The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #43304

Board 292: General-Purpose Artificial Intelligence Approaches for Intelligent Tutoring

Mr. Ryan Hare, Rowan University

Ryan Hare received his B.S. in Electrical and Computer Engineering from Rowan University in 2019. He is currently pursuing his Ph.D. in Electrical and Computer Engineering at Rowan University. His current research focus is applying machine learning and gamification to create engaging and adaptive tutoring systems using games and virtual environments to improve students' educational experiences.

General-purpose Artificial Intelligence Approach for Intelligent Tutoring

Ryan Hare, Dr. Ying Tang

Abstract

The recent increase in interest in virtual education, automated educational systems, and online virtual environments such as the metaverse highlights the need for general-purpose approaches to automated student education. The presented work comprises a yearly update from our continued development of a general-purpose personalized education system. Our system is composed of a modular, general-purpose architecture for tracking and controlling student progress within any serious game. We also apply reinforcement learning agents as a brain of our system to automatically direct students through the system and, by extension, the game. To show educational merit of last year's developments, we demonstrate our system within Gridlock, a game for first-year computer engineering students.

To show educational merit and student opinion on our system, we present results from our implementation of Gridlock within Introduction to Digital Systems, a relevant course at Rowan University. For educational merit, we compare pre- and post-intervention content tests between students who did and did not interact with Gridlock, showing an improvement in students who engaged with our system. We also show student surveys taken before and after intervention to gauge student attitudes about their program as a whole, showing that students who interacted with our system seemed to show more confidence in their own ability. Finally, we analyze student actions within the game itself to show that the PING system helps them complete content sections faster and with fewer attempts.

1. Introduction

With the advent of new methods and approaches in virtual education, automated systems, and online learning, there is a need for general-purpose approaches to help automate student education. These new approaches can serve not to replace traditional education, but to augment it through automated student assistance, easier classroom operation for instructors, and better support for under-performing students [1]. In turn, the implementation of more automated systems in a classroom helps to free up instructor time and resources, and to help raise overall classroom performance.

To achieve an automated educational support system that can stand without instructor intervention, intelligent tutoring systems (ITSs) offer a valuable avenue of research [2]. These systems are well-established in the field, but have seen a surge in development in recent years due to advancements in large language models like ChatGPT [3], better artificial intelligence methods [4], wider technology adoption, and the recent boom in e-learning [5]. However, a key aspect of computer- or web-based ITSs often remains unaddressed; they are boring.

For ITSs to function properly, it is necessary to perform regular student evaluations while students engage with educational content [6]. This regular evaluation gives the system a solid idea of the student's performance, from which new decisions can be made on how best to support them. In doing so, the system can often lose student interest, or cause students to become frustrated or disengaged with the regular testing [7]. To address this, ITSs can be combined with another promising approach; serious games.

Serious games (SGs) are virtual or physical games that focus on education, training, therapy, or other non-entertainment purposes [8]. In this case, we focus on educational serious games [9]. In educational SGs, gamification allows them to be more enjoyable than traditional lessons while still delivering the same content, supporting student learning [10],[11],[12], and offering a platform for evaluations [13],[14],[15]. Thus, the combination of ITSs and SGs provides the enjoyable and engaging presentation that SGs provide, while also allowing the ITS side of the combined system to measure student performance and offer support to those students.

As a proposed method for combining ITSs and SGs, this paper presents a general-purpose architecture of tracking and controlling student progress within a serious game. With this approach, serious games are divided up based on educational content, and an ITS is integrated with the serious game to direct students to different content based on their performance, while also providing individualized support when students engage with specific content. To make our system intelligent and adaptive, we adapt reinforcement learning, a machine learning paradigm [16]. This allows our system to base decisions off of student performance data, and to iteratively learn what support students require based on interactions with prior students. We refer to our complete system as the personalized instruction and need-aware gaming (PING) system.

To demonstrate the PING system, it has been fully integrated into a serious game designed for first-year engineering students called *Gridlock* [17]. *Gridlock* instructs students on the basics of digital logic and digital circuit design, both core principles in electrical and computer engineering. We use *Gridlock* as a test bed to verify the PING system. To that end, section 2 provides the framework of the PING system, as well as a brief overview of our adapted reinforcement learning method. Section 3 provides results from in-classroom testing, including pre-post-intervention content tests and student opinion surveys. Finally, section 4 provides conclusions and future research directions.

2. Adaptive Game System Framework

Since the main purpose of the PING system is to act as standalone classroom assistance, it was important in our design that the system requires little to no instructor intervention. Furthermore, to make the system general-purpose, the decision was made to create a system that operates on segmented, subject-specific modules within a serious game. Separating subjects into modules aligns with a divide-and-conquer approach to education, where students tackle individual subjects to build up their overall knowledge. To offer an example, a game that deals with mathematics might have two different modules for addition and multiplication. Ror the PING system's implementation, we would then expect that a serious game collects student data for each module. Data collection is discussed further in a later section.

Figure 1 shows the PING system's components. As stated, we expect a certain level of data

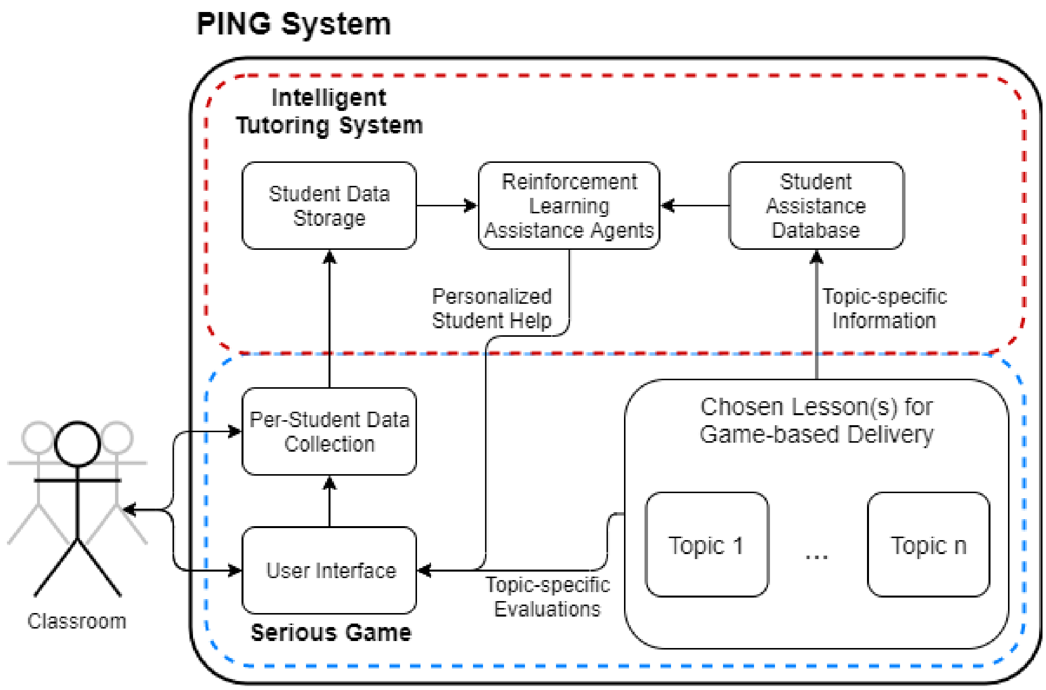

Figure 1: PING system architecture

collection from the connected serious game to inform the PING system's decision-making. Within the ITS side of the system, student data are collected and stored, and relevant topic-specific information is also initialized into the student assistance database. Student data are then passed to the reinforcement learning agents, with one agent for each topic a game teaches. The agents then pull assistance from the assistance database, deciding what assistance a student needs based on their incoming data before finally passing that assistance back to the game.

To offer a more detailed example implementation, we focus on our game, *Gridlock*. *Gridlock* contains a virtual environment where students assume the role of an engineer in a digital city. At the start of the game, students witness a traffic collision caused by a malfunctioning traffic light. Students are then tasked with learning necessary information before creating a new logic controller for the traffic light. The PING system is used to help students learn various topics, such as binary logic, logic gates, finite state machines, and basic programming in Verilog, the language used by the assignment. The PING system in *Gridlock* then contains a reinforcement learning agent for each of these topics.

When students go to learn various topics, they are given a few options. For students who prefer structured learning, the game contains written study materials, complete with figures and practice problems. Each topic also contains recorded video lectures on the same material. Finally, students can instead explore and move directly to applying their knowledge, ”learning on the job” so to speak. *Gridlock* is notable from an educational standpoint because it is a digitized form of a standard lab assignment. However, instead of simply coding their logic controller and testing it digitally, students do all coding and testing in the context of the virtual environment. This gives them visible feedback on their design in the form of a virtual traffic light.

2.1. PING System Data Requirements

As students interact with a PING-integrated game, the PING system must collect informative data on their performance. However, serious games offer a wide range of possible data collection as all data on student interactions is readily available. Games can directly assess student knowledge, or games can indirectly assess students by recording their behavior and interactions. To keep up engagement, *Gridlock* uses a balance of both assessment methods. For more technical topics, we use direct assessment, with students taking multiple choice quizzes. But for some topics, we were able to gamify our assessment into integrated mini-games; for example, one topic in *Gridlock* deals with the order that the traffic light should change in, shown in Figure 2. For this, we use an interactive simulator to let students freely explore and test their own ideas on how the light should operate.

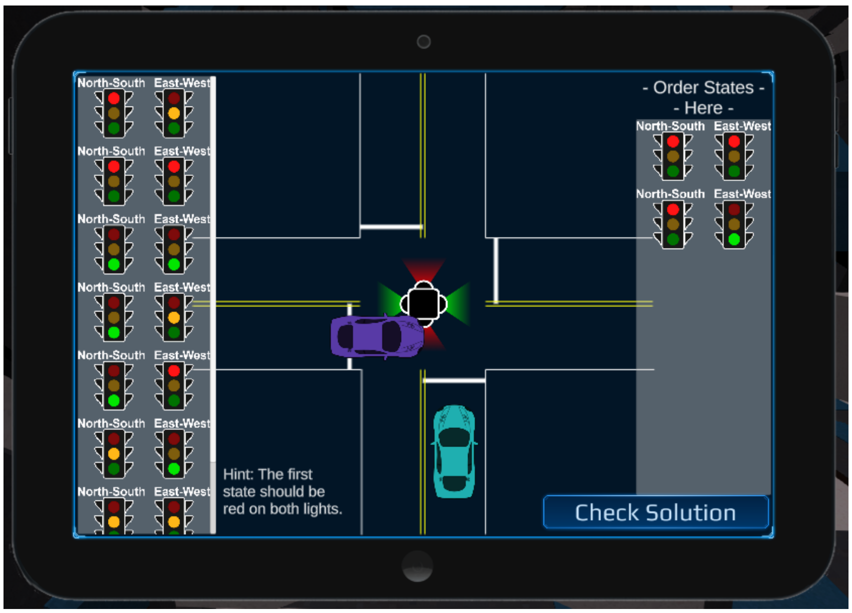

Figure 2: The in-game traffic light simulator that tests students on the correct order the light should change in.

The design of the PING system necessitates that all topics within the same game collect similarly formatted data. In our case, *Gridlock* collects direct performance data as percentage of successes (either on a quiz or in a mini-game). Students can also rate their confidence on answers, or how well they felt about their mini-game performance. We also collect timing data on how long students took to progress, as well as emotional data through webcam imaging [18]. In this case, participating students give consent before the game can access their webcam, and all data collection is IRB-approved. Finally, we collect information on key presses and mouse movements, as we've previously found that frustrated students sometimes rapidly move the mouse or press random keys.

All these numerical metrics are then fed into the PING system to inform decision-making. With the reinforcement learning, a student's updated data is passed in, and the agent for a given topic then chooses what assistance a student should be given. Then, given a chosen piece of assistance from a pool of possible decisions, the game can then pop-up some study materials, change the difficulty of the game, offer a new suggestion, or any other assistance relevant to the specific game and topic.

2.2. Reinforcement Learning

This section provides a high-level overview for the functionality of the PING system's reinforcement learning agents. We have also published a technical article for a more in-depth overview [19],[20]. Reinforcement learning is a machine learning paradigm that creates the "brain" of an agent. An agent, in this case, is responsible for any activities, and for observing the environment around it. In our case, the PING system uses multiple agents, each responsible for observing our recorded student data. By creating a specialized agent for each topic, we can have unique behavior for each topic in a given game. Then, our "environment" is the students and the assistance the game can give them.

Based on the name, reinforcement learning is learning by repeatedly reinforcing concepts. In the case of machine learning, this means trying out actions repeatedly to observe the result. Then, as more actions are tried in new situations, the agent gradually gains an understanding of how those actions influence the environment (in our case, the student). The agent also receives a numerical "reward" when it does well. So, to consider an example, if the agent chooses assistance for the student that ends up helping, the agent receives a reward. If the agent's decision negatively affects a student, the agent receives a punishment. In this way, the "brain" of the PING system can improve its own behavior by interacting with students.

3. System Evaluation

To show educational merit of the PING system, we focus on results from in-classroom testing over the previous year. Classroom testing was done in Introduction to Digital Systems; a required course for first-year electrical and computer engineering students at Rowan University. All participating students signed consent forms, and all testing was approved by Rowan University's IRB. First, we focus on pre- and post-intervention content tests between students who did and did not interact with the PING-integrated game, shown in Figure 3. As shown, students who engaged with the PING-integrated game showed greater improvement compared to students who used either the standard game or no game at all. As other educational content was kept the same, students who had concepts reinforced by the PING system's decision-making achieved a better grasp on the concepts presented, as indicated by their post-test results.

We also tested student attitude with regards to their program and personal performance before and after intervention, as shown in Figure 4. Students were tested using the questions in Table 1. As shown, some questions have a negative connotation, while some have a positive connotation. In Figure 4, negative connotation questions have been inverted to align all the data, so an increase on the plot is always considered a positive trend.

For the most part, students showed a negative trend in their responses, if any change at all. The negative trend could be due to the course's position as a first-year engineering course, which could cause student attitudes to shift as they become accustomed to their new program. However,

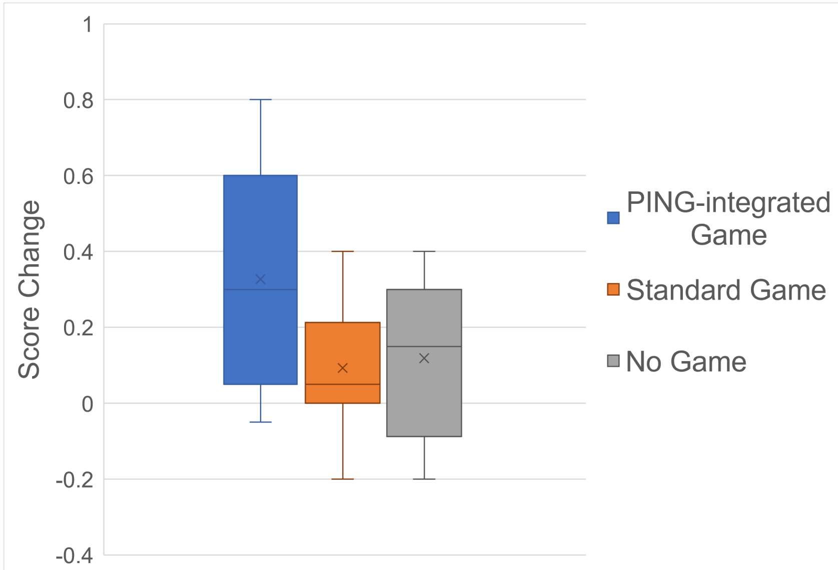

Figure 3: Average improvement in student test scores from pre-test (before intervention) to post-test (after intervention).

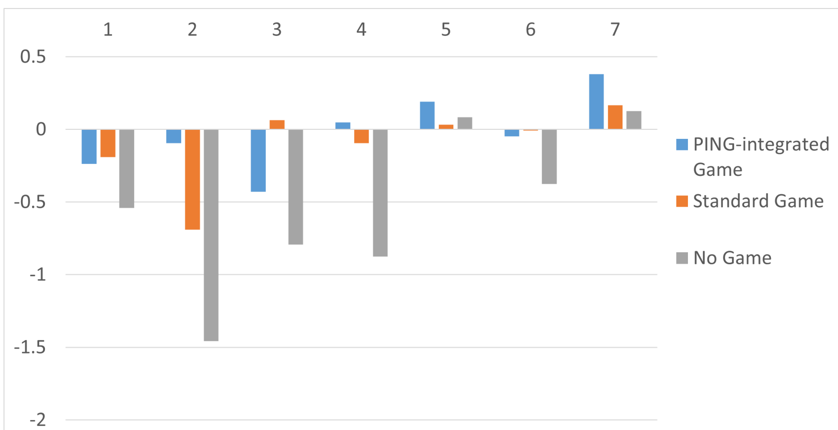

Figure 4: Average change in student attitude from mid-semester (before intervention) to end of semester (after intervention).

it can be seen on a few questions that students who did not interact with the game had worse trends than students who did. This is especially prevalent on questions 2 and 4. It is possible that students' interactions with peers while playing the game helped them feel more confident in their own abilities.

Finally, to compare the PING-integrated game with a non-adaptive game, we measured the amount of time it took students to complete the game, shown in Figure 5(a). We also measured the number of steps students took to complete the game between the two game versions, shown in Figure 5(b). In this case, steps refers broadly to attempts on quizzes, retries on mini-games, and viewing help documentation. As shown, students in the PING-integrated game took less time overall to complete the game, even though the educational content was identical between the two

Number	Question
1	I feel confident that I have the ability to do well in engineering.
2	I feel that I have less ability than others in engineering.
3	I feel comfortable in engineering.
4	Other people understand more than I do about what is going on in engineering.
5	I think in the same way as do people who do well in engineering.
6	It is a mystery to me how engineering at my university works.
7	I feel alienated from engineering at my university.

Table 1: Example path chosen by the reinforcement learning system with the student's grade for each section.

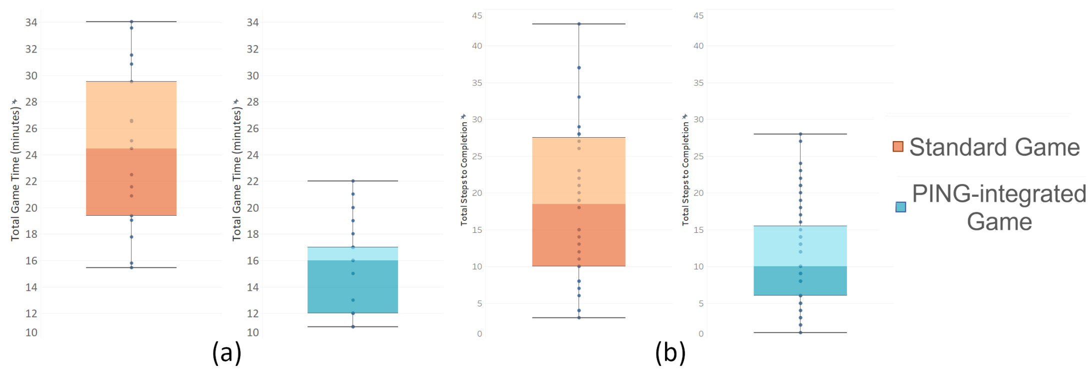

Figure 5: a. In-game data from participating students showing total completion time (left). b. In-game data from participating students showing total number of steps needed to complete the game, including attempts on quizzes, retries on mini-games, and viewing of help documentation.

versions. Additionally, students in the PING-integrated game required fewer attempts on average to clear sections of the game, likely contributing to their lower overall times.

4. Conclusion

This paper discusses our general-purpose personalized education system. Our personalized instruction and need-aware gaming (PING) system is modular, and designed for implementation within any educational serious game. Using reinforcement learning, the PING system can make live adjustments to its recommendations based on student feedback, ensuring that students always receive appropriate educational support. From our in-classroom testing, we show that students who interacted with a PING-integrated game performed better overall on pre-post-intervention content tests. Finally, we showed the effectiveness of the PING system, showing that students who played the PING-integrated game required fewer attempts and less time on average to clear the game compared to students in a standard version of the serious game. Moving forward, we intend to further standardize the PING system and, ultimately, release it in an easily-integrated form for wider use by researchers and serious game developers.

5. Acknowledgement

This work was supported in part by the National Science Foundation under Grant 1913809.

References

[1] A. Mousavi, M. Schmidt, V. Squires, and K. Wilson, "Assessing the effectiveness of student advice recommender agent (sara): the case of automated personalized feedback," *International Journal of Artificial Intelligence in Education*, vol. 31, pp. 603–621, Sep 2021.

[2] A. C. Graesser, X. Hu, and R. Sottilare, "Intelligent tutoring systems," in *International handbook of the learning sciences*, pp. 246–255, Routledge, 2018.

[3] C. K. Lo, "What is the impact of chatgpt on education? a rapid review of the literature," *Education Sciences*, vol. 13, no. 4, p. 410, 2023.

[4] F. AlShaikh and N. Hewahi, "Ai and machine learning techniques in the development of intelligent tutoring system: A review," in *2021 International Conference on innovation and Intelligence for informatics, computing, and technologies (3ICT)*, pp. 403–410, IEEE, 2021.

[5] I. A. Mastan, D. I. Sensuse, R. R. Suryono, and K. Kautsarina, "Evaluation of distance learning system (e-learning): a systematic literature review," *Jurnal Teknoinfo*, vol. 16, no. 1, pp. 132–137, 2022.

[6] E. Mousavinasab, N. Zarifsanaiey, S. R. Niakan Kalhori, M. Rakhshan, L. Keikha, and M. Ghazi Saeedi, "Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods," *Interactive Learning Environments*, vol. 29, no. 1, pp. 142–163, 2021.

[7] T.-d. Kim, M.-y. Yang, J. Bae, B.-a. Min, I. Lee, and J. Kim, "Escape from infinite freedom: Effects of constraining user freedom on the prevention of dropout in an online learning context," *Computers in Human Behavior*, vol. 66, pp. 217–231, 2017.

[8] W. Ravyse, S. Blignaut, V. Leendertz, and A. Woolner, "Success factors for serious games to enhance learning: a systematic review," *Virtual Reality*, vol. 21, 03 2017.

[9] S. De Freitas, "Are games effective learning tools? a review of educational games," *Journal of Educational Technology & Society*, vol. 21, no. 2, pp. 74–84, 2018.

[10] Y. Tang, K. Jahan, and T. Bielefeldt, "The effectiveness of an adaptive serious game for digital logic design," *2015 ASEE Annual Conference and Exposition Proceedings*, 2015.

[11] H. Dehghanzadeh, M. Farrokhnia, H. Dehghanzadeh, K. Taghipour, and O. Noroozi, "Using gamification to support learning in k-12 education: A systematic literature review," *British Journal of Educational Technology*, 2023.

[12] Y. Tang, C. Franzwa, T. Bielefeldt, K. Jahan, M. S. Saeedi-Hosseiny, N. Lamb, and S. Sun, "Sustain city," *Design, Motivation, and Frameworks in Game-Based Learning Advances in Game-Based Learning*, p. 57–91, 2017.

[13] V. J. Shute, L. Wang, S. Greiff, W. Zhao, and G. Moore, "Measuring problem solving skills via stealth assessment in an engaging video game," *Computers in Human Behavior*, vol. 63, p. 106–117, 2016.

[14] V. Shute and S. Rahimi, "Review of computer-based assessment for learning in elementary and secondary education," *Journal of Computer Assisted Learning*, vol. 33, no. 1, p. 1–19, 2017.

[15] R. Sottilare, A. Graesser, X. Hu, and K. Brawner, *Design Recommendations for Intelligent Tutoring Systems - Volume 3: Authoring Tools and Expert Modeling Techniques*. 06 2015.

[16] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[17] Y. Tang, C. Franzwa, and A. Johnson, "Ci-team demonstration: Interactive and collaborative learning environment using virtual reality games promoting metacognition for science and engineering design in context," 2013.

[18] D. Mehta, M. F. H. Siddiqui, and A. Y. Javaid, "Facial emotion recognition: A survey and real-world user experiences in mixed reality," *Sensors*, vol. 18, no. 2, 2018.

[19] R. Hare and Y. Tang, "Reinforcement learning with experience sharing for intelligent educational systems," *2023 IEEE International Conference on Systems, Man, and Cybernetics*, 2023.

[20] R. Hare and Y. Tang, "Hierarchical deep reinforcement learning with experience sharing for metaverse in education," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 4, pp. 2047–2055, 2023.