УДК 004.891

КЛЮЧЕВЫЕ ОБЛАСТИ ВНИМАНИЯ НА СТЫКЕ МНОГОАГЕНТНЫХ СИСТЕМ И КИБЕРБЕЗОПАСНОСТИ

И.В. Котенко

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук
Россия, 199178, Санкт-Петербург, 14-я линия, 39
E-mail: ivkote@comsec.spb.ru

Ключевые слова: многоагентная система, кибербезопасность, атака, защита от атак, атака на отказ в обслуживании, обманная атака, атака на агент.

Аннотация: Многоагентные системы (МАС) становятся все более популярными благодаря их успешному внедрению в ряде отраслей. Однако существует множество угроз, которые могут подорвать правильное выполнение МАС своих целевых функций. Обмен информацией, расширение возможностей по взаимодействию агентов и особенности функционирования отдельных агентов представляют возможность злоумышленникам реализовать различные классы атак, тем самым делая проблему обеспечения кибербезопасности МАС фундаментально важной. Кроме исследования кибератак и защиты от кибератак в МАС, критически важными являются такие области пересечения исследований и разработок в области МАС и кибербезопасности как использование МАС для кибербезопасности (как для защиты, так и реализации кибератак), кибератаки на консенсус по кооперативному управлению агентами, состязательные атаки на МАС и защита от них. Поскольку кибератаки становятся все более сложными и скоординированными, механизмы защиты, основанные на многоагентных подходах, могут стать ключом к борьбе с кибератаками. В докладе делается анализ исследований на стыке МАС и кибербезопасности. Представляются и классифицируются наиболее распространенные атаки на МАС, реализуемые на различных уровнях представления. Рассматриваются различные модели и механизмы защиты для противодействия этим атакам. Анализируются подходы как к предотвращению, так и обнаружению и обеспечению устойчивости МАС, основанные на репутации и доверии. Представляются исследования автора по использованию МАС для кибербезопасности.

1. Введение

Многоагентные системы (МАС) предоставляют эффективные средства координации пространственно распределенных и сетевых агентов (узлов, подсистем), так что совместные задачи могут быть выполнены с требуемой эффективностью [1, 2].

Фундаментальной проблемой в литературе по МАС является достижение скоординированных целей агентов в некоторой сетевой среде путем разработки оптимальных стратегий распределенного управления [3]. В последние десятилетия распределенной координации в МАС (в частности задаче устойчивого консенсуса) уделяется все больше внимания, о чем свидетельствует ее широкий спектр приложений, таких как формирование групп беспилотных беспилотных летательных аппаратов [4], управление взаимодействием роботов [5], распределенное обнаружение в беспроводных сенсорных сетях [6].

Одним из ключевых вопросов в изучении МАС является разработка протоколов и алгоритмов распределенного совместного управления и работы агентов в команде, которые зависят от локального обмена информацией в реальном времени между

взаимодействующими агентами [7, 8]. Однако совместная работа агентов и обмен информацией через сеть в МАС представляют возможность злоумышленникам для реализации различных классов атак, тем самым делая задачу обеспечения кибербезопасности МАС фундаментально важной [9].

Кроме того, учитывая тот факт, что кибератаки становятся все более сложными и скоординированными, злоумышленники осваивают новые методы, которые позволяют увеличить изощренность, скорость и масштабность реализации атак (обходя при этом традиционные механизмы защиты, благодаря так называемому наступательному искусственному интеллекту), механизмы защиты основанные на многоагентных подходах, могут стать ключом к борьбе с кибератаками [10, 11].

Исходя из данных посылок и учитывая текущие исследования в области кибербезопасности и МАС в настоящем докладе выделяются и характеризуются следующие области пересечения исследований и разработок в области МАС и кибербезопасности: (1) кибератаки на МАС; (2) защита МАС от кибератак; (3) кибератаки на консенсус по кооперативному управлению; (4) состязательные атаки на МАС и защита от них; (5) использование МАС для реализации кибератак; (6) использование МАС для кибербезопасности.

Чтобы получить базу публикаций в данных областях была использована поисковая система Google Scholar. Поиск в Google Scholar работ по различным ключевым словам дал следующие результаты: кибератаки против МАС (cyber attacks against multi-agent systems) - 32 900 статей; защита МАС от кибератак (multi-agent systems defense from cyber attacks) - 22 100 статей; кибератаки на консенсус по кооперативному управлению (cyber attacks against cooperative control consensus) - 380.000 статей; состязательные атаки на МАС и защита от них (adversarial attacks on multi-agent systems) - 17 600; использование МАС для реализации кибератак (multi-agent systems for cyber attacks) - 41.200 статей; использование МАС для кибербезопасности (multi-agent systems for cyber security) - 21.800 статей. Вполне объяснимо, что наибольшее количество работ посвящено консенсусу по кооперативному управлению, так как это одно из основных направлений исследований в области построения многоагентных систем. Поиск также показал, что работы по анализу кибератак превалируют над работами по защите от кибератак.

В последние несколько лет много усилий было посвящено разработкам методов безопасного управления в МАС и их практической реализации. Об этом свидетельствует ряд обзоров, опубликованных в литературе [12-18]. Однако, насколько известно авторам, существует недостаточно подробных обзоров современных решений по безопасному управлению МАС с учетом конкретных классификаций, характеристик и моделей атак, а также соответствующих стратегий и проблем разработки средств безопасного управления [19-20].

В докладе вначале представляются и классифицируются наиболее распространенные атаки на МАС, затем рассматриваются различные механизмы защиты от атак, а также представляются исследования автора по использованию МАС для кибербезопасности.

2. Атаки на многоагентные системы

С ростом развития искусственного интеллекта, мобильных телекоммуникаций, сенсорных сетей, граничных вычислений и робототехники использование МАС появляется во многих критически важных приложениях, включая военную, космическую, производственную, электронный бизнес, управление цепочками поставок и многие другие. Однако эти агенты действуют в неопределенной и

враждебной среде, что делает их деятельность ненадежной и небезопасной. Крайне важно обеспечить механизмы безопасности, гарантирующие такие принципы безопасности, как конфиденциальность, целостность, доступность, подотчетность и неотказуемость различных агентов и МАС в целом [12-14]. Из-за пространственного распределения агентов и требований обмена информацией между агентами в реальном времени МАС более восприимчивы к атакам, нацеленным как на физических агентов, так и на каналы связи [21]. Например, Stuxnet [22] был разработан для нарушения работы системы промышленного контроля завода по обогащению урана путем периодического увеличения и уменьшения скорости центрифуг. Вирус Havex атаковал SCADA-систему, что привело к разрушению плотин гидроэлектростанций и перегрузке атомной электростанции [16].

Одна из наиболее полных классификаций атак на МАС по принципу «черного ящика» выполнена в зависимости от вида атаки, источника, цели, частоты и воздействия на жертву [19].

Выделены такие виды атак как атаки на раскрытие информации, мутационные атаки, отказ в обслуживании, выдача себя за другое лицо. Например, атака на раскрытие информации — это атака на конфиденциальность, то есть злоумышленник получает доступ к конфиденциальной информации от агентов. Очевидный способ добиться этого — перехват конфиденциальных сообщений. Злоумышленник может получить авторизованный доступ к данным агента, таким как состояние и внутренний код агентов. Атаки на основе происхождения - раскрытие информации о маршруте мобильного агента. При зондирующих атаках злоумышленник исследует частную базу данных жертвы с конфиденциальной информацией. Одним из конкретных типов зондирующих атак является онтологическая атака - злоумышленник может получить доступ к личным локальным знаниям агента-жертвы (например, правилам и политикам принятия решений). Источник атаки — это точка уязвимости, через которую злоумышленники могут получить несанкционированный доступ к приложению МАС. Несанкционированный доступ может быть осуществлен путем компрометации одного или нескольких агентов или реализации атаки «человек посередине».

Злоумышленники могут атаковать одного агента или всех агентов в МАС. Также могут быть атакованы хост системы или промежуточное программное обеспечение, управляющее МАС. Атака может быть непрерывной в течение определенного периода времени или организованной через определенные промежутки времени. Внешнее событие, такое как изменение времени, местоположения или действие определенного агента, запускает атаку, вызванную событием, также известную как логическая бомба. Наконец, последствия атаки могут включать ухудшение производительности, утечку конфиденциальной информации, изменение конфигурации данных, агента или всей МАС, что приводит к тому, что система не ведет себя должным образом. Атака может привести к увеличению времени, необходимого агентам для достижения консенсуса, что приведет к нежелательным последствиям, включая полный отказ системы.

3. Защита многоагентных систем от атак

Методы защиты разделяются, как правило, на следующие основные классы: предотвращения, обнаружения, реагирования и обеспечения устойчивого функционирования [12-20]. Под методами предотвращения подразумеваются методы, реализуемые на основе проектных решений, исключающих возможность атак конкретных классов (шифрование, аутентификация, механизмы разграничения доступа). Методы обнаружения позволяют выявить атаку. Методы реагирования позволяют остановить атаку или смягчить ее последствия. Методы обеспечения

устойчивого функционирования предполагают мониторинг конкретных классов атак после развертывания и смягчения последствий.

Например, одним из ключевых вопросов обеспечения безопасности MAC является разработка эффективной стратегии защиты против атак на отказ в обслуживании (DoS-атак). С одной стороны, для конкретных случайных DoS-атак, когда сбои связи следуют с некоторым распределениям вероятностей, можно использовать несколько эффективных надежных стратегий управления для борьбы с ограниченным количеством атак, характеризующихся заданным ожиданием и дисперсией сбоев. С другой стороны, для распределенных DoS-атак некоторые типичные результаты [23-25] кратко рассматриваются со следующих двух точек зрения: устойчивый контроль, основанный на непрерывной связи, и устойчивый контроль, основанный на событийно-ориентированной коммуникации.

4. Использование многоагентных технологий для кибербезопасности и моделирования кибератак

В работе развивается подход к исследованию противоборства в компьютерных сетях на основе моделирования антагонистического взаимодействия команд агентов, представляющих злоумышленников и компоненты систем защиты, первоначально предложенный в [10, 11, 26]. Агенты различных команд соперничают для достижения противоположных намерений. Агенты одной команды сотрудничают для осуществления общего намерения (по реализации угрозы или по защите компьютерной сети). Выбор сценария поведения каждой из команд зависит, прежде всего, от выбранной цели функционирования, а конкретная реализация сценария определяется, в первую очередь, непосредственной реакцией противоположной команды. Выбор очередного шага поведения каждой из команд должен определяться динамически в зависимости от действий противоположной команды и состояния среды. Основным вкладом является имитационная модель, основной целью которой является создание общего фреймворка для реализации и оценки типов агентов киберзащиты и агентов кибератак в одних и тех же сетевых средах.

5. Заключение

В докладе представлены основные области внимания на стыке исследований и разработок в области MAC и кибербезопасности. Представлены наиболее распространенные атаки на MAC, рассмотрены различные механизмы защиты от атак, а также текущие исследования по использованию MAC для кибербезопасности и моделирования атак. Направления дальнейших работ связаны с исследованием механизмов защиты от различных типов атак против MAC, а также с использованием многоагентных технологий для построения перспективных систем защиты и моделирования киберпротивоборства.

Список литературы

1. Dorri A., Kanhere S. S., Jurdak R. Multi-agent systems: A survey // IEEE Access. 2018. Vol. 6. P. 28573-28593.
2. Gorodetsky V., Kotenko I., Karsayev O. The Multi-agent Technologies for Computer Network Security: Attack Simulation, Intrusion Detection and Intrusion Detection Learning // The International Journal of Computer Systems Science & Engineering, 2003, No.4. P. 191-200.

3. He W., Qian F., Lam J., Chen G., Han Q.-L., Kurths J. Quasi-synchronization of heterogeneous dynamic networks via distributed impulsive control: Error estimation, optimization and design // Automatica. 2015. Vol. 62. P. 249-262.
4. Yu X., Zhang Y. Sense and avoid technologies with applications to unmanned aircraft systems: Review and prospects // Progress in Aerospace Sciences. 2015. Vol. 74. P. 152-166.
5. Hausman K., Mueller J., Hariharan A., Ayanian N., Sukhatme G.S. Cooperative multi-robot control for target tracking with onboard sensing // The International Journal of Robotics Research. 2015. Vol.34, No.13. 2015, P. 1660-1677.
6. Ge X., Han Q.-L., Zhong M., Zhang X.-M. Distributed Krein space- based attack detection over sensor networks under deception attacks // Automatica. 2019. Vol. 109. P. 108557.
7. Gorodetsky V., Kotenko I. Scenarios Knowledge base: A Formal Framework for Proactive Coordination of Coalition Operations // Knowledge Systems for Coalition Operation / M. Pechoucek, A.Tate (Eds.). Third International Conference on Knowledge Systems for Coalition Operations (KSCO-2004). Pensacola, Florida. 2004. P. 83-97.
8. Kotenko I., Konovalov A., Shorov A. Agent-based simulation of cooperative defense against botnets // Concurrency and Computation: Practice and Experience. 2012. Vol. 24, No. 6. P. 573-588.
9. Leszczyna R., Kotenko I. Security and anonymity of agent systems // NATO Science for Peace and Security Series, Software Agents, Agent Systems and Their Applications. 2012 / Edited by M. Essaaidi, M. Ganzha, M. Paprzycki. Amsterdam, Netherlands: IOS Press. 2012. Vol. 32. P. 157-177.
10. Коновалов А.М., Котенко И.В., Шоров А.В. Исследование бот-сетей и механизмов защиты от них на основе имитационного моделирования // Известия РАН. Теория и системы управления. 2013. № 1. С. 45-68.
11. Котенко И.В. Интеллектуальные механизмы управления кибербезопасностью // Управление рисками и безопасностью. Труды Института системного анализа Российской академии наук (ИСА РАН). 2009. С. 74-103.
12. Cavalcante R.C., Bittencourt I.I., Da Silva A.P., Silva M., Costa E., Santos R. A survey of security in multi-agent systems // Expert Systems with Applications. 2012. Vol. 39, No. 5, P. 4835-4846.
13. Cetinkaya A., Ishii H., Hayakawa T. An overview on denial-of-service attacks in control systems: Attack models and security analyses // Entropy. 2019. Vol. 21, No. 2. P. 210:1-210.
14. Dibaji S., Pirani M., Flamholz D., Annaswamy A., Johansson K., Chakrabortty A. A systems and control perspective of CPS security // Annual Reviews in Control. 2019. Vol. 47. P. 394-411.
15. Ding D., Han Q.-L., Xiang Y., Ge X., Zhang X.-M. A survey on security control and attack detection for industrial cyber-physical systems // Neurocomputing. 2018. Vol. 275. P. 1674-1683.
16. Ding D., Han Q.-L., Xiang Y., Ge X., Wang J. Secure state estimation and control of cyber-physical systems: A survey // IEEE Transactions on Systems, Man, and Cybernetics. 2021. Vol. SMC-51, No. 1. P. 176-190.
17. Giraldo J., Urbina D., Cardenas A., Valente J., Faisal M., Ruths J., Tippenhauer N.O., Sandberg H., Candell R. A survey of physics-based attack detection in cyber-physical systems // ACM Computing Surveys. 2018. Vol. 51, No. 4. P. 76:1-76:36.
18. Peng C., Sun H., Yang M., Wang Y.-L. A survey on security communication and control for smart grids under malicious cyber attack // IEEE Transactions on Systems, Man, and Cybernetics, 2019. Vol. SMC-49, No.8. P. 1554-1569.
19. Owoputi R., Ray S. Security of Multi-Agent Cyber-Physical Systems: A Survey // IEEE Access. 2022. Vol. 10. P. 121465-121479.
20. Wangli H., Wenying X., Xiaohua G., Qing-Long H., Wenli D., Feng Q. Secure Control of Multiagent Systems Against Malicious Attacks: A Brief Survey // IEEE Transactions on Industrial Informatics. 2022. Vol. II-18, No. 6. P. 3595-3608.
21. Falliere N., Murchu L., Chien E. Stuxnet dossier // White Paper, Symantec Security Response. Ver.1.4, Mountain View, CA, USA. Feb. 2011.
22. Maitra A. K. Offensive cyber-weapons: Technical, legal, and strategic aspects // Environment Systems and Decisions. Vol. 35, No. 1, 2015. P.169-182.
23. Xu W., Hu G., Ho D.W.C., Feng Z. Distributed secure cooperative control under denial-of-service attacks from multiple adversaries // IEEE Transactions on Cybernetics. 2020. Vol. 50, No. 8. P.3458-3467.
24. Senejohnny D., Tesi P., De Persis C. A jamming-resilient algorithm for self-triggered network coordination // IEEE Transactions on Control of Network Systems, Vol. CNS-5, No. 3, 2018. P. 981-900.
25. Xu W., Ho D.W.C., Zhong J., Chen B. Event/self-triggered control for leader-following consensus over unreliable network with DoS attacks // IEEE Transactions on Neural Networks and Learning Systems. 2019. Vol. NNLS-30, No. 10. P. 3137-3149.

26. Городецкий В.И., Котенко И.В. Концептуальные основы стохастического моделирования в среде Интернет // Труды института системного анализа РАН. Фундаментальные основы информационных технологий и систем / Под ред. С.В.Емельянова. М.: URSS, 2005. Т. 9.