探测模块、作战模块、决策模块这 3 个核心组成模块进行评估分析，多 Agent 系统体系评估组成框图如图 3 所示。

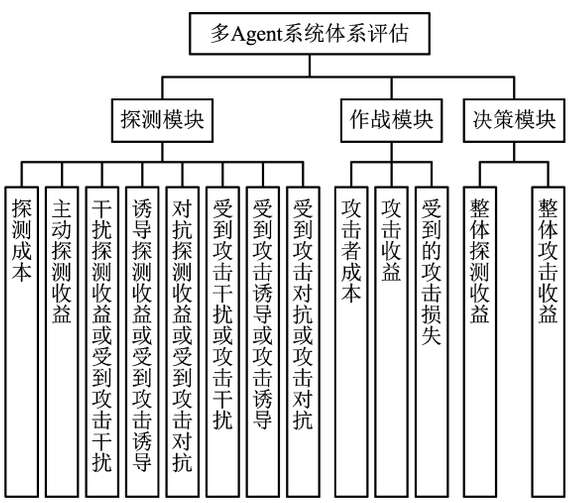

图 3　多 Agent 系统体系评估组成框图

Fig. 1　Evaluation formation box diagram of multi-Agent system system

4.1　探测模块

在海、陆、空、宇宙空间等实体战场，网络空间战场、信息通信战场等非实体战场多类型复杂战场的情况下，探测模块的信息感知功能可归结为战场信息获取、精确信息控制和战场空间理解。战场信息获取是指信息感知系统要具有能够及时、充分、准确地提供敌、我、友各方的状态、行动等战场态势信息的能力；精确信息控制是指信息感知系统要具备能够对所有感知系统中的作战资源进行准确调度和整合的能力；战场空间理解是指参战人员要具备对战场感知的信息、态势、知识等具有一致性理解能力。同时，探测模块也应具有欺骗、干扰、诱导、对抗的能力。

在防御者探测攻击者 Agent 过程中，设防御者探测攻击者成本为 $c_p(A_\varepsilon)$，防御者主动探测产生的收益为 $g_p(A_\varepsilon)$，干扰攻击者的探测收益为 $d_p(A_\varepsilon)$，诱导攻击者的探测收益为 $i_p(A_\varepsilon)$，对抗攻击者的探测收益为 $a_p(A_\varepsilon)$，则防御者简单探测收益(不考虑敌方探测干扰对我方造成损失所产生的探测收益)为：

$$G_p(A_\varepsilon)=c_p(A_\varepsilon)+g_p(A_\varepsilon)+d_p(A_\varepsilon)+i_p(A_\varepsilon)+a_p(A_\varepsilon) \tag{1}$$

对于攻击者，其探测成本为 $c_p(A_u)$，探测收益为 $g_p(A_u)$，受到防御者的攻击干扰为 $d_p(A_u)$，攻击诱导为 $i_p(A_u)$，攻击对抗为 $a_p(A_u)$。同时攻击者对防御者进行攻击干扰为 $d_p(A_\varepsilon)$，攻击诱导为 $i_p(A_\varepsilon)$，攻击对抗为 $a_p(A_\varepsilon)$，所以攻击者综合探测收益(考虑敌方探测干扰对我方造成损失所产生的探测收益)为：

$$G_p(A_u)=G_{p_1}(A_u)+G_{p_2}(A_u) \tag{2}$$

式中，$G_{p_1}(A_u)$为攻击者自主探测的探测收益，$G_{p_2}(A_u)$为防御者探测干扰、诱导、对抗为攻击者造成的损失，攻击者综合探测收益为：

$$G_p(A_u)=G_{p_1}(A_u)+G_{p_2}(A_u)=c_p(A_u)+g_p(A_u)-d_p(A_\varepsilon)-i_p(A_\varepsilon)-a_p(A_\varepsilon)+d_p(A_u)+i_p(A_u)+a_p(A_u) \tag{3}$$

防御者受到攻击者的攻击干扰为 $d_p(A_u)$，攻击诱导为 $i_p(A_u)$，攻击对抗为 $a_p(A_u)$，则防御者综合探测收益为：

$$G_p(A_\varepsilon)=G_{p_1}(A_\varepsilon)+G_{p_2}(A_\varepsilon) \tag{4}$$

式中，$G_{p_1}(A_\varepsilon)$为攻击者自主探测的探测收益，$G_{p_2}(A_\varepsilon)$为防御者探测干扰、诱导、对抗为攻击者带来的损失，防御者综合探测收益可扩展为：

$$G_p(A_\varepsilon)=G_{p_1}(A_\varepsilon)+G_{p_2}(A_\varepsilon)=c_p(A_\varepsilon)+g_p(A_\varepsilon)+d_p(A_\varepsilon)+i_p(A_\varepsilon)+a_p(A_\varepsilon)-d_p(A_u)-i_p(A_u)-a_p(A_u) \tag{5}$$

4.2　作战模块

通过信息感知和指挥控制获得的信息优势和决策优势，最终需通过火力优势体现出战场环境下的作战效果。因此对火力打击仿真建模重点在于火力打击网络的建模。

结合探测 Agent 探测到的结果以及敌、我、友多方战斗力量等信息，用仿真模型描述火力打击网络，形成火力打击综合集成力量，对战场火力打击策略制定至关重要，尤其可以帮助制定各

类火力打击系统节点的火力打击策略。例如，舰载防空导弹火力单元遂行火力打击任务可以抽象地描述为对目标的感知、跟踪、识别、发射决策、发射、引导拦截导弹飞向目标、爆破杀伤目标、进行毁伤效果判断等行动过程，如何制定这些行为的顺序、目标、策略是作战模块需要解决的主要问题。

同时，在现代战场上，信息对抗 Agent 对战场形势至关重要。其主要形式有物理摧毁与反摧毁对抗、电子对抗、网络对抗等。在信息化战争系统对抗条件下，电子对抗 Agent 中的电子侦察与截获模块、分析识别模块、电子对抗决策模块、实施干扰对抗模块互相配合展开信息对抗；学习模块用于给电子对抗目标信号的分析识别和电子对抗方案决策的知识库增加知识规则；数据库、模型库、知识库作为对目标信号进行分析识别和判断依据。

综合上述信息，将火力打击和信息对抗统一为攻击者 Agent，在防御者攻击过程中，防御者对攻击者进行攻击的成本为 $c_a(A_\varepsilon)$，将防御者攻击的收益表示为 $g_a(A_\varepsilon)$，则防御者简单攻击收益（不考虑敌方攻击干扰对我方造成损失所产生的攻击收益）为：

$$G_a(A_\varepsilon)=c_a(A_\varepsilon)+g_a(A_\varepsilon) \tag{6}$$

对于攻击者，其攻击成本为 $c_a(A_u)$，攻击收益为 $g_a(A_u)$，受到防御者的攻击造成损失为 $g_a(A_\varepsilon)$，所以攻击者综合攻击收益（考虑敌方攻击干扰对我方造成损失所产生的攻击收益）为：

$$G_a(A_u)=c_a(A_u)+g_a(A_u)-g_a(A_\varepsilon) \tag{7}$$

对于防御者，其受到攻击者的攻击为 $g_a(A_u)$，则防御者综合攻击收益为：

$$G_a(A_\varepsilon)=c_a(A_\varepsilon)+g_a(A_\varepsilon)-g_a(A_u) \tag{8}$$

4.3 决策模块

在决策模块中，指挥控制是将信息优势转化为决策优势的核心环节，是指挥人员和指挥控制系统共同作用来达到决策优势获取、发挥和转化的过程。指挥人员可以在指挥控制系统的辅助下完成对态势的评估、预测和作战计划的制定，同时指挥控制系统可以对某些特定的作战场景进行自动化智能指挥控制。决策系统同时可以帮助指挥人员对所属部队行使指挥、发布命令、监督执行和监督完成任务等功能。决策模块的目的是指挥控制探测 Agent 和作战 Agent 行动，来获取防御者最大的收益。

对于防御者来说，防御者的决策目的应为最大化探测收益与攻击收益的和，防御者探测—攻击整体收益可以表示为：

$$\begin{aligned}G_o(A_\varepsilon)=&G_p(A_\varepsilon)+G_a(A_\varepsilon)=\\&c_p(A_\varepsilon)+g_p(A_\varepsilon)+d_p(A_\varepsilon)+\\&i_p(A_\varepsilon)+a_p(A_\varepsilon)-d_p(A_u)-i_p(A_u)-\\&a_p(A_u)+c_a(A_\varepsilon)+g_a(A_\varepsilon)-g_a(A_u)\end{aligned} \tag{9}$$

即为探测收益和攻击收益的和。

5 算法优化模型与仿真

5.1 问题描述

在敌我双方多 Agent 对抗攻防过程中，模拟敌我双方都拥有攻击 Agent、防御 Agent、探测 Agent，双方指挥各自的 Agent 开展对敌方的攻击，并在己方损失最小的情况下摧毁敌方的武装力量。

在每一轮攻击中，防御者的防御 Agent 可以对防御者多个攻击、防御和探测 Agent 进行防御，以抵抗攻击者的攻击，同时攻击 Agent 可以对攻击者多个 Agent 进行攻击，如果防御者攻击力大于攻击者在该 Agent 上的防御力，那么可以对攻击者造成攻击力减防御力的损失。同理，在攻击者攻击轮次中，攻击者可以对防御者进行相同形式的打击。

5.2 多 Agent 系统贪婪算法优化模型

5.2.1 算法设计

在与攻击者的攻防博弈中，为了对抗攻击者，防御者可设计一种贪婪攻击策略，使用朴素攻击策略发动攻击，其核心部分是一种启发式的贪婪算法选出 A'_ε（即攻击的目标 Agent 集）。攻击者计算 $|R'_\varepsilon|/|R_\varepsilon|$，量化对攻击者 Agent($A$) 造成的攻击严重程度，即对方被攻击的 Agent 数量，$|R'_\varepsilon|/|R_\varepsilon|$ 越大，则攻击者受到的损失越大。通过引入启发式贪婪算法能够优化原有的朴素攻击策略，使得防御者能够在每轮次和攻击者的攻防博弈中，选择当前情况下能够造成最严重攻击程度的 Agent 进行攻击，从而提高防御者在攻防博弈中获胜的概率，并且有效缩短攻防博弈的轮次。

在朴素攻击策略中，攻击者反复地在每次循环中找到 $a\in A_\varepsilon\backslash A'_\varepsilon$，以最大化选择每一个 Agent 的收益，$\delta_a(A'_\varepsilon)=|A'_\varepsilon(A'_\varepsilon\cup\{a\})|/|A_\varepsilon|-|A'_\varepsilon|-|A'_\varepsilon(A'_\varepsilon)|/|A'_\varepsilon|$，然后把 a 插入 A'_ε，将

$A'_\epsilon(\cdot)$定义为计算因为攻击了一个 Agent 集合而导致攻击者受到损失的函数。注意到每个循环在挑选出拥有最大收益的 Agent 后，对数据集中的每个 a 总是重新计算 $A'_\epsilon(a)$。这个循环挑选的过程一直进行直到 $|A'_\epsilon|/|A_\epsilon|$ 达到预先设定的目标 Δ。

基于贪婪策略的多 Agent 优化算法(Multi-agent optimization algorithm based on greedy strategy)如下所示：

Input:$R_\epsilon, A_\epsilon, \Delta, \phi$
Output:A'_ϵ
$A'_\epsilon=\phi, R'_\epsilon=\phi$
for $a\in A_\epsilon$ do
　　Calculate $R'_\epsilon(a)$
end
while $|R'_\epsilon|/|R_\epsilon|<\Delta$ do
　　$a=\mathrm{argmax}_{a\in A_\epsilon/A'_\epsilon}\delta_a(A'_\epsilon)$
　　if $|A'_\epsilon|=\phi$ then
　　　break
　　end
　　$A'_\epsilon=A'_\epsilon\cup a$
　　$R'_\epsilon=R'_\epsilon\cup R'_\epsilon(a)$
　　for $a\in A_\epsilon$ do
　　　Calculate $R'_\epsilon(a)$
　　end
end

5.2.2　实验仿真

实验仿真环境为 Matlab R2021a，在 Matlab 环境中创建 Agent 类来模拟多 Agent 系统中 Agent 在敌我双方开始攻防之前，首先创建 Agent 类表示敌我双方的攻击、防御和探测 Agent，Agent 类包含属性见表 3 所列。

表 3　Agent 类型的属性

Tab. 3　Attribute table of Agent type

属性	类型	备注
序号	实数	Agent ID
类型	字符串	Agent 类型(探测、攻击、防御)
状态	字符串	状态(active、idle、damaged)
能力	实数	攻击力、探测力，防御 Agent 值为 0
剩余力	实数	再受多少攻击会被摧毁
防御力	实数	防御力(攻击、探测 Agent 值为 0)
重要性	实数	重要性(战略重要性)
角色	字符串	所属阵营(攻击者、防御者)

然后，为敌我双方创建各自的攻击、防御、探测 Agent，在本次仿真攻防过程中，分别为敌我双方初始化了 18 个攻击 Agent、18 个防御 Agent 和 18 个探测 Agent，并分别对敌我双方的攻击 Agent 设置 1～10 随机的攻击力，1～10 随机的剩余力，0 的防御力；为防御 Agent 设置 1～10 随机的防御力，1～10 随机的剩余力，0 的攻击力；为探测 Agent 设置 1～10 随机的探测力，1～10 随机的剩余力，0 的防御力。

攻防双方开始攻防，双方分别根据贪婪策略选中要攻击的 Agent(包括攻击 Agent、防御 Agent 和探测 Agent)，分别发动攻击，每轮攻击过程中双方都可发动攻击。依照以上过程，可以模拟敌我双方攻防过程如图 4 所示。图 4 中不同类型的 Agent 下方显示了该类型的 Agent 依次被攻击的 ID 顺序。根据仿真结果可以发现，攻防双方共进行了 6 轮攻防工程，并在第 6 轮攻防过程中，防御者摧毁了攻击者所有的攻击 Agent，攻击者丧失攻击力，防御者获胜。

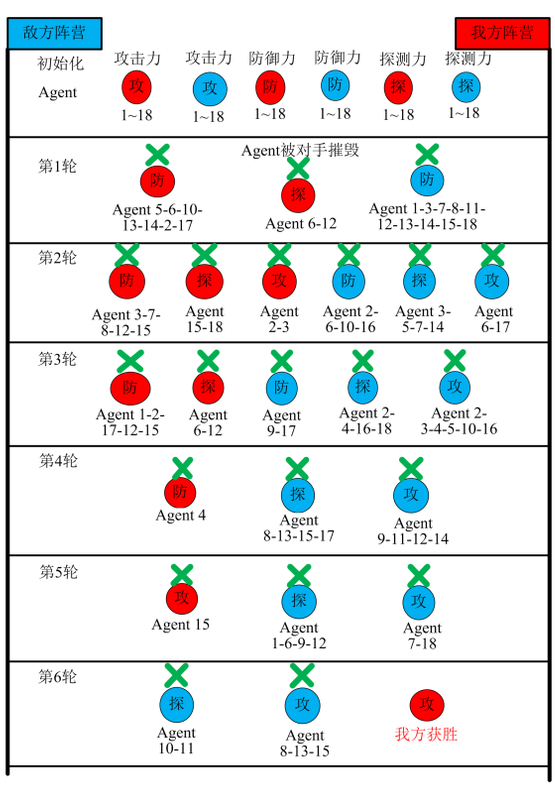

图 4　基于贪婪算法的敌我双方攻防过程结果

Fig. 3　Results of the offensive and defensive process of the enemy and us based on the greedy algorithm

5.3　多 Agent 系统遗传算法优化模型

5.3.1　算法设计

优化问题描述同 5.1 节。对于防御者来说，攻击最佳收益应为在攻击消耗最少的情况下，对

攻击者造成最大的伤害，即在攻击最少轮次的情况下，对攻击者造成最大程度的损伤，由此本文定义了攻击绩效来衡量对攻击者的攻击效率，首先给出攻击一个 Agent 的攻击绩效：

$$S_{E_a}=G_{\text{overall}}(a) \tag{11}$$

防御者攻击的目的是在每一轮的攻击中取得最大的攻击绩效。为了进一步提高朴素策略可以达到的攻击绩效，设计了一种核心是改良遗传算法的攻击策略来选出要攻击的 Agent(A'_ϵ)。通过引入改良遗传算法来优化朴素策略的攻击效果，基于改良的遗传算法，防御者在攻防博弈的过程中会逐渐演化到选择具有更大攻击绩效的 Agent 方向上，提高防御者在每轮进行攻防博弈选择具有更大攻击绩效的 Agent 的概率，从而提高防御者在攻防博弈中获胜的概率以及有效缩短攻防博弈的轮次。具体来说，改良遗传算法策略初始化 n 个不同基因作为第 1 代种群。一个基因是指对应一个目标 Agent 集合的一个二进制向量，这里用 gene2set 表示将二进制向量转换为目标 Agent 集合的函数，用 set2gene 表示将目标 Agent 集合转换为二进制向量的函数。所有基因都是基于 A_ϵ 制造出来的，不同基因之间的不同点在于各自选出的目标 Agent 子集(A'_ϵ)不同。

在第 1 代种群中，所有的基因中有一个基因是根据朴素 Agent 选取策略选择出来的 A'_ϵ 制造出来的。其他的基因的选择都是根据随机生成的 A'_ϵ 制造。这里用 g_1 表示第 1 代种群，其后不断进化来的每一代种群序列用 $g_1,g_2,g_3,\cdots$ 表示。基于遗传算法的多 Agent 优化算法见表 5 所列。在种群的进化过程中，分别有选择、交叉和变异长期在种群中存在，采用了轮盘赌算法来选择基因。在每一次执行轮盘赌算法过程中，本文概率性地在每一代种群的基因中选择一个基因，在轮盘赌算法进行的过程中，有更大攻击绩效 $S_E(g)$ 的基因，应该被赋予更大的选择概率。

5.3.2 实验仿真

实验仿真环境为 Matlab R2021a，在 Matlab 环境中创建 Agent 类来模拟多 Agent 系统中的 Agent 实体。在敌我双方开始攻防之前，首先创建 Agent 类属性。

基于遗传策略的多 Agent 优化算法(multi-agent optimization algorithm based on genetic strategy)如下所示：

Input:$R_\epsilon,A_\epsilon,n,A'_\epsilon$
Output:$A_\epsilon{}'$
$G_1=U_{(i=1,2,\cdots,n)}\text{set2gene}(X_i)|X_i=$
  randomagent$\cup$set2gene(A'_ϵ)
$S_{E_1}\max=\max S_E(g)|g\in G_1$
$j=1$
while $S_{E_j}\max\leqslant S_{E_1}\max$ do
  $j++$
  $G_j=\phi$
  while $|G_j|<n$ do
  select g_k and g_q over $\{S_{SSE_{(j-1)(z)}}\in G_{(j-1)}\}$
    $G_j=G_j\cup\{g_k,g_q\}$
    generate a random float $r,0<r<1$
    if $r<P_c(g_k,g_q)$ then
      $g'_k,g'_q=\text{crossover}(g_k,g_q)$
      $G_j=G_j\cup\{g'_k,g'_q\}$
    end
  end
  for $g\in G_j$ do
    if $r<P_m(g)$ then
      $g_m=\text{mutation}(g)$
      $G_j=G_j\backslash\{g\cup g_m\}$
    end
  end
  $S_{E_j}\max=\max\{S_E(g)|g\in G_j\}$
  if $S_{E_j}\max\leqslant S_{E_1}\max$ then
    $G_j=G_1$
  end
end

攻防双方开始攻防，双方分别攻防博弈过程中，防御者 Agent 将会有更高的胜率。根据表示敌我双方的攻击、防御和探测 Agent，Agent 类遗传策略选中要攻击的 Agent(包括攻击 Agent、防御 Agent 和探测 Agent)，遗传策略不同于贪婪策略，每轮攻击都要尽可能打击攻击者更多、更重要的 Agent，来尽快摧毁攻击者的战斗力，敌我双方分别根据遗传策略，分别发动攻击，每轮攻击过程中双方都可发动攻击。

依照以上过程，可以模拟敌我双方攻防过程如图 5 所示。图 5 中不同类型的 Agent 下方显示了该类型的 Agent 依次被攻击的 ID 顺序。可以发现，当敌我双方都使用遗传算法来优化攻防过程时，双方可在每轮攻击中最大化攻击的效率，从而在更少的轮次实现制敌取胜。依照仿真结果可以发现，攻防双方共进行了 4 轮攻防过程，并

在第 4 轮攻防过程中，防御者摧毁了攻击者所有的攻击 Agent，攻击者丧失攻击力，防御者获胜。

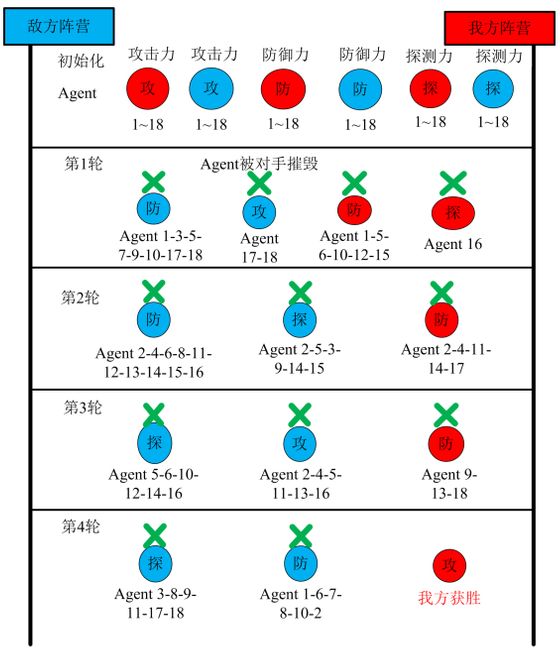

图 5　基于遗传算法的敌我双方攻防过程结果

Fig. 5　Results of the offensive and defensive process of the enemy and us based on the genetic algorithm

通过实验仿真结果，可以得出引入遗传策略，能够使得防御者在攻防博弈的过程中会逐渐演化到选择具有更大攻击绩效的 Agent 方向上，提高防御者在每轮攻防博弈时选择具有更大攻击绩效的 Agent 的概率，从而提高防御者在攻防博弈中获胜的概率以及减少攻防博弈的轮次。

6　结束语

本文提出一种多 Agent 系统安全的保护方式。对多 Agent 系统中不同 Agent 间的通信采取数据加密措施，利用公、私钥机制，实现了不同 Agent 间的身份认证以及安全通信机制，有效地保证了通信这一重要环节的系统安全，杜绝 Agent 代理身份冒用在上述基础上对多 Agent 体系进行系统的评估，并利用贪婪算法和遗传算法对多 Agent 系统进行模型建立、优化安全策略以及实验仿真。利用贪婪算法和遗传算法对多 Agent 系统的安全策略优化从而有效保证了通信环节的防御安全，防止 Agent 代理被攻击窃取相关信息。

参 考 文 献

[1] 沙基昌，毛赤龙，石建迈，等. 对抗性复杂系统的研究框架[J]. 科学技术与工程，2009，9(4)：815-820.
SHA Jichang，MAO Chilong，SHI Jianmai，et al. Framework of researching on antagonistic complex systems[J]. Science Technology and Engineering，2009，9(4)：815-820.(in Chinese)

[2] 刘晓平，唐益明，郑利平. 复杂系统与复杂系统仿真研究综述[J]. 系统仿真学报，2008，20(23)：6303-6315.
LIU Xiaoping，TANG Yiming，ZHENG Liping. Survey of complex system and complex system simulation[J]. Journal of System Simulation，2008，20(23)：6303-6315.(in Chinese)

[3] VINNAKOTA T. A conceptual framework for complex system design and design management[C]//Proceedings of 2016 Annual IEEE Systems Conference (SysCon). [S. l.]：IEEE，2016：1-6.

[4] 王亚康，郭晶，江汀，等. 基于 Agent 的复杂系统建模与仿真研究[J]. 电子设计工程，2011，19(9)：100-103.
WANG Yakang，GUO Jing，JIANG Ting，et al. Research on Agent-based modeling and simulation of complex system[J]. Electronic Design Engineering，2011，19(9)：100-103.(in Chinese)

[5] DORRI A，KANHERE，JURDAK R. Multi-agent systems：a survey[J]. IEEE Access，2018(6)：28573-28593.

[6] SARIKA S，PAUL V. AgentTab：an agent based approach to detect tabnabbing attack[J]. Procedia Computer Science，2015，46：574-581.

[7] MECHTRI L，TOLBA F D，GHANEMI S. MASID：multi-agent system for intrusion detection in MANET[C]//Proceedings of the 9 International Conference on Information Technology：New Generations. [S. l.]：IEEE，2012：65-70.

[8] GORODETSKI V，KOTENKO I. The multi-agent systems for computer network security assurance：frameworks and case studies[C]//Proceedings of 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002). [S. l.]：IEEE，2002：297-302.

[9] 张少苹，戴锋，王成志，等. 多 Agent 系统研究综述[J]. 复杂系统与复杂性科学，2011，8(4)：1-8.
ZHANG Shaoping，DAI Feng，WANG Chengzhi，et al. Summary on research of multi-agent system[J]. Complex Systems and Complexity Science，2011，8(4)：1-8.(in Chinese)

[10] SAEED I A，SELAMAT A，ROHANI M F，et al. A systematic state-of-the-art analysis of multi-agent intrusion detection [J]. IEEE Access，2020，8：180184-180209.

[11] CHAE C J，CHOI K N，CHOI K. Information interoperability system using multi-agent with security[J].

Wireless Personal Communications, 2016, 89(3): 819-832.

[12] CENTENO R, FAGUNDES M, BILLHARDT H, et al. Supporting medical emergencies by mas[C]//Proceedings of KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications. Germany, Berlin: Springer, 2009: 823-833.

[13] GREENBERG M S, BYINGTON J C, HARPER D G. Mobile agents and security[J]. IEEE Communications Magazine, 1998, 36(7): 76-85.

[14] WANG S, HU J, LIU A, et al. Security frame and evaluation in mobile agent system[J]. IEE Mobility Conference, 2005: 15-17.

[15] JIN X, LV S, DENG C, et al. Distributed adaptive security consensus control for a class of multi-agent systems under network decay and intermittent attacks [J]. Information Sciences, 2021, 547: 88-102.

[16] 丁俐夫,颜钢锋.多智能体系统安全性问题及防御机制综述[J].智能系统学报,2020, 15(3): 425-434.
DING Lifu, YAN Gangfeng. A survey of the security issues and defense mechanisms of multi-agent systems [J]. CAAI Transactions on Intelligent Systems, 2020, 15(3): 425-434.(in Chinese)

[17] BASHEER G S, AHMAD M S, TANG A Y C, et al. Certainty, trust and evidence: towards an integrative model of confidence in multi-agent systems[J]. Computers in Human Behavior, 2015, 45: 307-315.

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

[19] LI X M, ZHOU Q, LI P, et al. Event-triggered consensus control for multi-agent systems against false data-injection attacks[J]. IEEE Transactions on Cybernetics, 2019, 50(5): 1856-1866.

[20] WANG S, ZHAO C, ZHANG B, et al. Event-triggered based security consensus control for multi-agent systems with DoS attacks [J]. Neurocomputing, 2022, 505: 214-224.

[21] SETHI K, MADHAV Y V, KUMAR R, et al. Attention based multi-agent intrusion detection systems using reinforcement learning[J]. Journal of Information Security and Applications, 2021, 61: 2214-2126.

[22] NARESH V S, NASRALLA M M, REDDI S, et al. Quantum diffie-hellman extended to dynamic quantum group key agreement for e-healthcare multi-agent systems in smart cities[J]. Sensors, 2020, 20(14): 3940.

[23] HUYNH T D, JENNINGS N R, SHADBOLT N R. An integrated trust and reputation model for open multi-agent systems [J]. Autonomous Agents and Multi-Agent Systems, 2006, 13(2): 119-154.

[24] ZIKRATOV I, MASLENNIKOV O, LEBEDEV I, et al. Dynamic trust management framework for robotic multi-agent systems[M]//SERGEY B, SERGEY A, YEVGENI K. Internet of things, smart spaces, and next generation networks and systems. [S. l.]: Springer, Cham, 2016: 339-348.

[25] MAJD E, BALAKRISHNAN V. A trust model for recommender agent systems [J]. Soft Computing, 2017, 21(2): 417-433.

[26] JUBAIR M A, MOSTAFA S A, MUSTAPHA A, et al. A survey of multi-agent systems and case-based reasoning integration[C]//Proceedings of 2018 International Symposium on Agent, Multi-Agent Systems and Robotics(ISAMSR). [S. l.]: IEEE, 2018: 1-6.

[27] HALINKA A, RZEPKA P, SZABLICKI M. Agent model of multi-agent system for area power system protection[C]// Proceedings of 2015 Modern Electric Power Systems(MEPS). [S. l.]: IEEE, 2015: 1-4.

[28] ZHANG J, TAN R, SU C, et al. Design and application of a personal credit information sharing platform based on consortium blockchain[J]. Journal of Information Security and Applications, 2020, 55: 2214-2126.

责任编辑 董 莉